Dynamic Laser Fault Injection Aided by Quiescent Photon Emissions in Embedded Microcontrollers: Apparatus, Methodology and Attacks

by

Karim Amin

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2020

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Domenic Forte
Associate Professor, Dept. of ECE, University of Florida

Supervisor: Catherine Gebotys
Professor, Dept. of ECE, University of Waterloo

Internal Member: Manoj Sachdev
Professor, Dept. of ECE, University of Waterloo

Internal Member: Gordon Agnew
Associate Professor, Dept. of ECE, University of Waterloo

Internal-External Member: Germán Sciaini
Associate Professor, Dept. of Chemistry, University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Internet of Things (IoT) is becoming more integrated in our daily life with the increasing number of embedded electronic devices interacting together. These electronic devices are often controlled by a Micro-Controller Unit (MCU). As an example, it is estimated that today's well-equipped automobile uses more than 50 MCUs. Some MCUs contain cryptographic co-processors to enhance the security of the exchanged and stored data with a common belief that the data is secured and safe. However many MCUs have been shown to be vulnerable to Fault Injection (FI) attacks. These attacks can reveal shared secrets, firmware, and other confidential information. In addition, this extracted information obtained by attacks can lead to identification of new vulnerabilities which may scale to attacks on many devices. In general, FI on MCUs corrupt data or corrupt instructions. Although it is assumed that only authorized personnel with access to cryptographic secrets will gain access to confidential information in MCUs, attackers in specialized labs nowadays may have access to high-tech equipment which could be used to attack these MCUs. Laser Fault Injection (LFI) is gaining more of a reputation for its ability to inject local faults rather than global ones due to its precision, thus providing a greater risk of breaking security in many devices.

Although publications have generally discussed the topic of security of MCUs, attack techniques are diverse and published LFI provides few and superficial details about the used experimental setup and methodology. Furthermore, limited research has examined the combination of both LFI and Photo-Emission Microscopy (PEM), direct modification of instructions using the LFI, control of embedded processor resets using LFI, and countermeasures which simultaneously thwart other aspects including decapsulation and reverse engineering (RE).

This thesis contributes to the study of the MCUs' security by analyzing their susceptibility to LFI attacks and PEM. The proposed research aims to build a LFI bench from scratch allowing maximum control of laser parameters. In addition, a methodology for analysis of the Device Under Attack (DUA) in preparation for LFI is proposed, including frontside/backside decapsulation methods, and visualization of the structure of the DUA. Analysis of attack viability of different targets on the DUA, including One-Time Programmable (OTP) memory, Flash memory and Static Random Access Memory (SRAM) was performed. A realistic attack of a cryptographic algorithm, such as Advanced Encryption Standard (AES) using LFI was conducted. On the other hand, countermeasures to the proposed attack techniques, including decapsulation/RE, LFI and PEM, were discussed. This dissertation provides a summary for the necessary background and experimental setup to study the possibility of LFI and PEM in different DUAs of two different technologies, specifically PIC16F687 and ARM Cortex-M0 LPC1114FN28102. Attacks performed on on-chip peripherals such as Universal Asynchronous Receiver/Transmitter (UART) and debug circuity reveal new vulnerabilities. This research is important for understanding attacks in order to design countermeasures for securing future hardware.

Acknowledgements

All my gratitude is due to Allah, first and foremost; and then I would like to express my gratitude to the people who inspired and encouraged me during my PhD research at University of Waterloo. Without their support, this thesis would not have been completed.

I would like to kindly thank the thesis committee Professors Manoj Sachdev, Domenic Forte, Germán Sciaini and Gordon Agnew for peer-reviewing and providing valuable comments on my thesis. Also, I would like to acknowledge Professor Donna Strickland, from the Physics and Astronomy department at the University of Waterloo, for her enormous knowledge and experience related to using lasers, which she shared with us very openly. Her help, support and continuous advice were of crucial importance in operating and upgrading our laser setup. In addition, I'd like to thank KA Imaging Inc. for facilitating X-ray imaging for our research purposes. I'd like to thank all my lab mates and colleagues on their positive impact on getting my PhD degree, namely, Bahar, Brian, Caio, Edgar, Haohao, Mahmoud and Mustafa. Thanks to all my friends near by and far away who kept me believing in myself and in my ability to acquire the PhD degree. I would like to express my deep gratitude and sincere thanks to my supervisor Prof. Catherine Gebotys for her patient guidance and enthusiastic support of this research work. With her insight based on industrial and theoretical background, my research could be guided in the right way. She was always so helpful and inspiring. She always met me with a smile in our regular meetings even through the toughest times and obstacles we faced while conducting the research work. She always tried to make my life easier, provided support whenever needed and gave me confidence in myself. Her sublime attention to detail along with having the bigger picture in mind for our research objectives always inspired me. Without her generous support, this thesis wouldn't have been possible.

Thanks to the University of Waterloo for immersing me in such learning environment through out those many years of my PhD journey. Also, thanks to Professor Serag El-Din Habib, Yehea Ismail, Alaa El-Rouby and Amr Wassal who guided me through my BSc and MSc back in Egypt and taught me how to keep learning, be more productive and develop on the personal level. Moreover, joining the squash varsity team enhanced my attention to detail, time management, consistency and research productivity.

Special thanks to my parents and family. All of my accomplishments through my life have been realized by endless love and support of my mother Mona and sister Mariam. My mother showed me her positive attitude towards life specially through difficult times. Also, I'd like thank also my wife Safy for being supportive. Her love and friendship are great gifts in my life. My love towards my kids, Malik and Adam, supported me during hard times in my PhD journey.

Dedication

This PhD thesis is dedicated to my mom Mona.

Table of Contents

List of Figures

List of Tables

List of Code Listings

List of Acronyms

I_{reg} instruction register

W_{reg} working register

ADU Analog-Digital Unit

AES Advanced Encryption Standard

AHB Advanced High-Performance Bus

ALU Arithmetic Logic Unit

APD avalanche photodiode

ASIC Application-Specific Integrated Circuit

BGA Ball Grid Array

CaLIAD calibratable lightweight invasive attack detector

CCD charge-coupled device

CNC Computer Numeric Control

CPA correlation power analysis

CPLD complex programmable logic device

DAP Debug Access Port

DIP dual in-line package

DLL dynamic-link library

DMC Design and manufacturing company

DND Department of National Defence

DOF Depth of Field

DUA Device Under Attack

ECU Electronic Control Unit

EEPROM Electrically Erasable Programmable Read-Only Memory

EM Electro-Magnetic

EMFI electro-magnetic fault injection

EOFM electro-optical frequency modulation

EOP electro-optical probing

EPROM erasable programmable read only memory

FD-SOI Fully Depleted Silicon on Insulator

FF flip-flop

FI Fault Injection

FIB Focused Ion Beam

FinFET fin field-effect transistor

FPGA Field-Programmable Gate Array

HW hardware

IC Integrated Circuit

IoT Internet of Things

IR Infrared

IS instruction skip

ISR Interrupt Service Routine

JTAG Joint Test Action Group

KPA known-plaintext attack

LAPD low area probing detector

LED light-emitting diode

LFI Laser Fault Injection

LIVA Light Induced Voltage Alteration

LSB Least Significant Bit

LUT look-up table

LVP laser voltage probing

MCU Micro-Controller Unit

MDK Microcontroller Development Kit

MOSFET Metal-Oxide-Semiconductor Field-Effect Transistor

MSB Most Significant Bit

NA Numerical Aperture

NaN Not-A-Number

Nd:YAG Neodymium: Yttrium Aluminum Garnet

Nd:YLF Neodymium: Yttrium Lithium Fluoride

NIR Near Infrared

NMI nonmaskable interrupt

NOP no operation

NUV Near Ultraviolet

NVIC Nested Vectored Interrupt Controller

OBIC Optical Beam Induced Current

OTP One-Time Programmable

PAD probe attempt detector

PCB Printed circuit board

PEA photon emission analysis

PEM Photo-Emission Microscopy

PIN Personal Identification Number

PKC public-key cryptography

PLL phase-locked loop

PUF physical unclonable function

QFP Quad Flat Package

RE reverse engineering

RISC reduced instruction set computer

RO ring oscillator

ROM read-only memory

S-box substitution box

SCA side-channel attack

SCM scanning capacitance microscope

SEE Single Event Error

SEL Single Event Latch-up

SEM scanning electron microscope

SET Single Event Transient

SEU Single Event Upset

SOI Silicon on Insulator

SRAM Static Random Access Memory

STI shallow trench isolation

SWD Serial Wire Debug

SWIR short-wave Infrared

TEM Transverse Electric and Magnetic

TSOP Thin Small Outline Package

UART Universal Asynchronous Receiver/Transmitter

UV Ultraviolet

UV-C Ultraviolet light in the C spectrum

VMU Video Microscope Unit

Chapter 1

Introduction

The wide integration of embedded devices into almost every aspect in our lives is forcing manufacturers of these devices to design secure products. Contactless credit cards, e-passports, automobiles, etc., are just few examples of the everyday " things" that have embedded electronics built inside. Ensuring the security of such devices is of crucial importance, especially with different types of known popular attacks which can exploit any ignored security gaps to conduct a security breach. For example, various security threats and protection mechanisms in the automotive industry are mentioned in [15]. The implementation of cryptographic algorithms could be attacked on known side channels [16], for example analyzing the device's power which unintentionally leaks secret information during the execution of the security algorithm. Injecting faults in electronic circuits through Electro-Magnetic (EM) waves or laser for example as in [17] and [18] may also leak secret information and thus demands higher levels of physical security. Also, reverse engineering (RE) [19], which is defined as the processes of extracting knowledge or design information from anything man-made and reproducing it or reproducing anything based on the extracted information, is a known methodology in de-assembling electronic circuits. For example, RE may leak details on contents of secure memories or security implementations.

It's worth noting that it's highly likely any embedded system may be reverse engineered with variation in the time consumed, expenses incurred and expertise required to do so. The security measures taken while designing each system determine the costs required for RE that system. Also, advancement in technology leading to reduced size of on-chip transistors and increased device complexity are challenges facing RE. Different classifications have been proposed for the types of the RE processes that can be followed [20]. Some classifications are based on the chip type; for example, as shown in Figure 1.1, Application-Specific Integrated Circuits (ASICs) and custom Integrated Circuits (ICs) require structural RE, smart cards and MCUs need both structural and program-code RE while structural and bit-stream RE are used to deal with complex programmable logic devices (CPLDs) and FPGAs. Structural RE [20] means extracting all the information about the internal structure of a chip such as locations of transistors and interconnections in-between. This process is very tedious, time-consuming and requires a certain level of experience. Program-code RE targets retrieving the memory contents to understand how the device works. Bit-stream RE is based on converting the acquired bit-stream to logic gates and equations that describe the functionality. Companies like *TechInsights* [21] provide structural RE, while *Bottomline* Technologies [22] provides bit-stream RE.

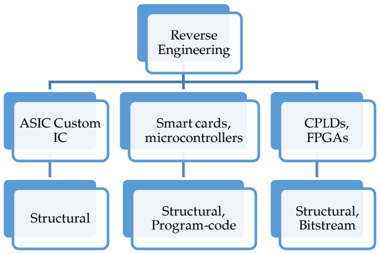


Figure 1.1: RE based on chip type.

RE is becoming a niche area for business since it can be used as a very strong aiding factor to reduce time-to-market because RE basically serves as a method to understand the function of a certain system and provides the ability to replicate a similar or better one.

1.1 Motivation

Many modern electronic devices such as computers, cell phones, smart cards and Electronic Control Units (ECUs) in automobiles have sensitive information stored inside which may need to be communicated securely. That need makes security a serious design concern for those devices. A lot of advancement in the field of encryption has been made. On the other hand, confidential information which can be encrypted securely using software algorithms could be easily bypassed through hardware attacks, for example like bit or row hammering [23, 24]. Physical access to the chip under attack increases the probability of breaking the security of an embedded system instead of trying to hack the chip remotely. Side channel analysis through EM emanations and power measurements could reveal important information about cryptographic activities inside the DUA. EM fault injection is another technique used to introduce global faults inside the DUA. Also, voltage variations and clock glitching could alter the normal behavior of the running DUA and lead to unauthorized access to confidential data.

During the beginning of this century, LFI was introduced as an accurate method for injecting local faults inside the DUA. Being more precise in injecting faults increases the ability of extracting secret data from inside the DUA. LFI requires the attacker to have access to the frontside and/or the backside of the chip. So LFI needs the packaged chip to become decapsulated. Mechanical milling, drilling, polishing and chemical etching are possible ways to have a decapsulated chip. The process of decapsulating the chip from the frontside/backside then identifying different structures under a microscope is called structural RE. Full RE by removing one layer after the other using suitable chemicals and imaging each layer before removal isn't needed when LFI will be used to introduce faults in the DUA. Large structures like SRAM and flash EEPROM (memories in general) are easily identifiable once you put the DUA under a microscope unit. LFI could then be used in theory to target bytes down to a single bit depending on different parameters such as the laser beam spot size, energy and duration of exposure of the DUA to the laser.

1.2 Objectives

This thesis has objectives that range from IC preparation, through to laser fault injection equipment setup and to laser fault injection analysis. The objective of analyzing the laser fault injection results is to better understand how to design countermeasures and thus create more secure embedded systems. This thesis aims to reveal the details of building a low cost LFI setup as well as a low cost methodology for both frontside/backside decapsulation. Also, this research analyzes various LFI attacks on two DUAs with different technology nodes, processor cores and design complexities. Combining static PEM emissions with LFI to provide fine tuning of faulty instructions in addition to reverse engineering the target clock cycle for successful LFI is another goal of this study; thus having a more controlled instruction replacement environment. This thesis intends to provide a practical setup to monitor the active laser pulse and laser induced current. This research tries to investigate real-life attacks such as an attack on AES encryption algorithm to reveal the secret key, attacking peripherals such as UART to inject faults in data sent off chip or disabling the code copy protection.

1.3 Thesis Overview

The remainder of this thesis is organized as follows:

Chapter 2 presents an introduction to the theory behind the laser operation and how it can be used to inject faults in ICs.

Chapter 3 offers a review for the published literature with respect to the conducted attacks, proposed countermeasures, equipment, targeted devices and the followed methodology starting from sample preparation for FI to successful FI.

Chapter 4 describes the proposed research methodology and the assembled setup for both LFI and PEM. Also this chapter demonstrates the custom-designed circuitry for triggering the LFI for ARM Cortex-M0 LPC1114FN28102 and gives a brief overview of the DUAs attacked during this research.

Chapter 5 lists the results achieved through applying the proposed methodology at different stages of the LFI process on the embedded MCU PIC16F687. Different attacks were conducted on PIC16F687 such as instruction skip/replacement, skipping configuration bits programming and retrieving the AES-128 key. In addition, PEM analysis was performed to read out the decoded instructions coming out of the flash memory.

Chapter 6 capitalizes on the knowledge acquired from attacking the PIC1F687 and applies it on a more complex MCU which is the ARM Cortex-M0 LPC1114FN28102. Structures other than memories were identified such as UART and debug circuitry. Instruction manipulation was also investigated.

Chapter 7 summarizes what has been achieved, lists the contributions to the state-of-the-art research, points out the limitations the assembled LFI and PEM setups. Possible future directions and ideas are proposed.

Chapter 2

Background

Let's begin by differentiating between a defect, fault and error. A defect is usually caused during the manufacturing process or through aging (i.e. physical) while a fault is caused through electrical manifestation. Therefore, an error could be defined as the manifestation of a fault [25, 26]. In the following section, a brief overview of different classes of fault injection attacks is given, then the sections to follow will focus on LFI.

2.1 Types of Attacks

Attacks can be categorized with respect to different metrics. According to [27], attacks could be classified into three main categories described next, specifically non-invasive, invasive, and semi-invasive.

2.1.1 Non-invasive

In these types of attacks the package of the chip is left unaltered and the chip is attacked while in operation. Examples of such types of attacks are mentioned below:

- Voltage glitching:

  Under-voltage and over-voltage attacks can disable or bypass protection circuits or even force incorrect execution of instructions by a processor. Therefore, some security processors have detection mechanisms for voltage variations.

- Clock glitching:

  Circuits that detect variation in voltage usually can't react to fast transients, so clock transients can be used to affect the decode and execution stages in some processors.

- Side channel analysis:

  Power, EM, or other measurements and analysis can lead to detecting different activities performed by a processor and hence attacking the target processor at specific locations during those activities. For example, SRAM write operations often generate strong signals when changing a single bit from '0' to '1' or vice versa.

- Test interface:

  Some MCUs and smart cards have a factory-test interface that provides access to on-chip memory and allows the manufacturer to test those devices. Normally, information on those test circuits is kept secret by the manufacturer, but the attacker can exploit different voltage levels to put the chip into test mode. Such circuits are usually destroyed in smart cards by the manufacturer after testing and before shipping the chips. Attacks, which exploit the testing interface in MCUs, usually have to bypass some sort of authentication in order to allow downloading from internal memory for example.

2.1.2 Invasive

These types of attacks start with the removal of the chip package. Once the die is exposed, direct contact or interface to circuit are possible. Also, altering the circuit directly through bulk silicon or metallization layers are other options. Examples of such types of attacks are mentioned below:

- Microprobing:

  Microprobing is a very important tool in both failure analysis and RE of chips [27]. Smaller feature size, multi-level metallization and polysilicon make microprobing more challenging. To establish probe contact, laser-cutting system and Focused Ion Beam (FIB) workstation are used for less sophisticated technology nodes and more sophisticated technologies less than $0.5\mu m$, respectively. A laser cutting system can be used to make precise cuts in the passivation layer for the sake of microprobing without short circuits. After establishing probe contact via the removal of the passivation layer, microprobing can be used for system-level functional analysis to understand the operation through injecting stimulus and observing reactions. Moreover, microprobing can be used for eavesdropping on on-chip signals and extraction of secret keys and memory contents. In [28], a technique is proposed to maintain the keys in the execution unit instead of a memory element to resist probing attacks of internal bus and memory locations. In addition, the vibrations of the probing needle can be used to remove the passivation layer for old technology nodes. A proper setup for a microprobing station will have a stable stage to hold a microscope (with 3 to 4 objective lens if it's optical), a device test socket connected to a computer, micromanipulators (basically micro-positioners with sub-micron resolution steps), and 6 to 8 probe tips which may be active for probing unbuffered internal signals or passive for eavesdropping, injecting signals and probing buffered signals.

- Short-circuiting or opening wires in metalization

  In the range of $nm/\mu m$, modifying a chip is often referred as performing an IC microsurgery. It's usually done after gaining some knowledge about the chip to know the right point for attacking. The tools used for chip modification are microprobe tip for less sophisticated attacks and larger technology nodes, FIB workstation for more sophisticated attacks, or laser-cutting. A microsurgery might be performed to modify, bypass or disable hardware locks either by cutting internal wires or completely destroying the module. For example, FIB can be used for establishing probe contact through probe (or test) point creation, imaging (down to $5nm$ resolution [28]), repairing and milling holes. In the laser-cutting system, the Near Ultraviolet (NUV) can be used to remove the passivation layer then green laser to cut the metal wires and NIR to cut through the top metal layer to access wires in the second metal layer.

2.1.3 Semi-invasive

These types of attacks need to have direct access to the silicon die and the ability to accurately target individual circuits so the chip package is removed. However, the circuit remains unaffected. The passivation layer remains the same, so the expensive microprobing tools aren't used. Examples of such types of attacks are mentioned below:

- Ultraviolet (UV) light can be used to disable security fuses in EEPROM or clear configuration bits in MCUs. Modern MCUs are less susceptible to this attack as they are designed with the UV attack in mind.

- Advanced frontside and backside imaging techniques such as laser scanning, IR and thermo-imaging can be considered as semi-invasive attacks.

- A powerful and accurate focused light spot is a simple method to induce alterations in the behavior of one or more logic gates of a circuit. For obtaining an accurately focused light beam from a camera flash, the use of a precision microscope is required. Otherwise, the use of low-quality lenses results in the diffraction of the light beam [18]. In [27, 29], successful targeting of the SRAM cells using this method was reported. The attack caused a bit-flip on the targeted SRAM cell of a MCU. In [30], by using a microscope, a modified camera flash and a computer, an effective attack against a cryptographic algorithm on an 8-bit smart card is reported. The attack was done on a decapsulated MCU running an embedded AES algorithm. Furthermore, the width of the gate dielectric in current fabrication technologies is more than 10 times smaller than the shortest wavelength of visible light. So, theoretically, it is no longer possible to hit a single SRAM cell on an up-to-date circuit with current etching technology.

- LFI in a decapsulated chip is another way to generate faults. Using a laser beam allows an opponent to target more precisely a small circuit area. It can be considered to be the most straightforward refinement of the previous technique [18]. The fault model is also similar to the previous one with more feasibility of creating faults and reproducibility.

Compared to non-invasive attacks, semi-invasive attacks are harder to implement as they require decapsulation of the chip. However, significantly less expensive equipment is needed than for invasive attacks. Semi-invasive attacks can be performed in a reasonably short period of time. Also they are scalable to a certain extent, and the skills and knowledge required to perform them can be easily and quickly acquired. Some of these attacks, such as an exhaustive search for a security fuse, can be automated. If compared to invasive attacks, the semi-invasive ones do not normally require precise positioning for success because they are normally applied to a whole transistor or even a group of transistors rather than to a single wire inside the chip. Other taxonomies of attacks and their corresponding countermeasures are presented in [31].

Section 2.2 will briefly discuss some laser-relevant definitions in optics and the theory behind the practical usage of lasers in injecting faults in ICs [32].

2.2 Laser Theory and Operation

Laser (Light Amplification by Stimulated Emission of Radiation) is a stimulated-emission EM radiation in the visible or the invisible domain. Laser is monochromatic, unidirectional, coherent and artificial (i.e. laser does not spontaneously exist in nature). Laser can be generated as a beam of very small diameter (a few μm). The beam can pass through various material obstacles before impacting a target during a very short duration.

A laser light can be produced in different wavelengths such as UV ($100 \sim 400nm$), visible colors ($400 \sim 700nm$) and NIR ($700 \sim 1400nm$) and IR domains ($1400nm \sim 1mm$).

2.2.1 Laser beams - Gaussian TEM

The Transverse Electric and Magnetic (TEM) mode means that both the electric field and the magnetic field (which are always perpendicular to one another in free space) are transverse (or perpendicular) to the direction of propagation. Different TEM modes are shown in Figure 2.1. The fundamental mode of TEM of a laser resonator has the same form as a Gaussian beam. Thus Gaussian beam propagation applies to most lasers, see Figure 2.3b. The Gaussian beam is unlike light coming from a light bulb or sunlight which are uniform (or has uniform intensity at any distance z). Furthermore, unlike light coming from a light bulb or sunlight, the laser wave has predictable phase and amplitude along its path. The total power is the same at all cross sections of the beam. Gaussian beams do not diverge linearly as do light beams.

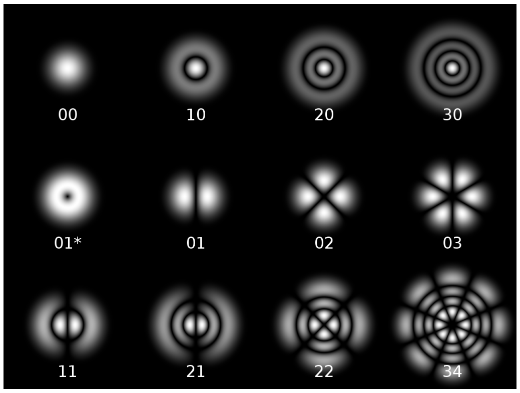


Figure 2.1: TEM modes [1].

For a time-harmonic wave of frequency, Maxwell's equations in free space are referred to as the Helmholtz equations, a time-independent form of the wave equation. Using a paraxial approximation of Helmholtz equations, one can describe the propagation of EM waves in the form of Gaussian beams. In particular, the paraxial approximation assumes the variation of the propagation is slow on the scale of wavelength and is slow on the scale of transverse extent of the wave. The solution to the Helmholtz equations is complex and involves Bessel functions, but leads to the important equations 2.1-2.4 in [3,33] describing the laser beam propagation in free space where $w(z)$ is the distance from the beam axis where the intensity drops to $\frac{1}{e^2}$ of its peak of the beam at location z; z_R is some constant to be defined later; w_o is the beam waist radius (or the minimum spot size) defined at $z = 0$; $R(z)$ is the beam wave front curvature radius at z; $\phi(z)$ is the Guoy phase shift ; and $I(r, z)$ is the beam intensity, in cross sectional plane with radial coordinate r.

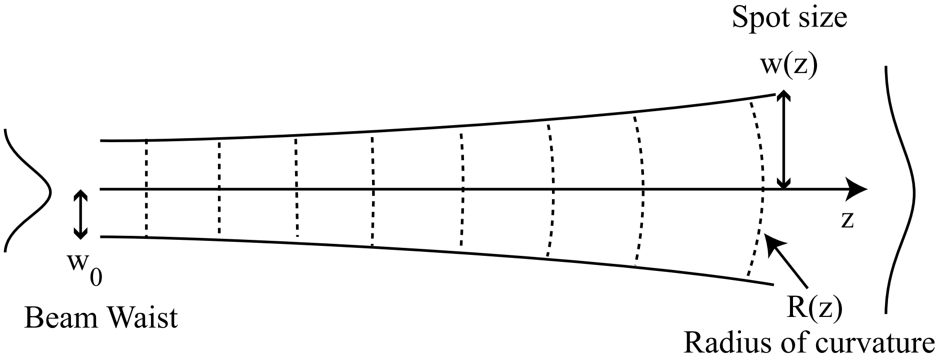


Figure 2.2: Illustration of the meaning of w_o, $w(z)$ and $R(z)$ [2].

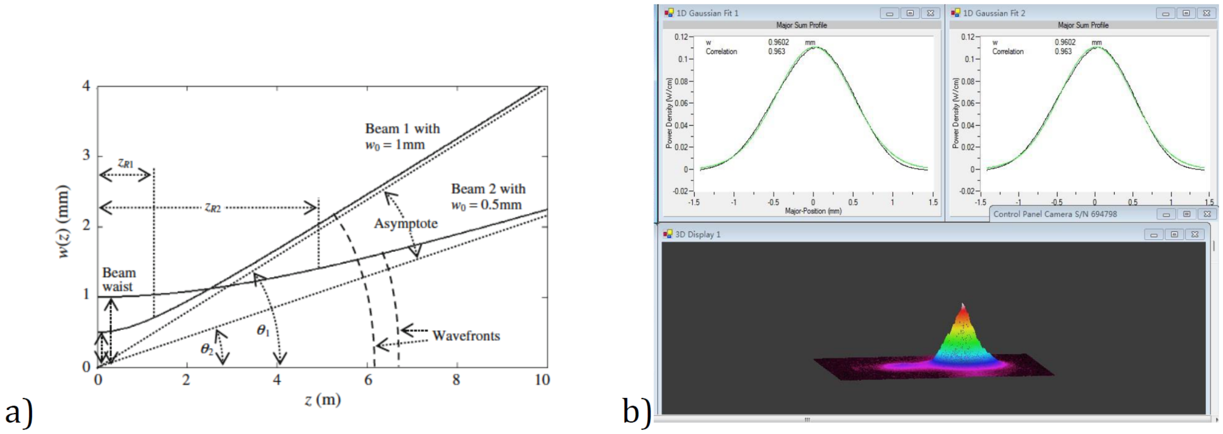
$$w(z) = w_o\sqrt{1 + (\frac{z}{z_R})^2} \tag{2.1}$$

$$R(z) = z\Big[1 + (\frac{z_R}{z})^2\Big] \tag{2.2}$$

$$\phi(z) = \arctan\frac{z}{z_R} \tag{2.3}$$

$$I(r, z) = I_o(z)e^{\frac{-2r^2}{w(z)^2}} \tag{2.4}$$

The parameter z_R is the Rayleigh range, where $w(z_R) = w_o\sqrt{2}$, from equation 2.1. The Rayleigh range defines the distance over which the laser's beam radius spreads by a factor of square root of 2. It is used as the separation between near field and far field analysis. The Rayleigh range can be further defined by equation 2.5 [3,33] which incorporates the wavelength, λ; index of refraction, n (ratio of velocity of light in free space to velocity of light in material through which it travels, where $n = 1$ for air); and M^2 parameter



w 0.9602 mm
Correlation 0.963

w 0.9602 mm
Correlation 0.963

Figure 2.3: Half angle divergence and Gaussian beams from [3], notice that values of w_o shown in labels in a) should be reversed and Gaussian beam intensity of laboratory laser in b).

(indicating the closeness of the beam to ideal Gaussian beam, ideally $M^2 = 1$, otherwise $M^2 > 1$).

$$z_R = \frac{n}{M^2}\frac{\pi w_o^2}{\lambda} \tag{2.5}$$

Note that the beam is totally characterized by the wavelength λ, M^2 and the beam waist radius w_o or Rayleigh range z_R. Near the beam waist, $z \ll z_R$, (and also at $z = \infty$), the wave is like a planar wave with no curvature, unlike $z \gg z_R$ or $z \sim z_R$ where there is curvature.

A Gaussian beam along its direction of propagation develops a phase shift, referred to as the Guoy phase shift. This is unlike a planar wave with the same optical frequency. For single wavelength lasers, typically TEM_{00} is assumed, however other TEM modes (TEM_{mp} where m, p parameters are dependent upon the Guoy phase shift) may be present in other applications such as lasers in multimode fiber optic cables. It is interesting to note that real lasers typically have some higher order modes and do not function exactly as the theoretical equations predict. In addition, scattering due to surface defects is typically more of a concern for lasers.

The divergence is a measure of how quickly the laser beam expands far from the beam waist. If a laser beam has low divergence it is referred to as being collimated. The divergence is defined by the change in beam diameter divided by the horizontal distance over which it travels. Near the laser the divergence angle is extremely small, whereas the far field divergence is typically measured at $10z_R$. Typically, far field distance for laser beams occurs meters from the laser. One can approximate the divergence of the beam in the far field using the relationship $\frac{z}{z_R} \gg 1$. Using the far field definition, equation 2.1 becomes equation 2.6 [3, 33]; and equation 2.7 [3, 33] illustrates the half angle divergence formulation using the far field assumption of equation 2.6.

$$w(z) \approx w_o \sqrt{\left[\frac{z}{z_R}\right]^2} = w_o \frac{z}{z_R} \tag{2.6}$$

$$\theta = \frac{w(z)}{z} \approx \frac{w_o}{z_R} = \frac{M^2}{n} \frac{\lambda}{\pi w_o} \tag{2.7}$$

For example, a $1064nm$ laser with a $1mm$ beam radius at beam width and a z_R of $3m$ in air, has a half angle divergence of approximately $0.33mrads$ which is 0.019°(multiply by 57°or $\frac{1}{2\pi}$), thus it can be considered collimated in a lab setup. In comparison the laboratory laser (CNI) has a calculated half angle divergence of $0.501mrads$ or 0.028°(using equation 2.7 with $M^2 = 1.38, n = 1, w_o = \frac{1.865}{2}mm$ from technical specifications), thus it is also assumed collimated.

One can rewrite the divergence as a full angle divergence θ_{00}, denoted by subscript 00, and utilize the beam waist diameter d_{00}, as shown in equation 2.8 [3, 33]. Two laser beams in Figure 2.3a [3] illustrate different divergences. Note that the larger beam waist has lower divergence for the same fixed z.

$$\theta_{00} \approx \frac{M^2}{n} \frac{4\lambda}{\pi d_{00}} \tag{2.8}$$

The Neodymium: Yttrium Aluminum Garnet (Nd:YAG) $1064nm$ laser and CO_2 $9.2 - 10.8\mu m$ lasers, continuous wave and pulsed, are widely used for research in semiconductor effects [34]. In particular, 1.06 and 1.08 μm pulsed lasers have penetration depths of about $700\mu m$ with almost uniform ionization of active regions through the depth (important for backside attacks) unlike non-uniform ionizations of lower wavelengths.

2.2.2 Optics and Laser beam spot size

It is important to note that all light, EM, particle beams, etc., exhibit the effects of diffraction. Highly focused laser beams used to create a minimum spot size are limited by diffraction. The highly focused laser beam will have far field distances from its waist (or focused spot) of only a few mm's or less. The diameter of a focused spot, d_{00}, is shown in equation 2.9 [3, 33], located at distance f (focal length) from the (thin) lens with beam width D_{00} entering the lens, as shown in Figure 2.4a) where $f \approx D_{00}$, the diameter of the spot size is approximately equal to the wavelength. Figure 2.4a) illustrates the focused spot size.

$$D_{00} \approx \frac{M^2}{n} \frac{4\lambda}{\pi d_{00}} f \tag{2.9}$$

Note that equation 2.9 illustrates that a wider beam D_{00} can be focused to a smaller spot size d_{00}, due to lower divergence.

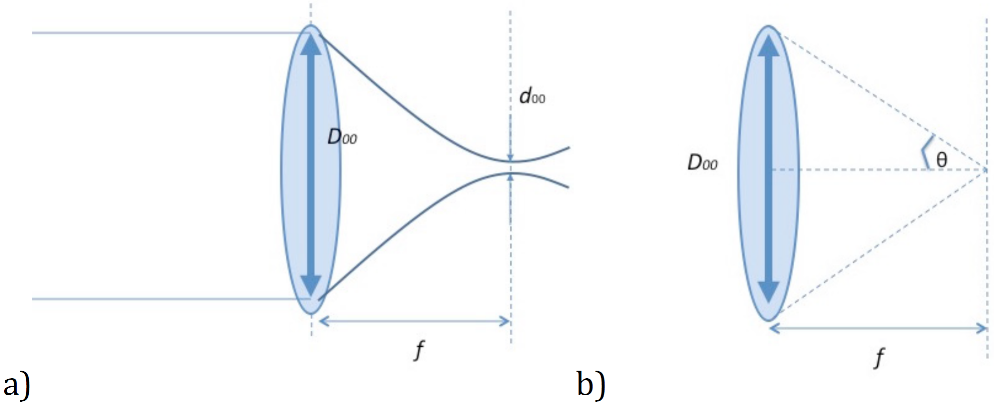


Figure 2.4: Focusing with thin lens to small spot size in a) and triangle of thin lens in b).

For example, consider using a 1064nm laser with beam width $D_{00} = 1.86mm$ (13.5% of peak) using Mitutoyo 50X objective (with focal length $f = 4mm$), the minimum spot size diameter can be calculated as $d_{00} \approx \frac{1^2}{1} \frac{4(1064nm)}{\pi(1.86mm)}(4mm) = 2.9\mu m$. Now using optics to expand the beam by 2 times[1] before focusing it into a small spot, the calculation of spot size produces a smaller spot size of $\frac{2.9}{2} = 1.45\mu m$. In addition, the beam width at the output of the beam expander must be less than the aperture of the focusing lens and alignment is important.

Equation 2.9 can be further illustrated using fiber coupling example. In order to couple a laser into a fiber, typically the beam is expanded with one lens (called a beam expander)

[1] Beam expander which can be screwed into laboratory laser head can be found at http://www.master-laser.cn/YAG%201064%20Beam%20Expander.html for 1064nm Nd:YAG lasers.

and then focused with a second in order to obtain a very small spot size equal to the fiber core diameter. The beam expander increases the beam width, D_{00}, so that it may be focused to a smaller spot size, d_{00}. Typically, this fiber coupler or fiber port is also often referred to as an up-collimator [35]. Additionally, it is critical for high energy pulses that the small spot size is exactly aligned into the fiber otherwise the result may be catastrophic (meaning the fiber may be destroyed). Alignment is usually performed with several mirrors and a pyroelectric sensor which measures laser energy. Typically, single wavelength lasers with fundamental TEM modes would utilize a single mode optical fiber cable (not multi-mode) in which case there is less loss of energy. Generally if the fiber has a core diameter more than the wavelength, there is the possibility of internal reflections causing a range of beam angles and velocities and thus causing multi-mode behavior ($TEM_{m,p}$), which has less peak power and wider beam diameter exiting the fiber. There is also a critical angle requirement which the entering laser beam must meet, otherwise there will be significant loss of energy or catastrophic results.

In microscopy applications, the f number, $f^{\#}$, is more frequently referred to rather than divergence. It's defined here by the input beam diameter in equation 2.10 [33] and provides another equation for spot size shown in equation 2.11.

$$f^{\#} = \frac{f}{D_{00}} \tag{2.10}$$

$$d_{00} \approx \frac{M^2}{n}\frac{4\lambda}{\pi} f^{\#} = 1.27\lambda f^{\#} \tag{2.11}$$

Numerical Aperture (NA) is another term frequently used for microscopy, light and imaging applications. It is defined in equation 2.12 as the acceptance angle of the objective, also indicating the resolving power of the lens. Angle θ is shown in Figure 2.4b). The NA may also be used to refer to lasers where if θ is the full angle, then the NA for $\frac{1}{e}$ drop is $n \sin\theta$. The spot size can also be formulated using NA. The Rayleigh (most commonly used), Abbe and Sparrow criterion are shown in equations 2.13 [36] respectively, for a circular aperture uniformly illuminated (light), producing an Airy disk (due to diffraction) with diameter d_{00} . For $NA = 1$, which is typical for air, the spot size is approximately equal to the wavelength. In modern optics typically $NA = 1.4$, hence the spot size is roughly proportional to half of wavelength. Since lasers are Gaussian and not uniform, the equations in 2.13 do not apply. By substituting the NA approximation of equation 2.12 [37] into equation 2.9, the laser spot size, or diameter d_{00}, becomes $0.637\frac{\lambda}{NA}$, performing slightly better. Note also that Depth of Field (DOF) in equation 2.14 [37], is proportional to the inverse of NA squared.

$$NA = n\sin\theta = n\sin\left(\arctan\left(\frac{D_{00}}{2f}\right)\right) \approx n\frac{D_{00}}{2f} \tag{2.12}$$

$$d_{00} \approx \frac{1.22\lambda}{NA}, \frac{\lambda}{NA}, \frac{0.94\lambda}{NA} \tag{2.13}$$

$$DOF = 2.5\lambda\left(\frac{f}{D_{00}}\right)^2 \tag{2.14}$$

In general lasers focused to small spots will spread out quickly as it moves away from the focus. In contrast, wider laser beams stay roughly the same size over long distances. Recent research has identified that significantly smaller spot sizes, are possible at regions within nm's away from the beam waist, such as [38].

An objective is typically preferred over a lens since it has less spherical aberration. Chromatic aberrations, where the focal length depends on the index of refraction which varies as a function of wavelength, are typically not an issue for single wavelength lasers. Achromatic or apochromatic lenses are corrected to bring more than one wavelength to focus at the same spot. Often the focal length for an objective is quoted for a specific wavelength. However the lens maker's equation, defined in equation 2.15 [39] (where r_1 and r_2 define the positive value of radius of curvature of each face of a converging thin lens), illustrates that the focal length is inversely proportional to the index of refraction minus 1.

For example, the Mitutoyo $5X$ objective has a specified focal length of $40mm$ for a wavelength of $500nm$ (Mitutoyo). Using the indexes of refraction of 2.34 and 2.23 for $500nm$ and $1064nm$ wavelengths respectively from the graph 6.6 in [40], one can correct this focal length for $1064nm$ wavelength. Thus the focal length of $40mm$ specified for a $500nm$ wavelength becomes approximately $43mm$ $\left(40\left[\frac{(2.34-1)}{(2.23-1)}\right]\right)$ for $1064nm$ wavelength. Similarly, in the previous example, the spot sizes of $2.9\mu m/1.45\mu m$ would be corrected to $3.12\mu m/1.56\mu m$ using the $50X$ objective with depth of fields $14.2\mu m/3.5\mu m$, respectively.

$$\frac{1}{f} = (n-1)(\frac{1}{r_1} + \frac{1}{r_2}) \tag{2.15}$$

2.2.3 Laser absorption and fault mechanisms

Whereas the laser beam and spot size equations utilize wave theory, the laser absorption into materials typically uses the quantum aspect where the particles are referred to as photons. Photons of the laser beam just inside the surface of the medium will couple to electronic or vibrational states depending on the energy of the photon. Photons absorbed into the medium, have energies near or greater than the medium's bandgap. The absorption refers to electronic coupling of the photon (or excited electronic states), specifically in semiconductors, the transition of electrons from valence to conduction bands (creation of electron-hole pairs) or within bands (inter-sub-band transitions). These electronic excited states may also then transfer energy to lattice phonons and thermalize within the medium. If the excitation rate is low compared to the thermalization rate, one may consider the absorbed laser energy being directly transformed into heat. If the photon energy is significantly less than the bandgap, no absorption will take place unless there are impurity or defect states to couple to [41]. The electron-hole pairs generated by absorbed photons will drift to create a transient current. This transient current changes the characteristics of the silicon structures, specifically delays, thus possibly causing faults to be injected. Typically pulsed lasers (on the order of ns pulses) are referred to as transient response (or dose-rate effects). Alternatively, if the duration of the laser irradiation is in the order of microseconds, e.g. using continuous wave (cw) lasers, lattice vibrations may dominate (equivalently absorption and creation of phonons) causing the silicon to heat. The impact of the heat is to reduce the bandgap energy and thus facilitate fault injection. Once the irradiation has ceased, the medium rapidly returns to equilibrium due to the carrier recombination rate (decreasing carrier densities) and high mobility. Each of these processes is discussed in more detail in this sub-section.

When the laser beam hits the silicon, reflection and absorption occurs. The reflection is caused by the discontinuity in the index of refraction (from air to silicon). The percent of power of the beam reflected is proportional to the reflectivity. The reflection depends on the index of refraction of the material, incident angle of beam and the beam's polarization. Generally for silicon (with wavelength around $1000nm$) it's typically 30% [41] or 41% [42]. It's important to note that after the beam enters the silicon it may travel through to the silicon substrate and be transmitted out of the opposite side of the silicon die (e.g. $42\% - 12\%$ transmittance for die thicknesses $20 - 260\mu m$ [42]). A metallic reflective plate on the other side of the silicon can be used to reflect back energy into the substrate [43]. For example, backside irradiation may use a metallic reflective plate on frontside of silicon die and vice versa. Typically, some energy is also naturally reflected from the back surface of the silicon substrate due to the discontinuity of the refractive index. For example, thinning of the backside of an IC can typically reduce thickness from $790\mu m$ to $30\mu m$ [44]. For a $362\mu m$ thick IC, the frontside includes $15\mu m$ metal layers (6 layers), $2\mu m$ active devices, $30\mu m$ FIB trenches and $315\mu m$ bulk silicon material [45].

Although some mechanisms of semiconductors can be characterized using linear relationships at low energies (up to $10^5 - 10^6 \frac{W}{cm^2}$ or equivalent dose rates up to $10^{12} rad(Si)s^{-1}$) and non-linear relationships at extremely high energies, the actual mechanisms are very complex and are affected by impurities, defects, circuit structures within the material, etc. Linear relationships will be covered next.

If the laser beam's energy is greater than the silicon's bandgap energy E_g, then each absorbed photon produces a single electron-hole pair, referred to as the photoelectric effect (e.g. liberates an electron). The photoelectric effect is energy and medium dependent. The silicon bandgap energy at room temperature is $1.12eV$. Lasers with $1064nm$ wavelength have energy just higher than the silicon band gap. For example, the energy of the liberated electron is $E_e = hv - E_g$. If the photon's energy is 3 to 4 times greater than the bandgap, the energy absorbed (or ionization energy, E_i) generally does not increase significantly, creating a threshold like effect. Specifically its value is shown in equation 2.16 [43].

$$E_i = 2.67E_e + 0.87\ eV \tag{2.16}$$

Also it's interesting to note that also at extremely high energy levels two other processes can occur, specifically Compton scattering (liberates an electron but also gives off a wave of a different wavelength) and pair production (produces an electron and a positron, requiring $MeVs$ of energy). These mechanisms are typically important to radiation simulation using lasers. All three mechanisms are referred to as ionization. Lasers with less than $10mW$ are insufficient to switch bits in memory but may change the gate switching time [46]. The corresponding penetration depth is about $700\mu m$ for neodymium lasers emitting at 1.06 or $1.08\mu m$ [43].

The absorption of the laser beam into the medium generally follows Beer-Lambert law [34]. Another more detailed formulation which incorporates time is shown in equation 2.17, and denoted as the depth profile of equivalent dose rate [43], where R is the reflection coefficient, I_{in} is the incident intensity, g_o is the electron-hole generation constant (4.3×10^{13}), α_f refers to fundamental absorption coefficient and hv is the energy of the photon.

$$P_e q(z,t) = \frac{(1-R)\alpha_f}{hvg_o} I_{in}(t) e^{-\alpha z} \tag{2.17}$$

The absorption coefficient strongly depends upon the wavelength but also depends upon the temperature, type/level of doping and laser energy [43]. For example, doping concentrations higher than $10^{18} cm^{-3}$, as typical in bulk and at the surface, lead to higher ionization.

Higher absorption occurs at higher temperatures due to reduction of the bandgap, however other effects may take place including latchup, resistivity, etc. If the laser energy exceeds the bandgap energy, the absorption coefficient, α, is defined in equation 2.18 [43], where α_f refers to fundamental absorption coefficient; σ_n, σ_p refer to the light scattering cross sections of free electrons and holes; the parameters N_n and N_p are the equilibrium free densities of free electrons and holes; and N_{FC} is the total density of generated and injected electron-hole pairs.

$$\alpha_t(z,t) = \alpha_f + \sigma_n N_n + \sigma_p N_p + (\sigma_n + \sigma_p) N_{FC}(z,t) \tag{2.18}$$

Parameters σ_n, σ_p vary as λ_y where y is between 2 and 3.5 depending upon the semiconductor [43]. The energy bandgap decreases as the doping concentration increases and this also raises the absorption coefficient (referred to as the Burstein-Moss shift).

It's well known that the bandgap energy of silicon reduces as the temperature rises. The reduction in bandgap leads to a laser absorption coefficient dependence shift. Experimental evidence has shown that this effect is stronger for wavelengths around $1.06\mu m$ where the charge tripled (although the theory predicted charge to double) [47]. In this experiment a pulsed laser, spot size $5\mu m$ ($8ps$ pulse width) was focused on the frontside of a pn-junction in $2\mu m$ CMOS and found to increase the charge over temperatures 22℃-110℃.

The spatial generation profile of electron-holes in silicon follows the laser profile which is a Gaussian distribution in this case represented in equation 2.19 [46] where η is photogeneration efficiency. This is similar to the intensity equation 2.4 and simulated model of laser current shown in equation 2.22.

$$G(r,z) = \eta\alpha \frac{I_o}{hv} \left[\frac{w_0}{w(z)} \right]^2 e^{-2\left(\frac{r}{w(z)}\right)^2} \tag{2.19}$$

Following an electron-hole pair generation, the electron-hole pair may recombine if there's no electric field. If an electric field is present (such as pn-junction or biasing) then the electron and hole will drift in opposite directions to create a transient current. If there is sufficient laser energy, e.g. over $10ns$, the transient current may change the transistor bias point and also change the parasitic capacitance characteristics. These changes can lead to delays or speed increases in the circuit. Charge collection mechanisms include charge drift, diffusion as well as funneling in sub-micron transistors where complex simulations have been performed to study possible mechanisms, see Figure 2.5. Drift (which also includes charge funneling, where the ion track distorts the electric field lines, allowing collection by drift beyond the depletion region) occurs quickly unlike diffusion which is slow.

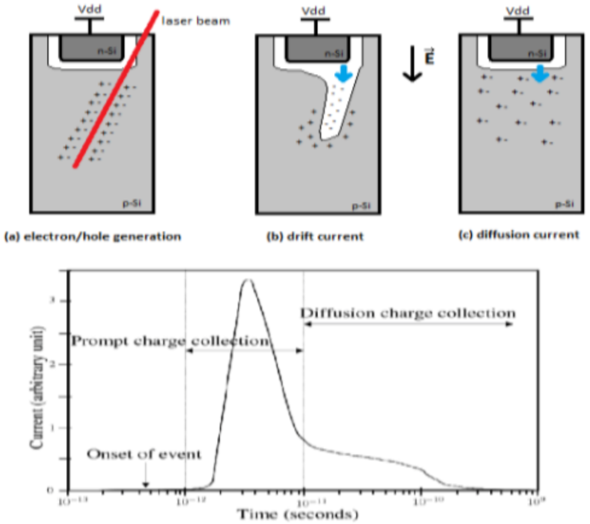


Figure 2.5: Laser pulse effect on semiconductor: electron-hole generation a), drift b), diffusion c) and resultant displacement current (below) [4].

For example, in a CMOS inverter if the input is a rising edge, and the PMOS transistor is irradiated, the output signal is delayed. If the input is a falling edge and the NMOS transistor is irradiated then the output is also delayed [46]. In general, it is more practical to affect the NMOS transistor since it has lower doping concentration (thus electrons/carriers have higher mobility), although PMOS transistors can also be affected but require higher irradiation. In deep sub-micron technologies, the induced current (normally a peak current caused by funneling followed by slower drop in current due to diffusion process) may also disturb the timing characteristics of other gates. The increase in delays help to induce a fault by violating the setup times, etc. Effect of laser pulses on memory has also been investigated and simulated using Spice [48], where the laser was modeled as a voltage controlled current generator.

The incident laser pulse energy can be represented in equation 2.20 if one ignores reflection and assumes uniform penetration depth [49], where h is Planck's constant, c is the heat conductivity of the material, LET is linear energy transfer and q is electronic charge. The LET is typically used to measure particle charge generation capability [50]. It has been suggested that angled laser strikes would increase the charge collection ($LET_{effective} = \frac{LET}{\cos\theta}$, where θ =0°for perpendicular strike onto surface of IC), however it's also likely that this charge would spread over more than one circuit node. This formulation is typically used in the field of Single Event Errors (SEEs) which studies effects of radiation on ICs (such as alpha particle radiation in satellites). However, LET is the action of radiation on matter (measured as $\frac{keV}{\mu m}$) and is typically simulated by laser pulses but does not apply to photons [29]. Nevertheless, LET and penetration depth is typically simulated in lasers by changing the beam spot diameter and wavelengths. Radiation from the laser is absorbed uniformly in a sphere of radius r_o and in steady state, the change in temperature, T, for all $r > r_o$ is shown in equation 2.21 along with the time to heat up a small region t [51]. The parameters are P, the power absorbed, and s, the specific heat. The device may be heated before thermal equilibrium occurs.

$$E_\Lambda = \frac{h\ c\ LET\ 10^{-15}}{\alpha\ \lambda\ q} \tag{2.20}$$

$$\begin{aligned} T &= \frac{r_o}{r}\Delta T_o \\ \Delta T_o &= \frac{P}{4\pi r_o c} \\ t &= r_o^2\frac{s}{3c} \end{aligned} \tag{2.21}$$

The voltage controlled current generation model used to simulate the induced current produced by laser irradiation in semiconductors is shown below in equation 2.22, where $p, q, s, c1, cc2, \beta, \gamma$ are constants [48]. The variables S, d, V represent the surface area under laser irradiation, the distance from center of laser spot to sensitive location of current generation (e.g. pn-junction location), and reverse bias voltage of pn-junction device, respectively.

$$\begin{aligned} I_{laser} &= P_{laser}\big[\big(pP_{laser} + q\big)V + s\big]\Omega_{laser}\ S \\ \Omega_{laser} &= \beta e^{-\frac{d^2}{c1}} + \gamma e^{-\frac{d^2}{cc2}} \end{aligned} \tag{2.22}$$

Research has shown that heat alone can be used to change bit values in memory. For example in [52], over 30/300 seconds are required to change a bit in EEPROM using $10mW/50mW$ $650nm/1065nm$ laser using frontside/backside illumination, respectively.

In continuous wave lasers, induced substrate currents were found to cause more perturbations to nearby gates and thus cause IC damage. In contrast researchers found that pulsed lasers were preferred since the substrate current was more limited and thus more local perturbations could be produced [46].

2.3 Photoexcitation Effect of Laser on Silicon

SRAM laser exposure is known to cause bit-flips in [29,31,53], a phenomenon called Single Event Upset (SEU). By tuning the beam's energy level below a destructive threshold, the target will not suffer any permanent damage. A conventional one-bit SRAM cell is made of two cross-coupled inverters as shown in Figure 2.6. Every cell has two additional transistors controlling the cell's content access during write and read. As every inverter is made of two transistors, an SRAM cell contains six Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs).

In each cell, the states of four transistors encode the stored value. By design, the cell admits only two stable states: a '0' or a '1'. In each stable state, two transistors are at an on state and two others are off.

If a laser beam hits the drain/bulk reversed-biased pn-junctions of a blocked transistor, the beam's energy may create pairs of electrons and holes as the beam passes through the silicon. The charge carriers induced in the collection volume of the drain-substrate junction of the blocked transistor are collected and create a transient current that inverts logically the inverter's output voltage. This voltage inversion is in turn applied to the second inverter that switches to its opposite state: all in all, a bit flip happens as reported in [31,53].

From the attacker's perspective, an additional advantage of LFI is reproducibility. Identical faults can be repeated by carefully tuning the laser's parameters and the target's operating conditions. Lasers are mainly divided into two categories: continuous and pulsed.

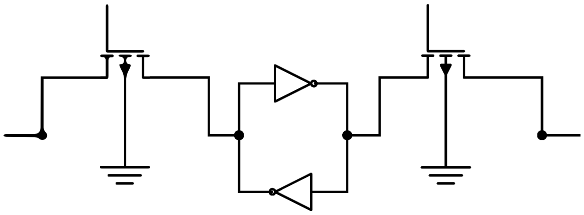


Figure 2.6: A typical one-bit SRAM cell.

2.3.1 Continuous vs pulsed laser

Pulsed laser refers to any laser not classified as continuous laser, where the optical laser pulses are of some duration at a certain repetition rate. A continuous laser source emits a laser beam of a constant power over time. Some laser can't simply run in continuous mode because it would require pumping the laser at a very high continuous power level which would be impractical or destroy the laser by producing excessive heat [33]. In our FI application, the pulsed laser is preferred to avoid damaging the DUA. There are different ways to make a laser source pulsed such as using Q-switching or mode-locking. Pulse durations from Q-switched lasers typically vary between $100ps$ and hundreds of ns. Mode-locked lasers can generate pulses with durations between $\approx 5fs$ and hundreds of ps.

2.4 Tunable Parameters in LFI

In a laser attack, the attacker usually controls the beam's diameter (AKA spot size), wavelength, amount of emitted energy, impact coordinates (attacked circuit part) and the exposure's duration (laser pulse width) or number of laser pluses. Also, the attacker may also synchronize the LFI with a given clock cycle of the target, or control the target's clock frequency, V_{dd}, temperature of DUA or alter the threshold voltage, V_{th}. According to [54], the pulse width can be controlled for both pulsed and continuous lasers.

The power of the laser beam could be tuned using a half-wave plate (or retarders) [55] and an attenuator. Power loss occurs for the beam when passing through the optics as noted in the manual for the Video Microscope Unit (VMU) [56]. For example, the transmission % for a $1064nm$ beam through the VMU-LB is 54% for a pulse width $\geq 10ns$ with an upper limit of laser input of $110\frac{mJ}{cm^2}$ per pulse for a Nd:YAG laser passing through an M Plan Apo NIR series objective. If the pulse width of the laser is to be shorted, reduce the upper limit of laser input density by the square root of the pulse width ratio. For example, when the pulse width is $2.5ns$, the upper limit of laser input (VMU-LB $1064nm$) becomes $55\frac{mJ}{cm^2}$.

Finally, laser attacks may attack the decapped chip from the frontside or backside. However, the chip's frontside and backside have different characteristics when exposed to a laser beam.

2.4.1 Frontside

The laser attacks on frontside decapsulated chips are particularly suited to green wavelength (532*nm*) but IR can be used as well. The visibility of chip's components makes positioning very easy in comparison to backside attacks. But, because of the metallic interconnects' reflective effect, it is difficult to target a component with enough accuracy. In addition, sometimes the chip's layout has a mesh protection for tamper resistance [57]. On the other hand, progress in manufacturing technologies results in both a proliferation of metal interconnects and much smaller chips. All in all, it becomes increasingly difficult to hit a target area from the frontside.

2.4.2 Backside

The silicon substrate is a transparent area for IR wavelengths [58]. According to this property, the backside laser attacks are more efficient as the IR rays (1064*nm*) enter deeply into the silicon and target components from the backside. Positioning may be more difficult due to the lack of visibility, unless sufficient NIR illumination and special NIR/SWIR cameras are available. Nevertheless, backside attacks avoid the problem of reflective metallic surfaces.

Figure 2.7 shows absorption coefficient for silicon at various doping levels of a p-type material. The photon energy is a function of laser's wavelength. The vertical solid and dashed lines mark the photon energy corresponding to the wavelengths $1.06\mu m$ and $0.80\mu m$ respectively.

In [58], the effective penetration depth in silicon, which is the average distance travelled by a photon before it gets absorbed, could be roughly estimated for different wavelengths knowing that the effective penetration depth is the inverse of the absorption coefficient. For example, the effective penetration depth in the silicon substrate is $\sim 1.27\mu m$ and 900 *nm* for 532*nm* (green) and 1064*nm* (IR) wavelengths respectively. Therefore, according to Figure 2.7, a 1064*nm* wavelength has a deep penetration in the silicon layer and is suitable for backside laser fault injection on the integrated circuits. Besides, for a frontside attack, wavelengths lower than 0.80 μm are suitable as they don't enter deeply into the silicon [5]. This range includes the 532*nm* green laser wavelength.

Before performing the laser fault injection, the target circuit must be prepared for such attacks. That process is called sample preparation. The sample preparation mainly consists of decapsulation and RE of different structures inside the DUA. Depending on the equipment and attack's requirements, the frontside or the backside of the circuit will need to be decapsulated. The decapsulation procedure will be further demonstrated in section 5.1.

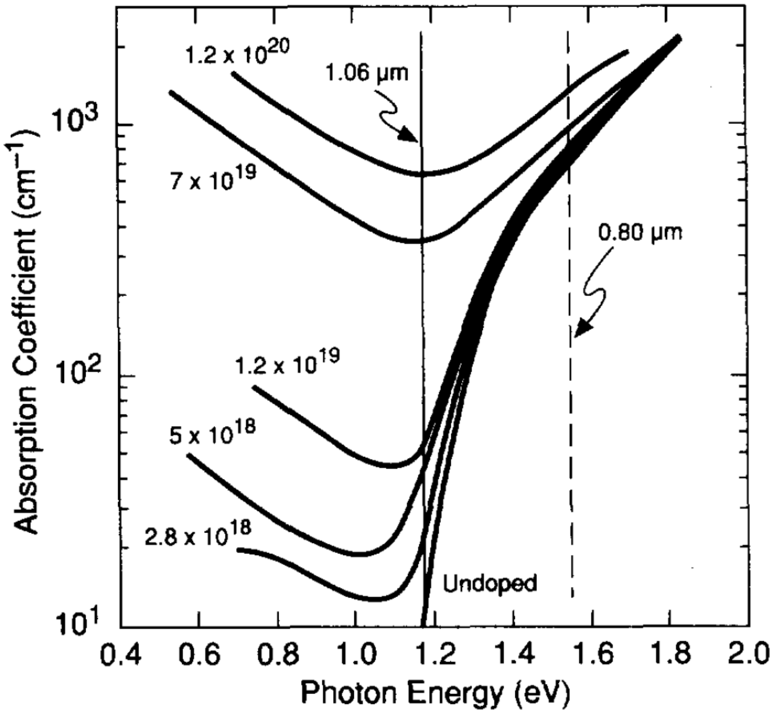


Figure 2.7: Absorption coefficient for silicon at various doping levels of a p-type material [5].

2.5 Summary

This chapter briefly reviewed some fundamental theories in laser beams and their application to FI. Both the wave and particle theories were utilized to explain effects of optics on the the laser beam and how the laser is absorbed when directed on a DUA. Different wavelengths were investigated to determine the suitability for either frontside/backside attacks. The next chapter will discuss previous research in these areas.

Chapter 3

Literature Review

Although the focus of this thesis is on decapsulation/LFI, other types of attacks and countermeasures are also relevant and/or applicable. For example, countermeasures for RE may also apply to decapsulation and LFI. Hence, this chapter further defines and examines previous research in a wide range including decapsulation, RE, microprobing, FI, and LFI.

3.1 Reverse Engineering

This section will focus on the structural RE of ASICs and custom ICs which is considered an invasive attack. The RE steps that can be followed to reconstruct a chip will be discussed. We'll mention the required tools for each step along the way. This section provides the flow for the RE steps. Section 3.2 will focus on microprobing as an important tool facilitating the RE process and also discusses several techniques for modifying the chip under analysis.

Structural RE requires a series of procedures [20] that can be summarized in Figure 3.1. The starting point for the RE is a high-end manufactured chip that the reverse engineer knows nothing about its internal structure and the outcome of the RE process is a netlist specifying every minute detail in the structure of that sophisticated chip. The final product of RE can be a complete set of functional analysis reports, architecture blocks, hierarchical schematics, hierarchical netlist, simulated waveforms, block diagrams, timing diagrams, and/or circuit equations. In the following subsections, a detailed description is presented for each RE step.

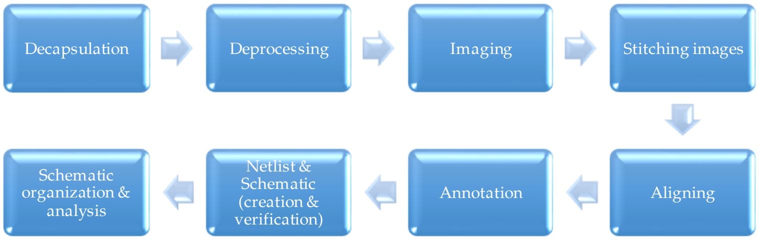

Decapsulation
Deprocessing
Imaging
Stitching images
Aligning
Annotation
Netlist & Schematic (creation & verification)
Schematic organization & analysis

Figure 3.1: RE flow.

3.1.1 Decapsulation

The very first step is to get the chip to be analyzed out of its package. This is called decapsulation or depackaging. We'll assume that nothing is known about that chip in terms of pin assignment or data sheets. It's recommended to do a functional analysis if the packing doesn't obstruct that. Getting some hints at this stage will be of a great help when trying to relate the chip functions with its internal structure. The tools needed for depackaging are usually affordable and cheap. Chemical acids, which are corrosive and dangerous, are needed as an etching agent for plastic packages. Safety-goggles, acid-resistant gloves and clothing should be worn. The chemical work should be done in a fume cupboard to limit the exposure to toxic fumes resulting of the package-etching acid (usually fuming nitric acid with a concentration $\geq 95\%$). A mixture of fuming nitric acid and concentrated sulphuric acid can be used to speed up the process and avoid reacting with the silver bonding pads. The process of decapsulation starts by milling a hole in the package then applying the acid on that area. The tools required for milling can be purchased from a "Do it yourself"shop. The etched plastic residues are removed through an ultrasound bath. Another interesting approach, instead of milling, is taping the package with an acid-resistant tape leaving the part to be etched untapped. The tapped chip is immersed in the acid that etches the uncovered part exposing the die. This method is used for Quad Flat Package (QFP), Thin Small Outline Package (TSOP) and Ball Grid Array (BGA) packages. Another method for manual etching without chemicals is performed through milling down the chip form the rear side to the copper layer which is removed mechanically. This method is used for smart cards. The Laboratory for Physical science [59] reported that any part in plastic and BGA interposer packages can be extracted with better than 90% yield. An IC (de)soldering station might be needed for decapsulation; oscilloscope, logic/spectrum analyzer and signal generator for system-level functional analysis.

In [60], various decapsulation techniques are discussed. Samples of *ARM Cortex-M3*, a 65nm FPGA and AVR MCUs were decapped from the frontside chemically as well as

AVR MCUs and the $65nm$ FPGA from the backside mechanically. The technology nodes of the devices under attack ranged from $250nm$ to $65nm$.

Backside decapsulation of chip is essential for photon emission analysis (PEA) [61, 62] which is used as a passive undetectable attack by capturing intrinsic emissions of switching transistors from the backside of a decapsulated chip. However, PEA requires millions of traces which could be thwarted by the number of executions of the security algorithm running on the DUA.

3.1.2 Deprocessing

After getting the multi-layered silicon die out of the package, the reverse engineer starts deprocessing or delayering the die. Deprocessing is actually the opposite process for chip fabrication. In the delayering process, the aim is to remove one layer after another while imaging each layer before removal until the last bottom layer is reached. Those images will be used together to reconstruct the whole netlist of the chip. The very top layer of a silicon die is a special layer called the passivation layer and its removal exposes the top metal layer allowing microprobing attacks. The deprocessing is either done mechanically through polishing or chemically through etching. Etching is either wet or dry (plasma). The mechanical polishing is time-consuming but useful for a multilayer interconnect fabricated with advanced planarization techniques. Wet etching is isotropic while dry etching is anisotropic. Removing each layer can make use of the proper etchant as mentioned in Table 3.1. Since delayering stared with the top layer so it's called frontside delayering. There's also backside delayering which applies to flip-chip circuit packaging. Backside delayering requires removing most of the silicon die thickness before the circuitry can be imaged and knowing the thickness of the remaining silicon after each thinning step is critical. Also backside access can be used to localize silicon access while preserving the electrical capabilities. The challenge is to enable a clean and fast localized access through silicon. Backside delayering could be useful to bypass the coating or upper mesh layer that signals an alarm if a wire is broken or two signals are shorted.

3.1.3 Imaging

Taking thousands of high-quality images from referenced locations to each layer of the chip throughout the delayering process is very crucial to reconstructing the whole chip. For technology nodes starting from $0.18\mu m$ and larger, a high-resolution optical microscope will be sufficient. Note that the depth of focus helps in separating different layers. For

Table 3.1: Proper etchant for different materials [13].

Material	Wet etching chemicals	Dry etching gases
Si	$HF + HNO_3, KOH$	CF_4, C_2F_6, SF_6
Poly Si	$HF + CH_3COOH + HNO_3$	CF_4, SF_6
SiO_2	$HF, HF + NH_4OH$	$CF_4, CF_4 + O_2, CHF_3$
Al	$HCl, H_2O_2 + H_2SO_4, HPO_3 + HNO_3 + CH_3COOH, KOH$	CCl_4, BCl_3
$W,\ Ti$	$HF + HNO_3, H_2O_2 + H_2SO_4, H_2O_2$	CF_4
Si_3N_4	$HF + HNO_3, HPO_3, Nitrietch$	CF_4
Polyimide	$H_2O_2, H_2O_2 + H_2SO_4$	$CF_4, CF_4 + O_2$

smaller technology nodes, a SEM, transmission electron microscope or scanning capacitance microscope (SCM) will be needed for a resolution $\leq 10nm$. The SEM uses electron particle gun for top and cross-sectional views with a resolution limit of $1nm$ while transmission electron microscope gives high resolution 3D images of the device structure with $1 - 3pm$ precision [63], and SCM allows seeing the positive and negative doping which form the actual working transistors, resistors, etc., in the silicon chip with a resolution limit of $2nm$. Layer imaging is usually done overnight using automatic 2D sub-micron steppers integrated with the microscope.

3.1.4 Stitching and Aligning

After layer imaging is done for each and every layer, a special developed software stitches thousands of images together. The alignment of images of multiple layers is critical to proceed through the image stitching process and is verified through the lining up of vias and contacts with the layers above and below. More software work will be needed for the aligning phase.

3.1.5 Annotation

At this stage, labeling transistors, on-chip capacitors, on-chip resistors, on-chip inductors, interconnects, vias, contacts, etc., in all layers is mandatory before we are able to read back the circuit. This labeling can be done automatically using a tool such as *Chipworks ICWorks Surveyor* [64] or manually.

3.1.6 Verification and schematic creation

The annotation is error-prone either because it's done manually by an extraction engineer or automatically with issues in the images themselves like bits of dust or the image stitching software introduced errors while integrating images. Therefore, verification is performed through design rule checks looking for any issues like spaces smaller than the feature size, hanging wires, shorted wires, etc. Then, the tool can extract the netlist creating a flat schematic. Also the netlist is checked for errors like shorted inputs/outputs/ supplies, floating gates/nets, etc.

3.1.7 Schematic organization and analysis

Now a certain level of expertise is required to organize different blocks in the schematic in a logical hierarchy that makes the design easy to understand. Irrelevant placement of blocks in the schematic will make the schematic look very strange and hard to understand. The organization and analysis phase is iterative. Analysis can be done in hand using the transistor and logic theory or using tools like *ICWorks* tools. The final product of the structural RE is a report comprising a number of hierarchical schematics/netlists, waveform, timing diagram, masks, layouts, etc., and everything else that enables a Design and manufacturing company (DMC) [65] to manufacture the analyzed chip.

3.2 Microprobing

A brief introduction to microprobing attacks and the required setup of such type of attack was presented in section 2.1.2. In this section, more microprobing attacks and countermeasures to those attacks are discussed.

In [66, 67], a theoretical foundation was proposed for protecting Boolean circuits from probing attacks as well as similar types of attacks that rely on leaked information from electrical circuits. The attacker is assumed to have limited capabilities observing up to any t internal wires. The proposed solution shows that any circuit can resist probing attacks using a generic defensive design strategy. That solution encodes the input/output/internal nodes in a certain way that makes the circuits perfectly secure against an attacker with t probes during a certain clock cycle at the expense of increased hardware overhead. In [67], the proposed solution was refined to be secure against unbounded number of reset-only wire faults but limited to only a bounded number of arbitrary wire faults per clock cycle. Since the faults per clock cycle might be permanent so the total number of introduced faults in more than one clock cycle is unbounded. One of the disadvantages of such countermeasure is the large overhead of the added circuitry. The added circuitry is a function in the number of adversarial probes the designer wishes to tolerate. Also the countermeasure is mainly designed to protect circuits against physical probing attacks.

In [45], the backside of modern ICs is considered the weakest link because it's not protected against fully-invasive attacks at all. Both backside probing and edit attacks have been carried out on two target ICs (Atmel AT90SC3232C and ATmega328P). The edit attack was to circumvent the copy protection by modifying the security fuse while the probing was to extract decrypted data. The DUAs were mechanically decapsulated from the backside using the *Ultratec ASAP-1* [68] down to $30\mu m$ substrate thickness. The Hammamatsu PHEMOS-1000 was used to produce reflective laser scanning images of the ATMega328P. These images were used to identify the points of interest to perform the attacks. For AT90SC3232C, optical microscope images were acquired. The FIB used a wavelength transparent to silicon i.e. 1000 nm. The probing attack was done using a $Karl\ Süss\ PA$ 150 probing station was used in conjunction with $Karl\ Süss\ PH$ 150 micro-manipulators and GGB T-4-10 probing needles.

In [57], a countermeasure to protect the IC from the backside attacks by stacking multiple dies inside a 3D mesh. To prove manufacturability, a passive cage was implemented using $130nm$ 6-metal-layers technology to protect an 8-bit register. The cage was a $26\mu m$ wide cube and provided very good RE protection. The cage design is based on a Hamiltonian path, which is an undirected path passing once through all the vertices of a graph. If any of the vertices is a transistor, then the cage is considered active otherwise it's a passive cage. Dynamic or active cages contained hardware canaries. A hardware canary (also defined as a configurable switch box) was defined as a binary constant placed between a buffer and stack data to detect buffer overflows. Upon buffer overflow, the canary gets corrupted and an overflow exception is thrown. The hardware canary is formed of a spatially distributed chain of functions positioned at the vertices of a 3D cage surrounding a protected circuit. A correct response from those functions to a challenge will attest the canary's integrity. One of the advantages of this countermeasure is that it's purely digital, not an analog sensor or a physical unclonable function (PUF), so it's portable across different technology nodes. Also, this countermeasure is effective against Focused Ion Beam (FIB) attacks as FIB attacks target interconnections of the cage not the values inside the switches so the FIB attack becomes very complex.Therefore, these 3D canaries act as cryptographically secure meshes. However, this countermeasure is costly when trying to thwart backside attacks because of the high cost of 3D IC manufacturing therefore it should be implemented only when truly needed.

In [69], Weiner et al. present a $65nm$ low area probing detector (LAPD) as an efficient approach to detect micro-probing through comparing delay differences between symmetric lines such as bus lines to detect timing asymmetries introduced by the capacitive load of a probe. Their simulation results show different LAPD implementations that can detect state-of-the-art commercial microprobes under both typical and worst case conditions. The area overhead for the added circuitry is relatively low compared to other protection mechanisms, such as the probe attempt detector (PAD) [70] or bus encryption. A calibratable lightweight invasive attack detector (CaLIAD) is proposed in [71] to account for manufacturing variations as well as small layout imbalances while maintaining the low rate of false positives. The resolution of detection is $5fF$ and $11fF$ under typical and worst case conditions respectively. The CaLIAD uses only digital components such as LAPD but has an area overhead in between LAPD and PAD so it can be considered the best tradeoff between LAPD and PAD.

In [72], the original protective (tough and chemically inert) coating was used as a fingerprint to detect any attacks from the frontside. This countermeasure is proven to be secure against FIB attacks. Physical attacks by making a hole through the coating are detected from the difference in capacitance measurements. The opaque protective coating was sprayed only on the frontside and it's questionable if it could be applied from the backside. This countermeasure acts as coating PUF covering the frontside of IC that needs protection. On-chip measurement circuit that measures capacitance values at several sensors was developed. To verify the countermeasure empirically, two holes were made by Gallium FIB leading to differences in capacitance measurements, thus the attack was detected.

In [73], optical interaction concepts are utilized for both security protection as well as introducing a security risk. For protection, the light emission from the operation of forward-biased p-n junction, which is considered as a LED, is detected by depleted areas like reverse p-n junctions, which are considered photocurrent detectors. The backside of the IC is coated with an optically active layer that reflects the light emitted by the LED, which in turn, is to be detected by the photodetectors. The backside protection layer provides angle dependent reflection and full transmission blocking so any damage in this protection layer will affect the angle of reflection for one or more of the LED light beams to be detected by the photodetectors. Therefore, the detected photocurrent can be used to verify the integrity of the backside protection layer and detect attacks. On the other hand, an attack on ring oscillator (RO) PUFs using electro-optical frequency modulation (EOFM) and electro-optical probing (EOP), a laser voltage probing (LVP) derivative that has been rarely used for security attacks, was presented. Power analysis of the RO in the frequency domain to estimate the approximate RO frequency followed by super-positioning of the enough near-by frequencies led to detecting all the ROs and registers switching with the same frequency through PEA. Finally according to [74], once the precise frequency of RO PUF is known, complete characterization and cloning of the PUF is possible.

In [75], the nanopyramid appears to be an effective countermeasure against backside optical attacks such as LVP. The nanopyramid is a passive CMOS compatible countermeasure that doesn't consume any energy and has no area overhead. Normally, the nanopyramids are inserted in between the transistor layer and the metal layer; while for a Silicon on Insulator (SOI) the nanopyramids are inserted beneath the transistors. According to the reported simulation results, the nanopyramid can disturb the optical measurements enough to make LVP attacks practically infeasible. The nanopyramid demonstrates a change in the reflection behavior of the active devices due to the enhanced scattering which arises from the pyramids. Due to the randomness of the nanopyramid in terms of size and location, the optical reflection collected from the nanopyramid devices is not reliable, thus an efficient LVP attack can be prevented.

3.3 Fault Injection and Laser-based Attacks

This section first generally reviews FI using other techniques such as glitching, EM FI, etc., and then focuses on the use of lasers and IR wavelengths in FI attacks and their countermeasures.

3.3.1 General Fault Injection

In general, FI has been used to launch attacks on crypto-systems through glitching the clock or V_{dd}, injecting an EM pulse or laser [18]. Two attacks utilize FI: one is instruction skipping and the other is injecting faults in crypto algorithms. For example, in [76], the full 128-bit AES key was extracted by means of a side-channel attack (SCA) in less than one hour leaving no physical traces and using low-cost equipment. Table 3.2 summarizes some known attacks on 128-AES as well as their requirements and shows targets of fault injections where faulty cipher text can be utilized to derive secret information such as the key; column 2 gives the target location (e.g., state or key byte), column 3 reports the requirement necessary to perform the attack in terms of spatial accuracy and the last column gives the number of distinct faults that must be injected to retrieve the secret key. For instance, in [77], a single bit in the plain-text M_0 has to be faulted. From Table 3.2, it is clear that simply being able to inject a single fault is far from being sufficient to perform an attack. The fault has to be injected at the right moment and has to affect either only one bit or only one byte without impacting other data. Although lasers have a sufficient timing accuracy to fulfill the first requirement, spatial accuracy and localization are more difficult to manage.

Table 3.2: Some Known Attacks on AES 128 [14].

Reference	Target	Focalization	Faults #
[77]	Data (M_0)	Bit	128
[78]	Data $7 * M_8$+Key $4 * K_9$	Byte	11
[30]	Data $16 * M_9$	Bit	16
[79]	Key $4 * K_9 + 4 * K_{10}$+Data $4 * M_8$	Byte	12
[79]	Data $4 * M_9$ or $1 * M_8$	Byte	4 or 1
[80]	Round counter	Round counter	1
[81]	Round counter	Byte	1
[82]	Data $4 * M_9$	Byte	4
[83]	Data before first sub-byte	Bit or byte	16
[84]	K_7	Byte	1
[85]	Data	Byte	1
[86]	First column of K_8	Byte	1

In [87], *PIC16F687* and *PIC16F866* MCUs were attacked where it was noted that MCUs generally exhibit a stronger side-channel leakage compared to FPGAs because the number of traces required to extract the key of an AES implementation using correlation power analysis (CPA) was one order of magnitude less in the case of MCUs. The configuration bits (or more specifically the erasable programmable read only memorys (EPROMs) bits) of a decapsulated *PIC16F687* were exposed to Ultraviolet light in the C spectrum (UV-C) to reset the security mechanism and allow code extraction. The UV-C attack is presumably applicable for the most of the *10F* to *16F* series of *Microchip* using UV-EPROM eraser at a certain angle to circumvent the metal shield covering the configuration fuses.

3.3.2 Laser-based Techniques

Several laser-based failure analysis techniques have been developed, such as Light Induced Voltage Alteration (LIVA) [88] and Optical Beam Induced Current (OBIC) [89]. LIVA is a technique that uses carrier stimulation to induce power alteration. It uses a constant current source to monitor changes in voltage and has been effective in imaging defects such as floating nodes and open conductors. Also, LIVA has been used for imaging transistor states and controlling logic states. OBIC takes advantage of photon generated electron-hole pairs to yield information about IC defects and functionality. It was shown in [27] that it is possible to read out masked read-only memory (ROM) by measuring the OBIC of the DUA *MC68HC705P6A*.

In [90], it has been noted that when the laser beam hits the pn-junction of a transistor, a current is created that might charge or discharge the output of the targeted gate. When targeting, e.g., SRAM or flip-flops (FFs), the state might be permanently altered. However, when shooting at general combinatorial logic, the effect caused by the laser is only transient which means that the circuit will restore its normal behavior, depending on its input, when the laser excitation is stopped. They investigated the transient effect of LFI on the substitution box (S-box) and noticed that increasing the intensity of the laser beam increased the number of affected bits. They successfully attacked an AES co-processor on an *Atmel ATxmega16A4U* with minimal feature size of $250nm$. The backside of the DUA was thinned to approximately $20\mu m$ remaining substrate. The used microscope was by *Opto GmbH*. The objective used was a 10x Mitutoyo Plan Apo NIR objective (with $NA = 0.26$). Two $975nm$ single mode fiber-coupled on-demand laser diodes were focused at the same spot with maximum power. The reported measured peak power was $0.52W$. The timing of the laser pulse was controlled by a Stanford Research Systems DG645 programmable delay generator based on a trigger signal at the start of each single AES encryption (a single AES encryption needed 375 cycles). The relation $\frac{(1.22\lambda)}{NA}$ was used to calculate the minimum spot size (which is typically used for light not for collimated beams i.e. lasers [56]). An Ophir Spiricon SP620U beam profiling camera (resolution of $4.4\mu m$) was used to measure the spot size. The laser pulse width was measured through a photo-diode in the optical path (converted to a voltage by a shunt and amplified using a Langer PA303 amplifier). It was mentioned that the laser pulse had to start before and end after the rising edge of the clock signal to attain an 80% successful LFI rate.

In [91], OBIC was used as an imaging technique to reduce the search space significantly and find areas of interest in the DUA. The same LFI setup in [90] was used.

In [92], $He-Ne$ laser ($0.63\mu m, 1mW$) was focused down on the silicon wafer surface to an area of $\sim 5\mu m$ spot size using a 10x objective in order to measure diffusion length in silicon. The laser beam was focused on the silicon substrate surface and excess carriers were generated at a known rate by the $1mW$ $He-Ne$ laser. Then the resulting total carrier number was detected by a novel IR emission technique proposed by the authors. The number of the detected carriers determine the diffusion length at that location and it was found that the number of carriers drop exponentially when moving laterally away from the laser spot.

3.3.3 Laser Fault Injection

In this subsection, we'll focus on the fault injection using laser, how it emerged and known countermeasures against such type of attack.

In [29], optical fault induction attacks have been introduced knowing the fact that optical radiation could ionize semiconductor regions if the radiation's energy exceeds the band gap of the semiconductor. A common *PIC16F84* MCU was decapped from the frontside and the SRAM was attacked with a flash lamp and a 1500x objective mounted on a Wentworth Labs MP-901 manual probing station. Single bit flips from 0 to 1 and from 1 to 0 were performed. The limitation was that the flash lamp doesn't produce even and monochromatic light and was difficult to be controlled; so using a laser came as an idea to overcome this limitation. A cheap class-II laser pointer ($< 1mW$ and 650 nm wavelength) was used then instead of the flash lamp and was operated with a supply current so that it can output up to 10 mW. The laser beam was focused down to $\sim 1\mu m$ spot. It has been noted that X-rays could penetrate most types of protective packaging likely to be encountered in practice therefore investigating whether the X-rays could be used to inject faults or not is worthwhile. Also, countermeasures such as top-layer metal shielding could be defeated by X-rays or IR laser; and bus encryption could be circumvented through direct register attacks. Therefore, those countermeasure aren't effective against LFI. Dual-rail logic was proposed as a countermeasure to flipping bits.

Insights about choosing the right laser setup are provided in [60] like choosing the suitable wavelength, spot size and laser power; then followed by some hands-on experience on device profiling for performing successful attacks such as counting for the attack location, pulse power, pulse duration (glitch length) and offset after triggering. The reported laser spot size varied between 10 and $800\mu m^2$ with $\geq 90\%$ success rates of bit flips. Every DUA required a different fault injection threshold which means that every DUA required different exposure time to the laser beam such that the state changes to a faulty one. So the safest option was to start at the lowest power then use small incremental steps to avoid destructive changes. Surprisingly, it was found that the percentage of power needed for fault injection in smaller technology nodes is higher than that for older technology nodes. A $20W$ $1064nm$ was used for backside FI and that power was reduced by the objectives to $\sim 7-8W$. For the $65nm$ FPGA, at least 80% of the laser power was needed to inject faults while for old MCU only $10-20\%$ was sufficient. An X-Y stage with $0.05\mu m$ resolution was used. The types of induced faults were bit flips, instruction skip (IS)/changes, execution disturbance and bit flips in slice registers in FPGA. A software countermeasure against IS proposed in [93] provides protection for embedded programs against multiple FIs attacks. The main idea of that countermeasure is that it provides a generic manipulation to the assembly code of the program executed on the DUA to protect it from IS attacks. The countermeasure divides the instruction set, which will be manipulated, to 3 different classes. The 1^{st} class is composed of idempotent instructions which can provide fault tolerance by simply duplicating those instructions. The 2^{nd} class contains the instructions that are not idempotent but can be replaced by an equivalent sequence of idempotent instructions. The 3^{rd} class collects some specific instructions that cannot easily be replaced by a list of idempotent instructions but for which a specific replacement sequence is possible. Moreover, the 3^{rd} class includes the instructions for which no replacement sequence, that ensures fault tolerance, can be provided. The solution for those instructions is either forcing the compiler to avoid using them or using a fault detection approach. The mentioned IS countermeasure doesn't claim complete protection but it claims reasonably good protection against IS attacks using reasonable cost equipment.

In [94], physical faults have been injected in more than a hundred units of 68000 MCUs through cutting either aluminum or poly tracks using laser. The beam accuracy was $0.1\mu m$ in the X-Y plane. Conformity and reduced conformity (functional and reduced functional) tests were used to detect injected faults. The reduced functional tests consisted of a random selection of test cases instead of exhaustive exploration. The 100% detection was attained only when the reduced conformity test program was looped because the length of the test program had an indirect effect on the detection of some faults. For example, stuck-at faults happen when a certain node in the circuit gets shorted either to V_{dd} or ground and the value can't be changed by software so a circuit edit would be essential to change its value. On the other hand, open circuit faults occur when wires are cut. Also bridge faults happen when shorts are created between wires. So LFI can be useful in pragmatic testing of stuck-at faults by artificial realization of such faults.

In [95, 96], the high energy disturbances caused by LFI and electro-magnetic fault injection (EMFI) were detected by a proposed phase-locked loop (PLL)-based sensor on multiple FPGA platforms. A $20W$ pulsed (10 MHz) diode *Riscure* laser system of a $64 \times 14\mu m^2$ spot size was targeted on a flip-chip $65nm$ Virtex-5 FPGA. The $20W$ power is reduced to $10W$ and $8W$ when a 5x and 20x objectives are used respectively, due to the losses of the used optics. The pulse length is adjustable with $1ns$ resolution and the offset from the trigger to laser pulse is $\leq 60ns$. During decapsulation, the heat sink of the DUA was removed and the substrate was thinned from $\sim 300\mu m$ to $\sim 100\mu m$. The motorized XY stage of the laser setup had a precision of $0.05\mu m$. The proposed sensor is sensitive to an energy level which was much lower than the required energy to inject the fault. The hardware (HW) cost of the proposed sensor is one PLL block and one look-up table (LUT) to implement a watch dog RO.

In [97,98], the authors were able to flip single bits/bytes in the SRAM of an 8-bit $0.35\mu m$ 16 MHz reduced instruction set computer (RISC) MCU. The LFI setup used was a Nd:YAG laser with a $532nm$ wavelength. The beam spot size and the energy per shot were $4\mu m$ and $\sim 15nJ$ at the laser source emitter before passing through a 20x Mitutoyo objective. A motorized programmable *Prior Scientific* [99] X-Y stage with $0.1\mu m$ resolution was used. A FPGA board was used to trigger the laser source. The DUA was decapsulated from the frontside using *Nisene JetEtch* device that can be programmed to use arbitrary ratios of nitric acid and sulfuric acid. A successful single-bit Giraud's LFI attack [30] was implemented on an AES round key by carefully controlling the laser's shooting time and location.

In [48], a $3W$ pulsed $1064nm$ laser source with a $50ns$ pulse duration was used to inject faults from the frontside. Three objectives were used and resulted in three spot sizes: $1\mu m$, $5\mu m$ and $20\mu m$. The laser was used at $1.6W$ with $1\mu m$ spot size to inject fault in the SRAM. The DUA was an 8-bit $0.35\mu m$ CMOS MCU.

In [100], a $532nm$ pulsed green laser with a $5ns$ pulse was used to attack the last AES-128 round running on a DUA from the frontside. The spot size was $125\mu m \times 125\mu m$ and the energy density was $17\frac{pJ}{\mu m^2}$. It was possible to target the last round of the AES with a jitter $\pm 5ns$. The DUA was a $0.13\mu m$ ASIC running at 25 MHz and had 6 metalization levels. Single bit flips were achieved using the refraction of the laser by the metalization levels.

In [101], a $1064nm$ pulsed laser was used to attack a $90nm$ chip from the backside. The laser pulse duration was variable and could be set as short as tens of ns and the output peak power varied from a few mW to $800mW$. The spot size was $2\mu m$ and the FFs that they attacked were $\sim 15\mu m^2$ and the SRAM cells were $3.25\mu m^2$.

In [102], individual bit flips were performed for both $45nm$ and $90nm$ different FPGAs with some limitations for the former. The two DUAs are Xilinx Spartan-6 and Xilinx Spartan-3A. The laser setup is based on a diode-pumped pulsed Nd:YAG solid-state laser source capable of emitting two different wavelengths, $532nm$ and $1064nm$. The laser pulse width is fixed at $800ps$ and has a maximum repetition rate of 1 kHz. A beam attenuator is used to control the amount of emitted beam energy. The DUAs were attacked from the backside using the $1064nm$ laser. The width of the beam is adjusted using a beam expander before the focusing 20x objective. A laser scanner is used to shift the beam position with a range of $\pm 0.25mm$ on the DUA at a step size of $100nm$. An xyz-motorized stage is used to position the area of interest in the DUA under the objective. The measured energy after all the losses from the optics ranged from $1nJ$ to $5\mu J$.

According to [14], frontside vs backside laser attacks required the characteristics mentioned in Table 3.3. However, [103] mentioned that no pre-thinning was required at all.

Table 3.3: Frontside and Backside Injection Characteristics according to [14].

	Frontside	Backside
Wavelength Used	Green ($532nm$) or IR	IR ($1064nm$)
Absorption Depth	$\approx 1\mu m$	$\approx 100\mu m$
Drawbacks	Presence of metal layers	Requires pre-thinning process

In [104–106], a high precision 6-axis translation table with $0.1\mu m$ step size was used to hold the DUA under a Neodymium: Yttrium Lithium Fluoride (Nd:YLF) $523nm$ laser to inject soft faults into electronic circuits by inducing transient errors. The used laser system can generate 10 $\frac{pulses}{sec}$ when performing automated attacking procedures. The pulsed laser was used to disable RO chains, flip latches repeatedly and inject transient errors in multiplying operations. Automating the process of translating the location to be attacked in the DUA, automating the alignment procedures, fine-tuning the LFI process according the technology node of the DUA, and synchronizing the observation of the system's response with the timing of the laser pulses were identified as challenges that need to be overcome should the LFI becomes automated.

3.4 Summary

The conducted literature review led to the observation of what the previous publications failed to study, which will be specified in this section. Most of the LFI setups mentioned in the literature aren't built from scratch and are either bought from *Riscure* or *AlphaNov* as a LFI solution. However, when a LFI system is built from scratch, the details of constructing such setup aren't clarified such as the presence of a collimator and/or usage of attenuators, utilization of TEM_{00}, alignment procedure of the laser beam with the optical path of the microscope, the effect of fiber optic cables on single-mode lasers, losses introduced to the laser beam through out its optical path from the laser source until hitting the DUA particularly when passing through different objectives, measuring the laser beam power before/after passing through the VMU and on the the surface of the DUA, etc. Also, tips and tricks in the decapsulation procedure, either from the frontside/backside, are not mentioned clearly while preparing the sample for LFI with the fact that some of the published literature outsource the decapsulation to companies like *Nisene*. Characterization and profiling for the used laser were not performed as well as proposing fault models for different DUAs or target areas in the DUAs. The number of pulses needed to do successful LFI for certain devices were discovered empirically and no fault model was formulated. Large laser beam spot sizes were touched upon in [90] but wasn't thoroughly investigated in the published literature.

A number of physical and logical countermeasures were compared in [107] in terms of the cost of implementation, production and testing, level of the protection provided against frontside and backside attacks, expertise and equipment required for a successful attack, impact on design flow and manufacturing process, reliability such as temperature effect on such countermeasures and the rate of false detections of attacks. Examples of those physical countermeasures are photodetectors, shielding, voltage/temperature sensors and other unique structures; while parity, majority and redundancy checks were presented as logical countermeasures. However, the path for these countermeasures to become practical was obstructed due to many reasons such as the large costly changes required in the fabrication process to implement such countermeasures. However it is important to fully understand the processing and experimental details of fault injection before one can understand how to develop countermeasures. The next chapter will detail the specific research objectives, methodology and setup.

Chapter 4

Research Objectives, Methodology and Setup

This chapter will list the research objectives, describe the methodology and provide an overview of the experimental setup. First, the objectives are stated followed by the methodology to be applied to achieve those objectives. Then the development of the laser bench setup is illustrated. This chapter also presents the custom-designed development PCB for ARM Cortex-M0 LPC1114FN28102, SEM measurements of die thickness, and illuminated images of lab decapsulated chips using SWIR and NIR backside imaging.

4.1 Objectives

Unlike previous research, this research tries to unfold the details of the sample preparation process for different DUAs, mainly the decapsulation steps for different types of packages and RE the structure of those DUAs. Also, this work precisely identifies the target clock cycle to inject faults in ICs using a class-IV laser so that one or more bit values are flipped to a desired value; for example, for the purpose of interrupting the boot up sequence to extract confidential information (e.g. secret key) or manipulating the execution order of a running program. A Demo LFI attack on a real application like AES and targeting different structures on chip such output of the Flash memory (i.e. instruction register), debug circuitry, UART and configuration bits were other objectives. Finally, this study proposes countermeasures to LFI, RE and/or sample preparation for the different structures mentioned above in commercial chips.

4.2 Methodology

Different types of hardware hacking may be performed according to how destructive or invasive the adversary can afford to be. Each path through the tree in Figure 4.1 (from the root to a leaf) indicates a possible scenario for the DUA hacking. For example, the left path in Figure 4.1 represents a context where the hacker has access to only one sole device, but no other devices or even similar devices are available. In this case, the hacker may not be able to perform invasive or destructive attacks such as RE since these attacks may harm the device. Alternatively, there may be only one device available however there may be some similar devices, see "Some Similar Devices" from the right branch at the "Only One Device" node of Figure 4.1. In this case the hacker may gain further information about part of the DUA from the similar devices. For example, a military communication device may be under attack and only one is available. However, the hacker may know or determine that a core in the military chip is also available commercially (for example a processor core). Thus RE of the similar device may reveal important information used to attack the single device. Alternatively, the hacker may gain access to only one device, yet also have access to another device which contains the same core. For example, an attacker may reverse engineer the similar device to determine the architecture of the processor core, thus enabling a targeted side channel analysis on the real DUA. On the right most branch at the root of the tree of Figure 4.1 is the opposite case where the hacker has access to "Many Devices". Identical devices are available for invasive destructive experimentations, in order to determine sufficient information to launch an effective attack on the device. Although the hacker may have access to multiple devices, these devices may either have different keys or a common key ("Key per device" or "Same Key" in Figure 4.1, respectively). Realistically multiple keys exist in the device, however we assume one key under attack in Figure 4.1 for illustration purposes only. Our LFI research will focus first on the "Many-Devices" approach till the LFI technique is mastered then "Only One/Some-Device(s)" approaches could be investigated using LFI. As an example, home IoT devices could be the DUAs as they are cheaply made and many of those IoT devices use the same symmetric key. Therefore, if the AES key is retrieved from one of those devices using LFI or any other attack method, it would impose a security threat/breach in all the other devices using the same key. On the other side, implementing public-key cryptography (PKC) or asymmetric encryption would support a unique private key per device and thus would protect other IoT devices if the private key of a single devices is compromised. So using PKC would allow more secured IoT devices, generally at the expense of slower execution speed of the asymmetric encryption algorithms. However, PKC won't protect a single device against LFI physical attacks which can be used to recover the private key for that single device.

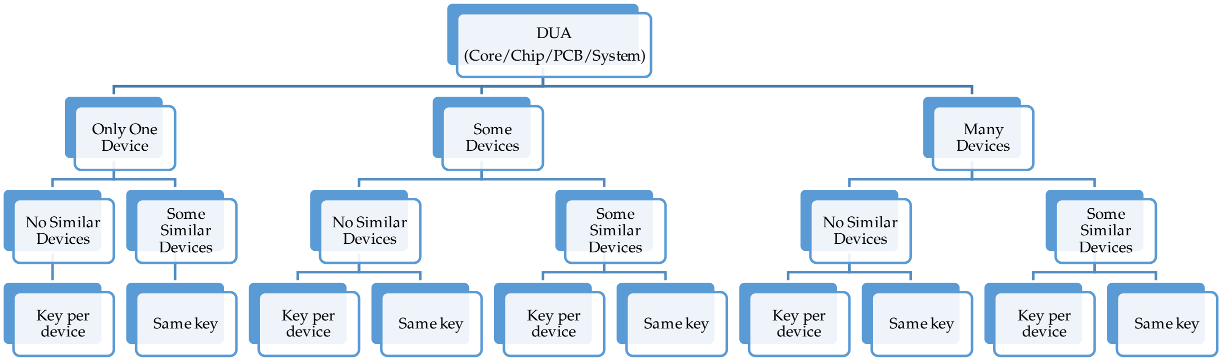


Figure 4.1: Scenarios of DUA (represented by paths through the tree).

4.2.1 General Flow

The flowchart shown in Figure 4.2 shows the main flow to successfully conduct a LFI attack on a sample DUA. Note that the DUA may be a standalone chip or already on a PCB.

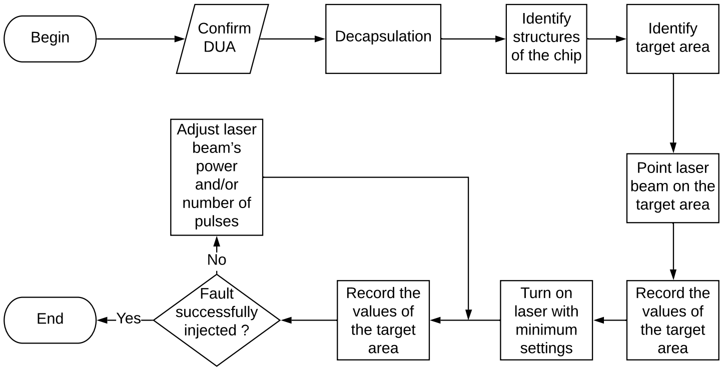

Begin
Confirm DUA
Decapsulation
Identify structures of the chip
Identify target area
Point laser beam on the target area
Record the values of the target area
Turn on laser with minimum settings
Record the values of the target area
Fault successfully injected ?
No
Adjust laser beam's power and/or number of pulses
Yes
End

Figure 4.2: General flow for LFI.

4.2.2 Flow for Decapsulation

The flowchart shown in Figure 4.3 demonstrates the proposed methodology for preparing a sample DUA through frontside or backside decapsulation. Further details and results of both frontside and backside decapsulation are presented in detail in section 5.1.

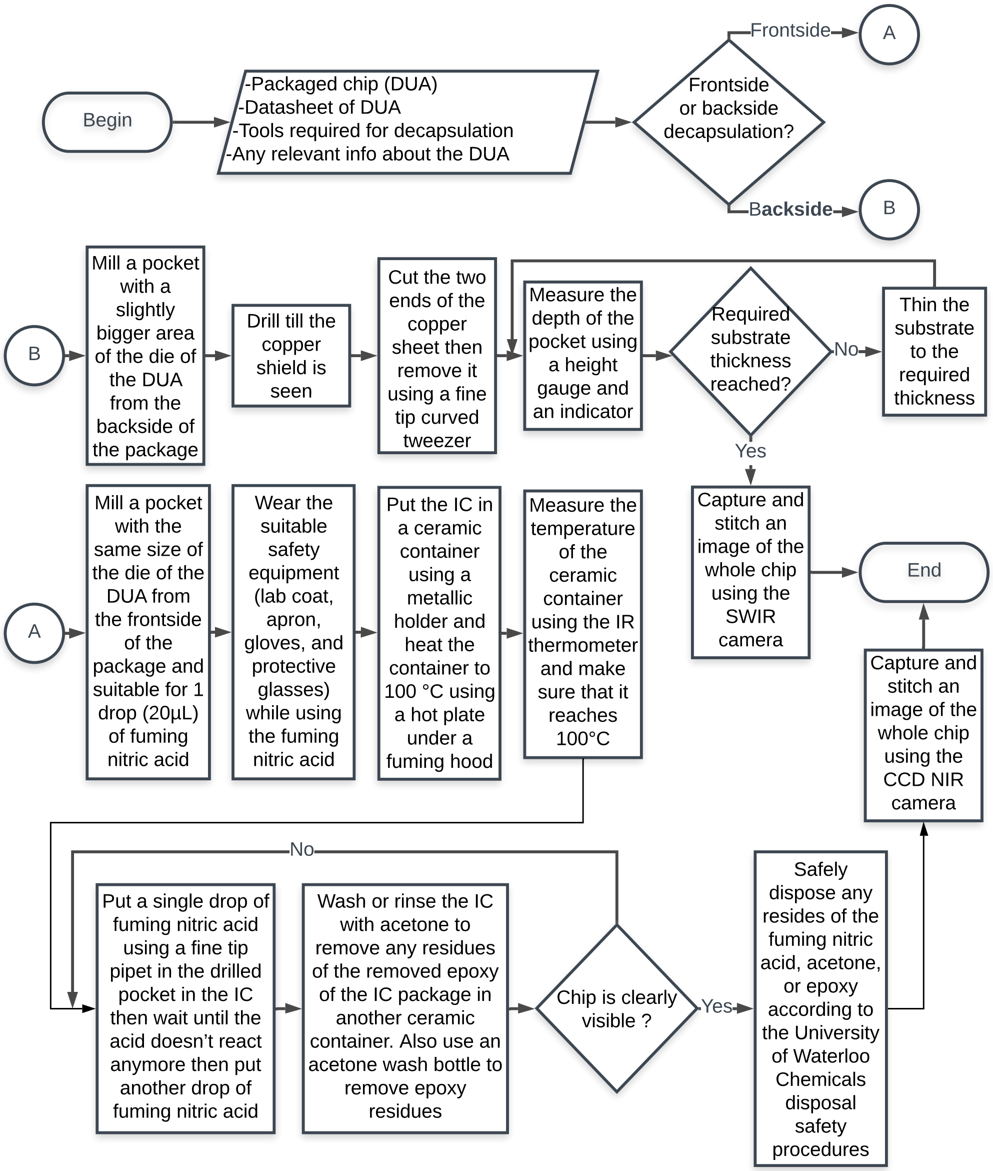

Begin
-Packaged chip (DUA)
-Datasheet of DUA
-Tools required for decapsulation
-Any relevant info about the DUA
Frontside or backside decapsulation?
Frontside
A
Backside
B
B
Mill a pocket with a slightly bigger area of the die of the DUA from the backside of the package
Drill till the copper shield is seen
Cut the two ends of the copper sheet then remove it using a fine tip curved tweezer
Measure the depth of the pocket using a height gauge and an indicator
Required substrate thickness reached?
No
Thin the substrate to the required thickness
Yes
Capture and stitch an image of the whole chip using the SWIR camera
End
A
Mill a pocket with the same size of the die of the DUA from the frontside of the package and suitable for 1 drop (20µL) of fuming nitric acid
Wear the suitable safety equipment (lab coat, apron, gloves, and protective glasses) while using the fuming nitric acid
Put the IC in a ceramic container using a metallic holder and heat the container to 100 °C using a hot plate under a fuming hood
Measure the temperature of the ceramic container using the IR thermometer and make sure that it reaches 100°C
Capture and stitch an image of the whole chip using the CCD NIR camera
No
Put a single drop of fuming nitric acid using a fine tip pipet in the drilled pocket in the IC then wait until the acid doesn’t react anymore then put another drop of fuming nitric acid
Wash or rinse the IC with acetone to remove any residues of the removed epoxy of the IC package in another ceramic container. Also use an acetone wash bottle to remove epoxy residues
Chip is clearly visible ?
Yes
Safely dispose any resides of the fuming nitric acid, acetone, or epoxy according to the University of Waterloo Chemicals disposal safety procedures

Figure 4.3: Methodology for decapsulation.

4.3 Experimental setup

The equipment setup for LFI and PEM, DUA sample preparation, laser pulse characterization and induced switching current setups will be described in this section.

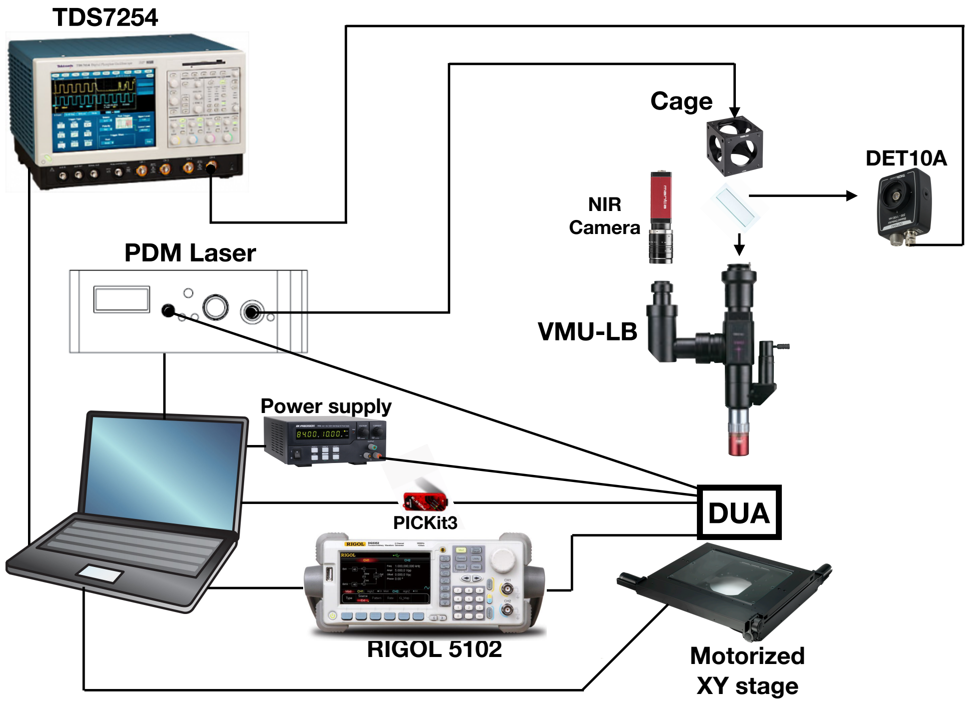


Figure 4.4: Experiment setup showing connections between PC and other devices such as XY-stage, RIGOL 5102, power supply, PDM laser, VMU, camera, photo-detector and oscilloscope in order to attack the DUA.

4.3.1 LFI setup

The laser is mainly composed of AlphaNov PDM+ laser module which is a 3W fundamental mode 1064nm laser beam using two channels (each channel delivers approximately 1.5W) [12]. The laser source has a maximum peak current of 4000mA [108] of which a certain percentage is used to change the amplitude of the tailored laser pulse according to the experiments conducted. We started only by using a single channel as it was sufficient to inject faults successfully in PIC16F687 through the 5X objective then used both channels with the 50X objective to be able to inject faults in the ARM Cortex-M0 LPC1114FN28102. The technology node listed in Table 4.1 and 4.2 for both DUAs is probably the reason for needing a higher power intensity to be able to inject faults in the latter DUA. The trigger pulse to the laser beam emission delay is $\sim 20ns$ as reported by Alpha. AlphaNov's PDM+ laser module is controlled via a script using a dynamic-link library (DLL) provided by AlphaNov.

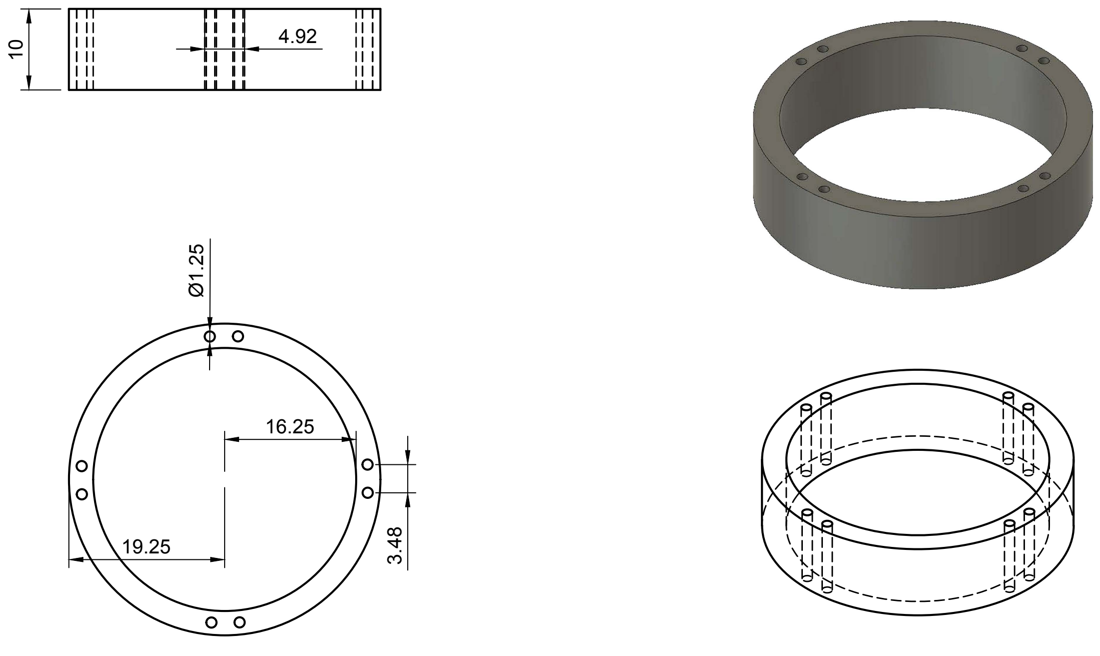


Figure 4.5: IR ring mechanical drawing with dimensions in *mm*. A real picture of the IR ring with the IR LEDs mounted through it is shown in Figure 5.13.

The laser bench design [109, 110] utilized an optical workbench with laser enclosure, AlphaNov PDM+ laser module, motorized XY-stage, Mitutoyo VMU and a 5X objective.

The laser enclosure has a glass window of optical density that protects the human eye from the 1064nm wavelength of the laser and allows looking at the laser while in operation although safety goggles are also worn while aligning the laser and during operation. The laser bench is Class 1 laser (which means that laser is safe under all conditions of normal use as specified by the IEC 60825-1 standard); hence when the enclosure is opened, an interlock mechanism turns the laser off. The laser source is connected to the Mitutoyo VMU using a fiber optic cable.

A Mitutoyo VMU is used to hold the 5X M Plan Apo NIR objective which allows the laser beam to be focused. The VMU is attached to a Z-Column that allows manual movement of the VMU in the vertical direction. A motorized XY-stage, with a precision of 0.1μm, is attached to the Z-Column. The movement of the motorized XY-stage is controlled through M-code commands. The motorized stage is used to adjust the position of the DUA under the VMU. The VMU has one laser mount and one C-mount for the camera. The laser beam is injected through the VMU and is viewed on a laptop screen which is connected to the camera. Two cameras were utilized at different times, a CCD NIR camera (the Allied Vision Manta G-145B NIR) for frontside/backside imaging and a SWIR camera (the Xeva-1.7-320 TE3) for PEM backside imaging.

As the beam passes through a choice of Mitutoyo lenses, it gets reduced by the lens' zoom factor and loses a big part of its energy. Since the 1064nm laser is not visible, we had to use a high power IR sensor card to view the laser beam spot as illustrated in Figure 4.6. The card we used has minimum detectable power of $1\frac{mJ}{cm^2}$ and a damage threshold of $10\frac{mJ}{cm^2}$ when it is exposed to a 1064nm pulsed laser. The transmission efficiency for the VMU for the 1064nm is $\sim$ 54% which was measured using a S146C power meter.

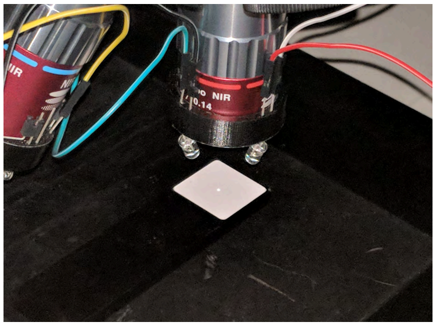


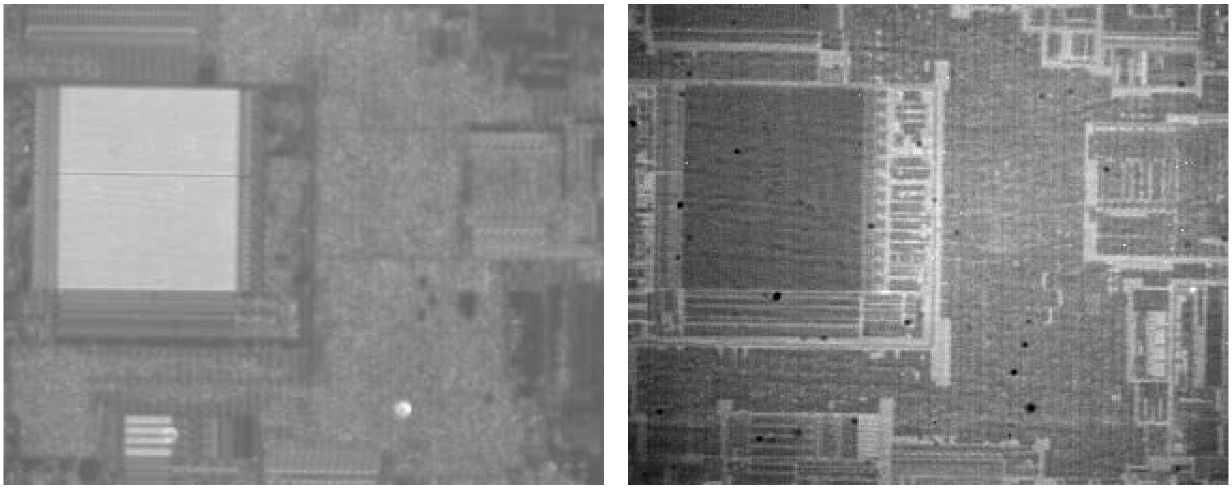
Figure 4.6: Laser spot using the Mitutoyo 5X M Plan Apo NIR objective on a high power IR sensor card.

Several experiments were conducted to measure various parameters controlling the timing and the energy of the laser pulse. A DET10A APD was used to capture a $\sim$ 4%

reflection of the laser beam. A microscope slide was used to reflect the $\sim$ 4% of laser beam out of the normal laser beam's path towards the photo-detector. The photo-detector was connected to a TDS7254 oscilloscope for various measurements such as the laser pulse amplitude, laser pulse width, delay between triggering the laser and laser pulse generation, etc. Both the fiber optic cable coming from the PDM+ laser source and the photo-detector are connected to a cage that contains a 45° tilted microscope slide as shown in Figure 4.4.

In contrast to the straight-forward imaging for a frontside decapsulated DUA (by just placing the DUA under the microscope and using the NIR camera), the backside imaging needs special IR illumination to make the silicon substrate transparent with respect to the SWIR or NIR cameras. An IR ring was designed to hold four MT51060-IR LEDs of a $1060nm$ peak emission wavelength to provide backside imaging. Figure 4.5 shows the 3D printed IR ring with the 8 holes to hold the LEDs. The LED ring is attached to the objective with the LEDs pointing towards the DUA.

Also an IR illuminator, MHAB-100W-IR, was used to compare its effect on the image captured by the SWIR camera under the Mitutoyo setup vs the custom-designed IR ring. The best achievable images for both IR sources are shown in Figure 4.7. The backside imaging for the RISC-V FE310-G002 using the illuminator is also shown in Figure 4.8. Details of decapsulation of this chip is given in the next chapter.



(a)

(b)

Figure 4.7: SWIR image of PIC16F687 under the Mitutoyo setup using a)custom-designed IR Ring and b)MHAB-100W-IR illuminator.

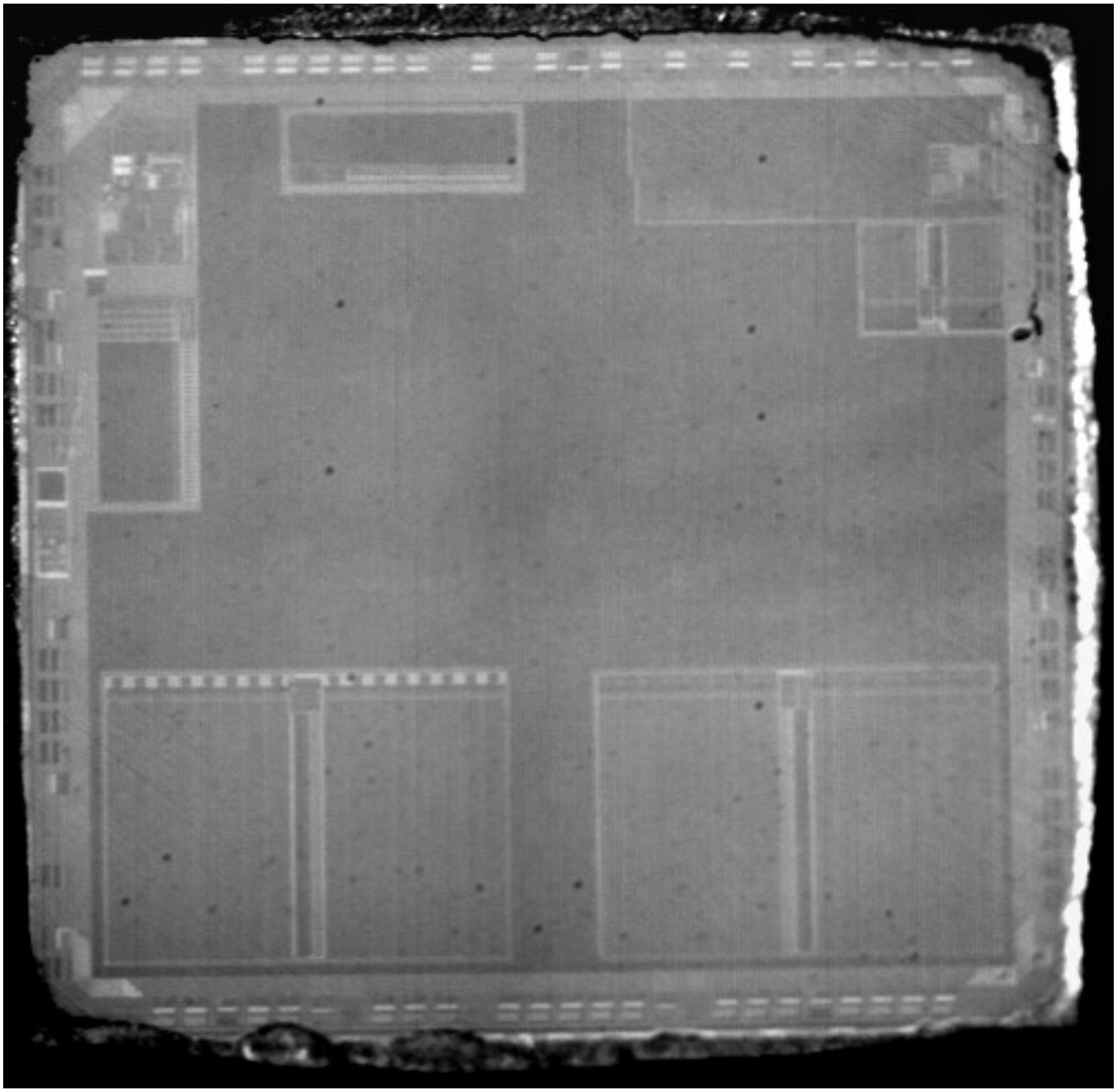


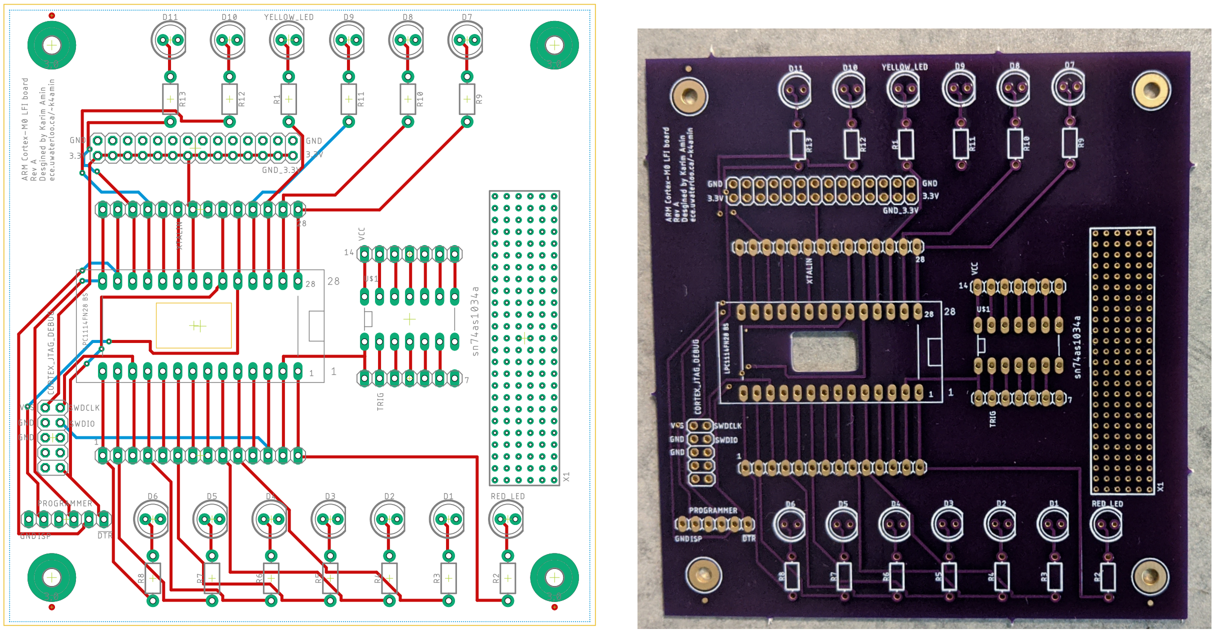
Figure 4.8: SWIR image of the RISC-V backside using the Mitotoyo setup with the MHAB-100W-IR illuminator.

A programmable power supply, BK PRECISION 9103, is used to control the power supplied to the DUA (power ON/OFF, over-voltage, under-voltage) which is mounted on a development board while running different experiments. Also, an arbitrary waveform generator, RIGOL 5102, is used to supply the DUA with the external clock needed to operate at a certain desired frequency or inject an arbitrary number of clock cycles. All the connections between the PC and other devices are through USB connections except the NIR camera is via LAN. A master Python script was used to control the equipment (erasing, programming and dumping the DUA's memory, move XY-stage, control power supply, arbitrary waveform generator, capture images using the NIR camera, capture wave-forms using TDS7254 and send them to the PC) and automate running different experiments.

LFI triggering for the ARM Cortex-M0 LPC1114

Unlike the PIC16F687 which was used to directly trigger the laser using one of its I/O pins, the ARM Cortex-M0 LPC1114 needed additional circuitry to be able to trigger the laser source. The ARM I/O pin could not be used since it did not have sufficient current. A hex driver, SN74AS1034AN, has been utilized to drive the 60Ω of the laser trigger in order to achieve the necessary current. The hex driver SN74AS1034AN offers high capacitive-drive capability to provide the sufficient current to trigger the laser source. The 60Ω laser trigger needed at least 37mA for a 2.2V drop and the SN74AS1034AN provided 48mA. The hex driver SN74AS1034AN delay is $\sim 8ns$ as per the data-sheet.

In order to make the breadboard setup more robust for the ARM, we designed a development board for the ARM Cortex-M0 LPC1114FN28102 that integrates the SN74AS1034AN hex driver and provides both programming and debugging headers. The layout of the design of the PCB, manufactured PCB and soldered PCB are shown in Figures 4.9a, 4.9b and 4.10 respectively. Moreover, to avoid the any unintentional movements of the DUA under the VMU, the PCB is fixed on motorized XY-stage using steel standoffs, magnets and a custom made steel plate that fits on the XY-stage as shown in Figure 4.10.



(a)

(b)

Figure 4.9: a) Design of the ARM Cortex-M0 LPC1114FN28102 PCB layout b) Manufactured PCB.

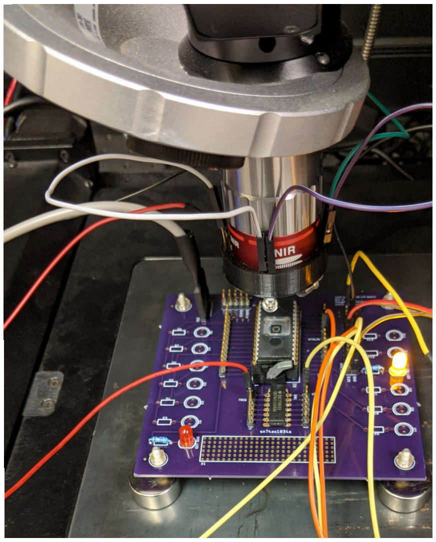


Figure 4.10: Soldered PCB with a functional backside decapped ARM LPC1114.

4.3.2 PEM setup

The PEM analysis of the DUA was investigated using the setup introduced in [6] which is depicted in Figure 4.11. The main different components from the LFI setup were an Olympus microscope (instead of the Mitutoyo VMU) with a manual XYZ-stage, a Xeva-1.7-320 TE3 SWIR camera, 10X objective (LMPLN10XIR) and a halogen lamp for illumination. The Olympus microscope has a single camera port, which is optimized for IR wavelengths, achieving a transmission efficiency up to $\sim 80\%$. This is considered a low cost setup used for PEM [6]. Thus unfortunately the DUA had to be removed from the LFI setup and transferred to the PEM setup whenever photon emissions were required since the VMU was not suitable for photon emissions (transmission efficiency up to only $\sim 54\%$).

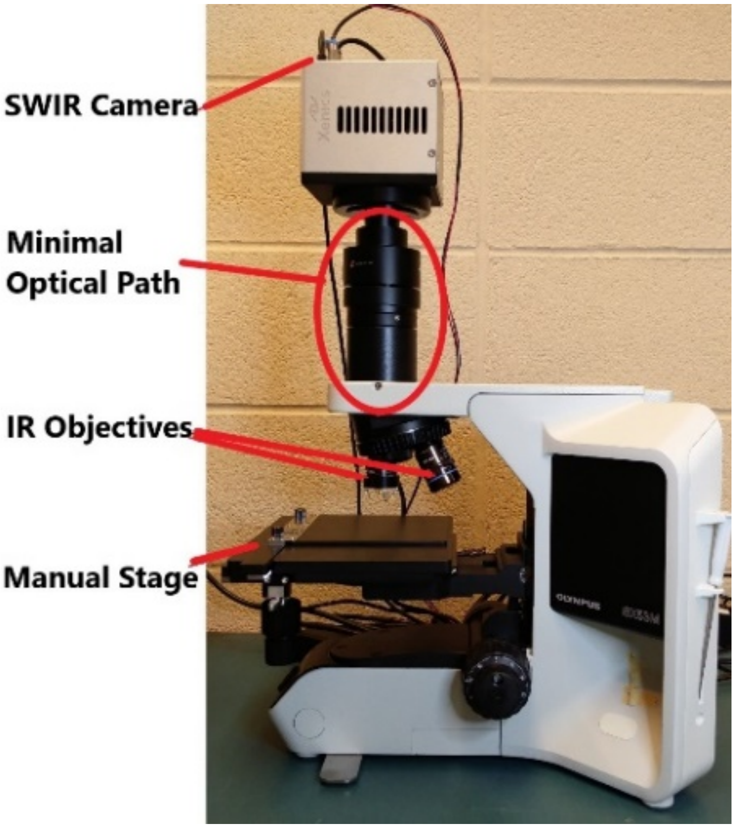


Figure 4.11: In house PEM setup [6] consisting of Xenics SWIR camera and Olympus IR optimized microscope and objective lenses.

4.3.3 Sample preparation

Before trying to inject faults in a sample DUA using LFI, the sample has to be prepared first through frontside/backside decapsulation depending on the desired method of attack and the available resources. According to the reviewed literature, the 1064nm laser is suitable to attack from the backside since the silicon is nearly transparent to that wavelength. So in our experiments the DUA was backside decapped mechanically using a low cost desktop CNC machine, Nomad 883 Pro, to remove first the epoxy then the copper shield and finally a glue layer before the silicon of the die is exposed. The glue (probably a thermal paste to provide better heat conduction between silicon and copper) could be removed using a plastic scrapper to avoid damaging the silicon. The final step is silicon polishing using a wool felt tool and a $20\mu L$ droplet of a $0.04\mu m$ Colloidal Silica solution so that the silicon surface is shiny and provides appropriate optical roughness for the 1064nm wavelength. Substrate thinning (100's μm) may be required before polishing (10's μm) but was not needed in our case. Enough photons were captured in the pre-descibed PEM experimental setup without any thinning. Figure 4.12 shows the backside of the DUA with some structures being identified. To measure the substrate thickness, a Hitachi S-3000N SEM was used. Figure 4.13a and 4.13b show the thickness of the die and the active layer respectively. The measurements acquired using the SEM were beneficial to the sample preparation process to know how much thinning and/or polishing of the DUA could be tolerated without inflicting damage to the active layer and hence to the functionality.

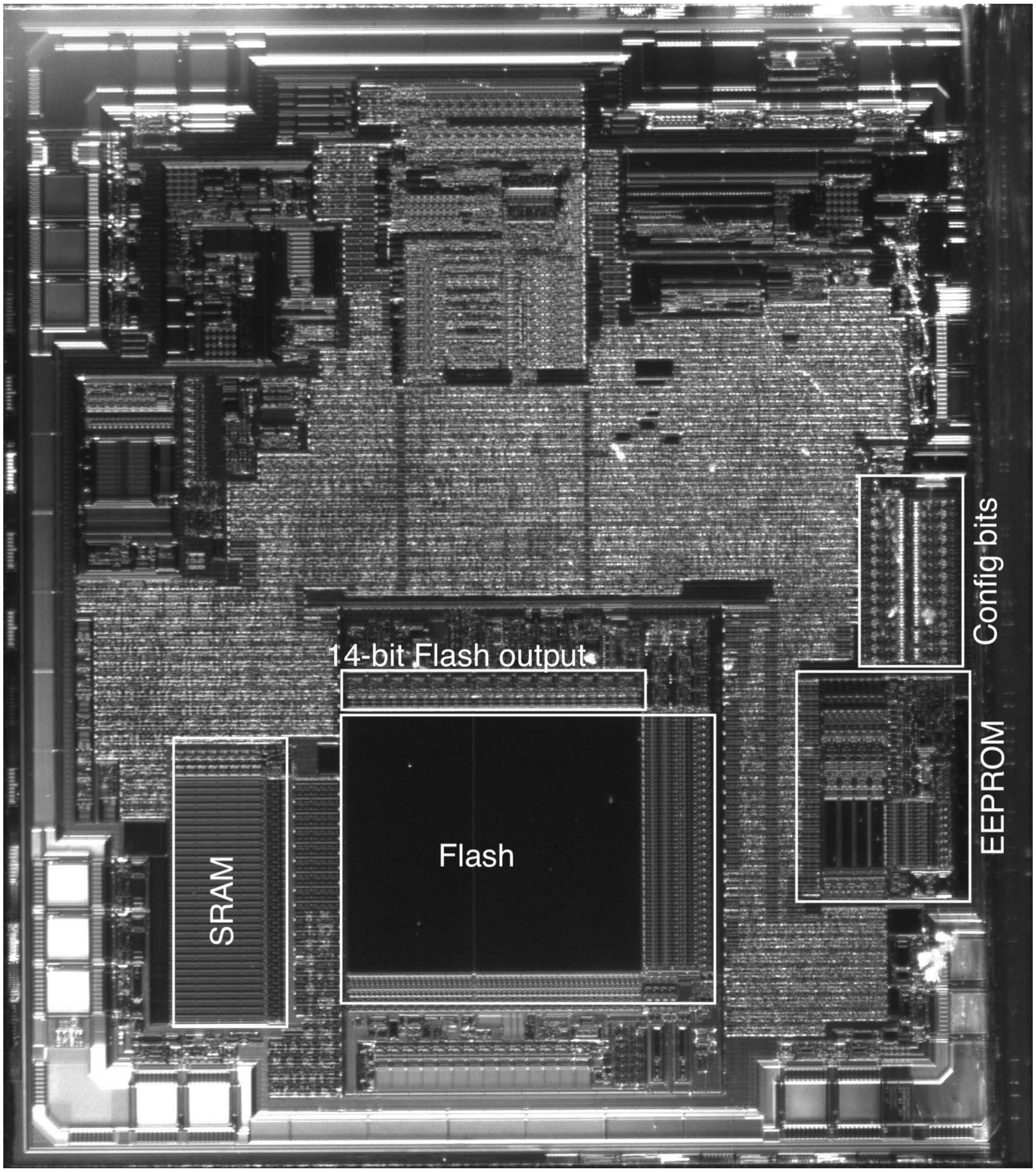
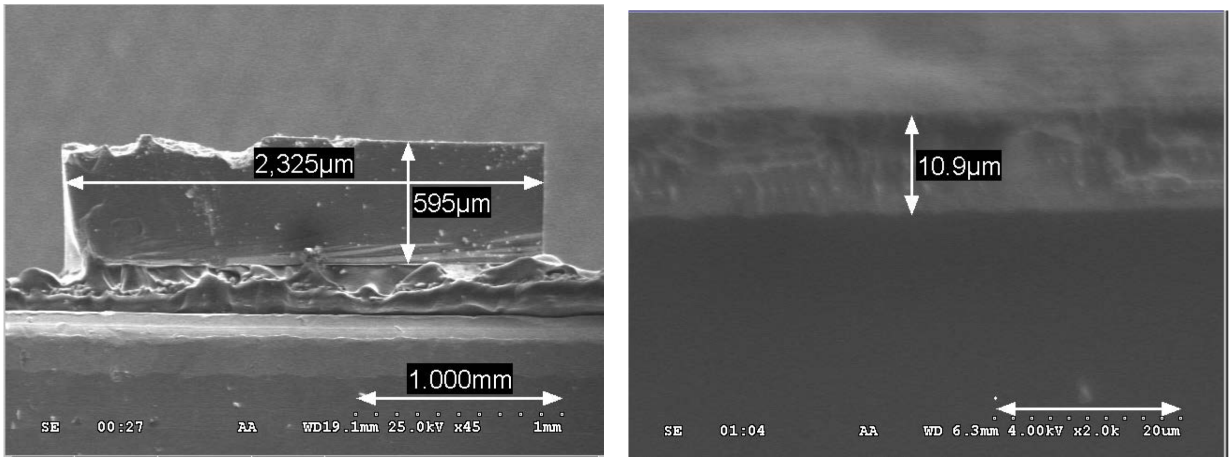


Figure 4.12: PIC16F687 backside view captured using the 5X Mitutoyo lens and the NIR camera with different structures identified such as SRAM, Flash, EEPROM and configuration bits.


2,325μm
595μm
1.000mm
SE 00:27 AA WD19.1mm 25.0kV x45 1mm
10.9μm
SE 01:04 AA WD 6.3mm 4.00kV x2.0k 20um

(a) (b)

Figure 4.13: a) Cross-sectional view of the substrate thickness captured using SEM. b) Active layer thickness is the top part of the die and represents only ∼ 2% of the whole die thickness.

4.3.4 Sample orientation

X-ray imaging of the DUA such as in Figure 4.14 was used alongside with backside imaging to know the orientation of captured backside image and how the die is connected to the pins. Further details of pin mapping to the captured backside image will be discussed in chapter 6.

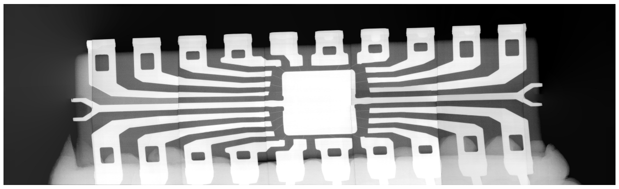


Figure 4.14: X-ray stitched image for PIC16F687 Frontside.

4.4 Devices under Attack

Basic introduction to the DUAs will be provided in this section. Further details about each DUA will be provided in chapters 5 and 6. The technology node information was obtained through contacting the manufacturers, Microchip Technology for PIC16F687-I/P and NXP for ARM Cortex-M0 LPC1114FN28102. In case of RISC-V FE310-G002, the technology node was open source [111] and provided by SiFive.

4.4.1 PIC16F687

The PIC16F687 [9] applies a 2-stage pipeline with a Harvard structure. The two-stage instruction pipeline overlaps the fetch and execution of instructions. Basic specifications of PIC16F687 are show in Table 4.1. The specific chip used was PIC16F687-I/P and it was a dual in-line package (DIP).

Table 4.1: Basic specifications for PIC16F687.

Parameter	Value
Technology node	$0.45\mu m$
Program Memory Type	Flash
Program Memory	3.5 KB
CPU Speed	5 MIPS
RAM	128 bytes
Operating Voltage Range	2 to 5.5V
Data EEPROM	256 bytes

4.4.2 ARM Cortex-M0 LPC1114FN28102

The ARM Cortex-M0 [112] processor is a 32-bit RISC processor, with a 3-stage (fetch, decode and execution) pipeline von-Neumann (load-store) architecture with a single bus interface. Basic specifications of LPC1114FN28102 [113] are show in Table 4.2. Further details on this processor are provided in chapter 6.

Table 4.2: Basic specifications for ARM Cortex-M0 LPC1114FN28102.

Parameter	Value
Part no.	NXP LPC1114FN28/102
Technology node	140nm
Maximum Frequency	50MHz
RAM	4kB
Operating Voltage Range	1.8 to 3.6V
ROM (Flash)	32kB

4.4.3 RISC-V

RISC-V [114] is a load–store architecture: instructions address only registers, with load and store instructions conveying to and from memory. Basic specifications of RISC-V are show in Table 4.3. The specific chip used was the SiFive's FE310-G002 [111] chip featuring E31 RISC-V core.

Table 4.3: Basic specifications for RISC-V with probable structures in Figure 4.8.

Parameter	Value
Technology node	180nm TSMC CL018G process, 6 metal layers, 1 poly layer
Instruction SRAM (icache)	16 KB (one of same sized blocks at bottom of pic - likely target)
Data SRAM	16 KB (tightly coupled data memory - located beside icache)
OTP	8 KB (maybe on top left)
boot code ROM	8 KB (maybe dual banks on top right)
Registers	likely at top middle
Flash	off-die

The next two chapters will provide details on the PIC16F687 and ARM Cortex-M0 LPC1114FN28102, specifically analyzing decapsulation and LFI. The decapsulation of the RISC-V and a FPGA are also briefly examined in chapter 5.

Chapter 5

Results for MicroChip PIC16F687

Although publications have generally discussed the topic of security of MCUs, attack techniques are diverse and published LFI work provides few and superficial details about the used experimental setup and methodology. This chapter contributes to state of the art LFI by:

- providing a detailed LFI experimental setup and methodology [109, 110].
- presenting a combined quiescent PEM and LFI methodology to provide exact laser positioning and laser pulse timing for controlled fault injections.
- providing a new attack on the AES revealing the secret key after only one short laser pulse.

In the following sections, we'll discuss the results for both frontside and backside decapsulation, measurements and effects of laser pulses on a DUA and attacks performed on that DUA.The nominal operating conditions for running different experiments are as follows unless otherwise mentioned within a certain experiment. The DUA analyzed in this chapter is the PIC16F687 [115] connected to a 5V supply voltage and running at 20MHz using an external clock supplied by a waveform generator. This chapter also provides a brief discussion of the decapsulation of two other chips, specifically the Spartan-6 FPGA XC6SLX9 and RISC-V FE310-G002 chips.

5.1 Frontside and backside decapsulation trials

The DUA used for initial decapsulation experiments was PIC16F687. Its package is similar to the drawing in Figure 5.1. The black material used in manufacturing the package is called epoxy. Pure epoxy is practically transparent with the laser at the 1064nm wavelength [116]. The optical absorption depth of cold epoxy resin was found to be approximately 4mm [116]. However, the package is manufactured using reinforced composites of epoxy hence impurities will burn when exposed to the laser. Thus it was necessary to remove the packaging. Also this supported imaging of the structure of the die. The die image helped in determining the target area either from the front or the backside for the laser spot.

The uniform pocket in the middle of the PIC, as illustrated in Figure 5.1, was required for both frontside and backside decapsulation using a milling or a Computer Numeric Control (CNC) machine. Regarding the frontside, the uniformity of the pocket was very crucial so that the fuming nitric acid etches the epoxy [117] uniformly at the middle of the packaged PIC and finally reveals the IC without dissolving the copper bonding wires. Different types of bonding wires might require different types of chemicals in order to be able to etch the epoxy without dissolving the bonding wires [7]. The frontside decapsulation video produced by a hardware engineer at Google Life Sciences [118] was extremely helpful in starting to develop a detailed decapsulation procedure. I took the technique in [118] as the basis for the procedure that I followed hereunder after making some modifications to the technique shown in the video.

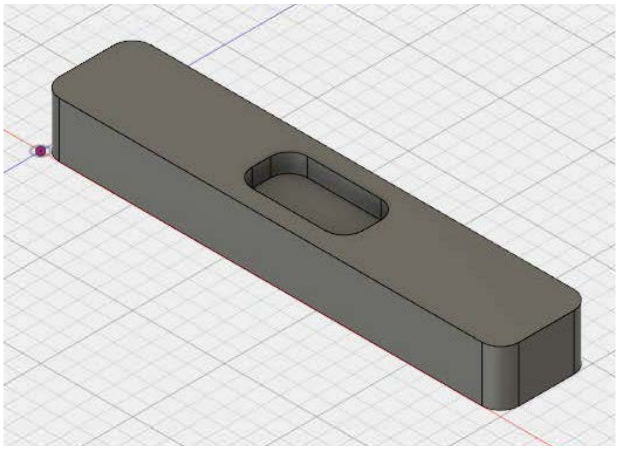


Figure 5.1: Chip pocket demonstration.

The following two sub-sections provide detailed steps to decapsulate chips from both front and backsides. Before going into the frontside decapping details, we would like to acknowledge both Richard Barber and Professor Peter Levine who allowed us to use the fuming hood in their lab at University of Waterloo to carry out the chemical etching process which involved using fuming nitric acid.

The frontside depackaging involves both mechanical drilling and chemical etching. Unlike previous research, the details of frontside decapsulation are exposed and the procedure for the frontside decapsulation is provided in the next subsection.

5.1.1 Frontside proposed procedure

1. Drill a pocket in the package of the DUA as shown in Figure 5.1 suitable for one droplet ($20\mu L$) of fuming nitric acid.
2. Wear the suitable safety equipment (lab coat, apron, gloves, and protective glasses) while using the fuming nitric acid.
3. Put the IC in a ceramic container using a metallic holder and heat the container to 100℃ using a hot plate inside a fuming hood.
4. Measure the temperature of the package of the DUA using an IR thermometer and make sure that it reached 100℃.
5. Put a single drop of fuming nitric acid using a fine tip pipette (shown in Figure 5.2) in the drilled pocket then wait until the acid doesn't react anymore (i.e starts to dry out) then put another drop of fuming nitric acid.
6. Wash or rinse the DUA with acetone to remove any residues of the removed epoxy of the package in another ceramic container.
7. Repeat 5 and 6 until the die is clearly visible.
8. Safely dispose any resides of the fuming nitric acid, acetone, or epoxy according to the University of Waterloo Chemicals disposal safety procedures.

Figure 5.2: Pipette used for frontside decap.

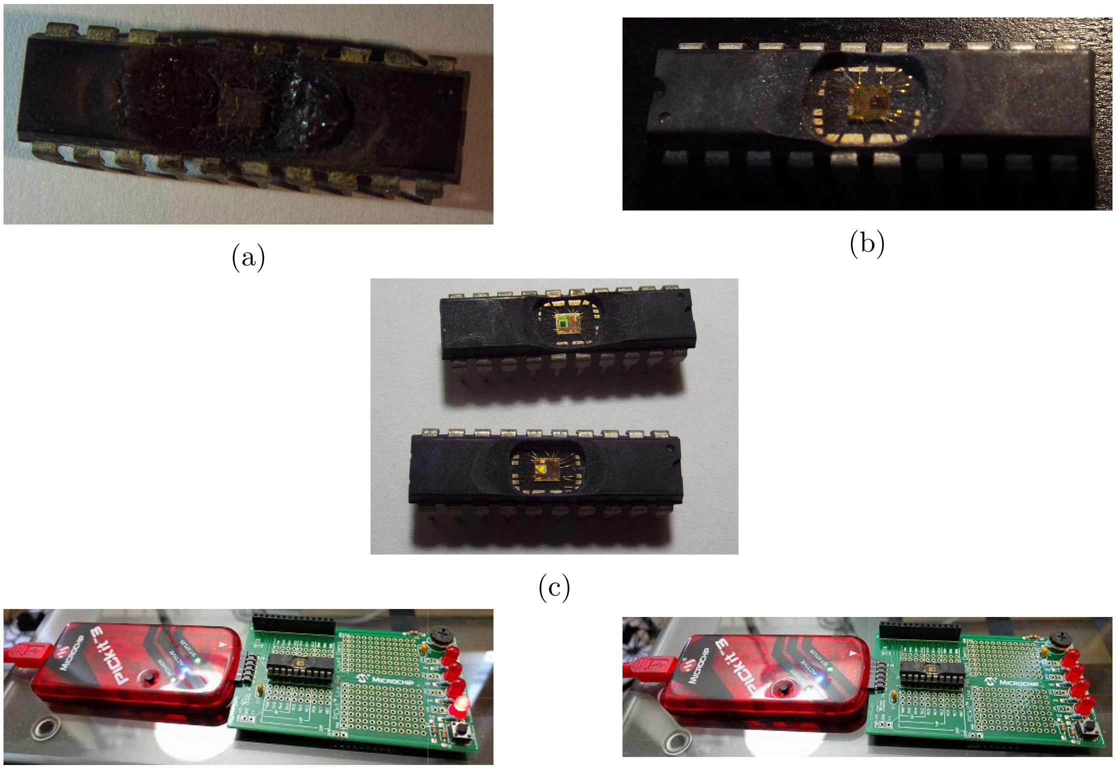
Pocket dimensions for two DUAs are mentioned in Table 5.1. Those dimensions were used to try different cavity sizes to improve the quality of the both the frontside and backside decapped samples. A $1mm$ depth for the cavity from the frontside was adequate. Moving down further might risk damaging the die if the cutting tool touches it or even breaking the bonding wires through the vibrations of the rotating cutting tool at 2500 rpm (rotation per minute). More precise pocket drilling using automatic CNC machines will lead to better chemical decap by only exposing the surface of the die with minimal exposure of the surrounding bonding wires. This was proven to be true in subsequent trials. Also, the dimensions mentioned in Table 5.1 were beneficial in estimating an approximate substrate thickness which lead to guided substrate thinning/polishing process carried in the machine shop at the University of Waterloo and at the Department of National Defence (DND) as well.

Early frontside decapsulation trials for PIC16F687 are shown in Figure 5.3. The first sample was totally damaged by the etching acid and wasn't functional as well as nothing of the die was barely visibly except some damaged bonding wires as shown in Figure 5.3(a). Then two samples were successfully decapped while still functional as shown in Figure 5.3(b) and Figure 5.3(c). In the 3^{rd} trial, four out of five samples were working.

Table 5.1: PIC pocket dimensions for different DUAs.

	PIC16F1827	PIC16F687
Original packaging thickness	$3.28mm$	$3.31mm$
Depth of pocket (frontside)	$1511\mu m$	$1511\mu m$
Depth of pocket (backside)	$1397\mu m$	$1397\mu m$
Width of the die	$2.4mm$	$2.0mm$
Length of the die	$3.2mm$	$2.2mm$

Since the VMU couldn't capture the whole area of the frontside of the decapped chip in one shot, individual images were captured as shown in Figure 5.4 then stitched together using a software program [119] offered by Microsoft Research. Figure 5.5 shows the whole stitched image. Figure 5.6 compares the level of detail gained from the 5x vs the 50x NIR objectives.



(a)

(b)

(c)

(d)

(e)

Figure 5.3: Unsuccessful frontside decapsulation trial in (a) and successful trials in (b) and (c). One of the samples was programmed and tested as shown in (d) and (e).

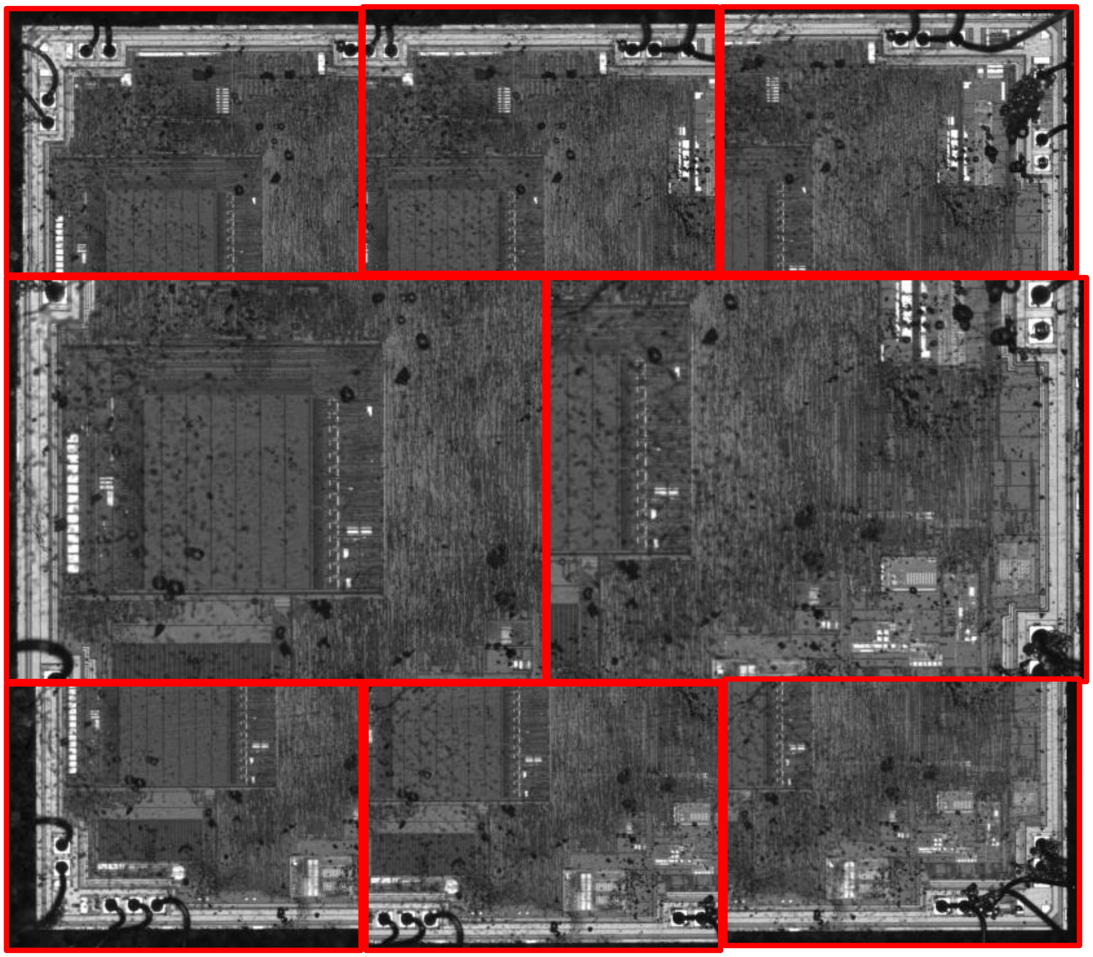


Figure 5.4: Individual frontside images for part of the PIC16F687 chip from using the CCD NIR camera. Captured 8 overlapping images of different parts of the de-packaged PIC16F687 using the 5X NIR objective.

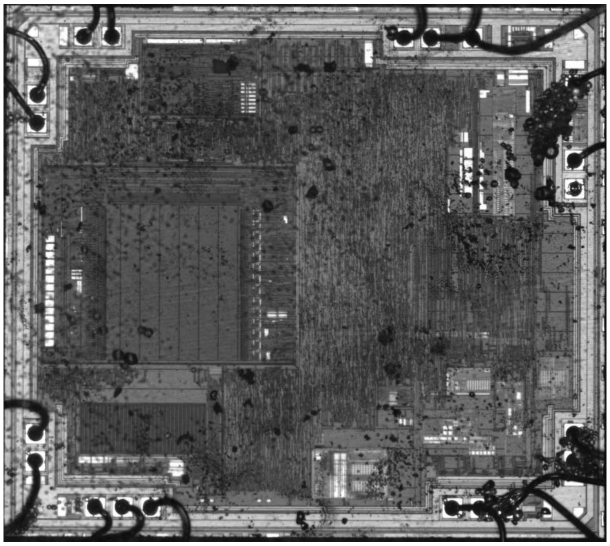


Figure 5.5: Stitched image of the whole frontside decapped PIC16F687 chip.

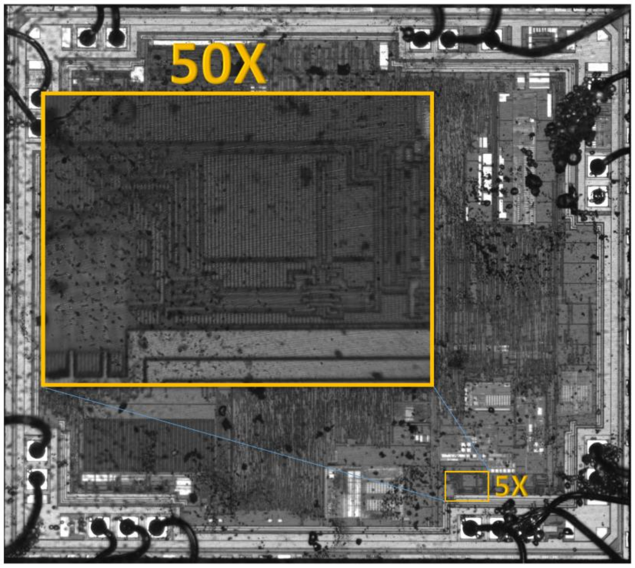


Figure 5.6: Frontside decapped PIC16F687 with M Plan APO NIR 5X and 50X objectives.

However, residues of the dissolved epoxy remained on the decapped chip and introduced a fire-hazard risk if exposed to the laser as well as obstructing the visibility of the structure of the die as shown in Figure 5.5. In the 4^{th} trial, while trying to improve the decapsulation technique, only one of five samples was functional with failure of removing the residues of the epoxy. It has been noticed that applying acetone through the nozzle at a higher pressure leads to removing epoxy residues as shown in Figure 5.7. An ultrasonic acetone path might not result in the central spot shown in Figure 5.7 (to be discovered).

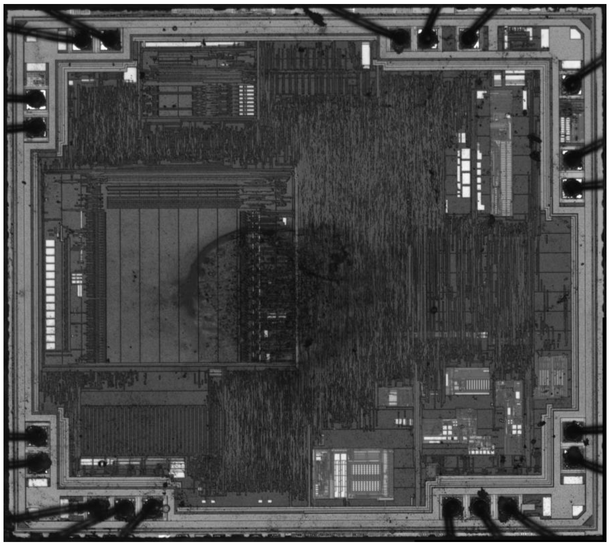


Figure 5.7: Clean sample of frontside decapped PIC16F687 with an acetone footprint in the center probably due to high pressure.

When the same frontside decapsulation procedure mentioned in section 5.1.1 was conducted on PIC16F1827 several times, the bonding wires always disappeared as shown in Figure 5.8 and green vapors appeared, most probably, while the fuming nitric acid was etching away the bonding wires. According to [7], bonding wires made of different materials require specific chemicals to avoid dissolving the bonding wires while etching away the epoxy. As a future work, we'll confirm the material of the bonding wire of PIC16F1827 which is suspected to be silver since the silver is dissolved when exposed to fuming nitric acid.

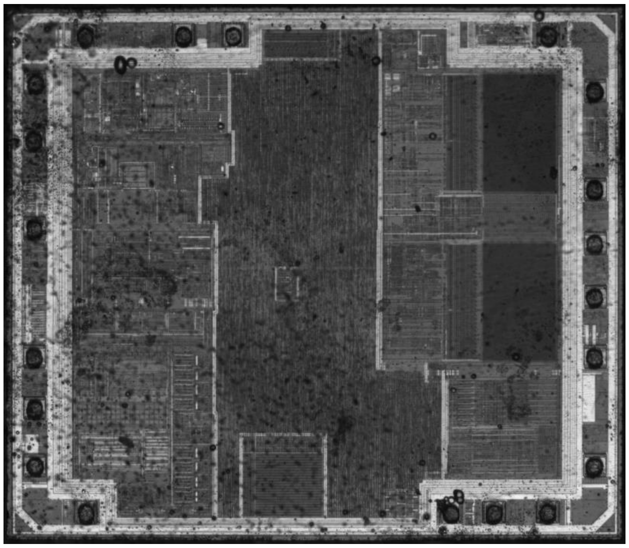


Figure 5.8: Bonding wires of the front side decapsulated PIC16F1827 chip were dissolved during the chemical etching process. Different chemicals may be needed to etch the epoxy without dissolving the bonding wires made of a material other than copper according to [7].

5.1.2 Backside proposed procedure:

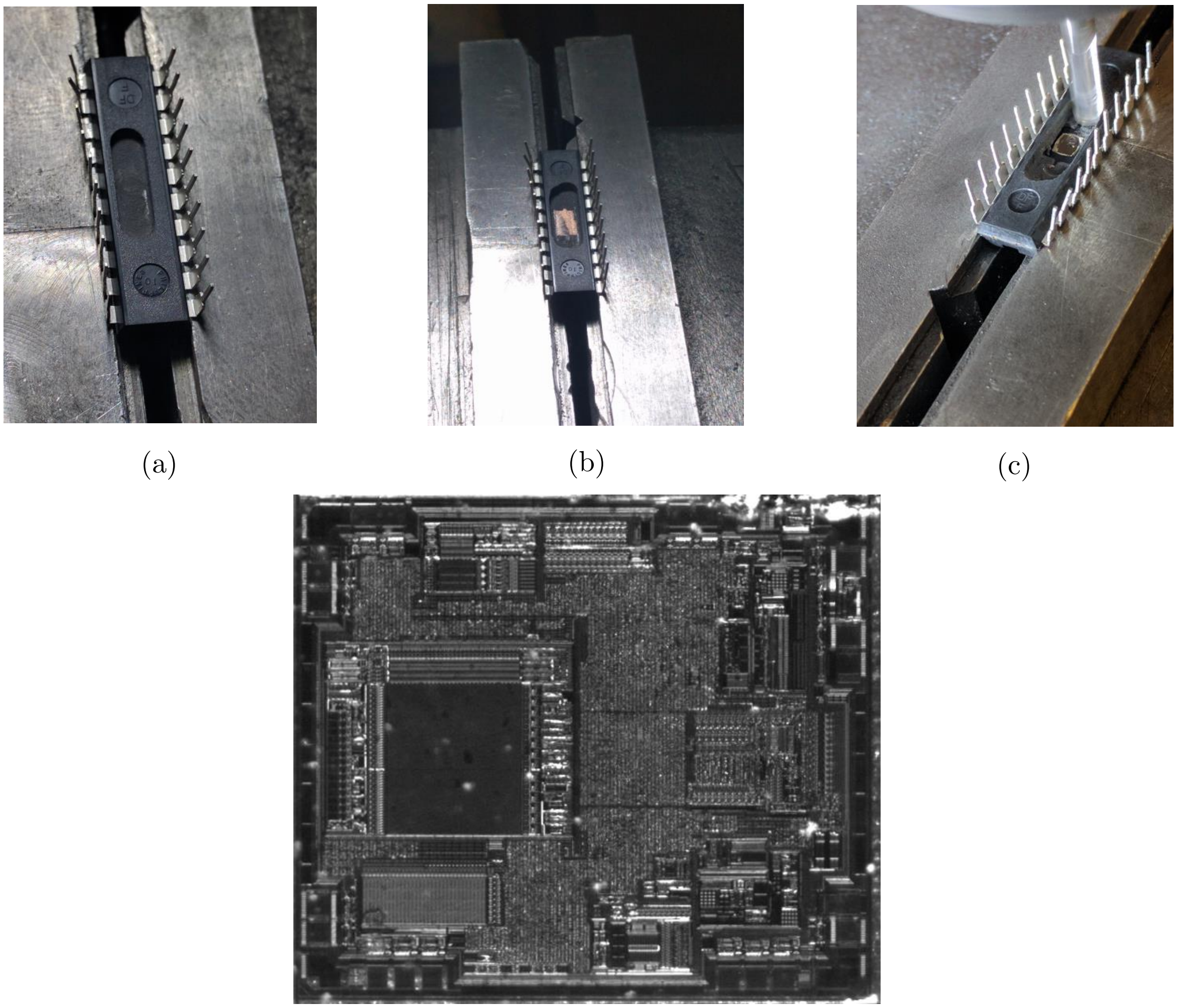
The backside depackaging, unlike the frontside, is only mechanical and doesn't involve any etching chemicals. Unlike previous research, the details of decapsulation are exposed below. The procedure for the backside decapsulation is listed in the following steps:

1. Drill a pocket in the backside of the DUA's package till the copper shield is visible.
2. Locate the two narrow copper connections at two side of the copper shield and cut those connections.
3. Remove the copper shield. The copper is glued to the substrate so use a small plastic tool and some lens cleaning liquid to remove the remaining glue till you have a shiny substrate surface.
4. Perform substrate thinning if needed. If the laser isn't powerful enough, substrate thinning will be required.
5. Do polishing to get clean structural image of the IC using the SWIR camera.

The captured images of the DUA shown in Figure 5.9 were drilled using milling machines (controlled manually) in the student machine shop at University of Waterloo. CNC machines can provide better accuracy and automation of the drilling operation. An inhouse desktop CNC machine [120] will allow greater flexibility in performing drilling and polishing for future decapped samples. The whole stitched and mirrored image of the backside is shown in Figure 5.9(d). The image was mirrored to be able to compare the structured components to those in the frontside captured view.

Also a Spartan-6 FPGA XC6SLX9 and a RISC-V ASIC were backside decapsulated as shown in Figure 5.10 and Figure 5.11 respectively. The thickness of the FPGA is smaller than that of the PIC MCU and the process of fastening the DUA while drilling from the backside was challenging in case of the FPGA. Beside that, the same steps mentioned for the PIC MCU were followed to decap the FPGA from the backside as well. The decapped Spartan-6 FPGA is shown in Figure 5.10a.

The SWIR camera, mentioned in subsection 5.1.2 in the last step in the backside decapsulation procedure, was used to capture the images from the backside. Details about using the SWIR camera in backside imaging have been presented in subsection 4.3.1.



(a)

(b)

(c)

(d)

Figure 5.9: Backside decapped chip at different stages (a) before getting to the copper shield, (b) With the copper shield should the attached ends, (c) Without the copper shield and (d) SWIR image of the backside decapped PIC16F687 chip.

(a)

(b)

Figure 5.10: (a) Backside of the decapsulated Spartan-6 FPGA and (b) its unaltered frontside.

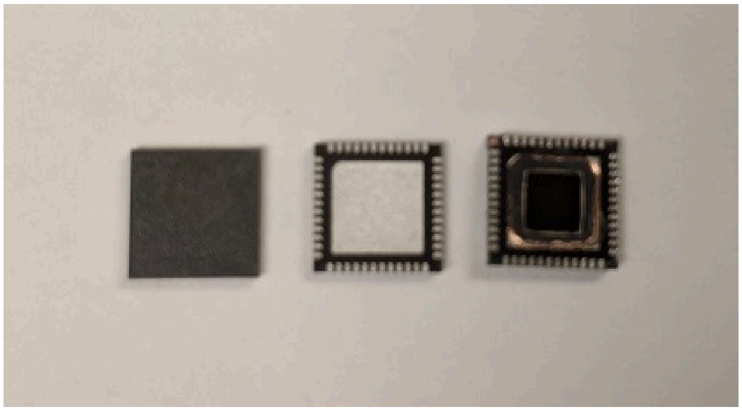
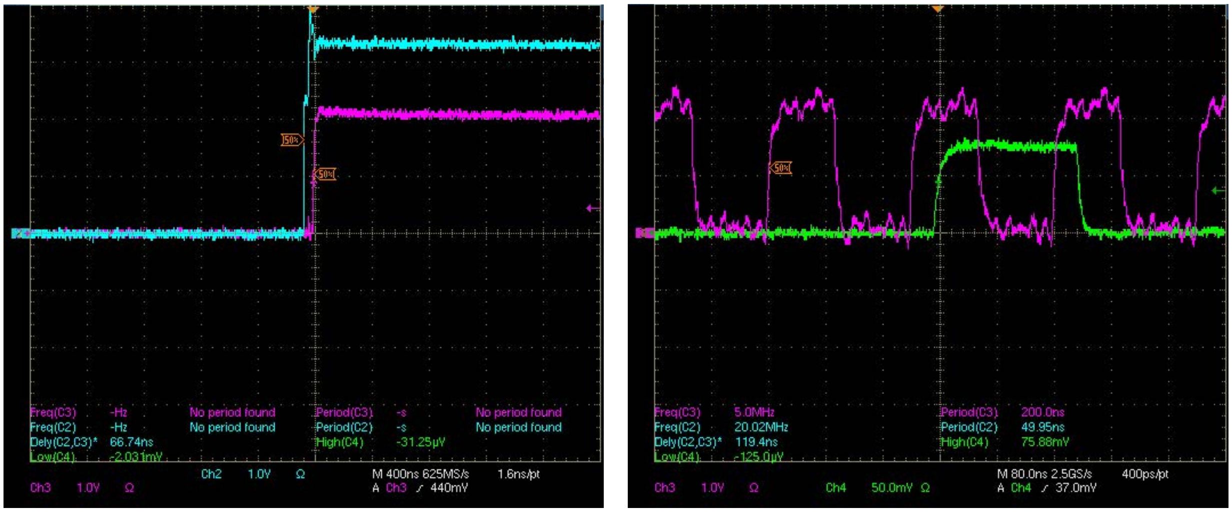


Figure 5.11: RISC-V FE310-G002 backside decapsulated chip with both frontside and backside before decapsulation.

5.2 Laser pulse measurements and characterization

Some experiments needed to be performed in order to tailor the energy of the laser pulse (width and amplitude) as well as the timing of the laser pulse so that it hits the DUA in the appropriate clock cycle to induce a current that successfully injects a fault. The delay between the external clock and the instruction clock (output pin of DUA) is depicted in Figure 5.12a. Figure 5.12b shows the delay between the instruction that triggers the laser and the appearance of the laser pulse from the APD output.



(a)

(b)

Figure 5.12: a) Clock-to-instruction cycle delay and b) Instruction cycle to laser pulse (green) delay with the DUA running @ 20MHz where the instruction frequency is 5MHz $(\frac{f_{osc}}{4})$.

5.2.1 Induced and switching currents

A Rohde & Schwarz RS H 2.5-2 H-field probe, connected to the oscilloscope, was placed on the V_{DD} pin of the DUA (shown in Figure 5.13) to analyze the effect of the laser pulse on the current drawn by the DUA. Figure 5.14a shows that there is a peak in the drawn current representing the induced current as a result of the injected laser pulse. On the other hand, Figure 5.14b shows the expected switching-current due to an I/O pin toggle.

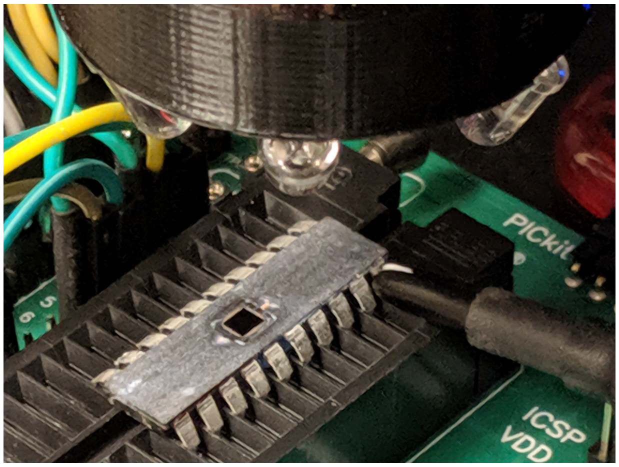
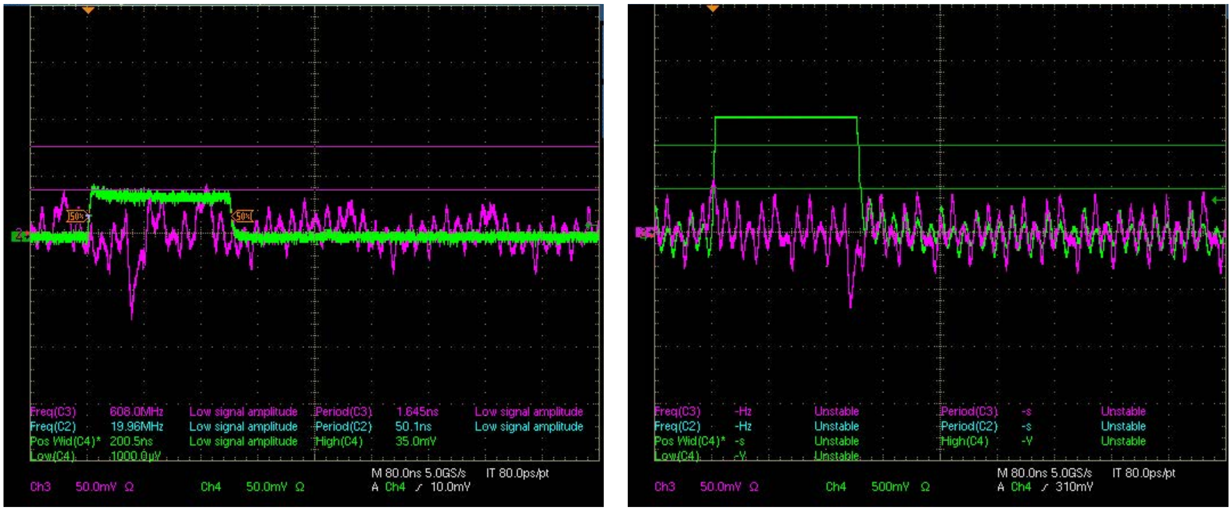


Figure 5.13: Rohde & Schwarz RS H 2.5-2 H-field probe [8] placed on the V_{DD} pin, as shown on the right hand side.



(a)

(b)

Figure 5.14: a) Induced current (purple) due to laser pulse (green), 200*ns* at 50% peak current, where the H-field probe is connected to CH3; and CH4 is connected to APD.b) Switching current (purple) due to triggering of the MCU where the H-field probe is connected to CH3; and CH4 is connected to trigger I/O pin (green).

5.3 Introduction to PIC16F687

The PIC16F687 [9] applies a 2-stage pipeline with a Harvard structure. The two-stage instruction pipeline overlaps the fetch and execution of instructions. It has a 3.5KB Flash program memory and 384 bytes data memories divided into 128 bytes SRAM and 256 bytes EEPROM. It has 14 configuration bits. The program memory code protection and data memory code protection are two special configuration bits that can be used to enable/disable reading program code from Flash memory or data code from EEPROM respectively. The data memory and the program memory have separate buses. The only available register in the data path is the working register (W_{reg}). Each instruction is a 14-bit word which gets fetched in one instruction cycle then executed in the next instruction cycle in the 2-stage pipeline. The instruction set has only 35 instructions. All instruction are executed in a single cycle except for branches. Each instruction cycle (fetch or execute) is divided in four Q cycles as shown in timing diagram in Figure 5.15. The Q cycles provide the timing/designation for the Decode, Read, Process Data, Write etc., of each instruction cycle. Q1 is used for decoding the instruction in the fetching phase or a forced no operation (NOP) in the execution phase. Q2, Q3, and Q4 are used for a read-modify-write procedure. Specifically, Q2 is for instruction read or NOP, Q3 to process the data and Q4 for instruction write or NOP [9].

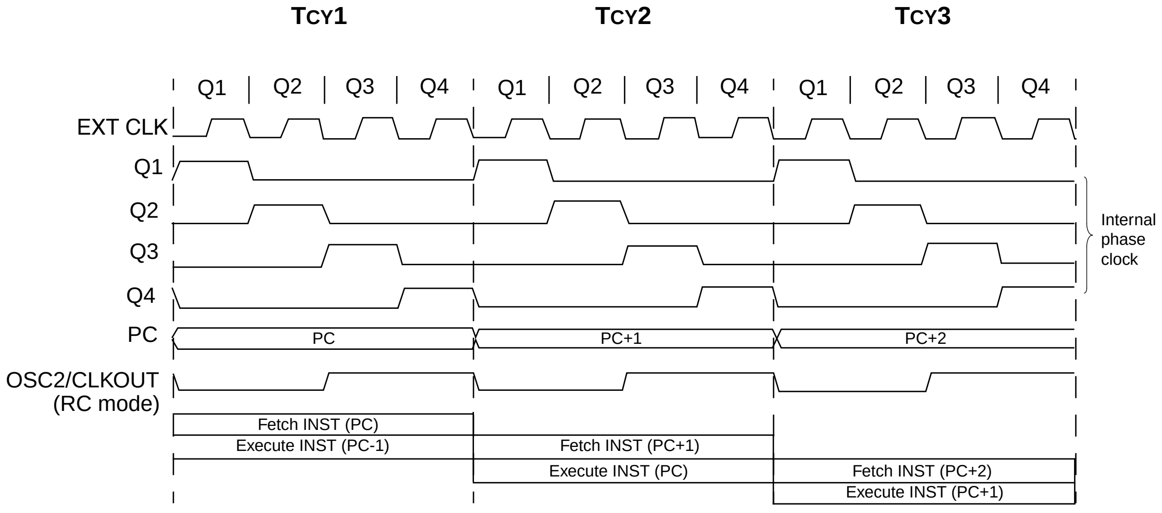


Figure 5.15: Fetch and execute timing diagram [9] for the two-stage pipelined DUA where the fetch and execution of instructions overlaps.

5.4 Attacks on PIC16F687

LFI experiments focused on the output from the flash memory, programming of configuration bits and attacking AES-128. The output of the flash will be referred to as the instruction register (I_{reg}).

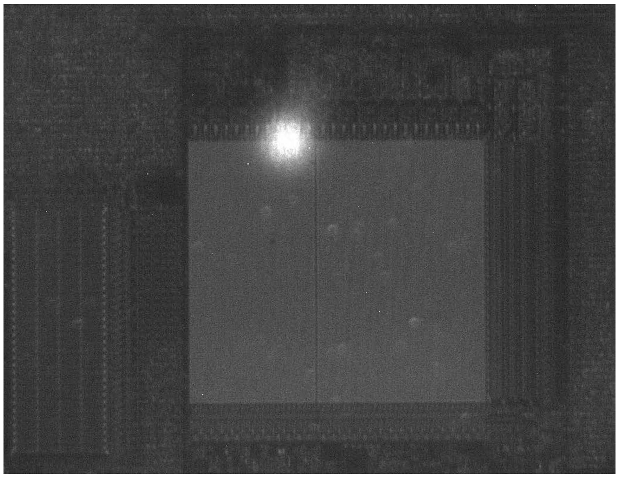


Figure 5.16: Laser beam spot target location captured using the NIR camera at 1.5% peak current near the I_{reg}.

Navigating the correct area to attack to achieve a successful intended attack could be a real challenge so the ability to identify different structures from the backside of the a decapsulated chip was very beneficial in reducing the time needed to find the sweet spot for a specific attack. For example, while sweeping across the DUA, some areas did not introduce any noticeable faulty behavior letting the chip run normally whereas other areas halted the chip when the laser pulse energy was increased.

Different types of attacks were performed on the DUA and will be explained in the following subsections. All the performed attacks proved to be repeatable under the same operating conditions. Unless otherwise stated the experimental settings involved one laser channel (with specified laser pulse width and percentage of laser peak current), 5X objective, and the PIC operated at 20MHz, 5V. In all cases the laser was triggered within the test program. All the attacks targeted the laser beam spot at the location depicted in Figure 5.16 unless otherwise mentioned in a certain experiment.

5.4.1 Instruction skip attack

In this attack, an arbitrary number of instructions were skipped depending on a tailored laser pulse whose laser peak current and laser pulse width were varied. The area of interest in this type of attack was the 14-bit flash read sensors as shown in Figure 4.12. The key concept here is to find the exact laser pulse energy (peak current and pulse width) needed to skip an arbitrary number of instructions without introducing any other faults to the executed instructions. This resembles replacing those skipped instructions with NOPs instructions. Examples of the skipped instructions are shown in Table 5.2.

Table 5.2: Effect of the laser source peak current on the number of skipped instructions for a 200ns laser pulse width.

Number of skipped instructions	Percentage of peak current (%)	Description of the instructions
2	33	addlw 0x01 addlw 0x01
4	36.25	addlw 0x01 addlw 0x01 addlw 0x01 addlw 0x01
1	31.25	andlw 0xFE
1	31.875	incf 0x20,f
2	33.75	incf 0x20,f incf 0x20,f
1	33.75	goto $

The assembly program that was written to test and verify the instruction skipping started by filling 96 bytes of the general purpose registers in the SRAM of the DUA with a certain known value (0xAA for example). It then performs some operations whose results are stored to chosen addresses. The stored result(s) should change if a fault is successfully injected. For example, the laser parameters required to skip one or more of the four successive instructions, *addlw 0x01* that increments the value of the W_{reg} [9], are shown in Table 5.2 to get a total count less than four. The laser spot was targeted towards the location shown in Figure 5.16.

5.4.2 Instruction replacement attack

This attack is the generic form of the instruction skip attack where one or more of the executed instructions can be replaced by another instruction including the NOP. We focus on providing examples of replaced instructions other than the NOP (since replacing the instructions with NOP was illustrated in subsection 5.4.1). It should be noted that when a laser pulse with sufficient energy is applied to a certain viable target area of the DUA, a certain number of instructions could be skipped. However if the laser pulse energy is a little bit more or a little bit less than the sufficient energy to skip a whole (integer) number of instructions, the last instruction to be faulted is replaced by a faulty non-NOP instruction. The faulty instruction value appears to be dependent on the energy of the injected laser pulse, target area location and the hamming weight of the original 14-bit instruction which was supposed to be executed if there was no LFI. A simple example of instruction replacement is altering the immediate value being added to the W_{reg} [9] (i.e. adding 0x02 to the W_{reg} instead of 0x01 by replacing addlw 0x01 with addlw 0x02). For example, due to the PEM results detailed in section 5.4.5 the immediate value of the instruction was located in the region of the laser pulse providing further evidence of fault injection causing the immediate value change. Since the PEM was a separate station (both LFI and PEM setups are described in sub-sections 4.3.1 and 4.3.2 respectively), it was not possible to inject a fault and carry the device over to the PEM to analyze the photon emissions or changes to the instruction. Thus based upon the test program and PEM identification of the instruction fields, the *addlw 0x01* instruction could be changed to two different values when targeting the beam at two different locations along I_{reg} using a (50ns, 34.625% peak current) pulse. Figure 5.16 shows one location of the beam to change *addlw 0x01* to *addlw 0x81* (one MSB changed in the immediate value) while Figure 5.17 depicts a different location of the beam to change *addlw 0x01* to *addlw 0xC1* (two MSBs changed in the immediate value).

It was also observed that two properly tailored laser pulses could produce the same faulty instruction. For example, a laser pulse with width 100ns with 17% of the peak current injected the same fault as a 200ns pulse with 9.375% of the peak current when targeting a certain location with nominal operating conditions.

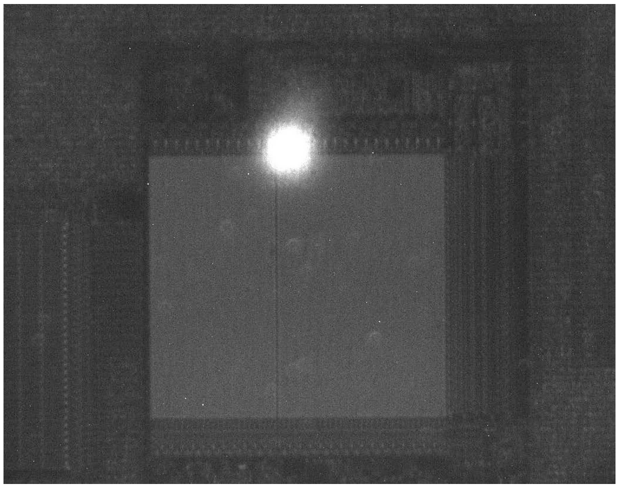


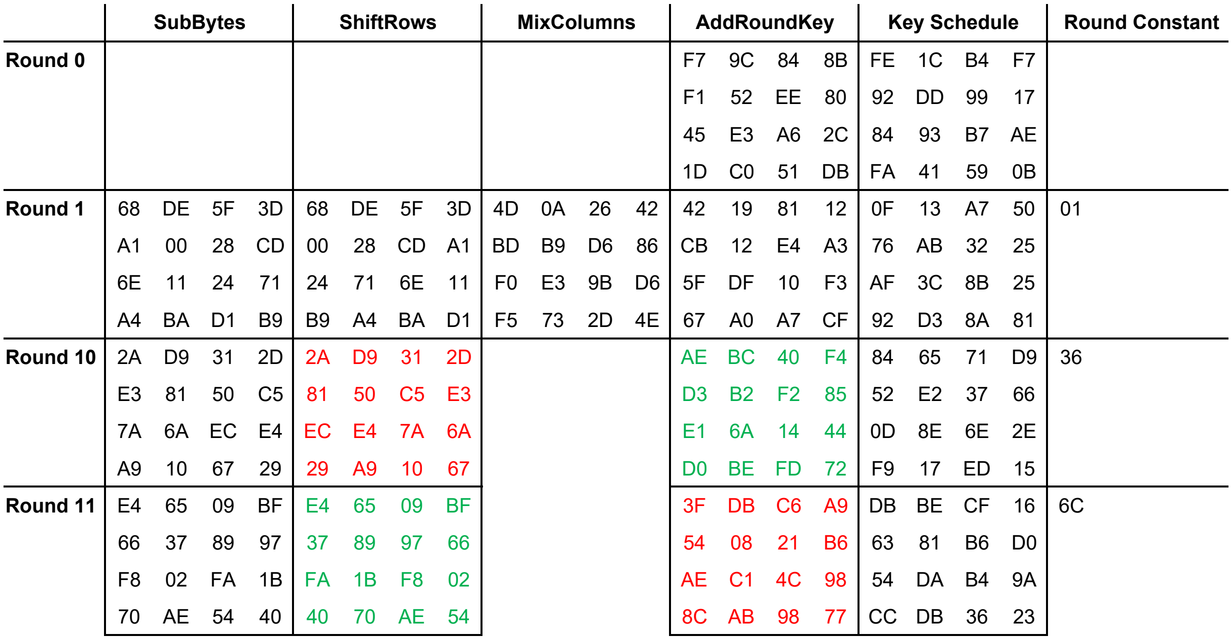
Figure 5.17: Laser beam spot target location captured using the NIR camera for a (50ns, 34.625% peak current) near the I_{reg} to alter the immediate value in *addlw 0x01*.

5.4.3 Configuration bits attack

One of the interesting attacks performed on the DUA was the successful skipping of the configuration bits (EEPROM) programming while flashing the DUA. The MCU was running at 20MHz so the instruction cycle was 5MHz and the laser was triggered only during the programming session then switched off. The laser spot was targeted over the center of the configuration bits area. The configuration bits area is annotated in Figure 4.12. The pulsed laser was configured to have a 100ns pulse width with a 500mA peak current at a 5MHz frequency (which is the same as the instruction cycle frequency for this attack to be successful) during the programming session of the DUA. After the programming session was completed, the laser was turned off. We wrote a simple assembly program to a write a certain value in the SRAM then several instructions were executed repeatedly to read, increment and store the value at the same SRAM location. The final value was written to EEPROM. The attack was as follows: when the laser was turned on during the programming session of the DUA while downloading an assembly program that enables code protection on DUA with code protection originally disabled (e.g. config bits = 0x33F7), it was found that config bits were not changed as the user intended to enable the code protection (e.g. using config bits = 0x3337) and the code was successfully read (i.e. the config bits remained as 0x33F7). When the same experiment was done with the laser turned off, the code protection was enabled and the code was not readable from the DUA because the new config bits were successfully changed to 0x3337.

5.4.4 AES-128 key retrieval

This section presents an attack [109, 110] that was conducted on the full assembly implementation of AES-128, provided by *Microchip*, while running on the DUA. The laser is triggered by adding two port set/clear instructions (*BSF*,*BCF*) before the target instruction (*decfsz*) as shown in the code in column 1 of Table 5.3. Table 5.3 focuses on the critical section in the assembly code where the laser source gets triggered to skip a certain target instruction (*decfsz* in this case). Also, Table 5.3 shows the number of clock cycles supplied by the external clock source needed to reach a certain Q cycle for a certain corresponding instruction either in the fetch or the execute stage of that instruction.



	SubBytes				ShiftRows				MixColumns				AddRoundKey				Key Schedule				Round Constant
Round 0													F7	9C	84	8B	FE	1C	B4	F7	
													F1	52	EE	80	92	DD	99	17	
													45	E3	A6	2C	84	93	B7	AE	
													1D	C0	51	DB	FA	41	59	0B	
Round 1	68	DE	5F	3D	68	DE	5F	3D	4D	0A	26	42	42	19	81	12	0F	13	A7	50	01
	A1	00	28	CD	00	28	CD	A1	BD	B9	D6	86	CB	12	E4	A3	76	AB	32	25	
	6E	11	24	71	24	71	6E	11	F0	E3	9B	D6	5F	DF	10	F3	AF	3C	8B	25	
	A4	BA	D1	B9	B9	A4	BA	D1	F5	73	2D	4E	67	A0	A7	CF	92	D3	8A	81	
Round 10	2A	D9	31	2D	2A	D9	31	2D					AE	BC	40	F4	84	65	71	D9	36
	E3	81	50	C5	81	50	C5	E3					D3	B2	F2	85	52	E2	37	66	
	7A	6A	EC	E4	EC	E4	7A	6A					E1	6A	14	44	0D	8E	6E	2E	
	A9	10	67	29	29	A9	10	67					D0	BE	FD	72	F9	17	ED	15	
Round 11	E4	65	09	BF	E4	65	09	BF					3F	DB	C6	A9	DB	BE	CF	16	6C
	66	37	89	97	37	89	97	66					54	08	21	B6	63	81	B6	D0	
	F8	02	FA	1B	FA	1B	F8	02					AE	C1	4C	98	54	DA	B4	9A	
	70	AE	54	40	40	70	AE	54					8C	AB	98	77	CC	DB	36	23	

Figure 5.18: AES-128 internals showing the 1st round till the 11th round where the 10th round check (*decfsz round_counter,f*) is skipped to add an additional 11th no-mix-column round.

We used the known-plaintext attack (KPA) model for cryptanalysis. Two different attacks were conducted by targeting two different instructions: *xorwf* and *decfsz*. The 1st attack skips the *xorwf* instruction (using laser pulse width 200ns and 9% percent of peak laser current) to extract a single byte out of the 16-byte key at a time, then repeating this to obtain the remaining 15 bytes of the key which is similar to [121] but using 16 short pulses to skip the *xorwf* instruction instead of a very long pulse to skip the whole 16-byte key addition in the 10th round. The 2nd attack skips the *decfsz* instruction (using laser pulse width 200ns and 9% percent of peak laser current) to add one more extra round with no mix-columns (11th round). This attack retrieves the whole 16-byte using a single fault injection. The later attack method is easier and more efficient and has been verified by generating 100 random generated keys, plain-text and cipher-text pairs. Figure 5.18 shows an example of the latter attack where the used plain-text was *0x0963c1e7808f7081307711087c9782d0*. The target location in this attack is shown in Figure 5.16.

In order to provide a formal proof for our new attack on AES-128 , let M denote the plain-text, K denote the AES key, K^i denote the i^{th} AES round key, C denote the correct cipher-text, M^i denote the temporary cipher result after the i^{th} round and D denote a faulty cipher-text. Each round is composed of 4 transformations: SubBytes, ShiftRows, MixColumns and a bit-per-bit XOR with a round key. The final 10^{th} round of the AES is composed of the same functions as a classical round except that it does not include the MixColumns transformation as thoroughly explained in [30].

Without LFI the attacker can get C mentioned in equation 5.1 using KPA and with LFI one more additional 11^{th} round is executed that does not include the MixColumns similar to the 10^{th} round and the attacker gets D as mentioned in equation 5.2. Equation 5.3 below shows how the attacker uses D and C to get the round key K^{11}. It is well known that by reversing the Rijndael's key schedule starting by K^{11}, the whole 16-byte key could be retrieved.

$$C = ShiftRows(SubBytes(M^9)) \oplus K^{10} \tag{5.1}$$

$$D = ShiftRows(SubBytes(C)) \oplus K^{11} \tag{5.2}$$

$$D \oplus ShiftRows(SubBytes(C)) = K^{11} \tag{5.3}$$

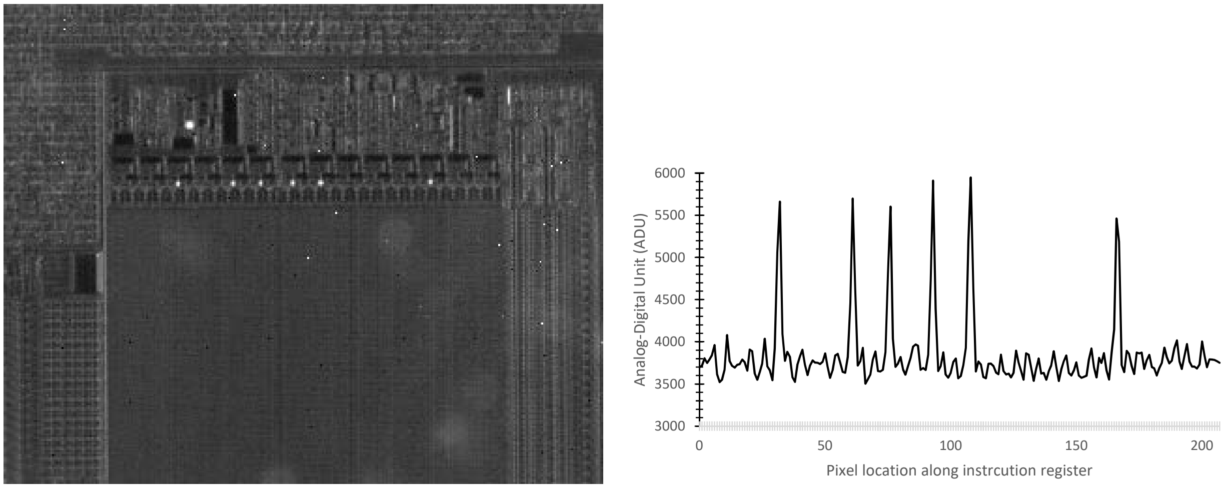
5.4.5 Instruction register PEM analysis

The SWIR camera and the waveform generator were the main tools used to analyze the I_{reg} (more specifically trying to locate the instruction register on the backside of the die). The external clock is internally divided by four to generate four non-overlapping quadrature clocks, namely Q_1, Q_2, Q_3, and Q_4. Internally, the program counter (PC) is incremented

Table 5.3: Injecting arbitrary number of clock cycles to reach the end of a certain Q cycle for a desired assembly instruction.

Instruction		$Q_{4_{\text{execute}}}$	$Q_{1_{\text{fetch}}}$	$Q_{2_{\text{fetch}}}$	$Q_{3_{\text{fetch}}}$
BSF	PORTC,RC4	23641	23634	23635	23636
BCF	PORTC,RC4	23645	23638	23639	23640
NOP		23649	23642	23643	23644
decfsz	round_counter,f	23653	23646	23647	23648

every Q_1, and the instruction is fetched from the program memory and latched into the I_{reg} in Q_4 [9]. The instruction is decoded and executed during the following Q_1 through Q_4. The Q clocks and instruction execution flow are illustrated in Figure 5.15.



(a)

(b)

Figure 5.19: At normal operating conditions without any over-voltaging ($V_{DD} = 5V$) for (*decfsz 0x44,f*) instruction (14-bit instruction word = *0b00101111000100*) a) I_{reg} emissions and b) ADU graph where the ADU values correspond to the emissions intensity along the I_{reg}.

The laser pulse target location for AES was believed to be close to the middle of the I_{reg}. Different locations along the I_{reg} were investigated and successful fault injection occurred with different laser pulse amplitudes at different locations. When moving away from the I_{reg}, no faults were injected.

A fixed number of clock cycles are launched, after which the clock is switched off to

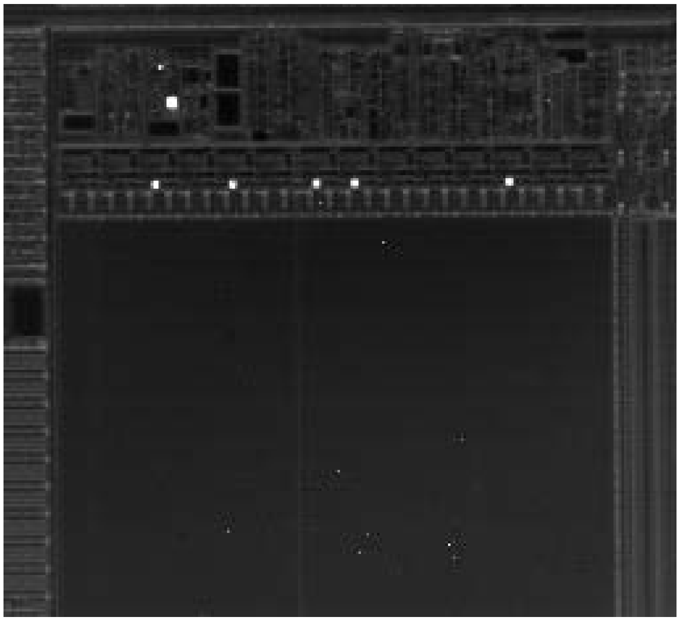


Figure 5.20: I_{reg} emissions for (*incf 0x44,f*) instruction (opcode = *0b00101011000100*) at V_{DD} = 10.5V.

capture a quiescent PEM image at the flash memory output of DUA. After the image is collected, the clock is injected again continuously to output data for analysis. Different emission intensities are shown in Figure 5.19a and 5.20 by varying the supply voltage. The emissions are only visible when stopping the external clock at the end of $Q_{1_{\text{fetch}}}$ and $Q_{2_{\text{fetch}}}$ (not the execution). At the end of $Q_{3_{\text{fetch}}}$, the emissions for the instruction being fetched disappear as likely the 14-bit instruction word will have been latched to the control circuitry beside the I_{reg}. The laser pulse effect (i.e. induced current) has to target Q_3 in the fetching phase of the target instruction (clock cycle# 23648 as mentioned in Table 5.3) to have an effect on a target instruction. For example, if the execution of the program is stopped at $Q_{1_{\text{fetch}}}/Q_{2_{\text{fetch}}}$ of the target instruction (I_{target}) with the laser being triggered to inject the pulse during $Q_{1_{\text{fetch}}}/Q_{2_{\text{fetch}}}$, the fault injection would be unsuccessful. On the other hand, if the execution of the program is stopped at the end of $Q_{3_{\text{fetch}}}$ (clock cycle# 23648 as mentioned in Table 5.3) of I_{target}, the effect of the transient induced current due to the laser pulse would have already been captured by the control circuitry beside the I_{reg} and the fault would be injected.

The quiescent (static) emissions of the I_{reg} could be captured within a 3 seconds integration time under the normal power supply (5V) as shown in Figure 5.19a. Figure 5.20

illustrates emissions from a different instruction again showing correct opcode readout, but using a higher 10.5V supply. The over-voltaging achieves higher intensities of emissions but as shown in Figure 5.19a over-voltaging is not necessary. By increasing the integration time of the SWIR camera (instead of increasing the supply voltage) we could capture more of those emitted photons from the I_{reg}. The intensity values of the emissions captured by the SWIR camera are shown in Figure 5.19b and are clearly correlated with the 14-bit *decfsz 0x44,f* instruction, whose value was *0b00101111000100*.

5.5 Countermeasures

Without a complete understanding of the FI impact on the MCU, countermeasures for FI attacks cannot be developed. However, any countermeasure comes at a price. For example, software countermeasures usually affect the performance because of added code overhead while hardware countermeasures usually increase the manufacturing cost. For example countermeasures were suggested for single-bit data corruption, although FI has also demonstrated multi-bit corruptions [122]. In addition, instruction skipping countermeasures were also developed, such as duplication in [123], although research also shows evidence of instruction replacement faults [124]. Software countermeasures such as duplication in [123] aren't effective against LFI since one or more instruction could be skipped using a customized LFI pulse as discussed in sub-section 5.4.1.

In this section, we discuss several LFI countermeasures specific to PIC16F687 but may be applied to similar targets. Also general physical and software countermeasures are discussed.

5.5.1 Configuration bits factory protection

Since the factory settings of the PIC16F687 chip is that the code protection is disabled, the user will want to enable the code protection when downloading on the DUA. Therefore, having the code protection enabled as a factory default setting would be a countermeasure against the attack that leads to skipping the programming of the configuration bits mentioned in sub-section 5.4.3.

5.5.2 Encrypted instruction set and dummy structures

Encrypting the instruction set will obfuscate both instruction alteration using LFI and instruction readout using PEM. In general the key could be implemented using fuses or as a PUF. The countermeasure can be thwarted if the attacker can determine the encryption key of the instruction set. Dummy structures may be added to the chip layout to mislead the attacker at the expense of some area/power overhead.

5.5.3 Thicker substrate to obstruct PEM

A thicker substrate will obstruct PEM and mandate a substrate thinning process which leads to a more complex sample preparation procedure by adding a thinning step during which the DUA might be damaged thus increasing the complexity of attack. Technology scaling tend mandate thicker substrates because of the increased wafer size [125] thus acts as a countermeasure for PEM.

5.5.4 Hardfault

Implementing an Interrupt Service Routine (ISR) that gets triggered when an invalid instruction, caused through LFI, enters the execution pipeline may be a valuable countermeasure instead of execution of a NOP which leads to a successful instruction-skip attack. For example, this type of ISR is referred to as a hardfault for a DUA like ARM Cortex-M0 LPC1114FN28102 which will be discussed in chapter 6.

5.5.5 Latchup

Countermeasures could involve a custom doping to the substrate and wells (either N or P) such that the substrate's sensitivity to latch-up increases against LFI.

5.5.6 Bulk current sensor and substrate coating

Sensors in the substrate [126] could be used to detect any current injected through LFI. In addition, the substrate can be coated with an optically active layer [73] that reflects the light emitted by the operation of forward-biased p-n junctions, which are considered as LEDs.The LEDs light can detected by depleted areas like reverse p-n junctions, which are considered photocurrent detectors.The backside protection layer provides angle dependent reflection and full transmission blocking so any damage in this protection layer will affect the angle of reflection for one or more of the LED light beams to be detected by the photodetectors. Therefore, the detected photocurrent can be used to verify the integrity of the backside protection layer and detect attacks.

It is also important to note that software countermeasures such as instruction duplication and error detecting/correcting codes have to take into account the multiple instruction skip attack; otherwise those countermeasures will be ineffective against LFI. Also, the area overhead for active devices to detect the laser induced current would also introduce increased energy consumption which is considered an important metric and limitation for low-power IoT devices.

5.6 Summary and comparison with previous research

Despite numerous publications in fault injection, laser fault injection methodologies remain diverse with limited details on equipment and setups. A new laser fault injection methodology is proposed which combines quiescent photon emissions with backside dynamic laser pulse profiling in time and space. Empirical results illustrate the impact of the laser on multiple-instruction fault injections, and controlled instruction replacement faults. Further details of comparative previous research is provided below.

LFI was introduced in 2002 by S. Skorobogatov and R. Anderson [29]. Several publications [18, 53, 90, 127] confirmed the local and temporal precision of LFI. Breier et al. [121] discussed the use of the laser pulse (glitch) to skip an arbitrary number of instructions executing on a $0.35\mu m$ ATmega328P. The laser beam was injected on the backside of the die at a single location found from scanning the entire die region (without die imaging). For example their attack on AES used one very long single laser pulse of $3\mu sec$ to skip the whole last round key addition for the 16-bytes. SEM imaging was combined with LFI in [128], to reduce the complexity of finding laser targets or flipflops on a 25MHz 90nm AES chip. Attacks using LFI on processors have also been empirically demonstrated however the target was data in the RAM blocks in the PIC16F84 and ATmega [129], or in the flipflops [127, 128]. Other LFI research such as [124] attacked the ChaCHa20 cipher executing on the ATmega328p. The fault injection attacks largely skipped instructions, with the exception of one attack which replaced addition instructions with subtraction instructions as mentioned in [124]. Other researchers have demonstrated an attack on the secure boot using static LFI [130]. This attack utilized a Quadcore A9 running at 1.4GHz and focused on using very long pulse widths and large spot size to flip bits in configuration registers. Roscian et al. [48] simulated the effect of the laser pulse on the DUA. LFI attack of a physical unclonable function was aided by PEM, FIB and Laser scanning microscope images to identify the location of the ring-oscillators and inverter chains on a complex programmable logic device [122]. However, the attack focused on changing the configuration bits in the lookup tables and utilized extremely expensive PHEMOS1000 equipment. In [123], the authors proposed a software countermeasure based on a fault model in which the attacker is able to skip a single assembly instruction. The single assembly instruction skip fault model has been observed on several architectures. Limited research has examined the analysis of dynamic LFI on processors utilizing quiescent photon emissions. In addition, there is limited analysis of the impact of laser parameter variations in time and space including pulse width and intensity variation effects on instruction fetches.

Unlike previous research, the results presented in this chapter show that quiescent photon emissions combined with laser fault injection provides fine tuning of faulty instructions in addition to reverse engineering within each clock cycle.

In the next chapter, a DUA, specifically ARM Cortex-M0 LPC1114FN28102 supplied by NXP, with a more recent technology node and higher design complexity will be investigated.

Chapter 6

Results for ARM Cortex-M0 LPC1114FN28102

This section describes the attack procedure and LFI results on the ARM Cortex-M0 chip. The attack procedure differs from that used in the previous chapter due to challenges with imaging the backside of the ARM using the laser Mitutoyo equipment setup. Results of PEM and a discussion of countermeasures is also covered.

6.1 Introduction to ARM Cortex-M0

The ARM Cortex-M0 processor is a 32-bit RISC processor, with a 3-stage (fetch, decode and execution) pipeline von-Neumann (load-store) architecture with a single bus interface. Basic specifications of LPC1114FN28102 are show in Table 4.2. The Cortex-M0 (for MCUs and ASICs) processor implements the ARMv6-M architecture (same architecture for Cortex-M1 used in FPGAs), which is based on the 16-bit Thumb instruction set and includes Thumb-2 technology (16-bit and 32-bit instructions). The ARMv6-M architecture inherits attractive features from various ARM architectures as shown in 6.1 which implies that attacks against the ARM Cortex-M0 LPC1114FN28102 could be likely successful against other Cortex-based MCUs. The 32-bit instructions are used when the 16-bit version can't carry out the required operations. This provides the exceptional performance expected of a modern 32-bit architecture, with a higher code density than other 8-bit and 16-bit MCUs and avoids the overhead of switching between two instruction sets. In total, the Cortex-M0 supports only 56 base instructions with some instructions having more than one form. The ARM Cortex-M0 LPC1114FN28102 instruction set has forward compatibility with the ARM Cortex-M3/M4.

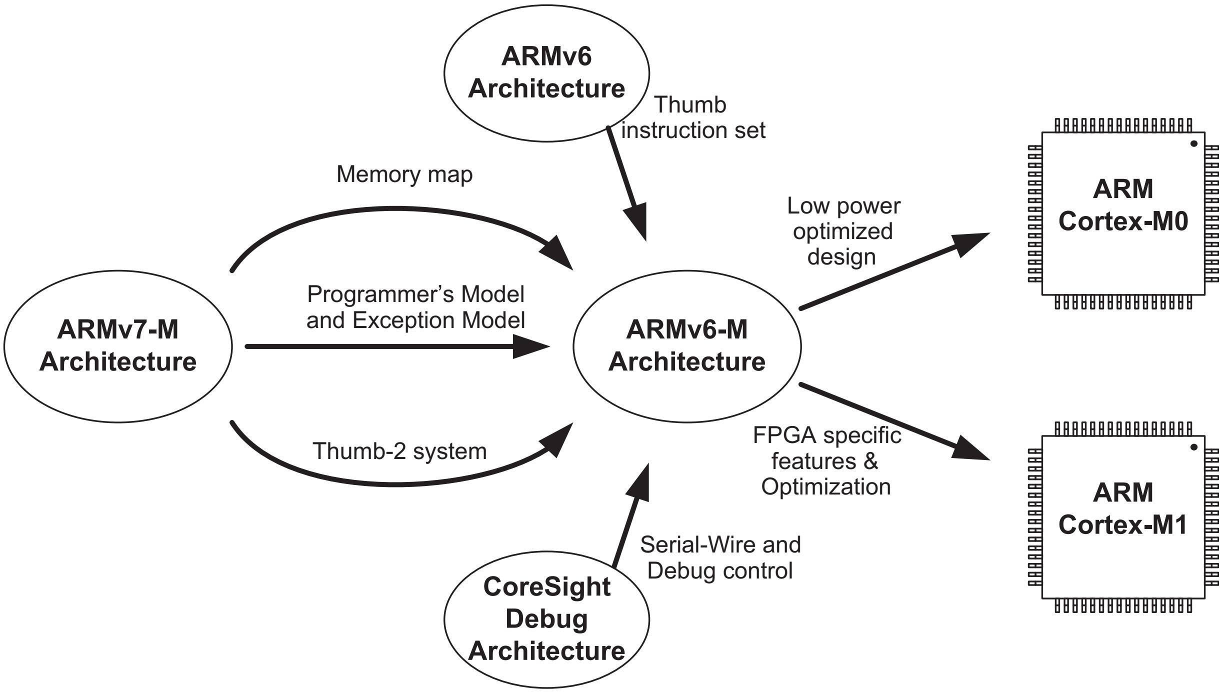

ARMv6 Architecture
Thumb instruction set
Memory map
Low power optimized design
ARM Cortex-M0
Programmer's Model and Exception Model
ARMv7-M Architecture
ARMv6-M Architecture
Thumb-2 system
FPGA specific features & Optimization
ARM Cortex-M1
Serial-Wire and Debug control
CoreSight Debug Architecture

Figure 6.1: ARMv6-M inherits attractive features from various ARM architectures [10].

The ARM Cortex-M0 LPC1114FN28102 processor core (12K gate count at minimum configuration) has register bank, Arithmetic Logic Unit (ALU), data path and control logic. The register bank has sixteen 32-bit registers (some of them are special registers). The data path and Advanced High-Performance Bus (AHB) are 32-bit wide. The addresses are linear and 32-bit with no memory paging. The processor has two modes: Thread(normal execution) mode and Handler (exception and interrupts) mode. The Nested Vectored Interrupt Controller (NVIC) supports up to 32 interrupt requests and a nonmaskable interrupt (NMI) input. If the interrupt is accepted by the processor according to the interrupt's priority level, the processor responds (minimum in 16 clock cycles but configurable) by executing the corresponding interrupt handler. For example, in Cortex-M0 programming the data variables stored in memory need to be stored at an address location that is a multiple of its size otherwise a hardfault will occur. Another reason to trigger the hardfault handler is a cleared Thumb-bit (it should always be set to indicate that ARM Cortex-M0 LPC1114FN28102 is executing Thumb instructions only) in the xPSR register. If the

T (Thumb) bit is cleared in the xPSR Register, then the hardfault was caused by an accidental switching to ARM state.

The Hardfault is one of the interrupts. Interrupts have seven priority levels: three fixed and four programmable. Under normal situations, the interrupt latency of the Cortex-M0 processor is 16 cycles. The interrupt latency is defined as from the processor clock cycle the interrupt is asserted, to the start of the execution of the interrupt handler. This interrupt latency assumes the following:

- The interrupt is enabled and is not masked by PRIMASK or other executing exception handlers.
- The memory system does not have any wait state. If the memory system has wait state, the interrupt could be delayed by wait states that occur at the last bus transfer before interrupt processing, stacking, vector fetch, or instruction fetch at the start of the interrupt handler.

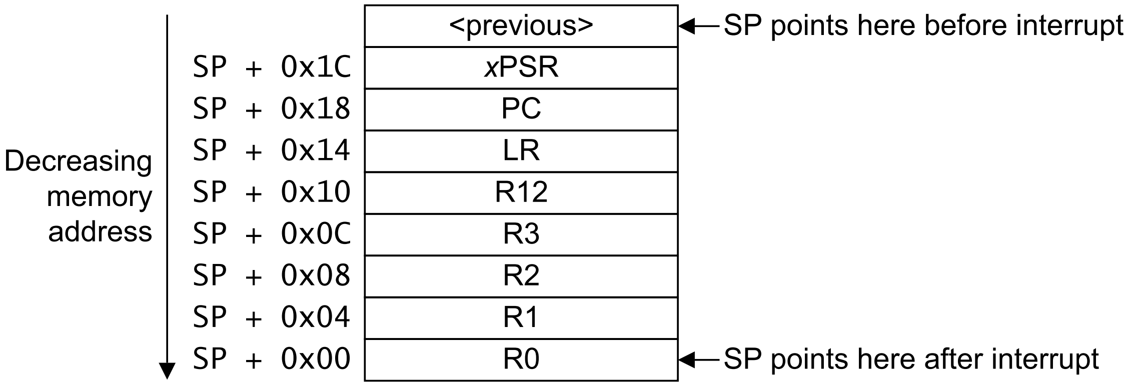


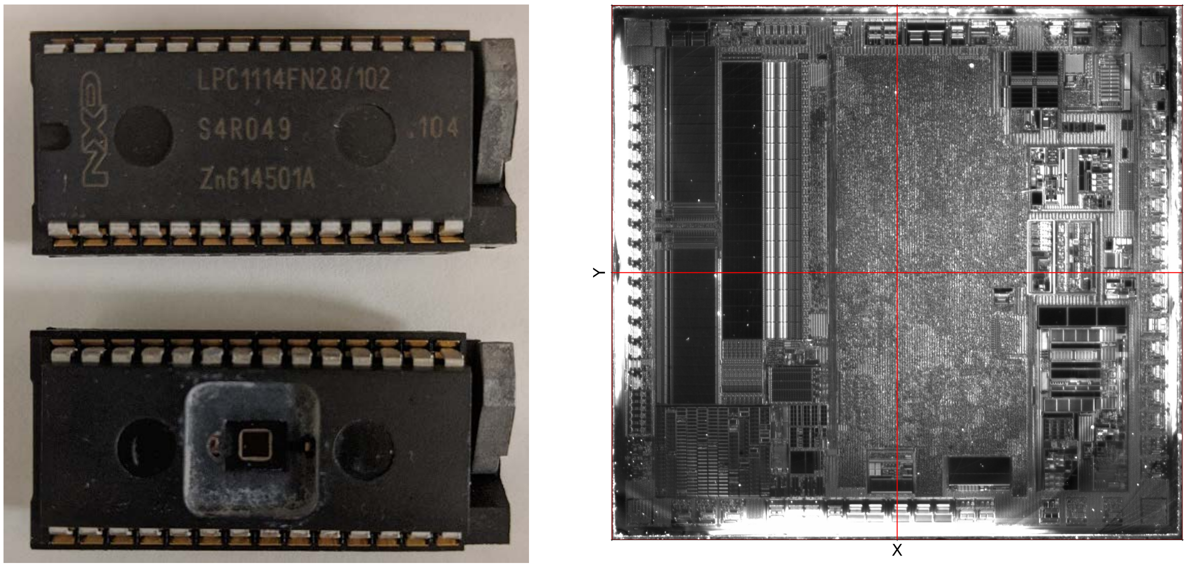
Figure 6.2: ARM Cortex-M0 LPC1114 stack frame.

On the Cortex-M0 processor, there is only one exception type that handles faults: the hardfault handler. The hardfault handler is almost the highest priority exception type, with a priority level of -1. Only the NMI can preempt it. When it is executed, we know that the MCU is in trouble and corrective action is needed. The hardfault handler is also useful for debugging during the software development stage. When a breakpoint has been set in the hardfault handler, the program execution stops when a fault occurs. By examining the contents of the stack frame shown in Figure 6.2, we can trace back to the location of the fault and try to identify the reason for the failure. This behavior is very different from that of most 8-bit and 16-bit MCUs. In these MCUs, often the only safety net is a watchdog timer. However, it takes time for a watchdog timer to be triggered, and often there is no way to tell how the program went wrong.

6.2 Attacking procedure

First the DUA was backside decapsulated as shown in Figure 6.3a using the same procedure for the PIC16F687 mentioned in subsection 5.1.2. The DUA was investigated under the Mitutoyo setup in order to be able identify uniform structures such as memories in order to make the attack easier by not having to scan the whole die area for sensitive locations for LFI. Backside imaging of LPC1114FN28102 wasn't straightforward as in the case of PIC16F687. Figure 6.4a shows how the LPC1114FN28102 looked under the best achievable backside imaging conditions with the SWIR camera. The NIR camera wasn't able to capture any structure beneath the substrate for the LPC1114FN28102. The settings used for getting a backside image for LPC1114FN28102 are shown in Figure 6.4b. The automatic dynamic range that was used in case of PIC16F687 yielded no image beneath the substrate surface of the LPC1114FN28102. Therefore, the dynamic range was manually adjusted until an image with identifiable structure could be captured. However, Figure 6.4a is still of a lower quality when compared to Figure 6.3b which was captured under the Olympus setup. This is probably due to the thicker substrate of the LPC1114FN28102 and the transmission efficiency of the VMU. The Olympus setup has higher transmission efficiency($> 80\%$) for the 1064nm wavelength than that of the Mitutoyo setup (54%). That efficiency increase is probably mainly due to that the Mitutoyo VMU has two paths (camera and laser ports) for the transmitted photons while only one path (camera only) in case Olympus setup. Therefore, in order to make use of the more interpretable image captured from the Olympus setup while attacking the DUA under the Mitutoyo setup, a grid similar to the coordinate system of the XY-stage was superimposed as shown in Figure 6.5. Using that superimposed grid avoided the difficulty of researching totally blind attacks. The grid in Figure 6.5 represents the coordinates of the XY-stage normalized by 1000 (i.e. 0 to 55,000 in the X direction corresponds to 0 to 2.2mm and 0 to 50,000 in the Y direction corresponds to 0 to 2mm while interpolating the values in the mentioned ranges). Also, dividing the die into four quads as shown in Figure 6.3b made it easier to attack structures that weren't identified at first glance.

The Mitutoyo setup in [109, 110] was re-used while only changing the DUA. The LPC1114FN28102 was tested on the PCB mentioned in subsection 4.3.1. However, us-

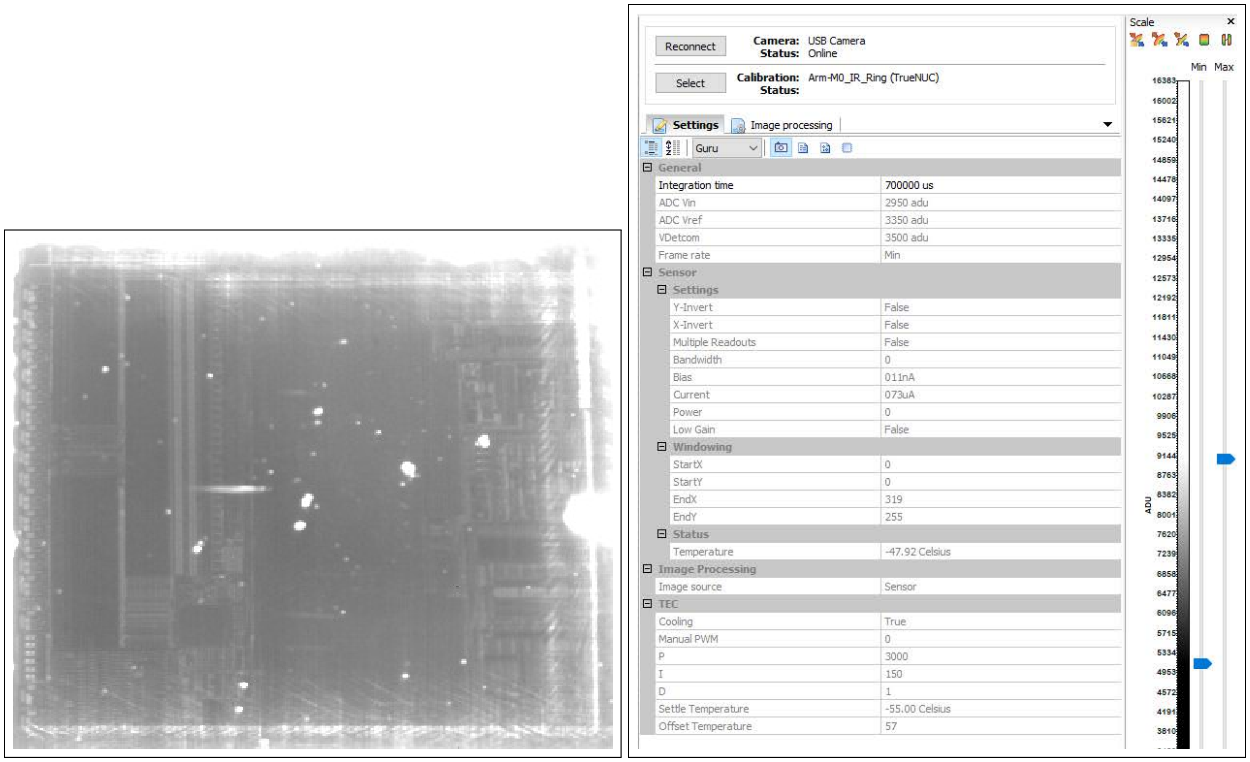


(a) (b)

Figure 6.3: ARM Cortex-M0 LPC1114FN28 a)backside decapsulated chip and b) backside image using SWIR Olympus.

ing one laser channel wasn't enough to successfully inject faults in different regions of the LPC1114FN28102 die (only the upper and lower right quads shown in Figure 6.3b were susceptible to successful LFI). Therefore, the test program and the PCB connections were modified in order to trigger both laser channels. When both channels were enabled, the bottom left quadrant shown in Figure 6.3b became susceptible to successful LFI. The upper left quadrant shown in Figure 6.3b only became susceptible to successful LFI when the 50X objective was used most probably due to the increased power intensity. The beam spot size was measured at 1.75% of the peak current with only one laser channel (CH1) turned on to be around $40\mu m$ and $4\mu m$ using the image acquired with 5X and 50X objectives respectively as shown in Figure 6.6. The 50X objective used was the Mitutoyo 50X M Plan Apo NIR objective. Also it's worth mentioning that when we tried to use the 50X objective early in the experiments, we were not able to capture the laser beam spot neither using SWIR nor the NIR camera. There was no recognizable circular illumination like the one shown in Figure 6.6b. We found that depending upon the silicon substrate surface smoothness the attacker may not see the spot size due to reflectance if the polishing quality was not good enough. Details about how different regions were susceptible to successful LFI will be further discussed in this chapter.

It might be worth mentioning that the second laser channel of AplhaNov PDM+ did not provide the same amount of power as the first channel. Also other laser parameters weren't identical. Details about both laser channels can be found in Appendix A.



(a)

(b)

Camera: USB Camera
Status: Online

Calibration: Arm-M0_IR_Ring (TrueNUC)
Status:

General	
Integration time	700000 us
ADC Vin	2950 adu
ADC Vref	3350 adu
VDetcom	3500 adu
Frame rate	Min
Sensor	
Settings	
Y-Invert	False
X-Invert	False
Multiple Readouts	False
Bandwidth	0
Bias	011nA
Current	073uA
Power	0
Low Gain	False
Windowing	
StartX	0
StartY	0
EndX	319
EndY	255
Status	
Temperature	-47.92 Celsius
Image Processing	
Image source	Sensor
TEC	
Cooling	True
Manual PWM	0
P	3000
I	150
D	1
Settle Temperature	-55.00 Celsius
Offset Temperature	57

Figure 6.4: a) ARM Cortex-M0 LPC1114FN28102 backside image using SWIR camera under Mitutoyo setup and b) Settings for such view.

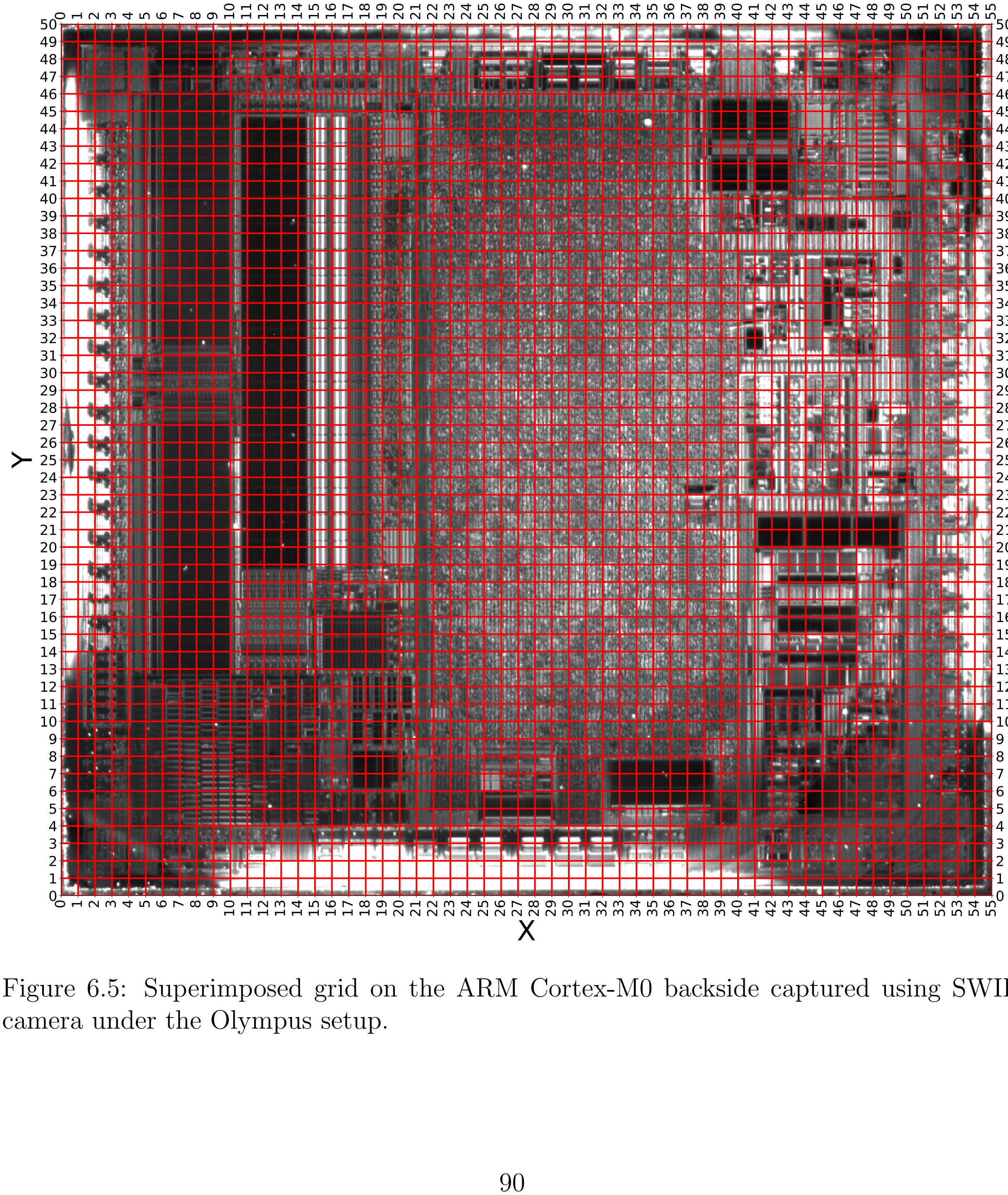
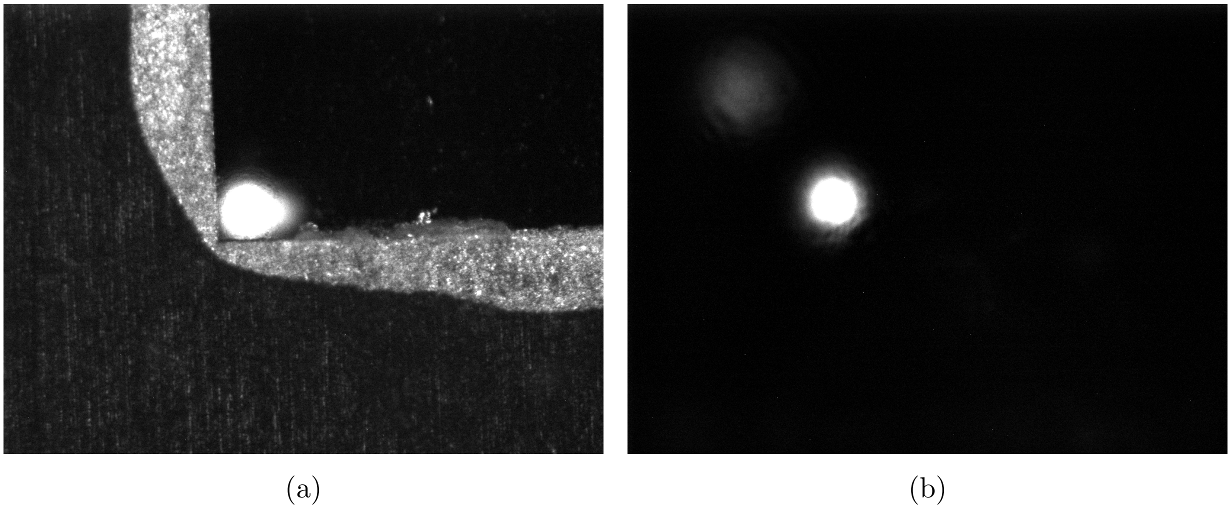


Figure 6.5: Superimposed grid on the ARM Cortex-M0 backside captured using SWIR camera under the Olympus setup.



(a) (b)

Figure 6.6: Laser beam spot with 1.75% laser intensity at the bottom left corner of the ARM Cortex-M0 LPC1114FN28102 using a) 5X and b) 50X objectives.

6.2.1 LFI firmware

The DUA was programmed with a firmware written in a way to target the manipulation of the decoded instructions coming out of the Flash memory. The firmware was written in assembly to ensure the correct timing of the laser pulse relative to the target instruction. The pseudo-code of the firmware is shown in Listing 6.1. In all code listings in this chapter, the text after the ";" symbol is a comment and the "..." denote that the instruction is repeated a number of times as noted in the comment. The firmware starts by properly initializing the DUA to ensure proper operation and the clock signal is supplied by an external wave generator. The system initialization is followed by the UART initialization as the UART will be used to send the data to the PC for analysis. Then each byte of the 4K bytes of the SRAM is initialized with 0xAA then certain arbitrary locations are initialized with specific values to detect if there was any address shifting when reading/writing data from/to SRAM. The reason of using 0xAA for initialization because it's an alternating pattern of 0's and 1's and would help in detecting bit sets and resets if there's any. After the system, UART and SRAM initialization, the laser is triggered just before the target instruction(s) as demonstrated in Listing 6.2. If the target instruction is manipulated, the value at the arbitrary location on line 1 in Listing 6.2 should be different from the expected value indicating a successful LFI. All the 4K bytes are sent to the PC using UART then stop execution using an infinite loop. If a hardfault is injected, the firmware keeps sending the 32 bytes of the stack frame depicted in Figure 6.2 repetitively (code elaborated in Listing B.3 in Appendix B) until the PC receive 4K bytes. If the DUA hangs or becomes locked-up (instruction execution stops because of LFI), the PC times out after 5 seconds instead of waiting indefinitely to receive the 4K bytes. The reason for the 5 seconds timeout is to be able to automate the experiments at different locations with different laser parameters (e.g. laser pulse width and amplitude) and operating conditions (e.g. supply voltage) then conduct the LFI analysis later. The mentioned LFI program in this subsection was run at each scanned target location and then the DUA was reset through a master Python script before re-running the program at a different target location or with different pulse parameters and/or operating conditions.

Listing 6.1: Pseudo-code for ARM Cortex-M0 LPC1114FN28102

```
BL      SYS_INIT        ;System configuration
BL      UART_INIT       ;Initialize UART
BL      SRAM_INIT       ;Initialize SRAM
BL      LASER_TRIGGER   ;Trigger laser
ADDS    R1, #0x01       ;Target instruction
STRB    R1, [R0]
BL      SRAM_readout    ;UART sends 4K bytes to PC
STOP    BL STOP
```

Listing 6.2: Code for LFI triggering using ARMv6-M instruction set

```
LDR     R0, =SRAM_arbitrary_location1
LDRB    R1, [R0]
 ;Trigger PDM laser CH1 and CH2 using PIO_7 and PIO_9
LDR     R2, =GPIO0DATA
LDR     R3, [R2]
LDR     R4, =((1<<PIO_7):OR:(1<<PIO_9))

 BICS   R3, R3, R4      ;Reset PIO_7 and PIO_9
STR     R3, [R2]        ;
EORS    R3, R3, R4      ;Set PIO_7 and PIO_9
STR     R3, [R2]        ;PIO_7 and PIO_9 are toggled (high)
ADDS    R1, #0x01       ;Target instruction repeated 85 times
 ...
```

```
ADDS    R1, #0x01
STRB    R1, [R0]
BL      SRAM_readout      ;UART sends 4K bytes to PC
STOP    BL STOP
```

6.3 Locating Debug Access Port (DAP) circuitry

During the design of the development PCB mentioned in subsection 4.3.1, a debugger was thought to be a very useful tool in investigating successful LFI. The ARM Cortex-M0 LPC1114FN28102 uses the SWD protocol to communicate with a compatible debugger. The SWD protocol requires two main signals to operate: data I\O and Clock. For ARM Cortex-M0 LPC1114FN28102, those two signals are known as: SWDIO and SWCLK as shown in Figure 6.7. The debugger ULINK2 was chosen over other debuggers from other vendors because it provides more insight when used with the Keil® Microcontroller Development Kit (MDK). Keil® MDK is the most comprehensive software development solution for Arm®-based MCUs and includes all components needed to create, build, and debug embedded applications. SWD protocol is a 2-pin debugging protocol which is better when compared to 5-pin debugging Joint Test Action Group (JTAG) protocol. SWD protocol can offer better performance than JTAG and can provide the same processor debug functionality.

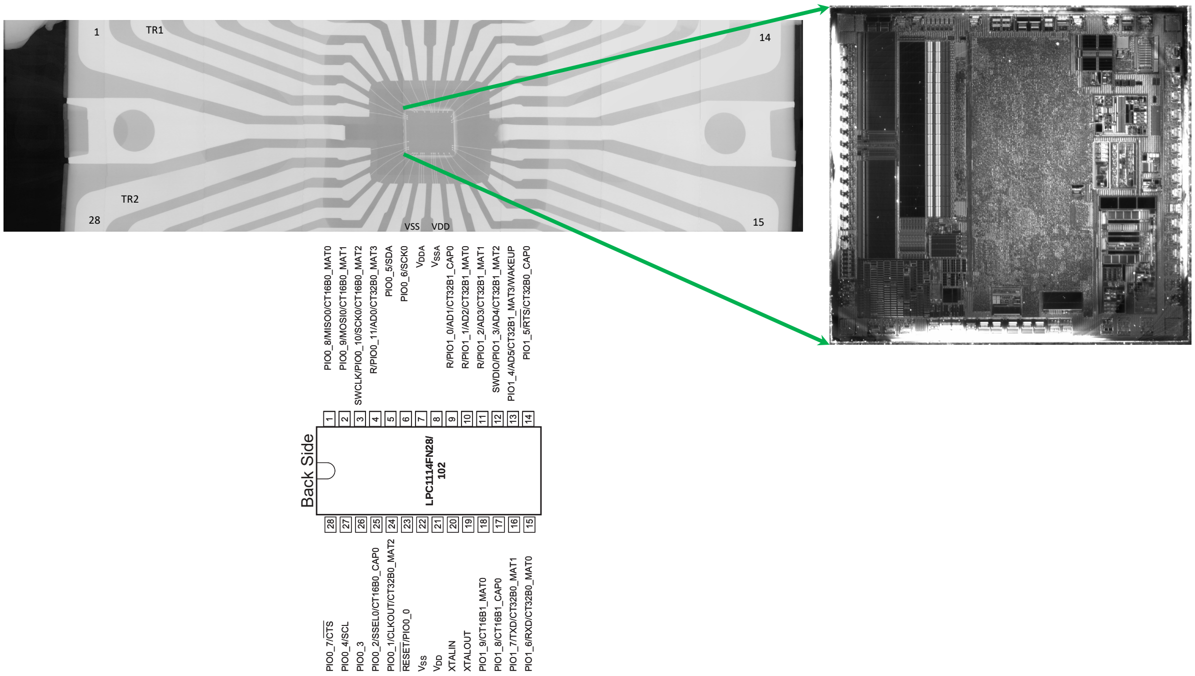


Figure 6.7: Pin mapping through X-ray scanning of a backside decapsulated ARM Cortex-M0 LPC1114FN28102 chip. TR1/TR2 are the pins used for triggering channels 1/2 of the laser source.

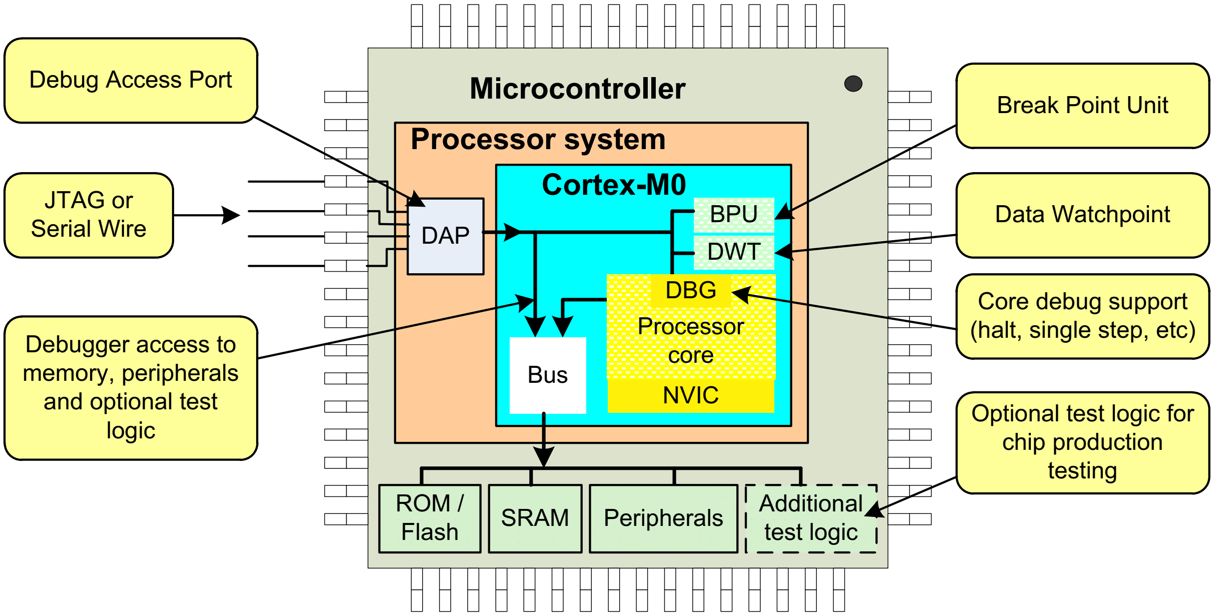


Figure 6.8: DAP in ARM Cortex-M0 LPC1114FN28102 [11].

The Debug Access Port (DAP) shown in Figure 6.8 provides up to 4 hardware breakpoints and 2 hardware watchpoints. The Breakpoint instruction *BKPT* could be used for unlimited number of software breakpoints.

Also, The Cortex-M0 processor has a halt mode, which stops program execution and allows the debugger to access processor registers and memory space. During halt mode, the following activities occur:

- Instruction execution stops.
- If the processor was in sleep mode, it wakes up from the sleep mode before halt. Two sleep modes can be entered by two instructions: *WFI*, *WFE*.
- Registers in the processor's register bank, as well as special registers, can be accessed (both read and write).
- Memory and peripheral contents can be accessed (this can be done without halting the processor).
- You can resume program execution, carry out single-step operation, or reset the MCU.

The DAP circuit location in the die was discovered when doing a LFI scan for the upper right quadrant in Figure 6.3b. The zeros, in the heat map in Figure 6.9 or any subsequent heat map, refer to positions where no faults were injected in the DUA. A reset behavior was detected as shown in the heat map in Figure 6.9 at the Not-A-Number (NaN)-labelled locations. The oscilloscope detected recurring laser pulses and the PC didn't receive any bytes that were supposed to be sent by the LFI firmware downloaded on the DUA. Also no hardfault was reported at that location therefore no bytes were sent by the UART. To further understand why and when the reset behavior happened at that location, the debugger was connected to the DUA and the LFI firmware was executed one assembly instruction at a time. When the laser triggering instruction (*STR R3, [R2]*) was reached (11th line in Listing 6.2), the DUA was still functioning as expected. As soon as the laser triggering instruction was executed, the debugging session was ended as shown in Figure 6.10. The ending of the debugging session due to the LFI effect is expected because the DUA was reset by an external factor other than the debugger itself. The debugger can be used to legitimately reset the DUA if desired during a debugging session without being kicked out of the debugging mode. The reset behaviour in the DAP area was 100% repeatable even after moving the laser back to the origin before re-doing the experiment. However some locations might not be detected as shown in Figure C.1 in Appendix C when the XY-stage resolution is changed due to the high spatial sensitivity of LFI to inject faults successfully and also due to the XY-stage position reproducibility. Also note that the 50X was used in Figure C.1 in Appendix C compared to the 5X in Figure 6.9. A possible cause of the reset behavior that leads to the recurring pulses is that the debugger, which can be used to legitimately request a reset through the DAP circuitry, is probably replaced by an illegitimate reset request induced by the LFI. Th effect of different utilized percentages of the peak current is illustrated in Appendix C in Figures C.7-C.11 where the sensitive spot is getting larger as the used percentage of the peak current increases.

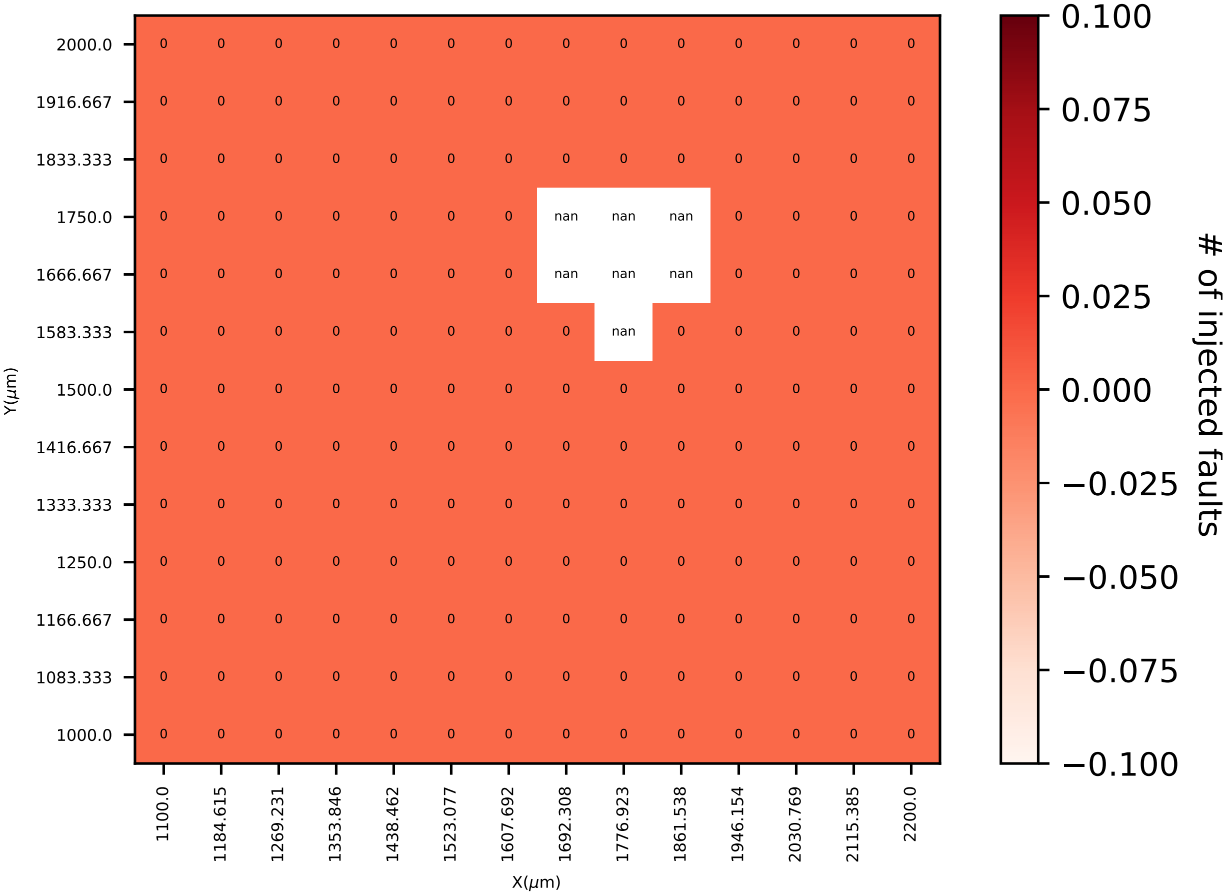
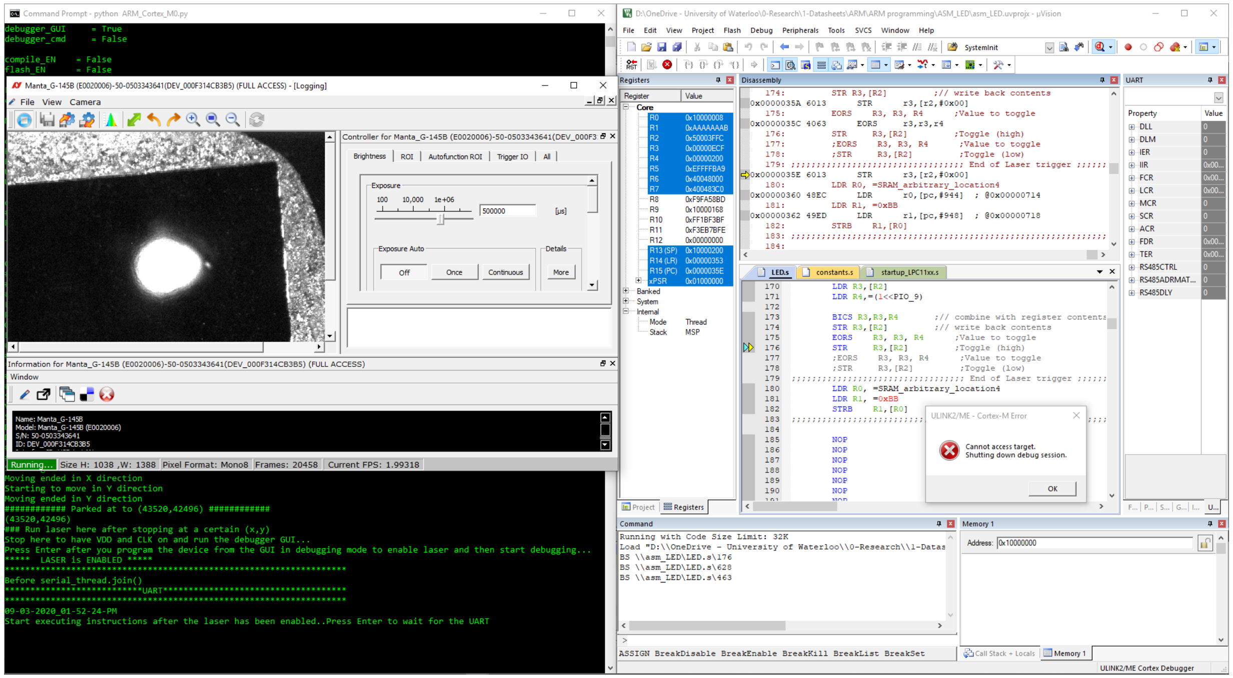


Figure 6.9: DAP location in heat map. 5X objective used at 182 locations scanned in the upper right quadrant at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 100.0% of peak current with both PDM CH1 and CH2 enabled. The seven sensitive locations (labelled as NaN) experienced a reset behavior.


Command Prompt - python ARM_Cortex_M0.py
Manta_G-145B (E0020006)-50-0503343641(DEV_000F314CB3B5) (FULL ACCESS) - [Logging]
File
View
Camera
Controller for Manta_G-145B (E0020006)-50-0503343641(DEV_000F3
Brightness
ROI
Autofunction ROI
Trigger IO
All
Exposure
100
10,000
1e+06
[µs]
Exposure Auto
Off
Once
Continuous
Details
More
Information for Manta_G-145B (E0020006)-50-0503343641(DEV_000F314CB3B5) (FULL ACCESS)
Window
D:\OneDrive - University of Waterloo\0-Research\1-Datasheets\ARM\ARM programming\ASM_LED\asm_LED.uvprojx - µVision
File
Edit
View
Project
Flash
Debug
Peripherals
Tools
SVCS
Window
Help
SystemInit
Registers
Disassembly
UART
LED.s
constants.s
startup_LPC11xx.s
Project
Registers
Command
Memory 1
Call Stack + Locals
Memory 1
ULINK2/ME Cortex Debugger

```
debugger_GUI     = True
debugger_cmd     = False

compile_EN    = False
flash_EN      = False
```

Exposure: 500000

```
Name: Manta_G-145B
Model: Manta_G-145B (E0020006)
S/N: 50-0503343641
ID: DEV_000F314CB3B5
```

Running... | Size H: 1038 ,W: 1388 | Pixel Format: Mono8 | Frames: 20458 | Current FPS: 1.99318

```
Moving ended in X direction
Starting to move in Y direction
Moving ended in Y direction
############ Parked at to (43520,42496) ############
(43520,42496)
### Run laser here after stopping at a certain (x,y)
Stop here to have VDD and CLK on and run the debugger GUI...
Press Enter after you program the device from the GUI in debugging mode to enable laser and then start debugging...
*****  LASER is ENABLED *****
*****************************************************************
Before serial_thread.join()
**************************UART***********************************
*****************************************************************
09-03-2020_01-52-24-PM
Start executing instructions after the laser has been enabled..Press Enter to wait for the UART
```

Register	Value
Core	
R0	0x10000008
R1	0xAAAAAAAB
R2	0x50003FFC
R3	0x00000ECF
R4	0x00000200
R5	0xEFFFFBA9
R6	0x40048000
R7	0x400483C0
R8	0xF9FA58BD
R9	0x10000168
R10	0xFF1BF3BF
R11	0xF3EB7BFE
R12	0x00000000
R13 (SP)	0x10000200
R14 (LR)	0x00000353
R15 (PC)	0x0000035E
xPSR	0x01000000
Banked	
System	
Internal	
Mode	Thread
Stack	MSP

```
   174:         STR R3,[R2]          ;// write back contents
0x0000035A 6013      STR      r3,[r2,#0x00]
   175:         EORS    R3, R3, R4        ;Value to toggle
0x0000035C 4063      EORS     r3,r3,r4
   176:         STR     R3,[R2]           ;Toggle (high)
   177:         ;EORS    R3, R3, R4        ;Value to toggle
   178:         ;STR     R3,[R2]           ;Toggle (low)
   179: ;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;; End of Laser trigger ;;;;;;
0x0000035E 6013      STR      r3,[r2,#0x00]
   180:         LDR R0, =SRAM_arbitrary_location4
0x00000360 48EC      LDR      r0,[pc,#944]  ; @0x00000714
   181:         LDR R1, =0xBB
0x00000362 49ED      LDR      r1,[pc,#948]  ; @0x00000718
   182:         STRB    R1,[R0]
   183: ;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
   184:
```

```
        LDR R3,[R2]
        LDR R4,=(1<<PIO_9)

        BICS R3,R3,R4          ;// combine with register contents
        STR R3,[R2]            ;// write back contents
        EORS    R3, R3, R4        ;Value to toggle
        STR     R3,[R2]           ;Toggle (high)
        ;EORS    R3, R3, R4        ;Value to toggle
        ;STR     R3,[R2]           ;Toggle (low)
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;; End of Laser trigger ;;;;;;
        LDR R0, =SRAM_arbitrary_location4
        LDR R1, =0xBB
        STRB    R1,[R0]
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

        NOP
        NOP
        NOP
        NOP
        NOP
        NOP
```

ULINK2/ME - Cortex-M Error

Cannot access target.
Shutting down debug session.

OK

Property	Value
DLL	0
DLM	0
IER	0
IIR	0x00...
FCR	0x00...
LCR	0x00...
MCR	0
SCR	0
ACR	0
FDR	0x00...
TER	0x00...
RS485CTRL	0
RS485ADRMAT...	0
RS485DLY	0

```
Running with Code Size Limit: 32K
Load "D:\\OneDrive - University of Waterloo\\0-Research\\1-Datas
BS \\asm_LED\LED.s\176
BS \\asm_LED\LED.s\628
BS \\asm_LED\LED.s\463
```

```
>
ASSIGN BreakDisable BreakEnable BreakKill BreakList BreakSet
```

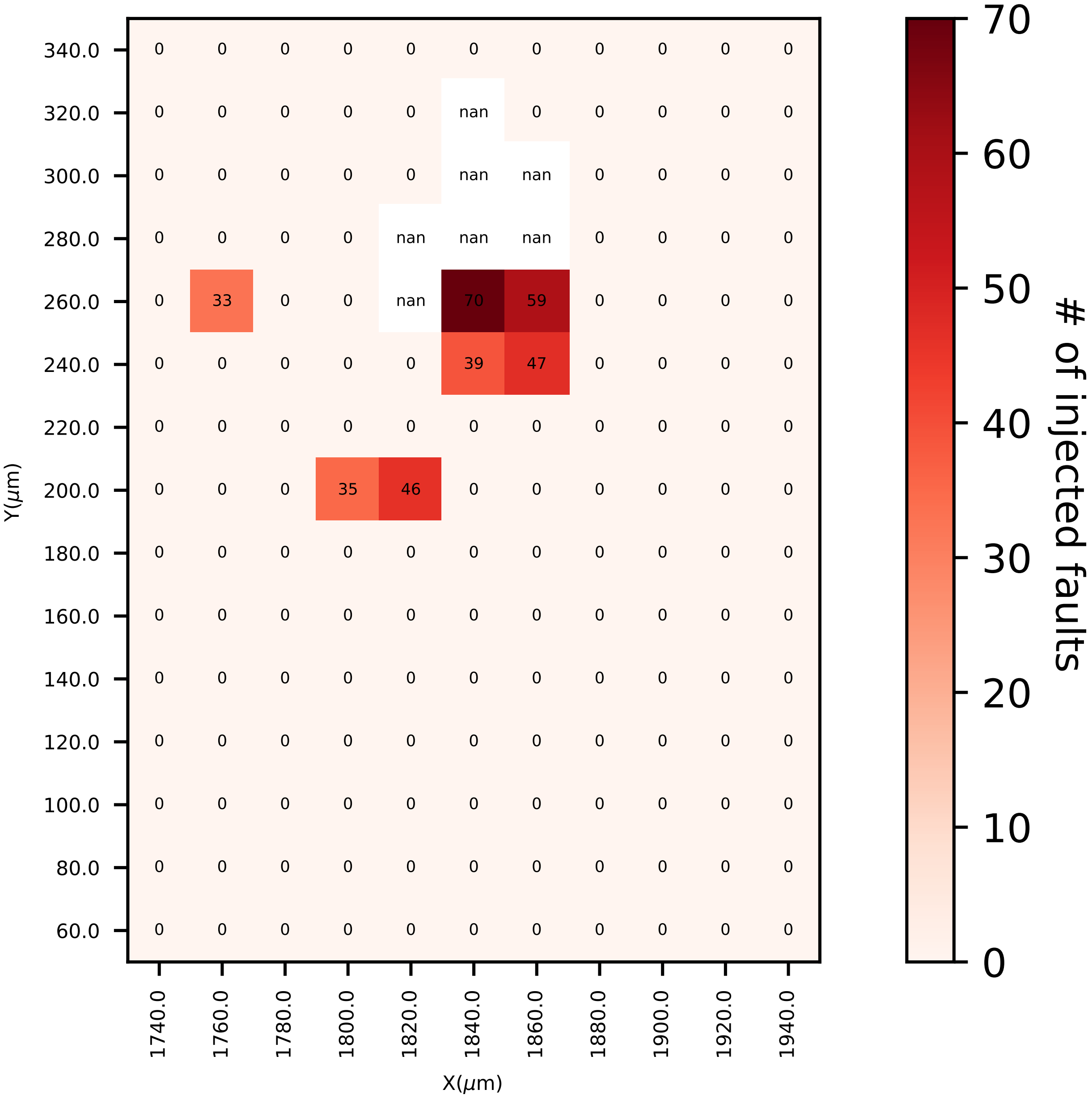
Address: 0x10000000

Figure 6.10: Debugger session ended as a result of LFI.

6.4 Attacking UART

One of the most common peripherals in every MCU is the UART. The UART location in the chip was discovered when the LFI firmware was modified to target seven different instructions instead of only targeting the *ADDS* instruction. Different target instructions were isolated in the assembly code from each other by inserting a series of *NOP* instructions as elaborated in Listing B.1 in Appendix B. The effect of the seven laser pulses injected in the DUA had an effect similar to the reset effect experienced while locating the DAP circuit. However, this time it wasn't a reset because laser pulses weren't recurring which was confirmed both visually using the NIR camera during the five seconds of the UART timeout and also confirmed through the oscilloscope that captures any laser pulses. The effect of those seven pulses is that the PC didn't receive any bytes from the UART. A possible explanation for these results of Listing B.1 in Appendix B is further elaborated in section 6.4.1.

In order to investigate the effect of the LFI on the UART, the firmware was modified to target only the *ADDS* instruction with seven identical consecutive laser pulses as elaborated in Listing B.2 in Appendix B. The result of such experiment is shown in Figure 6.11. Instead of causing a hanging effect, only partial hanging occurred where the PC received 1440 bytes out of the expected 4K bytes but all of the received bytes had the value 0x*FF*. The NaN-labelled locations are those locations where the partial hanging fault was experienced. Also the location of the UART RX and TX pins show in Figure 6.7 are in the neighborhood of the discovered location for the UART circuitry shown in Figure 6.15. Out of the 14 sensitive locations, seven locations (labelled as NaN) experienced a partial hanging fault, where only 1440 bytes were sent by the UART as 0x*FF* (these results may indicate a possible location for the transmission register U0THR). The remaining seven locations experienced shifting faults, where likely data was read from the wrong addresses. This was probably a bus error due to LFI because the 4K bytes of the SRAM were read but with some shifting in the order of the bytes. However the observed shifting faults (throughout this chapter) were not repeatable. This may be explainable due to the cumulative effect of the injected laser pulses at the previous location despite waiting one second between each location after resetting the DUA through the reset pin controlled via the master Python script. The annotated numbers in Figure 6.11 indicated the number of discrepancies between the data read and data expected from the SRAM, due to the shifting faults.



	1740.0	1760.0	1780.0	1800.0	1820.0	1840.0	1860.0	1880.0	1900.0	1920.0	1940.0
340.0	0	0	0	0	0	0	0	0	0	0	0
320.0	0	0	0	0	0	nan	0	0	0	0	0
300.0	0	0	0	0	0	nan	nan	0	0	0	0
280.0	0	0	0	0	nan	nan	nan	0	0	0	0
260.0	0	33	0	0	nan	70	59	0	0	0	0
240.0	0	0	0	0	0	39	47	0	0	0	0
220.0	0	0	0	0	0	0	0	0	0	0	0
200.0	0	0	0	35	46	0	0	0	0	0	0
180.0	0	0	0	0	0	0	0	0	0	0	0
160.0	0	0	0	0	0	0	0	0	0	0	0
140.0	0	0	0	0	0	0	0	0	0	0	0
120.0	0	0	0	0	0	0	0	0	0	0	0
100.0	0	0	0	0	0	0	0	0	0	0	0
80.0	0	0	0	0	0	0	0	0	0	0	0
60.0	0	0	0	0	0	0	0	0	0	0	0

Figure 6.11: 5X objective used at 182 locations scanned in a specific area in the bottom right quadrant at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 100.0% of the peak current with only PDM CH1 enabled.

6.4.1 Lockup state

The Cortex-M0 processor can enter a lockup state if another fault occurs during the execution of a hard fault exception handler or when a fault occurs during the execution of an NMI handler. This is because when these two exception handlers are executing, the priority level does not allow the hard fault handler to preempt. During the lockup state, the processor stops executing instructions and asserts a LOCKUP status signal. Depending on the implementation of the MCU, the LOCKUP status signal can be programmed to reset the system automatically, rather than waiting for a watchdog timer to time out and reset the system. The lockup state prevents the failed program from corrupting more data in the memory or data in the peripherals. During software development, this behavior can help us debug the problem, as the memory contents might contain vital clues about how the software failed. A number of conditions can cause lockup in the Cortex-M0 processor (or ARMv6-M architecture):

- A fault occurred during the execution of the NMI handler.
- A fault occurred during the execution of the hard fault handler (double fault).
- There was an SVC instruction execution inside the NMI handler or the hard fault handler (insufficient priority).
- A bus error response during reset sequence (e.g. when reading initial stack pointer value).
- There was a bus fault during the unstacking of the xPSR during the exception return using the main stack pointer (MSP) for the unstacking.

When the LFI firmware targeted seven instructions separated by a series of NOPs to isolate any LFI effects as elaborated in Listing B.1 in Appendix B, this led to experiencing a full hanging where the UART didn't send any bytes to the PC. This effect is similar to the Lockup state detailed earlier in this section.

6.5 Resetting without reset

A location in the chip near the reset pin shown in Figure 6.7 was found to be exploitable by LFI. The effect of this laser injection was to reset the chip as shown in Figure 6.12. The reset effect was confirmed because instead of injecting only one pulse at each location during the scanned target area, the laser beam spot appeared constantly during the 5 seconds timeout allocated for each location and the LFI firmware didn't send any of the 4K bytes to PC through the UART. Besides, the TDS7254 oscilloscope shown in Figure 4.4 was used to capture the recurring laser pulses resulting from the reset of instruction execution after the laser pulse hits the DUA. This experiment was 100% repeatable.

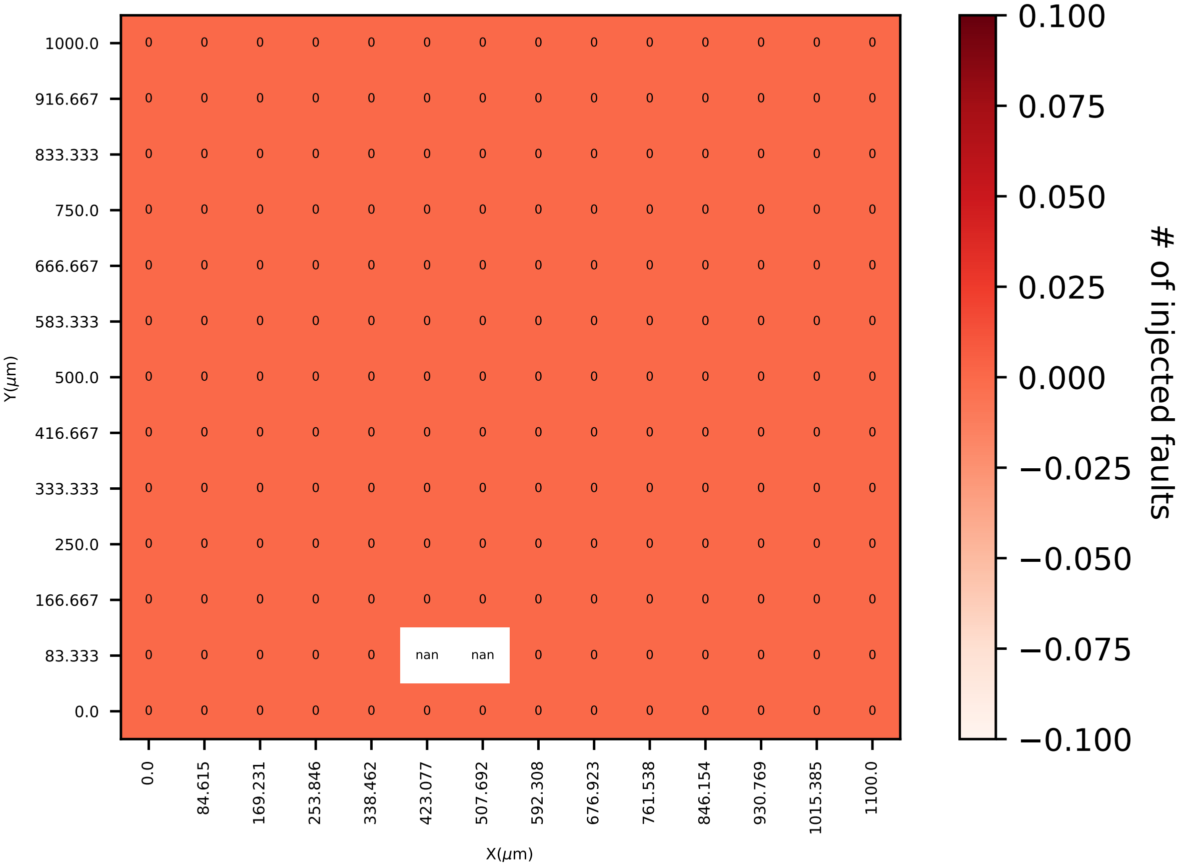


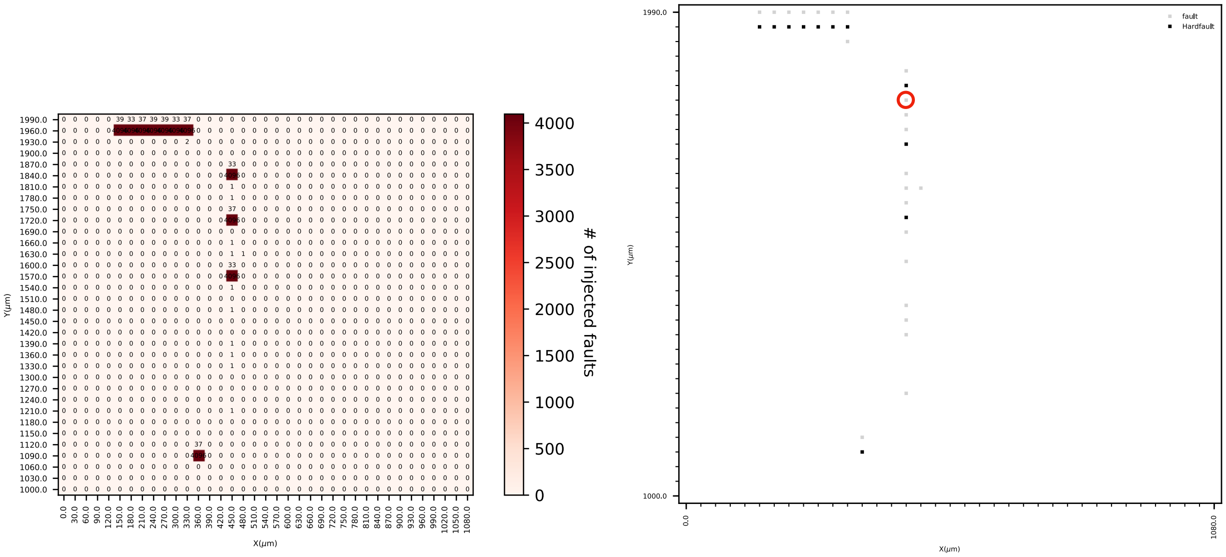
Figure 6.12: 50X objective used at 182 locations scanned in the bottom left quadrant at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 100.0% of peak current with both PDM CH1 and CH2 enabled. Only two locations (labelled as NaN) experienced a reset when the shown resolution was used for scanning.

6.6 Instruction manipulation

The instruction register location (output of the Flash memory) was guessed at first based on the backside image of the ARM Cortex-M0 LPC1114FN28102 as shown in Figure 6.15 then confirmed by scanning such area and analyzing the scanning results. Figure 6.13 and Figure 6.14 show that most of the induced hardfaults were in the upper part of the instruction register with the exception of the first fault at the top of Figure 6.13b (marked inside the red circle). This location is probably a valid bit as the number of uniform structures in the instruction register is 17 while the opcode is only 16 bits. Since most of the hardfaults are in the upper part of the instruction register, it's probably because the Most Significant Bit (MSB) of the opcode is located at the top of the instruction register below the valid bit and the Least Significant Bit (LSB) may be located at the bottom of the instruction register in the bottom left quadrant of the DUA. It's believed to be a valid bit because when it is hit by the laser it leads to an instruction skip effect (as if a *NOP* was executed). The hardfault was confirmed to be triggered at 16 cycles after the target instruction by reading out the program counter in the dumped stack frame illustrated in Figure 6.2.

Successful manipulation of the immediate value in the 'ADDS R1, #0x01' instruction by getting a value of 0x*FE* instead of 0x*FF* at the SRAM memory address 0x*10000801* was achieved only through targeting the beam at the red-circled location in Figure 6.14b. This effect was experienced by resetting the 1-bit in the immediate value in the 'ADDS R1, #0x01' instruction (12$^{\text{th}}$ line in Listing 6.2), likely creating a faulty instruction 'ADDS R1, #0x00'. Other locations which had only one fault in Figure 6.13a was also a faulty value at the SRAM address 0x*10000801* probably due to changing either the immediate value or the source/destination register used in the 'ADDS R1, #0x01' instruction to another register other than R1. Locations with more than one faulty value in Figure 6.13a were shifting faults. When the blue-marked location in Figure 6.14b (probably part of the address decoding circuit or the register file) was targeted under the same settings, the faulty value at the same SRAM memory address 0x*10000801* was 0x*CC* instead of 0x*FF*. The reproducibility of the instruction manipulation attack was affected by the repeatability of the XY-stage movement accuracy affecting laser positioning and the chosen scan resolution as shown in Figure 6.13a compared to Figure C.2 and C.3 in Appendix C. Note that a higher scan resolution was used in Figure C.2. Also in Appendix C, Figures C.4, C.5 and C.6 depict the effect of the repeatability of the XY-stage on the successful LFI at a single fine line scan of the output of the Flash memory area annotated in Figure 6.15.

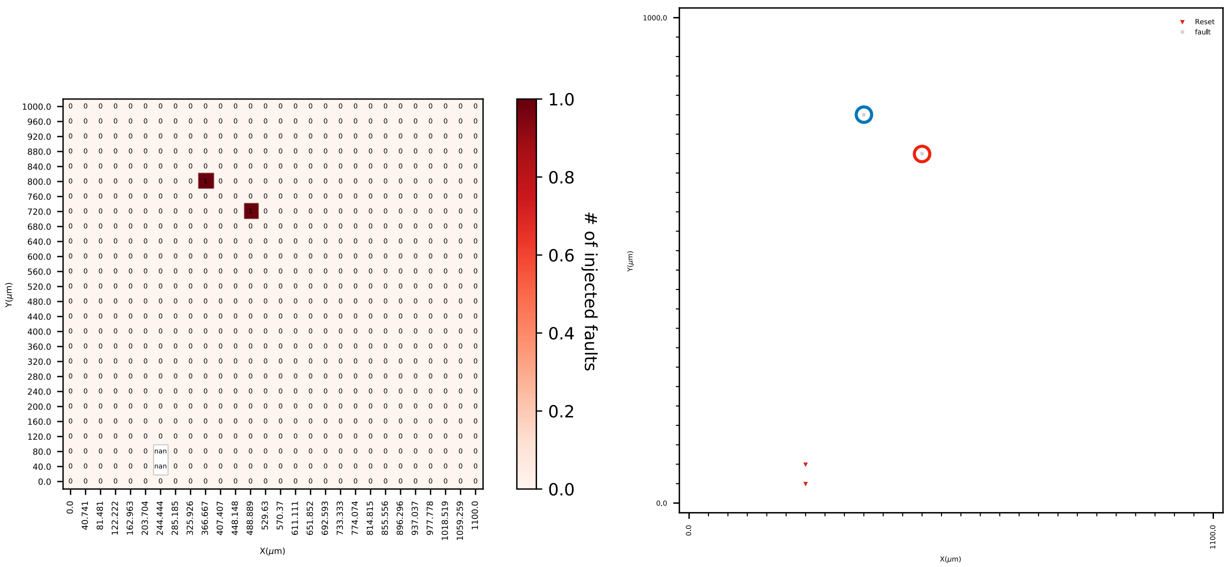
Note that the dashed shapes in Figure 6.15 indicate rough locations of LFI but details are found in specific Figures 6.11-6.14 and Appendix C.



(a)

(b)

Figure 6.13: a) Heatmap and b) scatter plot for ARM Cortex-M0 LPC1114FN28 upper left quadrant scan using 50X objective at 1258 locations at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 100.0% of the peak current with both PDM CH1 and CH2 enabled. 10 out of shown 34 sensitive locations produced hardfaults, whereas other locations having more than one fault were shifting faults.



(a)

(b)

Figure 6.14: a) Heatmap and b) scatter plot for ARM Cortex-M0 LPC1114FN28 bottom left quadrant scan using 50X objective at 728 locations at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 100.0% of the peak current with both PDM CH1 and CH2 enabled. Out of the four sensitive locations two triggered chip reset and the other two injected faults in the target instruction.

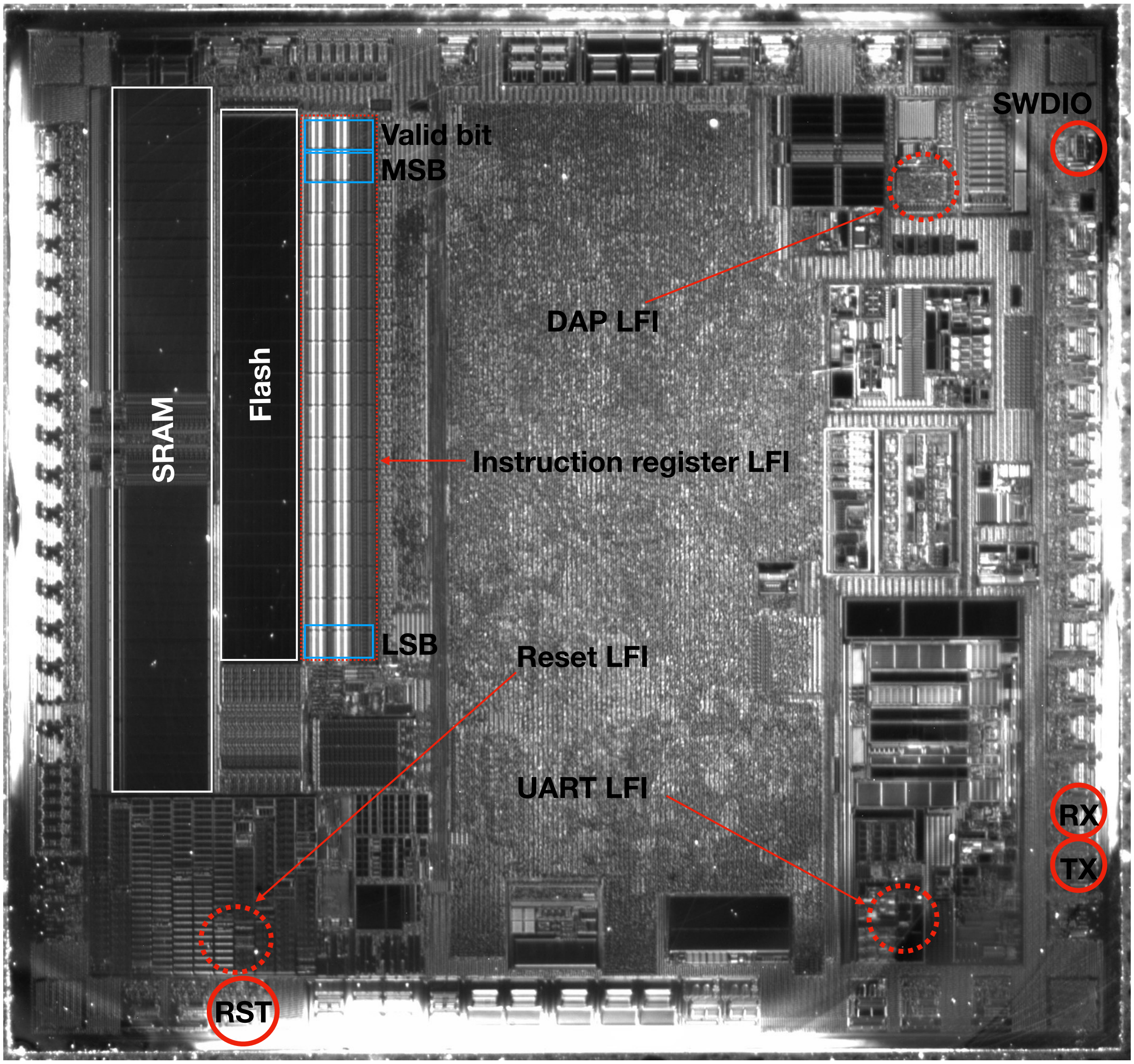


Figure 6.15: Annotated probable structures in ARM Cortex-M0 LPC1114FN28102. Solid circles indicate pins while dashed circles indicate rough LFI areas. The UART pins (RX,TX), Reset pin (RST) and SWD debugger IO pin (SWDIO) marked in Figure 6.15 could be mapped to the X-ray image in Figure 6.7.

6.7 PEM

In Figure 6.16 we could capture some PEM emissions using the setup described in subsection 4.3.2 but most of them are in the peripherals and we could not capture the emissions at the instruction register even with over-voltaging and single instruction execution.The instruction register emissions are probably blocked due to the thickness of the substrate and some substrate thinning might be required to capture those emissions with the SWIR camera. The thinning was not performed due to limitations of the used desktop CNC machine in the vertical direction (i.e. Z-axis could move in increments of $\sim 13\mu m$ only). The emissions shown in Figure 6.16 were acquired with core/IO supply voltage equal to 3.6V and offered a starting point for exploitable target areas and proved to be in close proximity to the probable identified structures in Figure 6.15. While several ARM Cortex-M0/M4 based MCUs were investigated in [131] and an ARM Cortex-M3 based MCU was attacked in [132], both [131, 132] didn't provide a PEM image where different structures of the DUAs could be identified through RE.

It might be worth mentioning that over-voltaging (up to 4V) didn't provide noticeable improvements regarding the captured emissions. On the other hand, under-voltaging (down to 1.65V) was expected to make the DUA more susceptible to LFI but it wasn't the actual case.

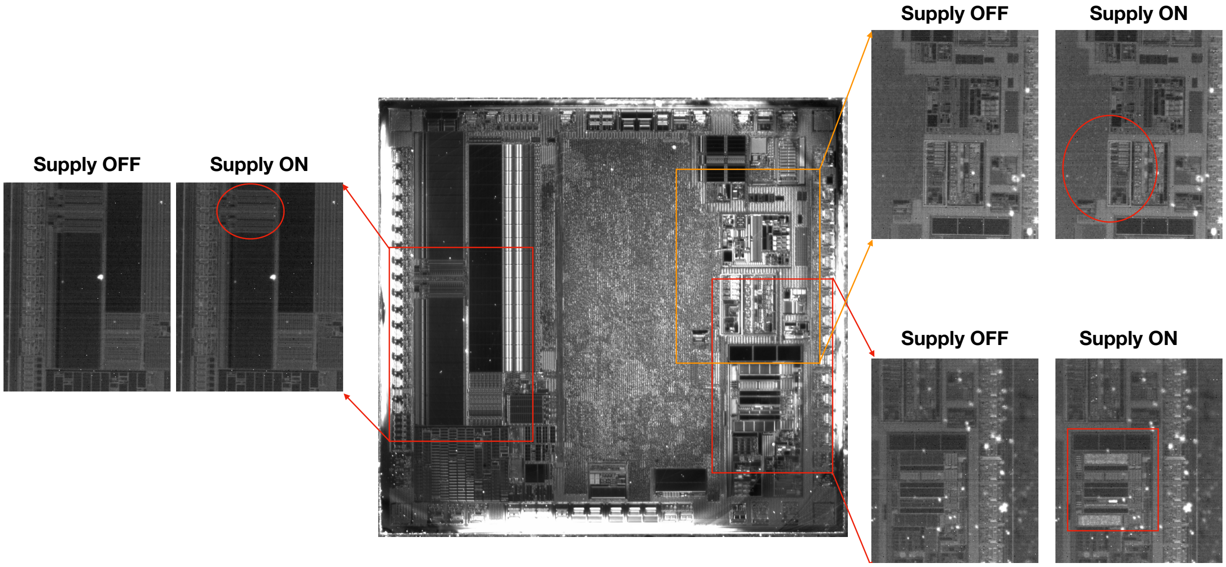


Figure 6.16: Captured PEM emissions in ARM Cortex-M0 LPC1114FN28102. Most of them are in the peripherals.

6.8 Countermeasures

Building on the discussion in section 5.5, the hardfault ISR was already implemented in ARM Cortex-M0 LPC1114FN28102. However, no latch-up effects were encountered in the ARM Cortex-M0 LPC1114FN28102 using the laser setup mentioned in sub-section 4.3.1. The hardfault hardware countermeasures which detect LFI could still be considered to be an effective but expensive countermeasure against LFI.

To counteract the reset attack presented in section 6.5, several locations over the chip could be used as trusted reset bit registers to differentiate between a legitimate reset request versus a LFI induced reset. Besides, on-chip LEDs could be implemented physically in the proximity of the debugging and UART circuitry to detect laser pulse(s) and shutdown the DUA as a countermeasure against the LFI attacks presented in sections 6.3 and 6.4 respectively.

6.9 Summary and comparison with previous research

During our various attacks on the ARM Cortex-M0 LPC1114FN28102, we didn't experience any Single Event Latch-ups (SELs) reported in [131] probably since the used AlphaNov PDM+ laser module used in our setup (maximum 3W for an adjustable laser pulse width between $5ns$-$500ns$) was not as powerful as the laser used in [131] ($1.2\mu J$ maximum laser energy for a fixed laser pulse width of $800ps$). However, we were able to cause Single Event Transients (SETs) through targeting the output of the instruction register in ARM Cortex-M0 LPC1114FN28102. Also we were not able to cause SEUs in the SRAM of ARM Cortex-M0 LPC1114FN28102 which was also reported by [131] totally for STM32F051R8T6 from ST Microelectronics and partially for LPC11E14FBD64 from NXP which are also both ARM-Cortex M0 based MCUs.

With both PDM channels enabled for a 200ns pulse through the 5X objective, the current percentage had to be above 65% in order to be able to inject faults. The objective also determines the spot size thus the power intensity. Based on our empirical results, from section 6.3 till section 6.6, it appears that given the best x-y-z position, it is the power intensity of the laser pulse that matters (both laser pulse width per unit area and laser pulse amplitude per unit area) as the most important metric to look for successful LFI.

Usually in consumer end products, manufacturers disable access to the debugging circuitry as a security measure. Therefore locating the debugging circuitry as described in section 6.3 could be an alarm for manufacturers to secure the debugging circuitry itself rather than just blocking access to the pins. In addition, if the attack mentioned in section 6.3 is conducted during an online circuit debugging session, it would trigger false failures which could be detrimental to the debugging process. The UART attack mentioned in section 6.4 could be used to fault desired bytes communicated by the DUA through the UART interface. For an example, it would be dangerous if the UART is attacked while sending critical medical information to hospitals. Also the reset attack mentioned in section 6.5 could be used to reset health devices as well as reset the counter for wrong Personal Identification Number (PIN) entries (e.g password guessing reset attack) when there's no access to the DUA's reset pin. The next chapter provides a brief summary as well as discussion of contributions and future work.

Chapter 7

Discussions, Conclusions and Future Work

This chapter will provide a brief summary of the research results. A discussion of the research limitations and contributions is also presented. Future work including a discussion of countermeasures is also presented.

7.1 Summary

In conclusion, the details of a LFI methodology are provided using both 5X and 50X objectives, quiescent PEM with SWIR camera (no over-voltaging), pulse timing (a glass slide plus APD) and imaging using 3D printed IR-LED ringlet. The laser thermal effect might be playing a small if not negligible part in creating the fault since the laser is pulsed and the thermal laser stimulation is typically done using a wavelength above $1300nm$ [133].

In one DUA, specifically the $450nm$ PIC16F687, only the 5X objective was required to precisely inject faults. Injecting the laser pulse at different locations along the Flash outputs (possible instruction register) provided direct instruction fault injection control and insertions. Additionally the single tuned laser pulse injection was shown to skip an arbitrary number of assembly instructions thus defeating assembly instruction replication countermeasures such as duplication, triplication, etc. The number of skipped instructions depended on several parameters such as the type of instruction, laser peak current, laser pulse width, frequency of the DUA, and location of the laser on the DUA.

For the other DUA, specifically the 140nm ARM Cortex-M0 LPC1114FN28102, both the 5X and 50X objectives were needed to be able to inject faults depending on the target area. Photon emissions were only captured at certain areas in the chip with no PEM emissions captured at the Flash memory output probably due to the increased substrate thickness for this more advanced technology node [125]. Despite having an automated LFI setup with a magnetic fixture for the DUA's custom-designed development PCB, equipment limitations such the repeatability of the motorized XY-stage affected the instruction manipulation experiment in section 6.6. However, target areas like the DAP and UART circuitry illustrated in section 6.3 and 6.4 respectively revealed new vulnerabilities for possible attack scenarios.

7.2 Contributions

The main contributions of this thesis are listed below:

- Sample preparation details for backside LFI using a relatively low cost CNC machine. Also sample preparation from the frontside was performed using low cost manual chemical etching unlike most previous research using very expensive equipment [68, 134–136] to prepare LFI-ready samples.

- Performed successful LFI analysis on two DUAs with different technology nodes, processor cores and design complexity. Most previous research utilized devices from consecutive technology nodes such as 90nm and 45nm in [102] or custom-made chips with full knowledge about the implementation and target areas such as in [127]. Also, in [131] the technology node was reported to be unknown for NXP LPC11E00 which is a very similar chip the NXP ARM Cortex-M0 LPC1114FN28102 investigated in this thesis.

- Analysis of dynamic LFI and quiescent photon emissions on an embedded processor aided by backside imaging to RE different structures on the decapsulated chip unlike previous research which only did laser analysis [127, 132] or only did PEM [6].

- Demonstrated a practical methodology to monitor the laser pulse during fault injection (using a microscope slide), including empirical analysis of the laser's pulse effect on the DUA using a field probe, data analysis and PEM. This is unlike previous research which analyzed pulse characteristics separate from experiments [137].

- Quiescent PEM with a 10X objective demonstrates that targeting different locations in the instruction register using the correct laser pulse energy (width and amplitude) for each location will result in controlled faults being injected into instructions. This is unlike previous research which utilized active (i.e. non-quiescent) PEM with no focus on instructions [122].

- Instruction-replacement fault injection supports new attacks such as adding one more round to AES-128 (11th round) through skipping the check for the 10th round, unlike previous research which largely focuses on attacks where instructions are skipped [115, 121] or bits are flipped in control registers using static faults [130].

- Ability to alter the immediate value of an instruction read from the Flash memory, unlike previous research which demonstrated entire instruction change [124] or attacks on flip-flops [90, 128] or SRAM [29, 48, 129].

- Reverse engineering the exact timing of the target clock cycle ($Q_{3_{\text{fetch}}}$) with quiescent PEM to determine when the laser should be effective for successful fault injection, unlike previous research [121].

- Robust automated LFI setup with magnetic fixture for DUA's custom developed PCB, scripted re-programming of the DUA for automated experiments and automatic active laser pulse capturing on the controlling PC through the oscilloscope and the APD detector/microscope slide unlike commercial LFI solutions [54].

- Analyzing the locations of inserted hardfaults which naturally support fault identification, thus acting as a countermeasure, unlike previous research [131] that only considered hardfaults as chip malfunction.

- Focusing on debug/UART circuitry as vulnerable areas for LFI attacks that were not previously discussed in the literature, thus attacking data sent off of chip and revealing new research vulnerabilities. This is unlike previous research which only reported latchups [131] or processor crashes [138] and didn't try to identify such LFI-sensitive areas.

7.3 Limitations and Future Work

Largely limitations are due to the available laboratory equipment. For example, the precision and repeatability provided by the used motorized XY-stage could be improved by using a higher-end motorized XYZ-stage, thus increasing the efficiency of the LFI which has high spatial dependency. Also, limited knowledge of the attacked processors, since they were not open source, was another aspect that limited the attacks through increasing the time needed to figure out sensitive areas of the DUA susceptible to LFI. Should the attacker have access to all the information of a chip (e.g. known layout structure, etc.), it is likely that an automated LFI attack using our setup could be performed in a matter of seconds; however we didn't have access to such a chip. While not having access to our own custom-designed chips was time-consuming to discover sensitive locations in a DUA, this lead to an understanding of obstacles for a real device attack thus fostering ideas on how to make it more difficult for the attacker. Analysing the data-sheets for the DUAs to understand the architecture, as well as interfacing and automating the hardware for equipment synchronization took a substantial amount of work. One key to improving the current LFI system would be to additionally automate the fault analysis process as well. Also one of the most important laser-independent parameters which should be reported is likely the energy in the laser spot per unit area on the DUA's surface. Unfortunately in this research we had no access to equipment to measure this parameter hence parameters were reported relative to inputs of the commercial laser system at hand.

Beside these equipment limitations, the LFI attack from the backside has a main limitation with respect to technology scaling even for a very highly equipped security lab. The beam spot size is limited by the wavelength of the laser beam, so $\sim 1\mu m$ is the smallest beam spot size for the $1064nm$ laser beam. The effect of the relatively large beam spot size on LFI for chips manufactured with different technology nodes was analyzed in [4, 90, 102, 127] starting from $350nm$ down to $90nm$, $45nm$ and $28nm$ where LFI still proves to be an effective and precise method to inject faults. It's unknown how the LFI will perform when attacking more advanced devices such as those manufactured using a $14nm$ process. Also, Fully Depleted Silicon on Insulator (FD-SOI) showed less sensitivity to LFI [4] and thus should be further explored for security dedicated circuits implementations. The shallow trench isolation (STI) and buried oxide (box) used in FD-SOI structures are made of an insulator. The FD-SOI structures channel is completely insulated from the substrate conversely to the bulk structure. Moreover, for the FD-SOI structure, the channel is made of intrinsic silicon. In a FD-SOI transistor, only two currents can be induced by the laser and those two currents are independent. The first current, which goes from source to drain, impacts the data path. The second current, which goes from N-well to substrate, changes the electric potential of the N-well. This change can alter the transistor function (e.g. threshold voltage). The threshold voltage modification modifies the transistors switching delay. This phenomenon may cause faults due to timing constraint violation in the circuits logical data path. Conversely, in a bulk structure without STI a laser beam generates many currents, which interact with each other. Thus there is a competition between all these currents. The result of this competition depends on experimental parameters such as the spot size and the lasers distance from the transistor [4]. Also, the SOI fin field-effect transistors (FinFETs) collected 270 times less charge than bulk FinFETs [139]. In terms of scalability, it's extremely unlikely that SOI FinFETs will replace MOSFETs for lower end devices because the MOSFETs are easy and cheap to produce. But at the high end, FinFETs are winning and Intel is already using them at their $22nm$ process node [140]. Therefore, SOI FinFETs may be more suitable for high end secure chips.

Without these limitations and given perfect equipment, LFI can be tailored to change code on the fly, control higher privileged functions like resets, modifying data sent off chip and changing addresses of the data being read out. Countermeasures will still remain difficult to design since lasers are dependent upon the specific location on the die. Therefore, the implementation of countermeasures remain challenging to thwart security attacks for important secure embedded systems. Integrity checks on instructions, authenticated-encrypted data sent off chip, and possibly other on-chip security-related data will be needed as countermeasures against successful LFI.

Future work would include focus on the following items in addition to further countermeasure development:

- Investigation of the effect of VMU objectives with higher magnification power on the beam spot size and successful LFI. Using 100X and 200X NIR objectives should result in $2\mu m$ and $1\mu m$ beam spot sizes respectively. A smaller beam spot size mean both higher precision and power intensity which proved to be very crucial in having successful LFI for ARM Cortex-M0 LPC1114FN28102.

- Using surface mount DUAs instead of through-hole as most MCU vendors offer up-to-date MCU in surface mount devices. Designing a PCB with a hole on the backside of the soldered DUA for different ARM Cortex-based MCUs should have very small variations since the ARM Cortex family have very similar interfaces.

- The increasing thickness's of modern chips [125] makes the PEM analysis almost impractical without any substrate thinning. Therefore innovating a low cost technique for thinning the substrate of up-to-date DUAs would be specially beneficial to the PEM research and embedded security field in general.

References

[1] "Laguerre-Gaussian transverse mode patterns." [Online]. Available: https://commons.wikimedia.org/wiki/File:Laguerre-gaussian.png (Accessed on 6/7/2017).

[2] "Phys 4510, chapter 5, course notes," 6/7/2017 2017. [Online]. Available: http://www.colorado.edu/physics/phys4510/phys4510_fa05/Chapter5.pdf

[3] H. Sun, *Laser Diode Beam Propagation Basics*, ser. Laser Diode Beam Basics, Manipulations and Characterizations. Dordrecht: Springer-Verlag, 2012, pp. 21–37. [Online]. Available: http://dx.doi.org/10.1007/978-94-007-4664-0_2

[4] S. d. Castro, G. D. Natale, M. L. Flottes, B. Rouzeyre, and J. M. Dutertre, "Figure of merits of 28nm si technologies for implementing laser attack resistant security dedicated circuits," in *2015 IEEE Computer Society Annual Symposium on VLSI*, 2015, pp. 362–367.

[5] J. S. Melinger, S. Buchner, D. McMorrow, W. J. Stapor, T. R. Weatherford, A. B. Campbell, and H. Eisen, "Critical evaluation of the pulsed laser method for single event effects testing and fundamental studies," *IEEE Transactions on Nuclear Science*, vol. 41, no. 6, pp. 2574–2584, 1994.

[6] M. Faraj and C. Gebotys, "Quiescent photonics side channel analysis: Low cost sram readout attack," in *Kangacrypt, Australian Workshop on Offensive Cryptography*, December 2018.

[7] S. Suzuki and M. Yamaguchi, "Acid decapsulation for silver wire bonded package," in *2015 IEEE 22nd International Symposium on the Physical and Failure Analysis of Integrated Circuits*, 2015, pp. 492–495.

[8] "Rohde & Schwarz, H field probe RS H 2.5-2." [Online]. Available: https://scdn.rohde-schwarz.com/ur/pws/dl_downloads/

dl_common_library/dl_brochures_and_datasheets/pdf_1/service_support_30/ HZ-15_16_17_bro_en_5213-6687-12_v0100.pdf (Accessed on 2017-07-15).

[9] "PICmicro Mid-Range MCU Family Reference Manual - Microchip." [Online]. Available: http://ww1.microchip.com/downloads/en/devicedoc/33023a.pdf (Accessed on 2017-07-15).

[10] J. Yiu, "Chapter 1 - introduction," in *The Definitive Guide to the ARM Cortex-M0*, J. Yiu, Ed. Oxford: Newnes, 2011, pp. 1 – 12. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780123854773100011

[11] J. Yiu, "Chapter 13 - debug features," in *The Definitive Guide to the ARM Cortex-M0*, J. Yiu, Ed. Oxford: Newnes, 2011, pp. 221 – 229. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780123854773100138

[12] AlphaNov, "AlphaNov PDM+ Testing report," April 1, 2018.

[13] F. Beck, *Integrated Circuit Failure Analysis : A Guide to Preparation Techniques.* Wiley, 1998.

[14] S. D. Castro, J.-M. Dutertre, B. Rouzeyre, G. D. Natale, and M.-L. Flottes, "Frontside versus backside laser injection: A comparative study," *J.Emerg.Technol.Comput.Syst.*, vol. 13, no. 1, pp. 7:1–7:15, nov 2016. [Online]. Available: http://doi.acm.org/10.1145/2845999

[15] I. Studnia, V. Nicomette, E. Alata, Y. Deswarte, M. Kaâniche, and Y. Laarouchi, "Survey on security threats and protection mechanisms in embedded automotive networks," in *2013 43rd Annual IEEE/IFIP Conference on Dependable Systems and Networks Workshop (DSN-W)*, 2013, pp. 1–12.

[16] J. Fan, X. Guo, E. D. Mulder, P. Schaumont, B. Preneel, and I. Verbauwhede, "State-of-the-art of secure ECC implementations: a survey on known side-channel attacks and countermeasures," in *2010 IEEE International Symposium on Hardware-Oriented Security and Trust (HOST)*, 2010, pp. 76–87.

[17] C. H. Kim and J. J. Quisquater, "Faults, Injection Methods, and Fault Attacks," *Design & Test of Computers, IEEE*, vol. 24, no. 6, pp. 544–545, 2007.

[18] A. Barenghi, L. Breveglieri, I. Koren, and D. Naccache, "Fault injection attacks on cryptographic devices: Theory, practice, and countermeasures," *Proceedings of the IEEE*, vol. 100, no. 11, pp. 3056–3076, Nov 2012.

[19] E. Eilam, *Wiley: Reversing: Secrets of Reverse Engineering.* Wiley, 2005.

[20] R. Torrance and D. James, "The state-of-the-art in semiconductor reverse engineering," in *Design Automation Conference (DAC), 2011 48th ACM/EDAC/IEEE*, 2011, pp. 333–338.

[21] TechInsights - Revealing the innovation others can't inside advanced technology. [Online]. Available: http://techinsights.com/ (Accessed on 2017-07-09).

[22] "Secure & Seamless Business Payments — Bottomline Technologies." [Online]. Available: https://www.bottomline.com/us (Accessed on 2017-07-09).

[23] Y. Kim, R. Daly, J. Kim, C. Fallin, J. H. Lee, D. Lee, C. Wilkerson, K. Lai, and O. Mutlu, "Flipping bits in memory without accessing them: An experimental study of DRAM disturbance errors," in *2014 ACM/IEEE 41st International Symposium on Computer Architecture (ISCA)*, 2014, pp. 361–372.

[24] ""Row hammering" – how to exploit a computer by overworking its memory – Naked Security." [Online]. Available: https://nakedsecurity.sophos.com/2015/03/12/row-hammering-how-to-exploit-a-computer-by-overworking-its-memory/ (Accessed on 2017-07-11).

[25] "Discussions with Prof. Manoj Sachdev at the University of Waterloo," September 22, 2017.

[26] Fault, Error, and Failure (based on slides by Prof. Lin Tan and others). [Online]. Available: https://ece.uwaterloo.ca/~agurfink/ece653/assets/pdf/W01P2-FaultErrorFailure.pdf (Accessed on 2020-05-13).

[27] S. P. Skorobogatov, "Semi-invasive attacks-a new approach to hardware security analysis," *Technical report, University of Cambridge, Computer Laboratory*, no. 630, p. 144, 2005. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.228.2204&rep=rep1&type=pdf

[28] O. Kömmerling and M. G. Kuhn, "Design Principles for Tamper-Resistant Smartcard Processors," *Proceedings of the USENIX Workshop on Smartcard Technology (Smartcard '99)*, pp. 9–20, 1999.

[29] S. P. Skorobogatov and R. J. Anderson, "Optical fault induction attacks," *International Workshop on Cryptographic Hardware and Embedded Systems (CHES 2002)*, vol. 2523, pp. 2–12, 2003.

[30] C. Giraud, "DFA on AES," in *Proceedings of the 4th International Conference on Advanced Encryption Standard*, ser. AES'04. Bonn, Germany: Springer-Verlag, 2005, pp. 27–41. [Online]. Available: http://dx.doi.org/10.1007/11506447_4

[31] H. Bar-El, H. Choukri, D. Naccache, M. Tunstall, and C. Whelan, "The sorcerer's apprentice guide to fault attacks," *Proceedings of the IEEE*, vol. 94, no. 2, pp. 370–382, 2006.

[32] C. H. Gebotys, "EM, Lasers and Methodologies," June 29, 2016, unpublished draft report.

[33] A. E. Siegman, *Lasers.* University Science Books, 1986.

[34] P. Amirtharaj and D. Seiler, *CHAPTER 36 OPTICAL PROPERTIES OF SEMICONDUCTORS*, 2nd ed., ser. Handbook of Optics, Vol II. North Holland, 1995, pp. 36.1–36.96.

[35] Fiber Optic Beam Delivery Systems, U.S. Laser Corporation. [Online]. Available: http://www.uslasercorp.com/envoy/fobd.html (Accessed on 2016-04-11).

[36] X. C. de Lega and P. J. de Groot, "Lateral resolution and instrument transfer function as criteria for selecting surface metrology instruments," in *Imaging and Applied Optics Technical Papers.* Optical Society of America, 2012, p. OTu1D.4. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=OFT-2012-OTu1D.4

[37] F. Träger, *Springer Handbook of Lasers and Optics.* Springer, 2007. [Online]. Available: https://link.springer.com/referencework/10.1007/978-0-387-30420-5

[38] K. Kitamura, K. Sakai, and S. Noda, "Sub-wavelength focal spot with long depth of focus generated by radially polarized, narrow-width annular beam," *Opt.Express*, vol. 18, no. 5, pp. 4518–4525, Mar 2010. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-18-5-4518

[39] C. E. Webb and J. D. C. Jones, *Handbook of laser technology and applications.* CRC Press, 2004.

[40] K. Rottwitt and P. Tidemand-Lichtenberg, *Nonlinear Optics: Principles and Applications.* CRC Press, 2014, vol. 3.

[41] M. S. Brown and C. B. Arnold, *Fundamentals of Laser-Material Interaction and Application to Multiscale Surface Modification*, ser. Laser Precision Microfabrication.

Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 91–120. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-10523-4_4

[42] A. R. Shayan, H. B. Poyraz, and J. Patten, "Laser absorption percent for si and sic," 6/7/2017 2008. [Online]. Available: http://wmich.edu/mfe/mrc/pdf/Amir%20and%20Bogac_Laser%20Absorption%20for%20Si%20and%20SiC_10-24-2008.pdf

[43] A. Y. Nikiforov and P. K. Skorobogatov, "Physical principles of laser simulation for the transient radiation response of semiconductor structures, active circuit elements, and circuits: A linear model," *Russian Microelectronics*, vol. 33, no. 2, pp. 68–79, 2004. [Online]. Available: http://dx.doi.org/10.1023/B:RUMI.0000018711.96346.1d

[44] G. Canivet, P. Maistri, R. Leveugle, J. Clédière, F. Valette, and M. Renaudin, "Glitch and Laser Fault Attacks onto a Secure AES Implementation on a SRAM-Based FPGA," *J.Cryptology*, vol. 24, pp. 247–268, 2011.

[45] C. Helfmeier, D. Nedospasov, C. Tarnovsky, J. S. Krissler, C. Boit, and J.-P. Seifert, "Breaking and Entering Through the Silicon," in *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, ser. CCS '13. New York, NY, USA: ACM, 2013, pp. 733–744. [Online]. Available: http://doi.acm.org/10.1145/2508859.2516717

[46] K. Sanchez, R. Desplats, F. Beaudoin, P. Perdu, J. P. Roux, G. Woods, and D. Lewis, "NIR laser stimulation for dynamic timing analysis," *Microelectronics Reliability*, vol. 45, no. 9, pp. 1459–1464, November 2005.

[47] P. K. Skorobogatov, A. Y. Nikiforov, O. B. Mavritsky, A. N. Egorov, and A. V. Kirgizova, "Influence of temperature on pulsed laser see testing," in *Proceedings of the 7th European Conference on Radiation and Its Effects on Components and Systems, 2003. RADECS 2003.*, 2003, pp. 153–155.

[48] C. Roscian, A. Sarafianos, J. M. Dutertre, and A. Tria, "Fault model analysis of laser-induced faults in sram memory cells," in *2013 Workshop on Fault Diagnosis and Tolerance in Cryptography*, 2013, pp. 89–98.

[49] A. H. Johnston, "Charge generation and collection in p-n junctions excited with pulsed infrared lasers," *IEEE Transactions on Nuclear Science*, vol. 40, no. 6, pp. 1694–1702, 1993.

[50] T. Wojcicki, *VLSI: Circuits for emerging applications.* CRC Press, 2015.

[51] F. B. A. Früngel, *Optical Pulses - Lasers - Measuring Techniques.* Elsevier Science, 2014. [Online]. Available: https://books.google.ca/books?id=pkaoBQAAQBAJ

[52] S. Skorobogatov, "Local heating attacks on flash memory devices," in *2009 IEEE International Workshop on Hardware-Oriented Security and Trust*, 2009, pp. 1–6.

[53] F. Darracq, T. Beauchene, V. Pouget, H. Lapuyade, D. Lewis, P. Fouillat, and A. Touboul, "Single-event sensitivity of a single sram cell," *IEEE Transactions on Nuclear Science*, vol. 49, no. 3, pp. 1486–1490, 2002.

[54] "Single-mode laser source for full temporal agility PDM+: Pulse-on-Demand Modules , Laser solutions for IC evaluation, Alphanov." [Online]. Available: http://www.alphanov.com/8-optoelectronics-systems-pulse-on-demand-modules.html (Accessed on 2017-07-15).

[55] "PHYS2212, Course notes, York University." [Online]. Available: http://www.yorku.ca/marko/PHYS2212/Lab8.pdf (Accessed on 2017-07-15).

[56] "Mitutoyo, MICROSOPE UNITS AND OBJECTIVES." [Online]. Available: http://www.mitutoyo.com/wp-content/uploads/2012/11/E4191-378_010611.pdf (Accessed on 6/7/2017).

[57] S. Briais, S. Caron, J.-M. Cioranesco, J.-L. Danger, S. Guilley, J.-H. Jourdan, A. Milchior, D. Naccache, and T. Porteboeuf, "3D hardware canaries," in *Cryptographic Hardware and Embedded Systems–CHES*, A. Karim, Ed. Springer, 2012, pp. 1–22.

[58] M. A. Green, "Self-consistent optical parameters of intrinsic silicon at 300 k including temperature coefficients," *Solar Energy Materials and Solar Cells*, vol. 92, no. 11, pp. 1305–1310, 11 2008.

[59] "The Laboratory for Physical science, Microelectronic Integration Group." [Online]. Available: http://www.lps.umd.edu/MicroelectronicsIntegration/MicroelectronicsCOTS.html (Accessed on 6/7/2017).

[60] J. Breier and C. N. Chen, "On determining optimal parameters for testing devices against laser fault attacks," in *2016 International Symposium on Integrated Circuits (ISIC)*, 2016, pp. 1–4.

[61] D. Nedospasov, J. P. Seifert, A. Schlösser, and S. Orlic, "Functional integrated circuit analysis," in *2012 IEEE International Symposium on Hardware-Oriented Security and Trust*, 2012, pp. 102–107.

[62] A. Schlösser, D. Nedospasov, J. Krämer, S. Orlic, and J.-P. Seifert, *Simple Photonic Emission Analysis of AES*, ser. Cryptographic Hardware and Embedded Systems – CHES 2012: 14th International Workshop, Leuven, Belgium, September 9-12, 2012. Proceedings. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 41–57. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-33027-8_3

[63] A. B. Yankovich, B. Berkels, W. Dahmen, P. Binev, S. I. Sanchez, S. A. Bradley, A. Li, I. Szlufarska, and P. M. Voyles, "Picometre-precision analysis of scanning transmission electron microscopy images of platinum nanocatalysts," *Nature Communications*, vol. 5, p. 4155, Jun. 2014.

[64] "ICWorks Surveyor." [Online]. Available: https://www.chipworks.com/images/content-documents/tablesofcontents-samples/ICWorks-English_datasheet_Surveyor.pdf (Accessed on 7/2/2017).

[65] I. McLoughlin, "Secure embedded systems: The threat of reverse engineering," in *14th IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, 2008, pp. 729–736.

[66] Y. Ishai, A. Sahai, and D. Wagner, *Private circuits: Securing hardware against probing attacks*, ser. Advances in Cryptology-CRYPTO 2003. Springer, 2003, pp. 463–481.

[67] Y. Ishai, M. Prabhakaran, A. Sahai, and D. Wagner, "Private circuits ii: Keeping secrets in tamperable circuits," in *Advances in Cryptology - EUROCRYPT 2006, 25th Annual International Conference on the Theory and Applications of Cryptographic Techniques*, ser. Lecture Notes in Computer Science, vol. 4004. Springer, 2006, pp. 308–327. [Online]. Available: http://www.iacr.org/cryptodb/archive/2006/EUROCRYPT/2508/2508.pdf

[68] "ASAP-1 — ULTRA TEC ." [Online]. Available: http://www.ultratecusa.com/asap-1 (Accessed on 7/2/2017).

[69] M. Weiner, S. Manich, R. Rodríguez-Montañés, and G. Sigl, "The low area probing detector as a countermeasure against invasive attacks," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 26, no. 2, pp. 392–403, 2018.

[70] S. Manich, M. S. Wamser, and G. Sigl, "Detection of probing attempts in secure ics," in *2012 IEEE International Symposium on Hardware-Oriented Security and Trust*, 2012, pp. 134–139.

[71] M. Weiner, W. Wieser, E. Lupon, G. Sigl, and S. Manich, “A calibratable detector for invasive attacks,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 27, no. 5, pp. 1067–1079, 2019.

[72] P. Tuyls, G.-J. Schrijen, B. . kori?‡, J. van Geloven, N. Verhaegh, and R. Wolters, *Read-Proof Hardware from Protective Coatings*, ser. Cryptographic Hardware and Embedded Systems - CHES 2006: 8th International Workshop, Yokohama, Japan, October 10-13, 2006. Proceedings. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 369–383. [Online]. Available: http://dx.doi.org/10.1007/11894063_29

[73] C. Boit, S. Tajik, P. Scholz, E. Amini, A. Beyreuther, H. Lohrke, and J. P. Seifert, “From ic debug to hardware security risk: The power of backside access and optical interaction,” in *2016 IEEE 23rd International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, 2016, pp. 365–369.

[74] H. Lohrke, S. Tajik, C. Boit, and J. P. Seifert, *No Place to Hide: Contactless Probing of Secret Data on FPGAs*, ser. Cryptographic Hardware and Embedded Systems – CHES 2016: 18th International Conference, Santa Barbara, CA, USA, August 17-19, 2016, Proceedings. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 147–167. [Online]. Available: http://dx.doi.org/10.1007/978-3-662-53140-2_8

[75] H. Shen, N. Asadizanjani, M. Tehranipoor, and D. Forte, “Nanopyramid: An optical scrambler against backside probing attacks,” in *ISTFA 2018:Proceedings from the 44th International Symposium for Testing and Failure Analysis. ASM International*, 2018, p. 280.

[76] D. Oswald, B. Richter, and C. Paar, *Side-Channel Attacks on the Yubikey 2 One-Time Password Generator*, ser. Research in Attacks, Intrusions, and Defenses: 16th International Symposium, RAID 2013, Rodney Bay, St. Lucia, October 23-25, 2013. Proceedings. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 204–222. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-41284-4_11

[77] J. Blömer and J. P. Seifert, *Fault Based Cryptanalysis of the Advanced Encryption Standard (AES).* Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 162–181. [Online]. Available: https://doi.org/10.1007/978-3-540-45126-6_12

[78] C. N. Chen and S. M. Yen, *Differential Fault Analysis on AES Key Schedule and Some Countermeasures.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 118–129. [Online]. Available: https://doi.org/10.1007/3-540-45067-X_11

[79] G. Piret and J. J. Quisquater, *A Differential Fault Attack Technique against SPN Structures, with Application to the AES and Khazad*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 77–88. [Online]. Available: https://doi.org/10.1007/978-3-540-45238-6_7

[80] H. Choukri and M. Tunstall, “Round reduction using faults,” in *In Proceedings of the 2nd International Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC’05)*, 2007, pp. 13–24.

[81] J. H. Park, S. J. Moon, D. H. Choi, Y. S. Kang, and J. C. Ha, “Differential fault analysis for round-reduced aes by fault injection,” in *ETRI Journal*, vol. 33, 06 2011, p. 434–442.

[82] P. Dusart, G. Letourneux, and O. Vivolo, *Differential Fault Analysis on A.E.S.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 293–306. [Online]. Available: https://doi.org/10.1007/978-3-540-45203-4_23

[83] B. Robisson and P. Manet, “Differential Behavioral Analysis,” in *Workshop on Cryptographic Hardware and Embedded Systems*, I. V. Pascal Paillier, Ed. Vienne, Austria: Springer Berlin Heidelberg, Sep. 2007, pp. 413–426. [Online]. Available: https://hal-emse.ccsd.cnrs.fr/emse-00481468

[84] C. H. Kim and J. J. Quisquater, *New Differential Fault Analysis on AES Key Schedule: Two Faults Are Enough.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 48–60. [Online]. Available: https://doi.org/10.1007/978-3-540-85893-5_4

[85] S. S. Ali and D. Mukhopadhyay, *Differential Fault Analysis of AES-128 Key Schedule Using a Single Multi-byte Fault.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 50–64. [Online]. Available: https://doi.org/10.1007/978-3-642-27257-8_4

[86] S. S. Ali and D. Mukhopadhyay, “A differential fault analysis on aes key schedule using single fault,” in *2011 Workshop on Fault Diagnosis and Tolerance in Cryptography*, Sept 2011, pp. 35–42.

[87] D. Strobel, D. Oswald, B. Richter, F. Schellenberg, and C. Paar, “Microcontrollers as (in)security devices for pervasive computing applications,” *Proceedings of the IEEE*, vol. 102, no. 8, pp. 1157–1173, 2014.

[88] J. C. H. Phang, D. S. H. Chan, M. Palaniappan, J. M. Chin, B. Davis, M. Bruce, J. Wilcox, G. Gilfeather, C. M. Chua, L. S. Koh, H. Y. Ng, and S. H. Tan, “A review of laser induced techniques for microelectronic failure analysis,” in *Proceedings of*

the 11th International Symposium on the Physical and Failure Analysis of Integrated Circuits. IPFA 2004 (IEEE Cat. No.04TH8743), 2004, pp. 255–261.

[89] E. I. Cole, J. M. Soden, J. L. Rife, D. L. Barton, and C. L. Henderson, "Novel failure analysis techniques using photon probing with a scanning optical microscope," in *Proceedings of 1994 IEEE International Reliability Physics Symposium*, 1994, pp. 388–398.

[90] F. Schellenberg, M. Finkeldey, N. Gerhardt, M. Hofmann, A. Moradi, and C. Paar, "Large laser spots and fault sensitivity analysis," in *2016 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*, 2016, pp. 203–208.

[91] F. Schellenberg, M. Finkeldey, B. Richter, M. Sch?pers, N. Gerhardt, M. Hofmann, and C. Paar, "On the complexity reduction of laser fault injection campaigns using obic measurements," in *2015 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC)*, 2015, pp. 14–27.

[92] J. C. White, T. F. Unter, and J. G. Smith, "Contactless nondestructive technique for the measurement of minority-carrier lifetime and diffusion length in silicon," *Solid-State and Electron Devices, IEE Journal on*, vol. 1, no. 5, pp. 139–145, 1977.

[93] N. Moro, K. Heydemann, E. Encrenaz, and B. Robisson, "Formal verification of a software countermeasure against instruction skip attacks," *CoRR*, vol. abs/1402.6461, 2014. [Online]. Available: http://arxiv.org/abs/1402.6461

[94] R. Velazco and B. Martinet, "Physical fault injection: a suitable method for the evaluation of functional test efficiency," in *Proceedings of International Workshop on Defect and Fault Tolerance on VLSI Systems*, 1991, pp. 179–182.

[95] W. He, J. Breier, S. Bhasin, N. Miura, and M. Nagata, "An fpga-compatible pll-based sensor against fault injection attack," in *2017 22nd Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2017, pp. 39–40.

[96] W. He, J. Breier, S. Bhasin, N. Miura, and M. Nagata, "Ring oscillator under laser: Potential of pll-based countermeasure against laser fault injection," in *2016 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC)*, 2016, pp. 102–113.

[97] M. Agoyan, J. M. Dutertre, A. P. Mirbaha, D. Naccache, A. L. Ribotta, and A. Tria, "How to flip a bit?" in *2010 IEEE 16th International On-Line Testing Symposium*, 2010, pp. 235–239.

[98] M. Agoyan, J. M. Dutertre, A. P. Mirbaha, D. Naccache, A. L. Ribotta, and A. Tria, "Single-bit dfa using multiple-byte laser fault injection," in *IEEE International Conference on Technologies for Homeland Security (HST)*, 2010, pp. 113–119.

[99] "Prior Scientific." [Online]. Available: http://www.prior.com/ (Accessed on 2017-07-13).

[100] C. Roscian, J. M. Dutertre, and A. Tria, "Frontside laser fault injection on cryptosystems - application to the aes' last round -," in *IEEE International Symposium on Hardware-Oriented Security and Trust (HOST)*, 2013, pp. 119–124.

[101] F. Courbon, P. Loubet-Moundi, J. J. A. Fournier, and A. Tria, *Adjusting Laser Injections for Fully Controlled Faults*, ser. Constructive Side-Channel Analysis and Secure Design: 5th International Workshop, COSADE 2014, Paris, France, April 13-15, 2014. Revised Selected Papers. Cham: Springer International Publishing, 2014, pp. 229–242. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-10175-0_16

[102] B. Selmke, S. Brummer, J. Heyszl, and G. Sigl, *Precise Laser Fault Injections into 90 nm and 45 nm SRAM-cells*, ser. Smart Card Research and Advanced Applications: 14th International Conference, CARDIS 2015, Bochum, Germany, November 4-6, 2015. Revised Selected Papers. Cham: Springer International Publishing, 2016, pp. 193–205. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-31271-2_12

[103] B. Selmke, J. Heyszl, and G. Sigl, "Attack on a dfa protected aes by simultaneous laser fault injections," in *2016 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC)*, 2016, pp. 36–46.

[104] W. A. Moreno, F. J. Falquez, J. R. Samson, and T. Smith, "First test results of system level fault tolerant design validation through laser fault injection," in *Proceedings International Conference on Computer Design VLSI in Computers and Processors*, 1997, pp. 544–548.

[105] J. R. Samson, W. Moreno, and F. Falquez, "Validating fault tolerant designs using laser fault injection (lfi)," in *1997 IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems*, 1997, pp. 175–183.

[106] J. R. Samson, W. Moreno, and F. Falquez, "A technique for automated validation of fault tolerant designs using laser fault injection (lfi)," in *Fault-Tolerant Computing, 1998. Digest of Papers. Twenty-Eighth Annual International Symposium on*, 1998, pp. 162–167.

[107] N. A. Anagnostopoulos, “Optical fault injection attacks in smart card chips and an evaluation of countermeasures against them,” Master’s thesis, University of Twente, 2014.

[108] AlpohaNov, “PDM operating manual,” March 27, 2018, unpublished operating manual.

[109] K. Amin, C. Gebotys, M. Faraj, and H. Liao, “Analysis of dynamic laser injection and quiescent photon emissions on an embedded processor,” in *PAINE, International Conference on Physical Assurance and Inspection of Electronics*, July 2019.

[110] K. Amin, C. Gebotys, M. Faraj, and H. Liao, “Analysis of dynamic laser injection and quiescent photon emissions on an embedded processor,” *Journal of Hardware and Systems Security*, vol. 4, no. 1, pp. 55–67, Mar 2020. [Online]. Available: https://doi.org/10.1007/s41635-020-00090-1

[111] “SiFive FE310-G002 Manual: v19p04.” [Online]. Available: https://sifive.cdn.prismic.io/sifive%2F9ecbb623-7c7f-4acc-966f-9bb10ecdb62e_fe310-g002.pdf (Accessed on 2020-05-15).

[112] J. Yiu, “Chapter 2 - cortex-m0 technical overview,” in *The Definitive Guide to the ARM Cortex-M0*, J. Yiu, Ed. Oxford: Newnes, 2011, pp. 13 – 24. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780123854773100023

[113] “LPC111x/LPC11Cxx User manual - NXP.” [Online]. Available: http://www.keil.com/dd/docs/datashts/nxp/lpc11xx/lpc11xx_um.pdf (Accessed on 2020-05-15).

[114] “RISC-V Specifications.” [Online]. Available: https://riscv.org/specifications/ (Accessed on 15/5/2020).

[115] “PIC16F687 datasheet.” [Online]. Available: http://ww1.microchip.com/downloads/en/DeviceDoc/40001262F.pdf (Accessed on 2017-07-26).

[116] T. Stratoudaki, C. Edwards, S. Dixon, S. B. Palmer, D. O. Thompson, D. E. Chimenti, C. Nessa, S. Kallsen, and L. Poore, “Optical absorption of epoxy resin and its role in the laser ultrasonic generation mechanism in composite materials,” *AIP Conference Proceedings*, vol. 657, no. 1, pp. 965–972, 03/27; 2017/06 2003. [Online]. Available: http://aip.scitation.org/doi/abs/10.1063/1.1570238

[117] S. Murali and N. Srikanth, “Acid decapsulation of epoxy molded ic packages with copper wire bonds,” *IEEE Transactions on Electronics Packaging Manufacturing*, vol. 29, no. 3, pp. 179–183, 2006.

[118] B. Krasnow, “Decapping ICs (removing epoxy packaging from chips to expose the dies).” [Online]. Available: https://www.youtube.com/watch?v=mT1FStxAVz4&t=2s (Accessed on 6/7/2017).

[119] “Image Composite Editor.” [Online]. Available: https://www.microsoft.com/en-us/research/product/computational-photography-applications/image-composite-editor/ (Accessed on 6/11/2017).

[120] “Carbide 3D - Nomad 883 Pro.” [Online]. Available: http://carbide3d.com/nomad/ (Accessed on 6/14/2017).

[121] J. Breier, D. Jap, and C.-N. Chen, “Laser profiling for the back-side fault attacks: With a practical laser skip instruction attack on aes,” in *Proceedings of the 1st ACM Workshop on Cyber-Physical System Security*, ser. CPSS ’15. New York, NY, USA: ACM, 2015, pp. 99–103. [Online]. Available: http://doi.acm.org/10.1145/2732198.2732206

[122] S. Tajik, H. Lohrke, F. Ganji, J. Seifert, and C. Boit, “Laser fault attack on physically unclonable functions,” in *2015 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC)*, Sep. 2015, pp. 85–96.

[123] N. Moro, K. Heydemann, E. Encrenaz, and B. Robisson, “Formal verification of a software countermeasure against instruction skip attacks,” *Journal of Cryptographic Engineering*, vol. 4, no. 3, pp. 145–156, Sep 2014. [Online]. Available: https://doi.org/10.1007/s13389-014-0077-7

[124] S. V. D. Kumar, S. Patranabis, J. Breier, D. Mukhopadhyay, S. Bhasin, A. Chattopadhyay, and A. Baksi, “A practical fault attack on arx-like ciphers with a case study on chacha20,” in *2017 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC)*, Sep. 2017, pp. 33–40.

[125] M. Goldstein and M. Watanabe, “450 mm silicon wafers challenges - wafer thickness scaling,” in *ECS Transactions*. ECS, 2008. [Online]. Available: https://doi.org/10.1149%2F1.2980288

[126] K. Matsuda, S. Tada, M. Nagata, Y. Komano, Y. Li, T. Sugawara, M. Iwamoto, K. Ohta, K. Sakiyama, and N. Miura, “An IC-level countermeasure against laser fault injection attack by information leakage sensing based on laser-induced opto-electric bulk current density,” *Japanese Journal of Applied Physics*, dec 2019. [Online]. Available: https://doi.org/10.7567%2F1347-4065%2Fab65d3

[127] J. Dutertre, V. Beroulle, P. Candelier, S. De Castro, L. Faber, M. Flottes, P. Gendrier, D. Hély, R. Leveugle, P. Maistri, G. Di Natale, A. Papadimitriou, and B. Rouzeyre, "Laser fault injection at the cmos 28 nm technology node: an analysis of the fault model," in *2018 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC)*, Sep. 2018, pp. 1–6.

[128] F. Courbon, J. J. A. Fournier, P. Loubet-Moundi, and A. Tria, "Combining image processing and laser fault injections for characterizing a hardware aes," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, no. 6, pp. 928–936, June 2015.

[129] T. Korak, "Investigation of parameters influencing the success of optical fault attacks," in *Foundations and Practice of Security - 6th International Symposium, FPS 2013, La Rochelle, France, October 21-22, 2013, Revised Selected Papers*, 2013, pp. 140–157. [Online]. Available: https://doi.org/10.1007/978-3-319-05302-8_9

[130] A. Vasselle, H. Thiebeauld, Q. Maouhoub, A. Morisset, and S. Ermeneux, "Laser-induced fault injection on smartphone bypassing the secure boot," in *2017 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC)*, Sep. 2017, pp. 41–48.

[131] B. Selmke, K. Zinnecker, P. Koppermann, K. Miller, J. Heyszl, and G. Sigl, "Locked out by latch-up? an empirical study on laser fault injection into arm cortex-m processors," in *2018 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC)*, 2018, pp. 7–14.

[132] B. Colombier, A. Menu, J.-M. DUTERTRE, P.-A. Moëllic, J.-B. Rigaud, and J.-L. Danger, "Laser-induced Single-bit Faults in Flash Memory: Instructions Corruption on a 32-bit Microcontroller," in *2019 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*. McLean, United States: IEEE, May 2019, pp. 1–10. [Online]. Available: https://hal.telecom-paris.fr/hal-02344050

[133] T. Krachenfels, H. Lohrke, J.-P. Seifert, E. Dietz, S. Frohmann, and H.-W. Hübers, "Evaluation of low-cost thermal laser stimulation for data extraction and key readout," *Journal of Hardware and Systems Security*, vol. 4, no. 1, pp. 24–33, Mar 2020. [Online]. Available: https://doi.org/10.1007/s41635-019-00083-9

[134] "Allied High Tech Products Inc., MetPrep 4 Grinder/Polisher with Power Head." [Online]. Available: http://www.alliedhightech.com/equipment/grinding-polishing (Accessed on 15/5/2020).

[135] “Allied High Tech Products Inc., X-Prep Precision Milling/Polishing System.” [Online]. Available: http://www.alliedhightech.com/equipment/mechanical-milling (Accessed on 15/5/2020).

[136] “Nisene JetEtch Pro.” [Online]. Available: https://www.nisene.com/products/jetetch-pro/ (Accessed on 15/5/2020).

[137] R. A. C. Viera, J. Dutertre, R. P. Bastos, and P. Maurine, “Role of laser-induced ir drops in the occurrence of faults: Assessment and simulation,” in *2017 Euromicro Conference on Digital System Design (DSD)*, Aug 2017, pp. 252–259.

[138] J. P. Walters, K. M. Zick, and M. French, “A practical characterization of a nasa spacecube application through fault emulation and laser testing,” in *2013 43rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, June 2013, pp. 1–8.

[139] F. El-Mamouni, E. X. Zhang, R. D. Schrimpf, R. A. Reed, K. F. Galloway, D. McMorrow, E. Simoen, C. Claeys, S. Cristoloveanu, and W. Xiong, “Pulsed laser-induced transient currents in bulk and silicon-on-insulator finfets,” in *2011 International Reliability Physics Symposium*, 2011, pp. SE.4.1–SE.4.4.

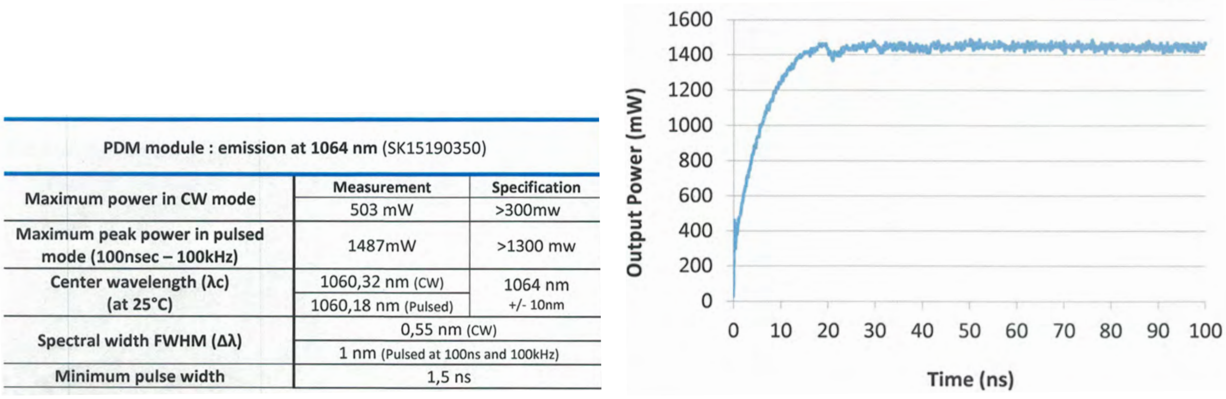
[140] H. . Lee, S. Rami, S. Ravikumar, V. Neeli, K. Phoa, B. Sell, and Y. Zhang, “Intel 22nm finfet (22ffl) process technology for rf and mm wave applications and circuit design optimization for finfet technology,” in *2018 IEEE International Electron Devices Meeting (IEDM)*, 2018, pp. 14.1.1–14.1.4.

APPENDICES

Appendix A

AlphaNov PDM+ measurements

This Appendix includes the main features and test measurements for the AlphaNov PDM+ laser module used for LFI throughout this thesis. The total peak power that can be outputted by both laser channels is around 3.1W with channel 1 providing around 1.5W and channel 2 providing 1.6W. To get the maximum combined power output, both channels' trigger signals have to be synchronized. Figures A.1, A.2, A.3 and A.4 depict different laser parameters and measurements for both laser channels as provided by AlphaNov [12].

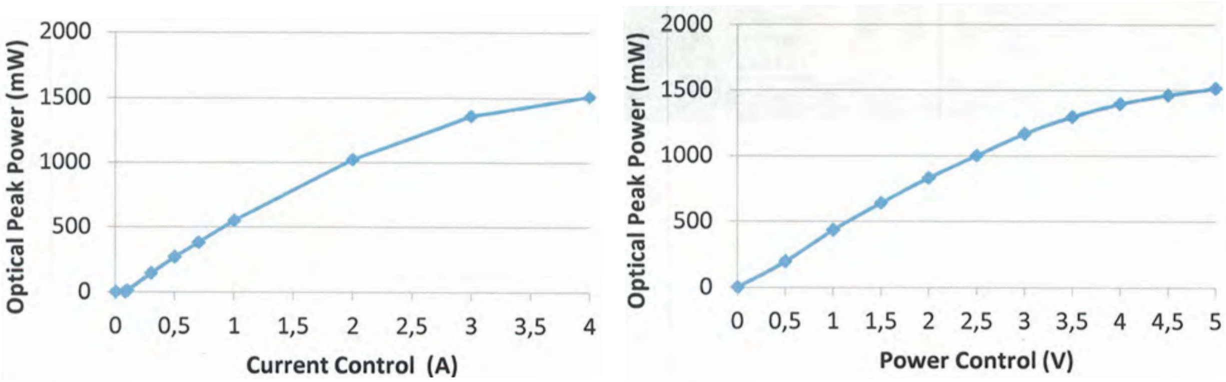


PDM module : emission at 1064 nm (SK15190350)

	Measurement	Specification
Maximum power in CW mode	503 mW	>300mw
Maximum peak power in pulsed mode (100nsec – 100kHz)	1487mW	>1300 mw
Center wavelength (λc) (at 25°C)	1060,32 nm (CW) 1060,18 nm (Pulsed)	1064 nm +/- 10nm
Spectral width FWHM (Δλ)	0,55 nm (CW) 1 nm (Pulsed at 100ns and 100kHz)	
Minimum pulse width	1,5 ns	

(a) (b)

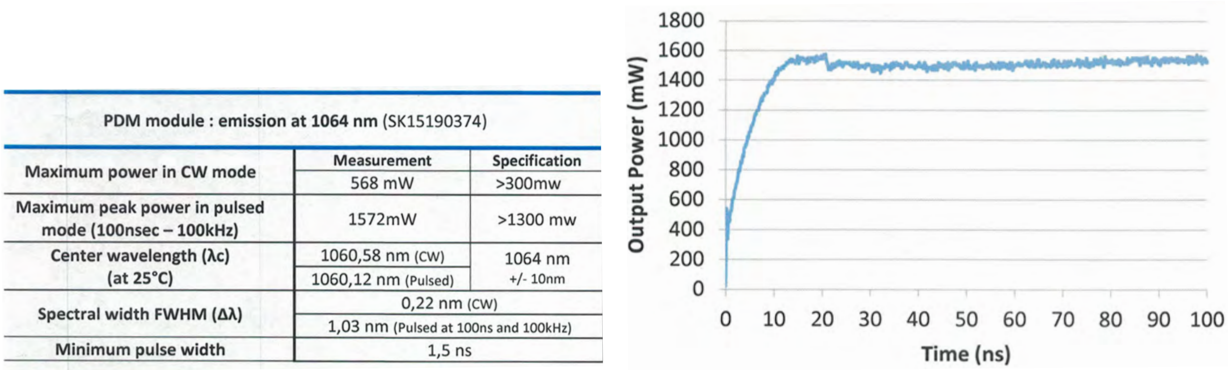
Figure A.1: Channel one (CH1) a) parameters and b) peak power [12].



(a)

(b)

Figure A.2: Channel one (CH1) peak power (100kHz/100ns) versus a) software current and b) analog voltage (BNC input) [12].

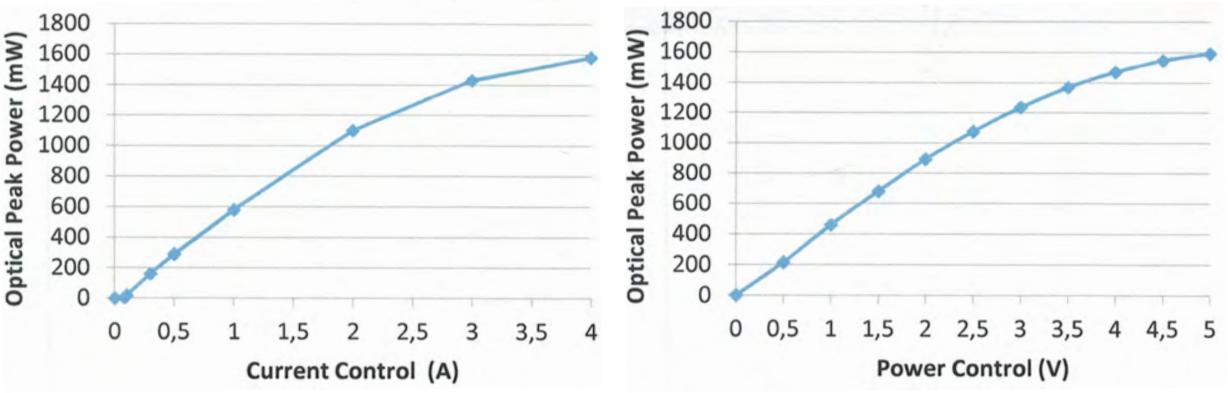


PDM module : emission at 1064 nm (SK15190374)		
	Measurement	Specification
Maximum power in CW mode	568 mW	>300mw
Maximum peak power in pulsed mode (100nsec – 100kHz)	1572mW	>1300 mw
Center wavelength (λc) (at 25°C)	1060,58 nm (CW) 1060,12 nm (Pulsed)	1064 nm +/- 10nm
Spectral width FWHM (Δλ)	0,22 nm (CW) 1,03 nm (Pulsed at 100ns and 100kHz)	
Minimum pulse width	1,5 ns	

(a)

(b)

Figure A.3: Channel two (CH2) a) parameters and b) peak power [12].



(a)

(b)

Figure A.4: Channel two (CH2) peak power (100kHz/100ns) versus a) software current and b) analog voltage (BNC input) [12].

Appendix B

Code snippets

This Appendix include selected parts of the source code used in the LFI experiments conducted on the ARM Cortex-M0 LPC1114FN28102 in chapter 6. In all code listings below, the text after the “;” symbol is a comment and the “...” denote that the instruction is repeated a number of times as noted in the comment.

Listing B.1: Different target instructions for ARM Cortex-M0 LPC1114FN28102

```
 ; Targeting 1st LDR instruction
 ; Laser trigger
 ; Trigger PDM laser CH1 and CH2 using PIO_7 and PIO_9
LDR     R2, =GPIO0DATA
LDR     R3, [R2]
LDR     R4, =((1<<PIO_7):OR:(1<<PIO_9))
BICS    R3, R3, R4      ;Reset PIO_7 and PIO_9
STR     R3, [R2]        ;
EORS    R3, R3, R4      ;Set PIO_7 and PIO_9
STR     R3, [R2]        ;PIO_7 and PIO_9 are toggled (high)
 ; End of Laser trigger
LDR R0, =SRAM_arbitrary_location4        ;target instruction
LDR R1, =0xBB
STRB    R1,[R0]
NOP ;repeated 20 times to isolate target instructions
 ...
NOP
```

```
 ; Targeting 2nd LDR instruction
LDR R0, =SRAM_arbitrary_location4
 ; Laser trigger
 ; Trigger PDM laser CH1 and CH2 using PIO_7 and PIO_9
LDR      R2, =GPIO0DATA
LDR      R3, [R2]
LDR      R4, =((1<<PIO_7):OR:(1<<PIO_9))
BICS     R3, R3, R4      ;Reset PIO_7 and PIO_9
STR      R3, [R2]        ;
EORS     R3, R3, R4      ;Set PIO_7 and PIO_9
STR      R3, [R2]        ;PIO_7 and PIO_9 are toggled (high)
 ; End of Laser trigger
LDR R1, =0xBB   ;target instruction
STRB     R1,[R0]
NOP ;repeated 20 times to isolate target instructions
 ...
NOP
 ; Targeting MOVS instruction
LDR R0, =SRAM_arbitrary_location4
 ; Laser trigger
 ; Trigger PDM laser CH1 and CH2 using PIO_7 and PIO_9
LDR      R2, =GPIO0DATA
LDR      R3, [R2]
LDR      R4, =((1<<PIO_7):OR:(1<<PIO_9))
BICS     R3, R3, R4      ;Reset PIO_7 and PIO_9
STR      R3, [R2]        ;
EORS     R3, R3, R4      ;Set PIO_7 and PIO_9
STR      R3, [R2]        ;PIO_7 and PIO_9 are toggled (high)
 ; End of Laser trigger
MOVS R1, #0xBB  ;target instruction
STRB     R1,[R0]
NOP ;repeated 20 times to isolate target instructions
 ...
NOP
 ; Targeting LDR in R1 inSTRuction and STM instruction
LDR R0, =SRAM_arbitrary_location7
 ; Laser trigger
 ; Trigger PDM laser CH1 and CH2 using PIO_7 and PIO_9
```

```
LDR     R2, =GPIO0DATA
LDR     R3, [R2]
LDR     R4, =((1<<PIO_7):OR:(1<<PIO_9))
BICS    R3, R3, R4        ;Reset PIO_7 and PIO_9
STR     R3, [R2]          ;
EORS    R3, R3, R4        ;Set PIO_7 and PIO_9
STR     R3, [R2]          ;PIO_7 and PIO_9 are toggled (high)
 ; End of Laser trigger
LDR R1, =DUMMY_DATA_WORD1              ;target instruction
LDR R5, =DUMMY_DATA_WORD2
LDR R6, =DUMMY_DATA_WORD3
LDR R7, =DUMMY_DATA_WORD4
STM    R0!,{R1,R5-R7}
NOP ;repeated 20 times to isolate target instructions
 ...
NOP
 ; Targeting LDM instruction
LDR R0, =SRAM_arbitrary_location7
LDR R1, =DUMMY_DATA_WORD1
LDR R5, =DUMMY_DATA_WORD2
LDR R6, =DUMMY_DATA_WORD3
LDR R7, =DUMMY_DATA_WORD4
STM    R0!,{R1,R5-R7}
LDR R0, =SRAM_arbitrary_location7
 ; Laser trigger
 ; Trigger PDM laser CH1 and CH2 using PIO_7 and PIO_9
LDR     R2, =GPIO0DATA
LDR     R3, [R2]
LDR     R4, =((1<<PIO_7):OR:(1<<PIO_9))
BICS    R3, R3, R4        ;Reset PIO_7 and PIO_9
STR     R3, [R2]          ;
EORS    R3, R3, R4        ;Set PIO_7 and PIO_9
STR     R3, [R2]          ;PIO_7 and PIO_9 are toggled (high)
 ; End of Laser trigger
LDM    R0!,{R1,R5-R7}     ;target instruction
NOP ;repeated 20 times to isolate target instructions
 ...
NOP
```

```
;Targeting STM  instruction
LDR R0, =SRAM_arbitrary_location7
LDR R1, =DUMMY_DATA_WORD1
LDR R5, =DUMMY_DATA_WORD2
LDR R6, =DUMMY_DATA_WORD3
LDR R7, =DUMMY_DATA_WORD4
 ; Laser trigger
 ; Trigger PDM laser CH1 and CH2 using PIO_7 and PIO_9
LDR       R2, =GPIO0DATA
LDR       R3, [R2]
LDR       R4, =((1<<PIO_7):OR:(1<<PIO_9))
BICS      R3, R3, R4        ;Reset PIO_7 and PIO_9
STR       R3, [R2]          ;
EORS      R3, R3, R4        ;Set PIO_7 and PIO_9
STR       R3, [R2]          ;PIO_7 and PIO_9 are toggled (high)
 ; End of Laser trigger
STM     R0!,{R1,R5-R7}      ;target instruction
NOP ;repeated 20 times to isolate target instructions
 ...
NOP
 ; Targeting ADDS instruction
LDR R0, =SRAM_arbitrary_location1
LDRB R1, [R0]
 ; Laser trigger
 ; Trigger PDM laser CH1 and CH2 using PIO_7 and PIO_9
LDR       R2, =GPIO0DATA
LDR       R3, [R2]
LDR       R4, =((1<<PIO_7):OR:(1<<PIO_9))
BICS      R3, R3, R4        ;Reset PIO_7 and PIO_9
STR       R3, [R2]          ;
EORS      R3, R3, R4        ;Set PIO_7 and PIO_9
STR       R3, [R2]          ;PIO_7 and PIO_9 are toggled (high)
 ; End of Laser trigger
ADDS      R1,#0x1      ; target instruction repeated 85 times
 ...
ADDS      R1,#0x1
STRB      R1,[R0]
NOP ;repeated 20 times to isolate target instructions
```

```
 ...
NOP
BL       SRAM_readout     ;UART sends 4K bytes to PC
STOP     BL STOP
```

Listing B.2: Seven consecutive pulses for ARM Cortex-M0 LPC1114FN28102

```
 ; Trigger PDM laser CH1 and CH2 using PIO_7 and PIO_9
LDR      R2, =GPIO0DATA
LDR      R3, [R2]
LDR      R4, =((1<<PIO_7):OR:(1<<PIO_9))
BICS     R3, R3, R4       ;Reset PIO_7 and PIO_9
STR      R3, [R2]         ;
EORS     R3, R3, R4       ;Set PIO_7 and PIO_9
STR      R3, [R2]         ;PIO_7 and PIO_9 are toggled (high)
 ;;;;; Additional 6 laser pulses
EORS     R3, R3, R4       ;Value to toggle
STR      R3,[R2]          ;Toggle (high)
EORS     R3, R3, R4       ;Value to toggle
STR      R3,[R2]          ;Toggle (high)
EORS     R3, R3, R4       ;Value to toggle
STR      R3,[R2]          ;Toggle (high)
EORS     R3, R3, R4       ;Value to toggle
STR      R3,[R2]          ;Toggle (high)
EORS     R3, R3, R4       ;Value to toggle
STR      R3,[R2]          ;Toggle (high)
EORS     R3, R3, R4       ;Value to toggle
STR      R3,[R2]          ;Toggle (high)
ADDS     R1,#0x1      ; target instruction repeated 85 times
 ...
ADDS     R1,#0x1
STRB     R1,[R0]
NOP ;repeated 20 times to isolate target instructions
 ...
NOP
BL       SRAM_readout     ;UART sends 4K bytes to PC
STOP     BL STOP
```

Listing B.3: Stack frame dumping for ARM Cortex-M0 LPC1114FN28102

```
 HardFault_Handler_Loop
   MRS R0, MSP ; stacking was done using MSP
   ; R0 as the start address of the stack frame
   ; Read out STACK FRAME and keep sending to UART
   LDR R1, =U0THR   ; Address of the transmission register
   LDR R2, =U0LSR   ; Register containing THRE
   MOVS  R3, #(1<<U0LSR_THRE) ; R3=0x20 (1<<U0LSR_THRE)
   MOVS  R5, R0
   ADDS  R5, #StackFrame_Size ; 8 registers each is 4 bytes
                              ; R0 = SP+0x00 till SP+0x1F + 0x1
NEXT_STACK_BYTE
   ;Start transmission by putting the desired data in register
   LDRB R4, [R0]    ; Data to be transmitted in U0THR register
   STRB R4, [R1]    ; [U0THR]=R4=[R0]=[SP+0x00]
   ; Wait till transmission is complete to avoid overwrite
   LDR R4, [R2]  ; R4 = [U0LSR]
   TST R4, R3    ; THRE = 0 so U0THR contains valid data.
                 ; THRE = 1 so U0THR is empty.
   BEQ still_transmitting_stack
   ADDS     R0,R0,#0x1
   CMP      R0,R5
   BNE      NEXT_STACK_BYTE
   B        HardFault_Handler_Loop
```

Appendix C

Additional experiments

This Appendix includes several experiments for LFI on ARM Cortex-M0 LPC1114FN28102. The experiments mentioned in this Appendix are discussed and compared to other experiments in chapter 6.

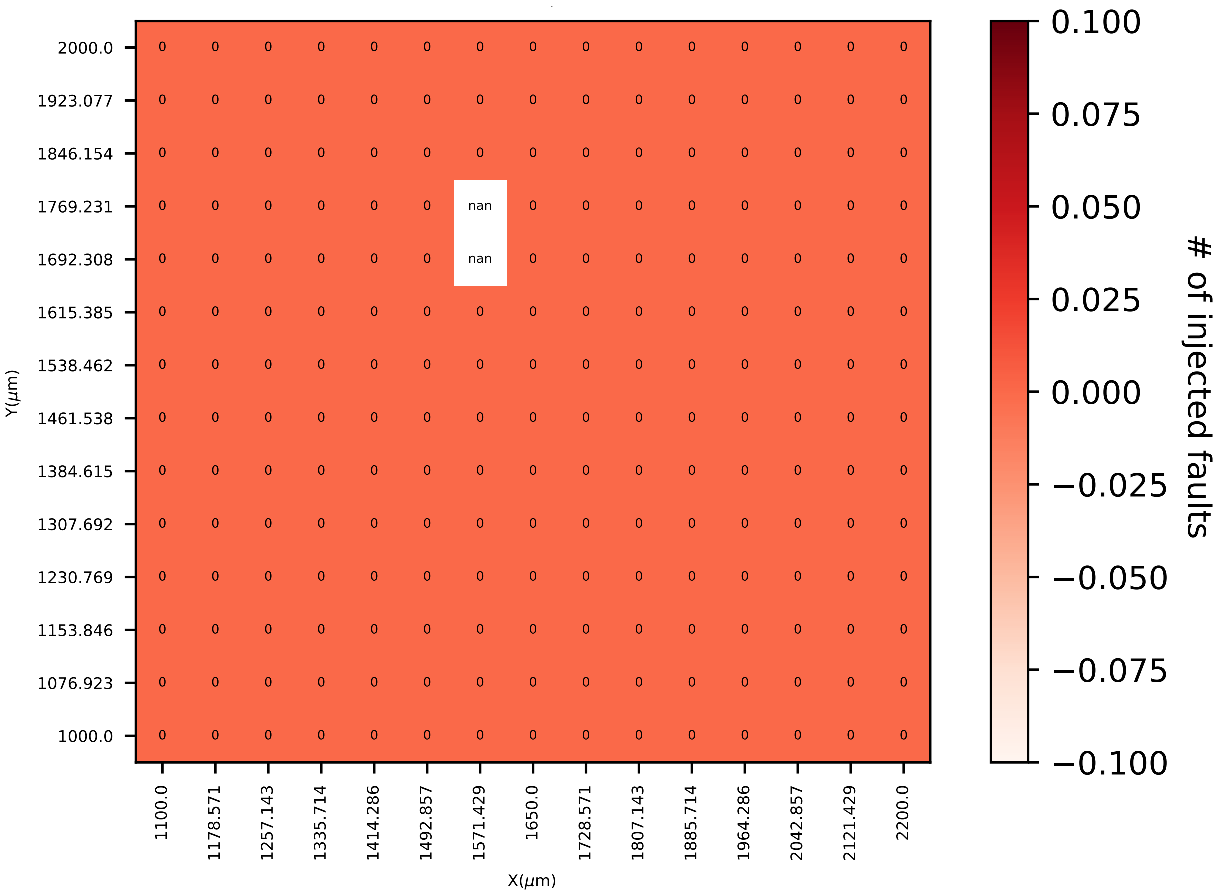
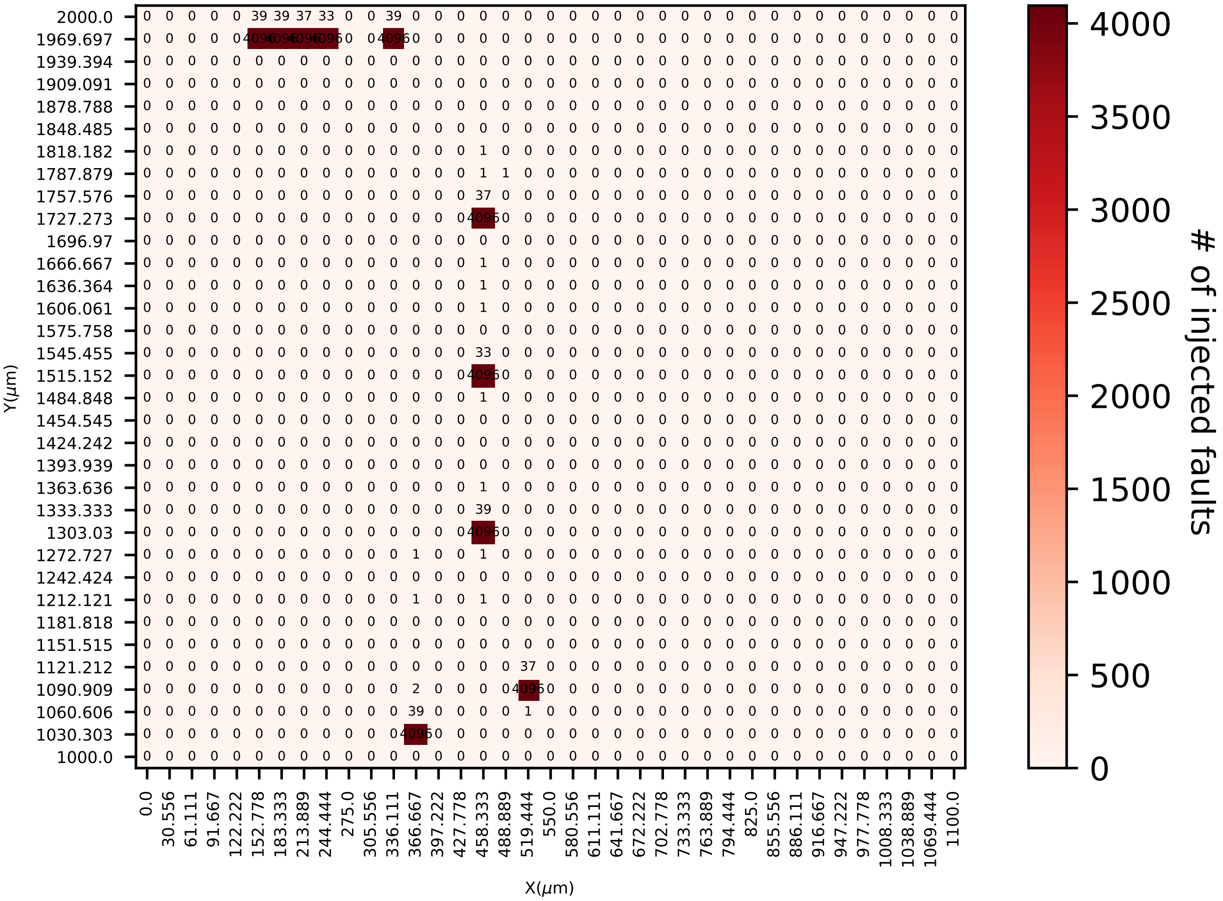
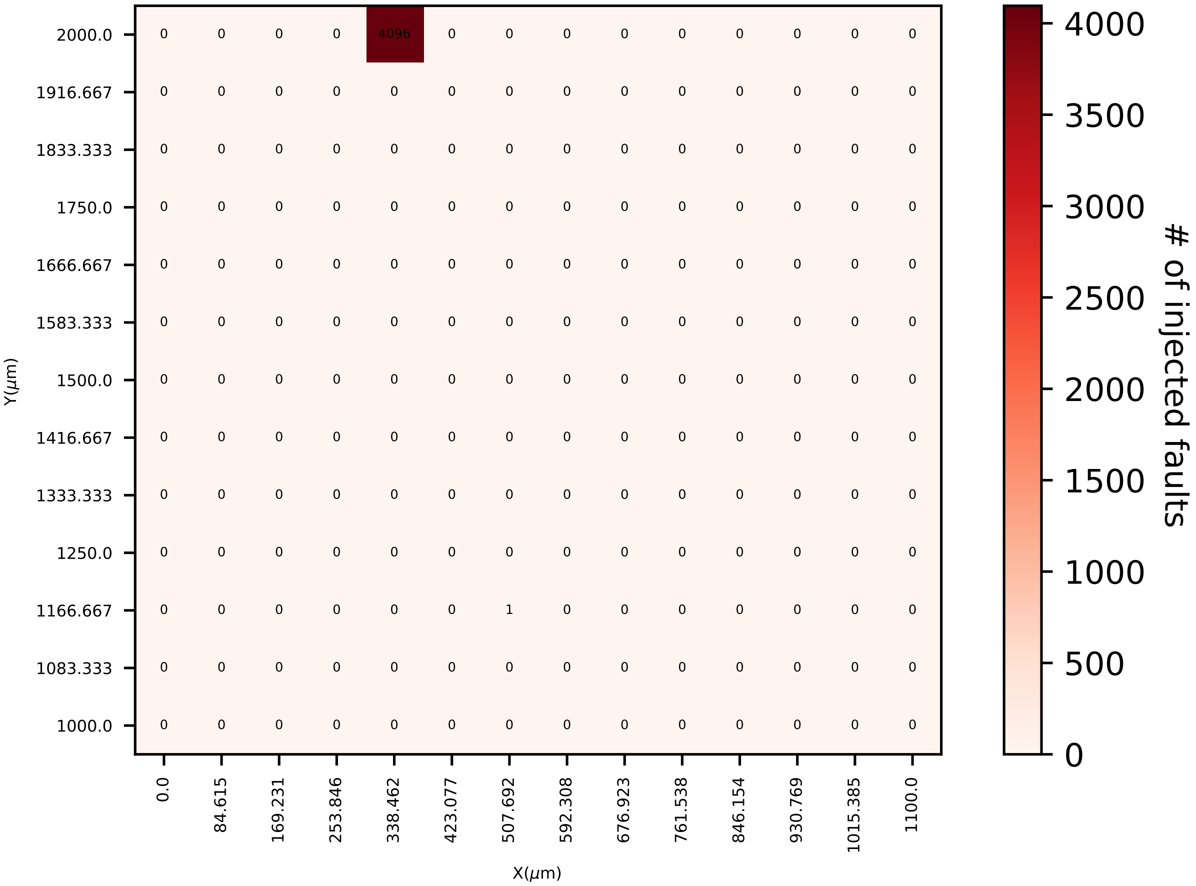


Figure C.1: DAP location in heat map. 50X objective used at 210 locations scanned in the upper right quadrant at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 100.0% of peak current with both PDM CH1 and CH2 enabled. Only two sensitive locations (labelled as NaN) experienced a reset behavior.



	0.0	30.556	61.111	91.667	122.222	152.778	183.333	213.889	244.444	275.0	305.556	336.111	366.667	397.222	427.778	458.333	488.889	519.444	550.0	580.556	611.111	641.667	672.222	702.778	733.333	763.889	794.444	825.0	855.556	886.111	916.667	947.222	977.778	1008.333	1038.889	1069.444	1100.0
2000.0	0	0	0	0	0	39	39	37	33	0	0	39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1969.697	0	0	0	0	0	4096	4096	4096	4096	0	0	4096	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1939.394	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1909.091	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1878.788	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1848.485	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1818.182	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1787.879	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1757.576	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1727.273	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4096	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1696.97	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1666.667	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1636.364	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1606.061	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1575.758	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1545.455	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1515.152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4096	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1484.848	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1454.545	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1424.242	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1393.939	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1363.636	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1333.333	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1303.03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4096	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1272.727	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1242.424	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1212.121	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1181.818	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1151.515	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1121.212	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1090.909	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	4096	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1060.606	0	0	0	0	0	0	0	0	0	0	0	0	39	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1030.303	0	0	0	0	0	0	0	0	0	0	0	0	4096	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1000.0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure C.2: 50X objective used at 1258 locations scanned in the upper left quadrant at 1.65V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 100.0% of peak current with both PDM CH1 and CH2 enabled.



Y(μm) \ X(μm)	0.0	84.615	169.231	253.846	338.462	423.077	507.692	592.308	676.923	761.538	846.154	930.769	1015.385	1100.0
2000.0	0	0	0	0	4096	0	0	0	0	0	0	0	0	0
1916.667	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1833.333	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1750.0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1666.667	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1583.333	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1500.0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1416.667	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1333.333	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1250.0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1166.667	0	0	0	0	0	0	1	0	0	0	0	0	0	0
1083.333	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1000.0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure C.3: 50X objective used at 182 locations scanned at upper left quadrant at 1.8V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 100% of peak current with both PDM CH1 and CH2 enabled. Only a hardfault at the top and a single instruction manipulation fault at the SRAM address 0x10000801.

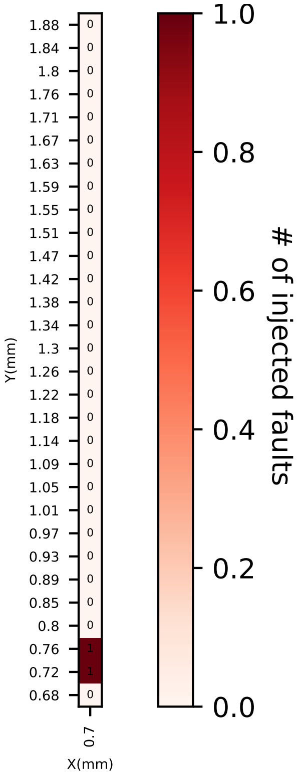


Figure C.4: 50X objective used at 30 locations scanned at Flash output at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 100% of peak current with both PDM CH1 and CH2 enabled. Both faults were instruction manipulation faults at the SRAM address 0x10000801.

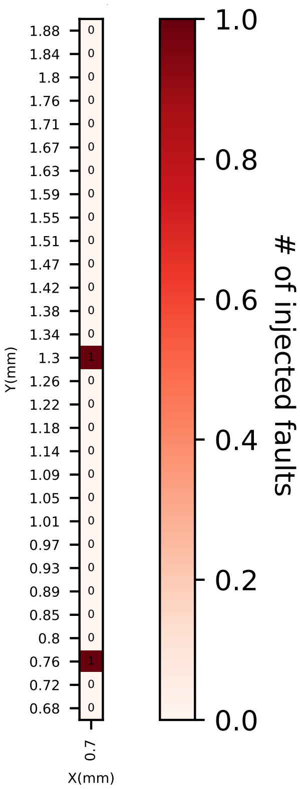


Figure C.5: 50X objective used at 30 locations scanned at Flash output at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 100% of peak current with both PDM CH1 and CH2 enabled. Both faults were instruction manipulation faults at the SRAM address 0x10000801.

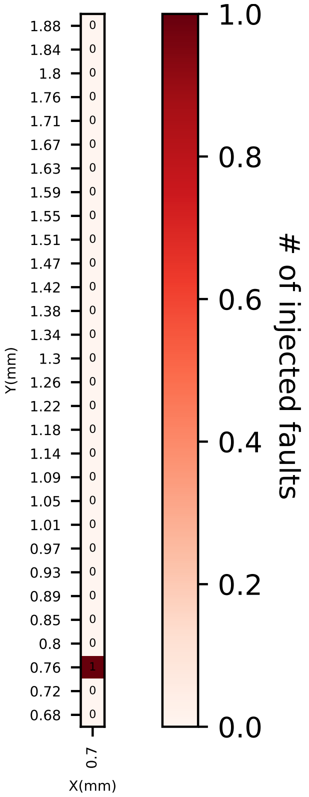


Figure C.6: 50X objective used at 30 locations scanned at Flash output at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 100% of peak current with both PDM CH1 and CH2 enabled. The fault was an instruction manipulation fault at the SRAM address 0x10000801.

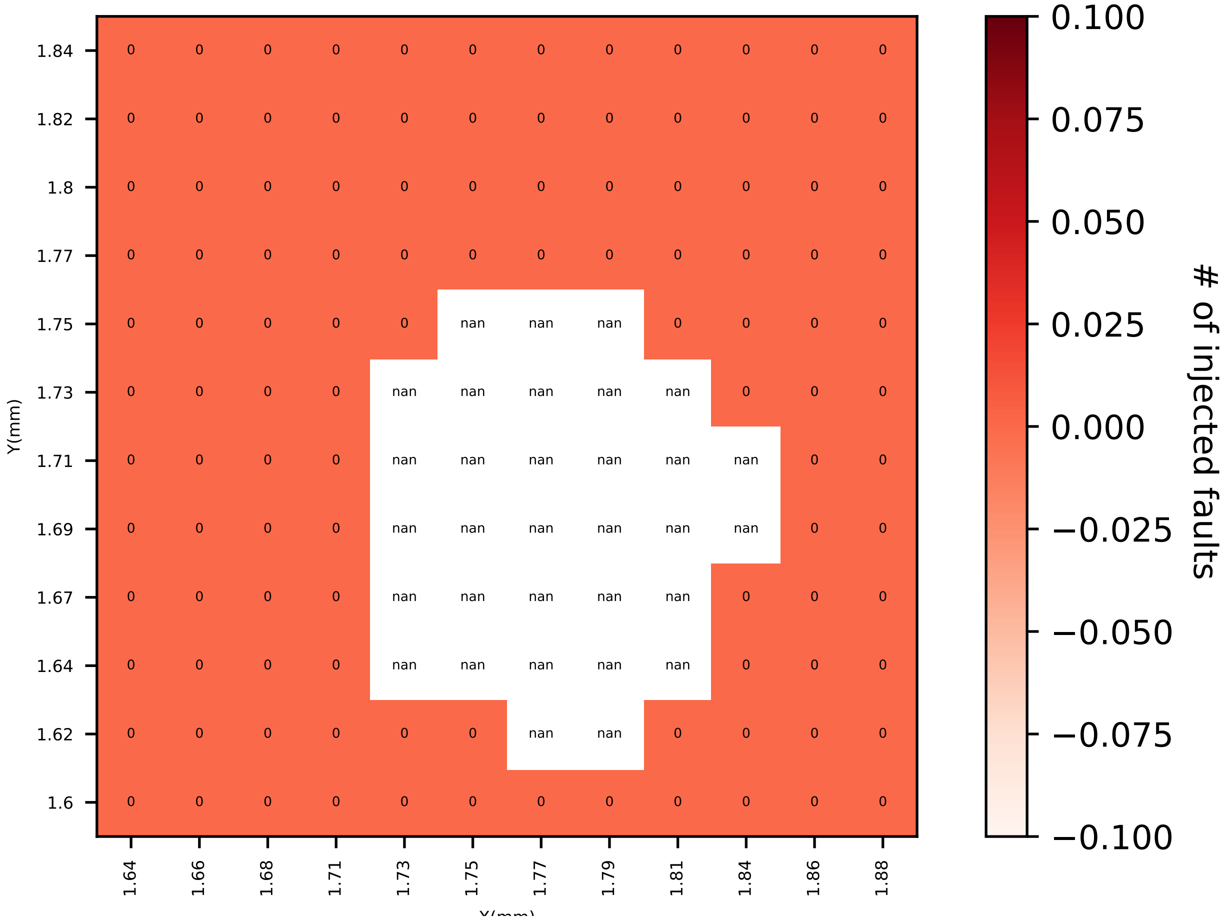


Figure C.7: 5X objective used at 144 locations scanned in the DAP area at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 12.5% of peak current with only PDM CH1 enabled.

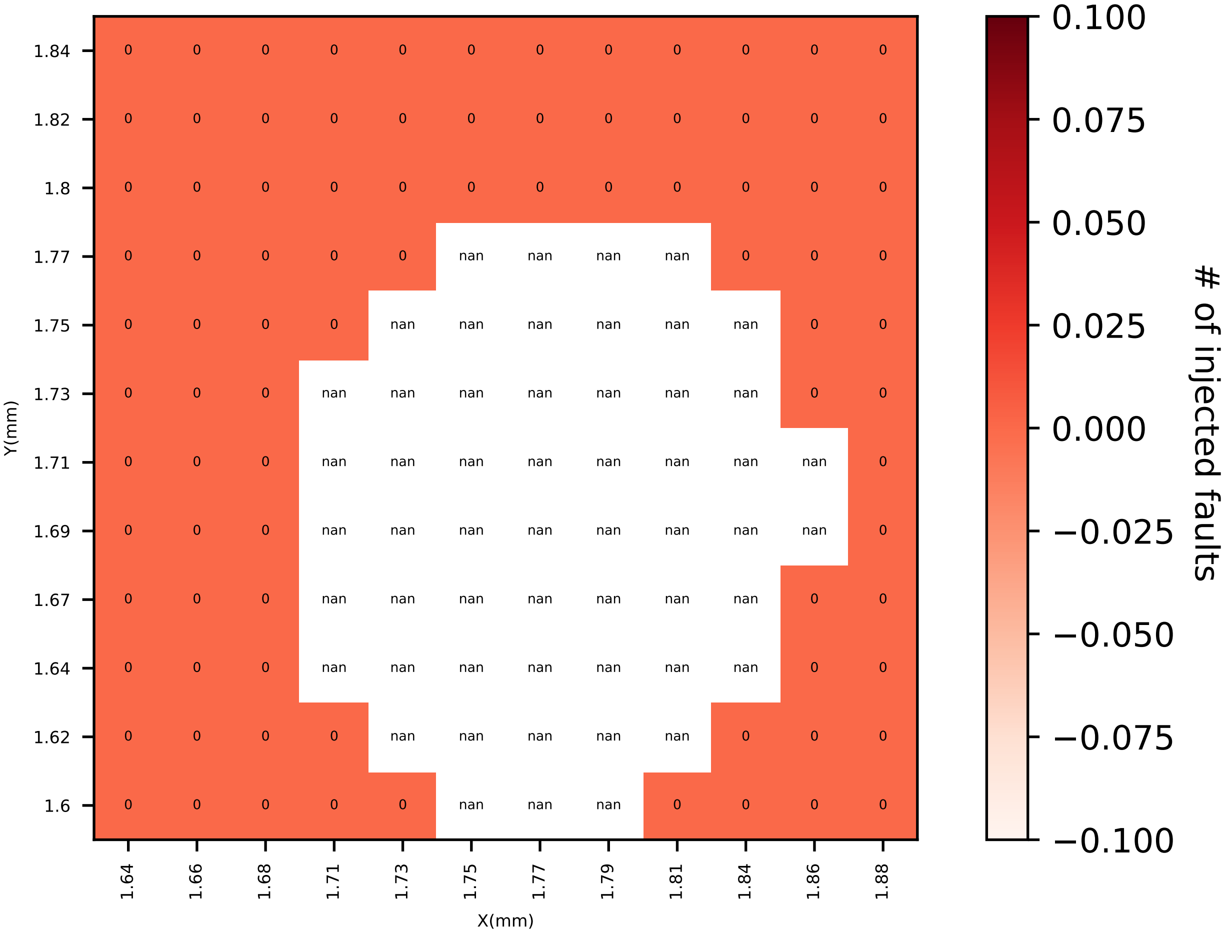


Figure C.8: 5X objective used at 144 locations scanned in the DAP area at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 25% of peak current with only PDM CH1 enabled.

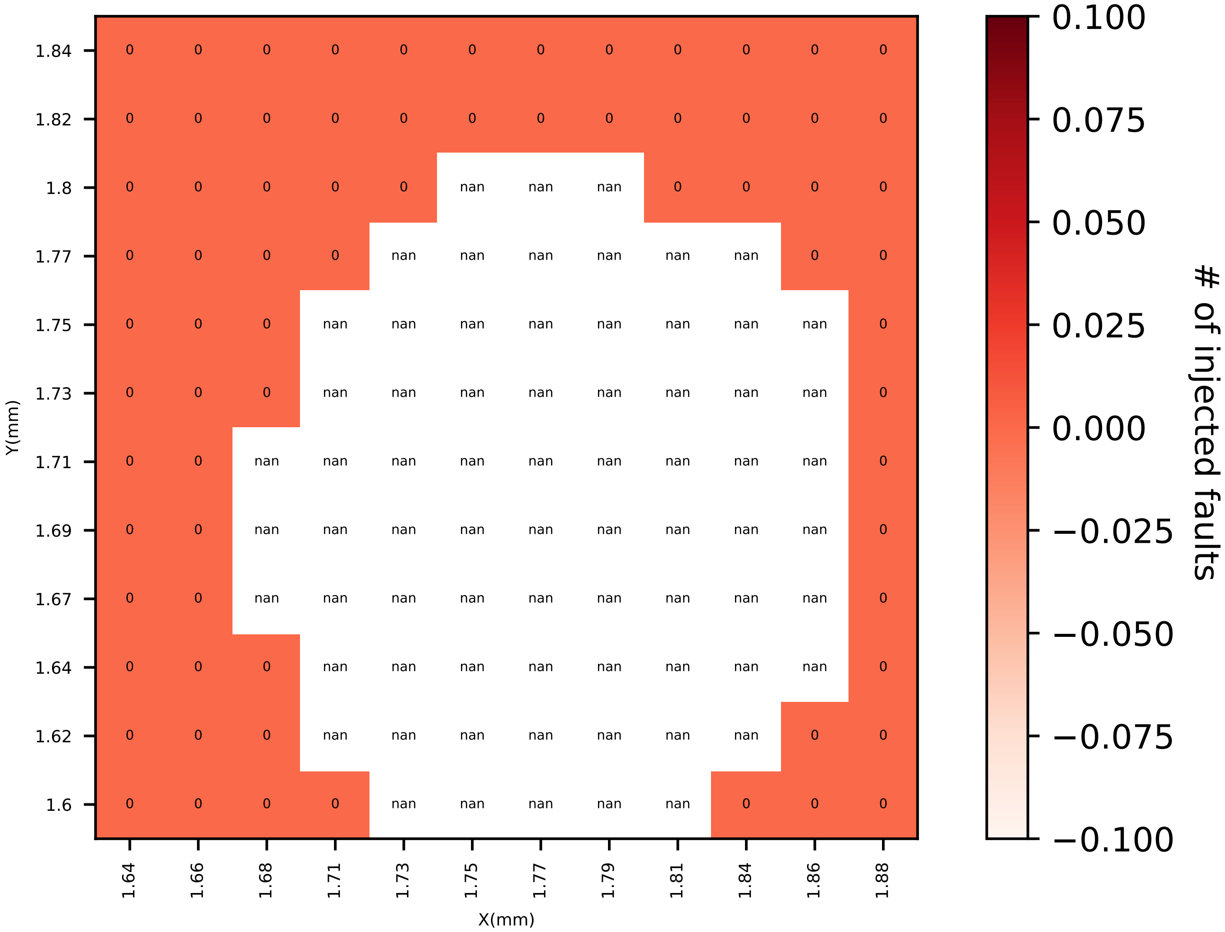


Figure C.9: 5X objective used at 144 locations scanned in the DAP area at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 50% of peak current with only PDM CH1 enabled.

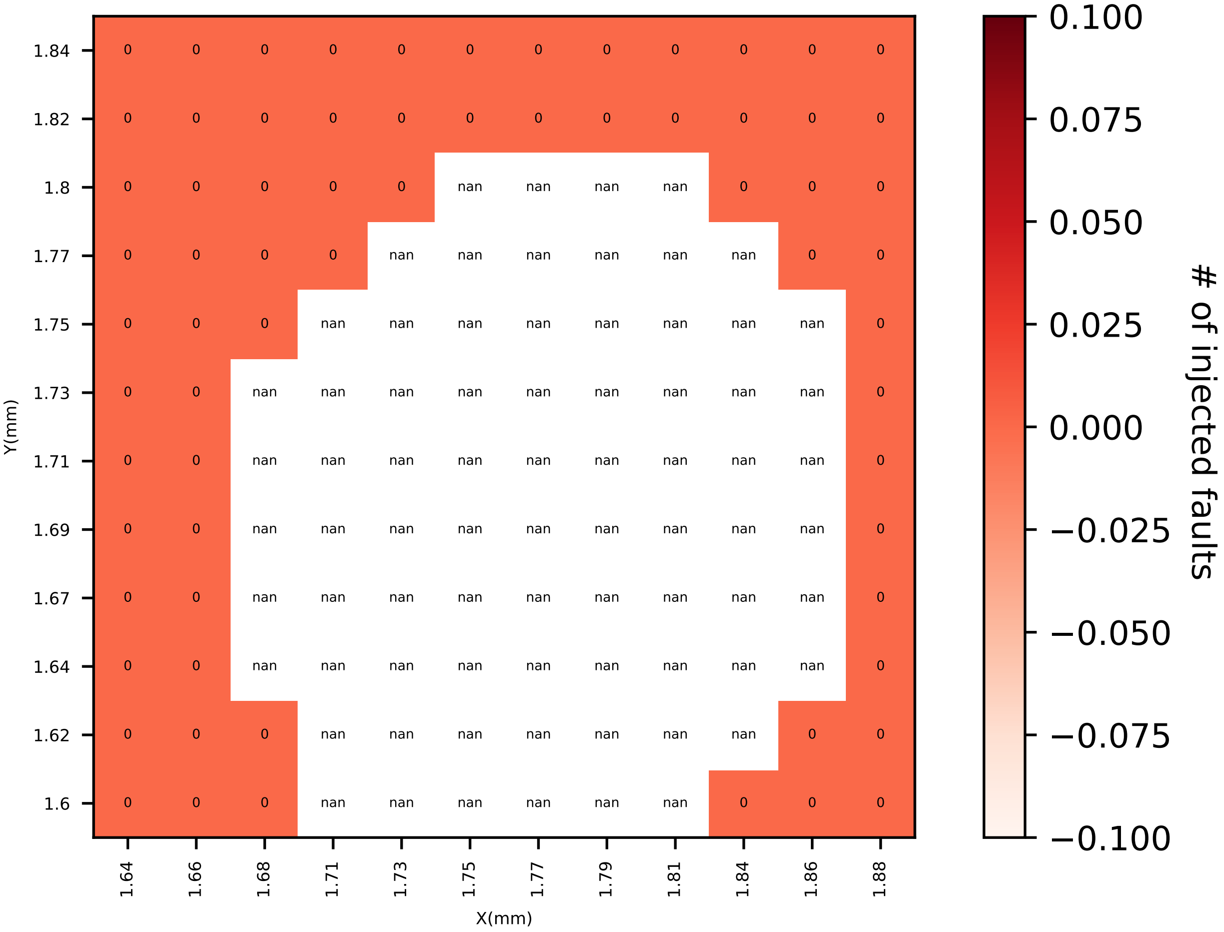
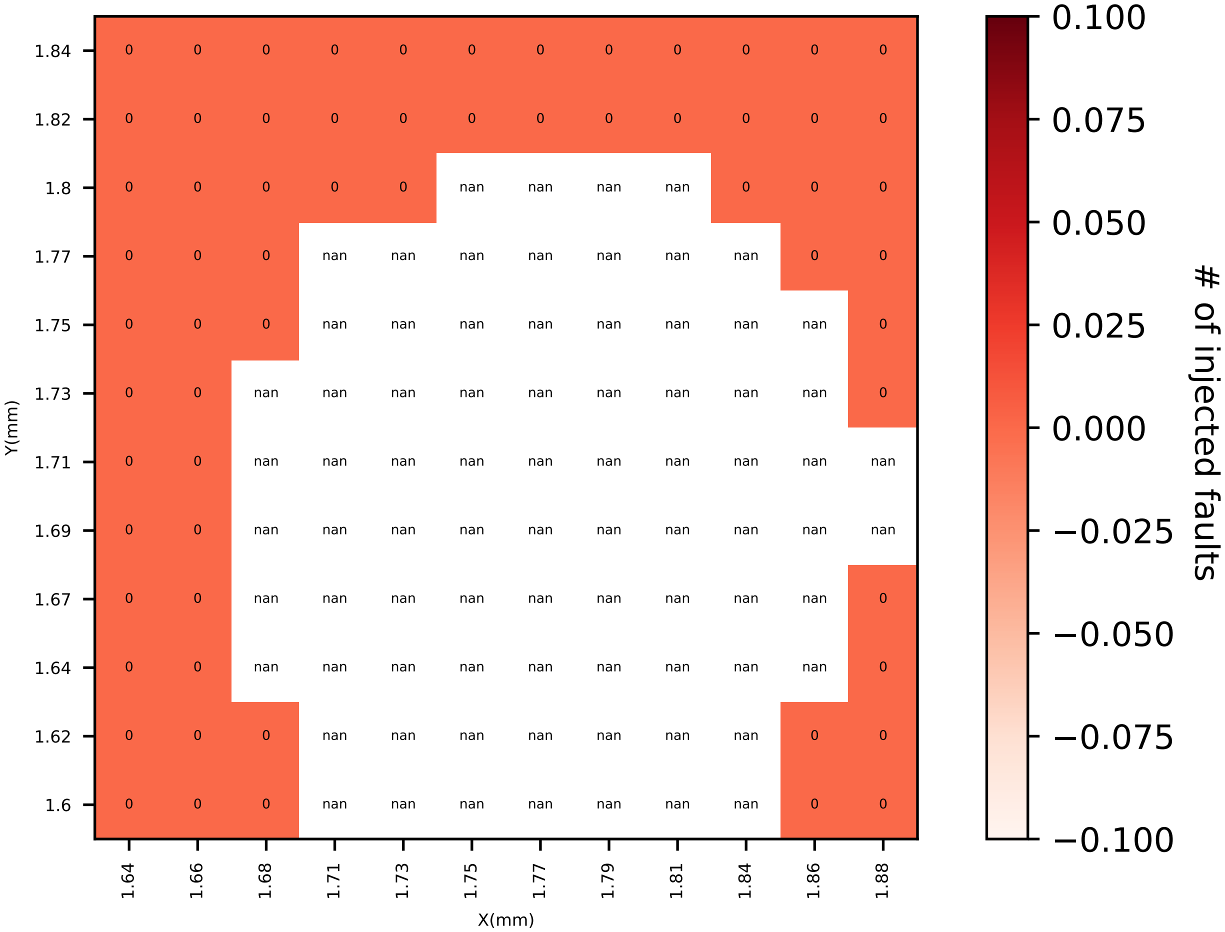


Figure C.10: 5X objective used at 144 locations scanned in the DAP area at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 75% of peak current with only PDM CH1 enabled.



Y(mm) \ X(mm)	1.64	1.66	1.68	1.71	1.73	1.75	1.77	1.79	1.81	1.84	1.86	1.88
1.84	0	0	0	0	0	0	0	0	0	0	0	0
1.82	0	0	0	0	0	0	0	0	0	0	0	0
1.8	0	0	0	0	0	nan	nan	nan	nan	0	0	0
1.77	0	0	0	nan	nan	nan	nan	nan	nan	nan	0	0
1.75	0	0	0	nan	nan	nan	nan	nan	nan	nan	nan	0
1.73	0	0	nan	nan	nan	nan	nan	nan	nan	nan	nan	0
1.71	0	0	nan	nan	nan	nan	nan	nan	nan	nan	nan	nan
1.69	0	0	nan	nan	nan	nan	nan	nan	nan	nan	nan	nan
1.67	0	0	nan	nan	nan	nan	nan	nan	nan	nan	nan	0
1.64	0	0	nan	nan	nan	nan	nan	nan	nan	nan	nan	0
1.62	0	0	0	nan	nan	nan	nan	nan	nan	nan	0	0
1.6	0	0	0	nan	nan	nan	nan	nan	nan	nan	0	0

Figure C.11: 5X objective used at 144 locations scanned in the DAP area at 3.3V supply voltage and DUA running at 20MHZ and a 200ns laser pulse width at 100% of peak current with only PDM CH1 enabled.