FOBOS 3: An Open-Source Platform for Side-Channel Analysis and Benchmarking

Eduardo Ferrufino
George Mason University
Fairfax, VA, USA
eferruf@gmu.edu

Luke Beckwith
George Mason University
Fairfax, VA, USA
lbeckwit@gmu.edu

Abubakr Abdulgadir
PQSecure Technologies
Boca Raton, FL, USA
abubakr.abdulgadir@pqsecurity.com

Jens-Peter Kaps
George Mason University
Fairfax, VA, USA
jkaps@gmu.edu

ABSTRACT

The lightweight cryptography (LWC) standardization process by the National Institute of Standards and Technology (NIST) of the US is the latest example of competitions that require benchmarking and side-channel leakage evaluation of hardware implementations of a multitude of candidate algorithms. A common hardware application programming interface (API) streamlines the development of a test harness. However, no existing platform is directly compatible with the LWC algorithms' hardware interface. Hence, a significant effort is needed to evaluate and benchmark a large number of candidates.

This paper presents an open-source, multi-user platform for side-channel analysis and benchmarking we call FOBOS 3. It contains its own measurement board (FOBOS Shield) and target board (FBD-A7 with Xilinx Artix-7-A12 FPGA) and enables side-channel leakage evaluation as well as measurement of power and energy consumption. Case studies are included to highlight both features.

CCS CONCEPTS

• **Security and privacy** → **Side-channel analysis and countermeasures**; • **Hardware** → *Reconfigurable logic applications*; • **Applied computing** → Computer-assisted instruction.

KEYWORDS

Side-channel Analysis; Benchmarking; Hardware Security; Lightweight Cryptography

ACM Reference Format:
Eduardo Ferrufino, Luke Beckwith, Abubakr Abdulgadir, and Jens-Peter Kaps. 2023. FOBOS 3: An Open-Source Platform for Side-Channel Analysis and Benchmarking. In *Proceedings of the 2023 Workshop on Attacks and Solutions in Hardware Security (ASHES '23), November 30, 2023, Copenhagen, Denmark.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3605769.3623987

ASHES '23, November 30, 2023, Copenhagen, Denmark

ACM ISBN 979-8-4007-0262-4/23/11.
https://doi.org/10.1145/3605769.3623987

1 INTRODUCTION

Security evaluation and benchmarking of cryptographic implementations are crucial phases in designing solutions capable of successful deployment. Side-channel attacks [12] pose a severe threat to the security of cryptographic functions, which lie at the core of security mechanisms protecting valuable data.

However, implementing side-channel-resistant cryptographic functions is a complex and error-prone process. Hence, validation steps must be performed to ensure that the design meets the intended security goals. Validation can take the form of mounting attacks at various attack points or performing statistical leakage assessments such as Test Vector Leakage Assessment (TVLA) [5, 22].

Power and energy consumption are also vital considerations, especially for battery-powered devices. Hence, they should be accurately measured to ensure the design is within the power and energy budgets for the application in mind.

Tools for performing security evaluation and benchmarking are already available and range from high-end and costly commercial equipment to ad-hoc setups built for a single experiment. However, in some situations, low-cost, highly accessible, and flexible tools are needed. This is where we see a gap in currently used platforms.

In this work, we present a significantly updated version of the FOBOS side-channel analysis platform with applications in research and education. This includes a new control board with an internal ADC capable of measuring traces using a sampling clock synchronous to the target clock, providing precise measurements. Additionally, the control board can measure the target's power consumption. We also introduce a new low-cost FPGA-based target board, providing another option for evaluating hardware designs. Furthermore, we present case studies illustrating the effectiveness of our platform in the security evaluation and benchmarking of finalists in the NIST lightweight cryptography standardization process.

Hence, we claim the following contribution:

(1) An upgraded side-channel analysis platform capable of mounting attacks, assessing leakage, and performing power measurements. The platform is open-source, enabling result replication and further improvement by the community.
(2) A new target board based on a small low-cost FPGA.
(3) Comparison of leakage assessment results of a protected implementation of the NIST LWC finalist Xoodyak using synchronous and asynchronous sampling.

(4) Comparison of power consumption measurements of the NIST LWC winner Ascon on two different FPGA boards with estimates generated by simulation.

2 PREVIOUS WORK

Several side-channel and benchmarking platforms are available and widely used in the cryptographic community. The Rambus DPA Workstation [19] and Riscure Inspector [20] are examples of commercial platforms. The SASEBO [10] and the SAKURA boards [6] have been used in many academic publications and provide FPGA and smart-card targets. The NewAE ChipWhisperer is an open-source platform widely used in academia and industry and provides target boards for microcontrollers and FPGAs. ChipWhisperers sampling clock can be synchronized with the target clock enabling more precise measurements [15].

FOBOS, which stands for "Flexible Opensource workBench fOr Side-channel analysis" and is loosely named after the Greek god Phobos ($\phi\acute{o}\beta o\varsigma$) who personifies fear and can pierce shields, was first presented in the "Work in Progress" session of COSADE 2012 [25]. Unlike the SAKURA, it de-couples the control device from the Device Under Test (DUT), also called *target*. It uses off-the-shelf FPGA boards as control and modified DUTs along with custom-built boards. Version 2 of FOBOS, presented at ReConFig [1] in December 2019, replaces the monolithic control logic design with a more flexible design based on functional modules and a Xilinx MicroBlaze processor. It provides a simple wrapper for FPGA implementations of cryptographic algorithms according to the CAESAR [7] and LWC Hardware API [9], which uses a FIFO-based approach. Additionally, FOBOS provides a simple Control-DUT protocol. Hence, we chose FOBOS 2 as the basis for our work.

3 FOBOS 3

In order to develop an efficient multi-user platform for side-channel analysis and benchmarking, several drawbacks of the FOBOS 2 platform had to be overcome. The main drawback is that FOBOS 2 uses an external oscilloscope which increases the setup cost. Synchronous sampling is only possible with advanced oscilloscopes that allow for an external reference clock. Furthermore, FOBOS 2 has no built-in capability to perform power consumption measurements and uses USB/UART communication which leads to a low data transfer speed and it lacks multi-user capability. To address these drawbacks, we performed a major update to the FOBOS 2 platform resulting in the FOBOS 3 platform that we discuss in detail in this paper.

3.1 Platform Overview

The main components of the FOBOS 3 platform are as follows:

(1) The *SCA Workstation* hosts capture and analysis scripts. In typical system usage, the user interacts with Jupyter notebooks to run capture scripts that send test vectors (e.g., key and plaintext) one at a time to the connected control board and receive the results (e.g., ciphertext) and power traces. Analysis scripts can be run on the saved results and traces or optionally executed simultaneously with trace collection.
(2) The *control board* receives a test vector from the SCA workstation and forwards it to the DUT. The control board asserts a trigger signal as soon as the DUT starts operating on the test vector (i.e., running the cryptographic function). This signal can trigger an external oscilloscope or the internal ADC hosted in the FOBOS Shield, which we will discuss in Section 3.3. The control board also generates the target clock. The internal ADC can sample changes in the power consumption of the DUT at a maximum rate of 100 M Sample/sec. The sampling clock and the target clocks are synchronized to achieve high-precision measurements.
(3) The *DUT board* hosts the cryptographic function that is being evaluated. This board is designed (or modified) to allow measurement of the power consumption using a shunt resistor or a current probe. We tested the platform using a modified Digilent Nexys 3 board featuring a Xilinx Spartan-6 FPGA, the NewAE CW305 Artix-7 target board, and our FOBOS Artix-7-based DUT board that is discussed in Section 4.

A typical FOBOS 3 setup is shown in Figure 1. The control board is based on the PYNQ-Z1 board, which features a Xilinx Zynq-7020 system-on-chip with a dual-core ARM processor running Linux and tightly coupled with an FPGA fabric. We chose the PYNQ board for FOBOS 3 as it has more flexibility and performance than a softcore processor or hardcoded logic. Furthermore, it supports segmentation of the control logic into software and hardware and its AXI bus allows us to re-use IP cores from FOBOS 2.

The custom FOBOS Shield is mounted on top of the PYNQ board and provides a 10-bit ADC, a pre-amplifier, power regulators, power measurement circuitry, and a target communication interface. We chose to develop a Shield for PYNQ-Z1 rather than a capture board complete with Zynq chip for simplicity. A version of the FOBOS Shield for PYNQ-Z2 is currently under development. The target device shown in Fig. 1 on the right is the FOBOS Artix-7-A12 DUT.

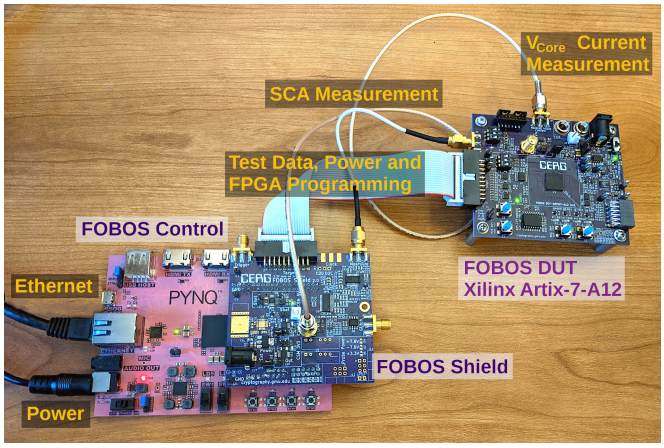

Figure 1: FOBOS 3 setup with FOBOS Artix-7-A12 DUT

The high-level block diagram of the FOBOS 3 system shown in Fig. 2 details the modules we built into the programmable logic on the Zynq chip. This includes modules to control and interact with the hardware of the FOBOS Shield. At the center is a clock wizard that generates the clock signals for the DUT (DUT Clock) and the ADC (ADC Clock) and ensures synchronization. The power module

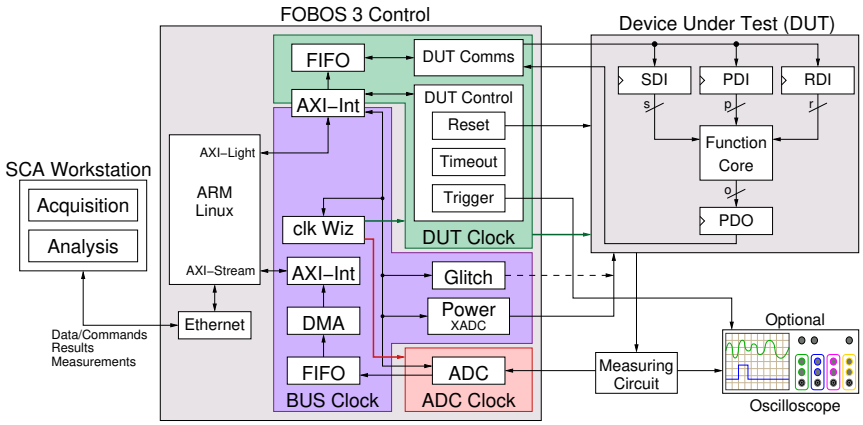

Figure 2: Block Diagram of FOBOS 3

controls the power regulators and uses the Xilinx ADC (XADC) of the Zynq chip to measure output voltages and currents. The DUT control module provides triggering, among other tasks, while the DUT communications module handles the communication with the DUT. The control board is connected to the SCA workstation using a Gigabit Ethernet port to ensure high-speed data transfer.

On the target board, we provide a flexible wrapper that receives data from the control board and distributes it to input FIFOs: SDI (secret data input) to store keys, PDI (public data input) for data such as plaintext, and RDI (random data input) for random data for masked implementations or seeds for stream ciphers. Once data is stored in the input FIFOs, the wrapper allows the cryptographic function to run and consume their data.During the operation of the cryptographic function, all I/Os except for the clock are idle to minimize noise. The wrapper accumulates the output in the data output FIFO (DO) and sends it back to the control board after the function core finishes its execution. Optionally, random bits needed for masked implementations can be produced in the target FPGA using the provided Trivium-based [4] random number generator with a seed from RDI.

3.2 Online Access and Application to Teaching

We use JupyterLab, an interactive web-based development environment, to provide access to FOBOS. JupyterLab is hosted on the SCA Workstation and communicates with a server program on the FOBOS control board. The protocol allows the sending of commands, test vectors, and FPGA configuration data for the FOBOS DUT and the receiving of the power traces from the ADC, power consumption data as well as results from the DUT. Figure 4 shows how FOBOS can be accessed remotely. All functions a user needs to interact with FOBOS are provided in a Python library (foboslib) including functions to handle multiple user requests for the same FOBOS setup. When a FOBOS setup is idle, a user can get exclusive access to configure the setup, program the DUT, and run measurements, after which the user should relinquish the access. In case the DUT becomes unresponsive, e.g., through glitching, or the user is idle for a long time, access is rescinded and the setup is brought back to a default state. This type of online access is not meant to offer SCA as a service to a wide audience, but to allow groups of people to run their experiments from home and to easily share equipment.

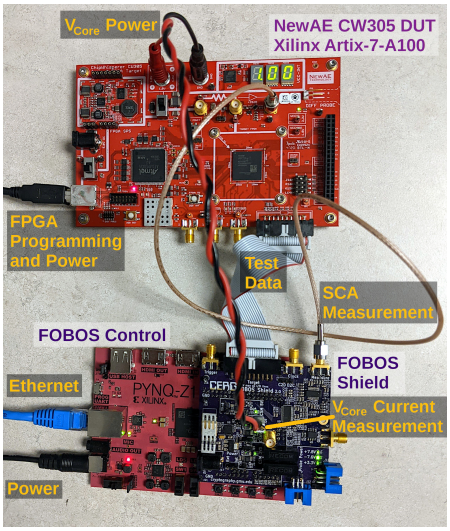

Figure 3: FOBOS 3 setup with CW305 Artix-7-A100 DUT

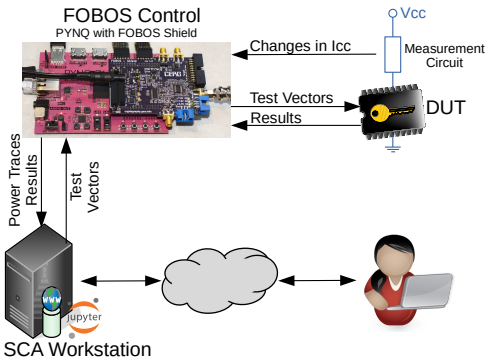

Figure 4: Remote Access to FOBOS 3

We started using an early version of FOBOS 3 for on-line labs in an undergraduate university course in Spring 2020 when COVID19 forced a shutdown of in-person classes at our university. Our experience with using on-line SCA labs in the subsequent years have been described in [2].

3.3 FOBOS Shield

The FOBOS Shield circuit board plugs into the Arduino/chipKIT connector of the PYNQ-Z1 board. It contains a version of the OpenADC [14] with a 10-bit ADC with 105 MS/s sampling rate and a software adjustable Low Noise Amplifier (LNA) with a gain of up to 55 dB and 40 MHz bandwidth. This is the same circuit used by ChipWhisperer-Lite and used to capture changes in power consumption or EM radiation of the target device. Communication and clocking of the target device is achieved through a 20-pin target connector, which is compatible with ChipWhisperer target boards. It provides additional functionality such as programming the FPGA of our FOBOS Artix-7-A12 DUT. The FOBOS Shield also contains a simple crow-bar voltage glitching circuit and outputs for clock and trigger signals on SMA connectors. Furthermore, it has a ChipWhisperer-compatible isolated probe power supply. In order to facilitate power consumption measurements for benchmarking, the FOBOS Shield contains power supplies that allow calibrated voltage and current measurements.

3.4 Power Measurements

The FOBOS Shield has three separate power outputs, a 5 V, a 3.3 V, and a variable voltage output which can supply between 0.90 V and 3.65 V in 0.05 V steps and is controlled through the FOBOS Python library. This enables it to provide V_{Core} for Xilinx Artix-7 FPGAs of 1 V. Each of these power sources has a measurement circuit composed of a shunt resistor and an INA225 current shunt monitor (CSM), which allows current and voltage to be measured. The CSMs have programmable gain, which is controlled through the FOBOS control board. The gain settings in (V/V) are 25, 50, 100, and 200. The error range for these settings are respectively ±0.15%, ±0.15%, ±0.2%, and ±0.3%. The diagram of these power sources and their corresponding measurement circuits are shown in Fig. 5.

The measurement circuits can be sampled either internally using the XADC or externally by connecting an oscilloscope to the Shield. The measurement using the XADC requires no additional hardware and can be easily controlled using software. The average and maximum values are calculated automatically within the FPGA while the target is running and can be read using the FOBOS Python library. The XADC samples current and voltage from each power source. The effective sampling rate for any single source is 137 kHz. The precision of samples is 12 bits. Up to 2^{20} samples can be taken before the averaging circuit overflows. Given that a sample is taken approximately every 7.3 microseconds, this corresponds to a maximum measurement time of 7.65 seconds. The primary drawback of the internal measurements is the low sample rate. As we will discuss in the case study in Section 5.2, this sample rate is acceptable for many applications. However if a higher sample rate is needed, all power outputs may be accessed externally. Drivers are provided to measure current externally using programmable oscilloscopes.

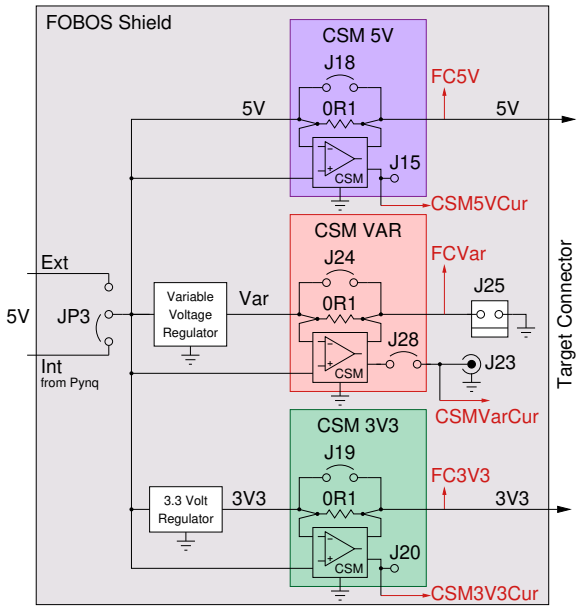

Figure 5: FOBOS 3 Power Supply and Measurement

3.5 Power Calibration

To account for the error tolerances in various components within the power measurement circuitry, a script for calibration is provided. For the variable voltage supply, the voltage is calibrated by connecting the variable power source output to a programmable DC Load and comparing the load's measured voltage against the FOBOS Shield's measured voltage at various voltages. A polynomial is then fit to the difference between these two measurements and used to correct any error in the measurements. For current calibration of all three voltage supplies, the programmable DC load runs through a series of currents. The same method to calculate an error correction polynomial is used for the current measurement. Since the voltage output of the 5 V and 3.3 V supplies cannot be modified, the voltage measurement is calibrated at the same time as the current. Since a wide range of current draw is used during calibration, the range of variation in voltages is representative of any voltage drop that may occur during normal use.

As an example of the need for calibration of the power measurement circuitry, we discuss results of calibration of the variable output current measurement. In this example, the current measurement is calibrated for the range [0 mA, 200 mA]. The DC load iterated from 0 mA to 200 mA in 5 mA steps. At each step the actual current consumption of the DC load is recorded as well as the current measured through the FOBOS Shield. With no calibration, on average the FOBOS Shield measure 15% lower than the actual current. With the error correction polynomial calculated from these measurements the error is reduced to ±2 mA.

4 FOBOS FPGA-BASED DUT

For our DUT we selected the XC7A12T-1CPG238C FPGA from the Artix-7 FPGA family due to its value-based positioning in the marketplace, with the A12T member of the family offering the

least amount of programmable fabric making it a likely processing element for a resource constrained IoT system. This FOBOS Xilinx ARTIX7-A12 DUT (or FBD-A7 for short) is shown in Figure 6.

4.1 Hardware Overview

The FBD-A7 enables direct low-noise sampling of V_{Core} current and voltage over SMA connectors. It also provides a low impedance interface to V_{Core} for voltage glitching. The 20-pin target connector is used to receive power and communication signals from the FOBOS Shield. A DIP switch allows for setting of the boot mode to enter the various programming modes. The board features a SPI Flash which can be used to store the configuration bitstream in non-volatile memory. A header provides access to the JTAG programming interface. The board features user buttons, LED's, and a PMOD connector. The general system architecture and power distribution is illustrated in Figure 7. The power subsystem is designed to be powered in three configurations, FOBOS Shield powered, FOBOS Shield + external supply for V_{Core}, or external supply only. The benchmarking work performed in this paper was performed using the FOBOS Shield as the primary power source. Each individual FPGA voltage rail has connection points that enable direct measurements.

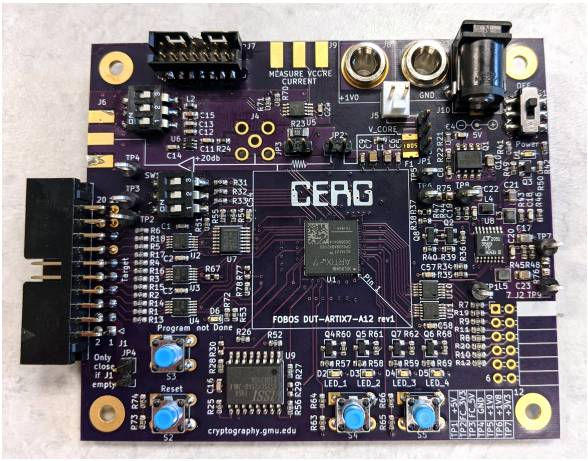

Figure 6: FOBOS Xilinx Artix-7-A12 DUT

FPGA Programming. The FBD-A7 board supports three different configuration modes - Slave Serial, Master SPI, and JTAG configuration mode. The primary configuration mode is Slave Serial mode, as this allows FOBOS Control to program the FPGA with a bitstream via the target connector as part of a side-channel evaluation application. The second FPGA programming method builds on the first method by initially programming the FPGA with a flash interface core and subsequently passing the bitstream through to on-board flash. Programming methods 1 and 2 have the added benefit of reducing the overall cost of the FOBOS workstation by not requiring the purchase of a USB JTAG programmer.

Power Consumption: A secondary purpose of the DUT board is to enable measurement of the V_{Core} current consumption. It contains the same INA225 current shunt amplifier as used on the FOBOS Shield, re-using the 0.1 Ω shunt used for power analysis measurements.

Power Glitching: The board provides access to V_{Core} Voltage via SMA connector for power glitching. Additionally all signals on the target connector have ESD protection. Voltage glitching involves momentarily shunting the V_{Core} rail to ground using a PowerFET or other low-impedance switching mechanism. Voltage glitching can result in corrupted register states, erratic state machine behavior, and other out-of-bound type errors.

Design for Cost: Cost was a primary consideration and challenge when designing FBD-A7. The Printed Circuit Board(PCB) layout has four layers, which keeps PCB fabrication costs reasonable. The selected XC7A12T FPGA is the lowest-cost Artix-7 variant offered that still has enough logic resources to implement protected LWC candidates. The total Bill of Material cost for FBD-A7 is $ 150 and PCB costs are $ 40 bringing total cost to less than $ 200 at the time of this publication.

4.2 Design for Side-channel Analysis

The primary purpose of the FBD-A7 board is to enable measurement of the V_{Core} voltage rail while the Artix-7 FPGA is performing cryptographic operations. The voltage fluctuations observed on the power rail are so minuscule that preliminary analysis would not detect any leaked information. In order to amplify the side-channel leakage, a BGA 2801 wide-band amplifier is used to provide 22.2 dB of gain. In order to improve the side-channel measurement Signal to Noise Ratio (SNR) performance, various methods are implemented at the PCB level to reduce noise on the V_{Core} voltage rail such as routing critical measurement traces as coplanar waveguides, providing low loss measurement connections, and amplifying the signal in near proximity to the measurement point. V_{Core} can be generated using an integrated switching power supply, whose output is filtered using a low-pass filter.Alternatively, an external power supply can also be used to power V_{Core}. Footprints for V_{Core} voltage rail bypass and decoupling capacitors are present, although component sites are not populated as the capacitors will filter side-channel leakage.

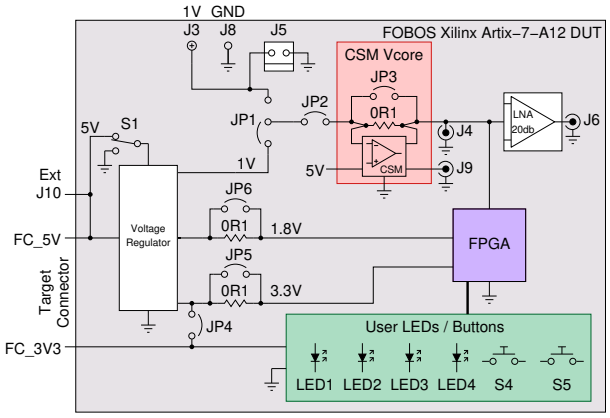

Figure 7: FBD-A7 Power Measurement

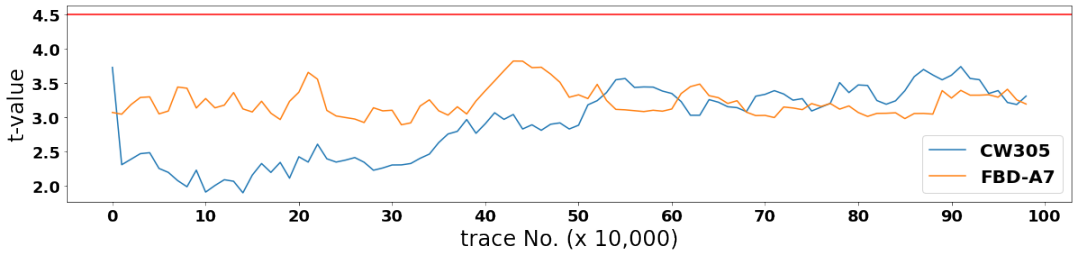

Figure 8: T-test measurements for LWC Candidate Ascon

4.3 Side-Channel Measurement Performance

Several FPGA side-channel target boards are available on the commercial market place such as the SAKURA-X and ChipWhisperer CW305. The latter presents a suitable reference point to compare the performance of FBD-A7 since both boards feature Artix-7 FPGA devices and are compatible with FOBOS 3. The CW305 uses an XC7A100T-2FTG256L FPGA while the FBD-A7 DUT uses the smaller XC7A12T-1CPG238C FPGA. In order to validate the side-channel measurement performance of our design we utilize experimental methods. All performance evaluations are performed without decoupling capacitors on the core voltage rail.

Method 1: SNR Measurement of Side-Channel traces. To validate the performance of FBD-A7 side-channel traces we leverage the method for measuring SNR performance of side-channel traces set forth by Iokibe et al. [8]. The signal in the context of a side-channel measurement is quantified by calculating the variance of side-channel traces while processing a random data set. Conversely, the noise can be quantified by calculating the variance of side-channel traces while processing a fixed data set. The cryptographic core used for this evaluation is the first-order protected Ascon developed by Amir Moradi's group at Ruhr University Bochum and available at [21].

Table 1: SNR Measurement Results.

DUT	FPGA	SNR(dB)
CW305	XC7A100T-2FTG256	1.816
FBD-A7	XC7A12T-1CPG238C	7.774

Method 2: Welch's T-Test Comparison. Welch's T-test determines if there is any correlation between the observed power traces and the processed data. The benchmark for performance is the CW305 board. The cryptographic core used for this evaluation is the same as used for Method 1. It was built using default Vivado settings. The front-end gain settings were adjusted until the signal exceeded the ADC range, and then reduced to provide some safety margin. The acquisition settings were set the same for both boards. The results of this experiment are displayed in Figure 8. For both target boards, the result is that after 1M traces, neither can detect a strong correlation between the processed data and collected power traces. Both target boards trend toward similar values over 1M traces.

Method 3: Measurements to Disclosure Tests. In this test series we run Correlation Power Analysis (CPA) attacks against a simple, not protected implementation of the Advanced Encryption Standard (AES). We implemented this 11-clock cycle per round AES with the FOBOS wrapper using Vivado with either the port mapping file (XDC) for the CW305 or the FBD-A7 board. Both implementations require 1328 LUTs, 923 Flip-Flops, and 6 Block RAMs. The control board for these tests is our FOBOS 3 and the measurements were made by the FOBOS Shield. After performing all tests on the FBD-A7 board, the CW305 board was connected to the same FOBOS Shield using the same SMA and ribbon cables to ensure that the setup, cable quality, and location does not influence the comparison. We determine for how many encryptions we have to measure the changes in power consumption in order to be certain that we recovered the correct key, known as Measurements To Disclosure (MTD). We perform this MTD test for each byte of the AES key and repeat this test for 10 different random keys. As a lot of factors such as the particular layout of the circuit, the value of the key byte, environmental noise, power supply noise, noise from other components on the DUT board, etc. influence how clean the measurement of changes in the current consumption are, i.e., how well they relate to the data being processed, we see large variations in the results.

Figure 9 shows the Box and Whiskers plot generated from the MTDs when performing CPA on both boards. The box starts at the 1st quartile till the 3rd quartile showing the median with the center line. The whiskers show the minimum and maximum MTDs. While the maximum is important for breaking all bytes of the key, the graphs are an indicator of the noise in the measurement. The smaller the box and the shorter the whiskers, the less the noise. The first two plots in Fig. 9 compare the FBD-A7 board with the CW305. Recovering all bytes of the keys requires less than 25,0000 encryptions on the FBD-A7 board while some key bytes on the CW305 board need almost 4 times as many encryptions. Even the difference between the medians is almost a factor 4. While the CS305 board contains a much larger FPGA than the FBD-A7, the AES implementations require the same amount of resources on the FPGAs of both boards. We think it unlikely that the larger amount of unused FPGA resources is the reason for the larger noise. However, the larger internal capacitances of the larger FPGA should remove noise. One possible source is the power supply of the CW305. When checking it for noise with an oscilloscope, we could find none. However, powering the core voltage (1 V) of the CW305 from an external lab power supply and repeating all tests we noticed that the MTD results improved as shown in the last plot in Fig 9. The maximum MTD is now less than 4 times that of the

FBD-A7 and the median is now less than 3 times larger. Supplying the core voltage from an external power supply also improves the performance on the FBD-A7 board but by less than 7%.

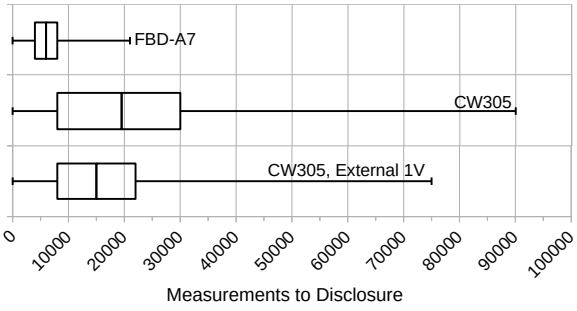

Figure 9: Comparison of MTD for different boards when performing CPA on an unprotected AES for all key bytes of 10 different random keys.

The results of all three methods let us conclude that the side-channel measurement performance of FBD-A7 is similar or better than that of the CW305.

5 CASE STUDIES

5.1 Side-channel Protection Evaluation of LWC Candidate Xoodyak

The effect of synchronizing the sampling and the target clocks on the number of traces needed for key recovery has been observed and discussed in the literature [15]. In this section, we show that performing sampling for leakage assessment using a synchronized clock is significantly more effective in leakage detection. In other words, using synchronized sampling and target clocks, one can detect leakage using significantly fewer traces than setups not using synchronized clocks.

In the following tests, we use a masked implementation of the NIST LWC finalist Xoodyak. This implementation is based on an unprotected design by the Xoodyak Team [24] and protected by Amir Moradi's group at Ruhr University Bochum with the help of the AGEMA tool [11]. While AGEMA is able to add protection to the datapath of the design, the control logic has to be manually modified. Our TVLA tests show leakage which was identified as an issue with providing random bits at some clock cycles, which was a bug in the control logic. It since has been fixed.

We performed TVLA tests on the masked implementation of Xoodyak using an external oscilloscope at a sampling rate of (test *A*) 1 GS/s with 8-bit resolution and (test *B*) 125 MS/s with 15-bit resolution. We repeated the same experiment, but this time, we used FOBOS 3 to capture traces at 50 MS/s with 10-bit resolution (test *C*). In the FOBOS 3 case, the ADC clock is synchronized with the target clock, while the external oscilloscope sampling clock is not. Table 2 shows the details of each experiment and the corresponding results. In all cases, we used the exact same Xoodyak implementation, which was instantiated in the NewAE C305 and ran at 10 MHz, and we used the same fixed-vs-random test vectors. Figure 10 shows the maximum t value in each experiment as a function of the number of traces processed. The red line marks the 4.5 thresholds with t values exceeding this threshold, indicating leakage detection. The figure shows that in test *C*, which uses the synchronous clock, the t values exceed the threshold after processing 1.3 million traces. For experiment *A*, which uses a much higher sample rate but less resolution, the test detects leakage after processing 1.6 million traces, while in experiment *B* which uses a higher sample rate **and** much higher resolution, the leakage is detected after 8.7 million traces are processed.

Test results from other teams for the above masked Xoodyak implementation and publicly available at [3] are listed as tests *D* and *E* in Table 2. No leakage was detected after processing 10 million traces in the two cases. In both tests, the clock and target clock are not synchronized. In the case of test *D*, the sampling rate is lower than in all of our experiments. The lack of clock synchronization combined with the lower sampling rate and resolution differentiates this test from test *C*. For test *E*, however, the sampling rate is similar to test *A*, but the hardware target used is different, so the comparison is not straightforward.

These results show that TVLA results should be taken cautiously since test parameters, particularly clock synchronization, significantly affect the test results. Additionally, using synchronized clocks can save the evaluation lab time since leakage is detected using fewer traces. At the same time, a lower sampling rate can be used, which translates to cheaper equipment.

5.2 Power and Energy Consumption of LWC Candidate Ascon

To show the power measurement capabilities of FOBOS, we provide an example of power measurement for the hardware reference implementation of the LWC winner Ascon [18]. We generated results for two different variants of this design - one lightweight design and one high performance design. The lightweight implementation, designated v1, requires 1,465 LUTs, has a maximum frequency of 191 MHz, and has a throughput of 1.53 MB/s. The high performance implementation, designated v5, requires 2,797 LUTs, has a maximum frequency of 150 MHz, and has a throughput of 2.4 MB/s. This experiment was performed using a CW305 target board as well as the FBD-A7 board discussed previously. Measurements were taken using both the XADC in the PYNQ's FPGA and an external oscilloscope with a sampling rate of 4 GSa/s and a resolution of 8 bits. Long messages with 8 KB of plaintext and 8 KB of associated data were taken to limit the impact of the transition period for the current consumption at the edges of the measurement. For the CW305 board, the power to the FPGA core is provided using the variable voltage output on the FOBOS Shield. For the FBD-A7 board, the current consumption for the FPGA core can be measured using the measurement circuit on the DUT.

First, we will discuss the comparison of results from the Shield versus the external oscilloscope. The traces were taken at the following frequencies: 5 MHz, 25 MHz, and 50 MHz. The average current values from both the Shield and oscilloscope are shown in Table 3. Power is calculated by taking the average current consumption and multiplying by the average voltage. Energy is determined using the calculated average power and latency, as well as using area under the curve for the oscilloscope trace. An example trace from

Table 2: TVLA results of masked Xoodyak implementation based on measurement setup, all ran for 10 Million traces

Lab	Scope	DUT	Resolution	Sync.	Sample Rate	DUT Freq.	Fails at Traces	Test
Ours	PicoScope 5244D	CW305	8 bit	No	1 GS/s	10 MHz	1.6 M	A
	PicoScope 5244D	CW305	15 bit	No	125 MS/s	10 MHz	8.7 M	B
	FOBOS 3	CW305	10 bit	Yes	50 MS/s	10 MHz	1.3 M	C
TU Graz [23]	PicoScope 6404C	CW305	8 bit	No	22 MS/s	1 MHz	Pass	D
Tsinghua [26]	LeCroy Wave Runner 8404M	SAKURA -X	8 bit	No	1 GS/s	6 MHz	Pass	E

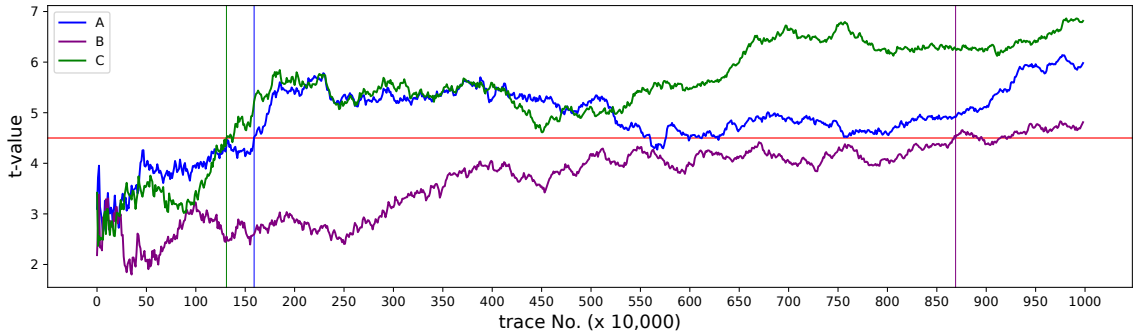

Figure 10: Maximum t-value vs. the number of processed traces. Vertical lines indicate the point at which the t-value exceeds the threshold in each test

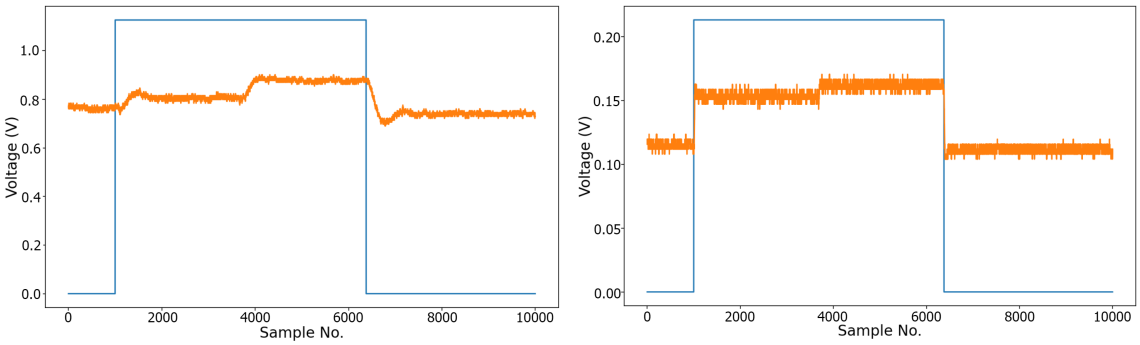

Figure 11: CSM voltage during Ascon-v1 encryption of 8 KB of AD and 8 KB of PT at 45 Mhz. Left: CW305. Right: FBD-A7.

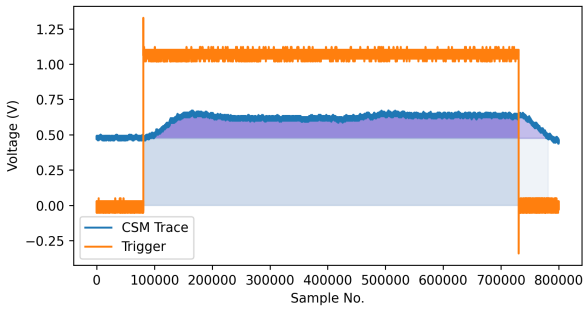

Figure 12: Example Power Trace

the external oscilloscope is shown in Fig. 12. The light shaded area represents the static energy, the dark area represents the dynamic energy. In the trace-aligned measurements, the power is calculated from the start of the trigger until the power returns to the static power level. The FOBOS measurements only support trigger-aligned measurements, and thus do not account for the discharge energy after the operation completes.

First, we note that the energy calculations using average power and area under the curve are near identical. We can also observe that there is a small, but noticeable, difference between the trace-aligned and trigger-aligned measurements. Even with the lower sample rate of the XADC, the average current is close to that of the external oscilloscope. As observed in Figure 11, the current consumption is smoothed out by the capacitance of the target system.

Table 3: Average Current Measurements of Ascon-v1 Encryption 8 KB of AD and 8 KB of PT Using XADC and Oscilloscope

Frequency	FOBOS (XADC)			Oscilloscope						
	Trace Average								Area Under Curve	
	Trigger-Aligned						Trace-Aligned			
	Power (mW)	Energy (μJ)	Num. Samples	Power (mW)	Energy (μJ)	Num. Samples	Power (mW)	Energy (μJ)	Energy (μJ)	Num. Samples
50MHz	33.63	109.32	446.6	32.886	106.87	1299933	32.887	106.88	106.87	1312390
25MHz	46.49	30.22	89.4	44.44	28.89	1299936	44.347	28.82	28.82	1350269
5MHz	61.45	19.97	44.8	59.39	19.3	699881	59.1	19.21	19.21	649971

Table 4: Power and Energy Results for FOBOS Artix-7 DUT and CW305 Targets

Target	Design	Frequency	Power (mW)		Energy (μJ)		Latency (ms)
			Measured	Estimated	Measured	Estimated	
CW305	Ascon-v1	5	29.99	62	97.46	201.44	3249
	Ascon-v1	45	45.62	76	16.47	27.44	361
	Ascon-v1	75	56.43	85	12.23	18.36	216
	Ascon-v5	5	33.00	93	53.68	151.22	1626
	Ascon-v5	45	86.18	208	15.58	37.44	180
	Ascon-v5	75	125.91	282	13.65	30.46	108
FBD-A7	Ascon-v1	5	11.19	87	36.37	282.66	3249
	Ascon-v1	45	26.23	102	9.47	36.82	361
	Ascon-v1	75	37.89	112	8.18	24.19	216
	Ascon-v5	5	15.65	121	25.46	196.75	1626
	Ascon-v5	45	66.15	221	11.96	39.78	180
	Ascon-v5	75	105.07	284	11.35	30.67	108

Thus instantaneous spikes in power are spread out allowing the lower sampling frequency of the FOBOS Shield to achieve similar accuracy to the oscilloscope with a much higher sampling rate. The FOBOS measurements are within 5% of the oscilloscope for all frequencies. However, we can observe that the number of samples for the 50 MHz design is already low. Thus for module with very low latency's, the Shield will not gather enough measurements to produce a meaningful result.

Next we will discuss the comparison of results between the CW305 and FBD-A7 DUT board. Since the FPGAs on these boards vary substantially in their size and speed grade, we do not expect similar power consumption from both devices. The traces for the 8 KB PT and AD on both target boards are shown in Figure 11. For power and energy results, we compare measurements taken using the external oscilloscope. We also generated power estimates using the power estimation feature of Vivado with inclusion of the post-implementation timing simulation signal activity.

Table 4 and Figure 13 show the power and energy results for each of the target boards. The relationship between power and frequency can be represented as a linear equation $p = a + b \cdot freq$ where a is the static power and $b \cdot freq$ is the dynamic power. For these two designs and platforms, the power is represented by the following equations:

$$p_{cw305\_v1} = 0.39 \cdot freq + 28.25\ mW$$
$$p_{cw305\_v5} = 1.33 \cdot freq + 26.4\ mW$$
$$p_{fbd-a7\_v1} = 0.38 \cdot freq + 9.22\ mW$$
$$p_{fbd-a7\_v5} = 1.28 \cdot freq + 9.1\ mW$$

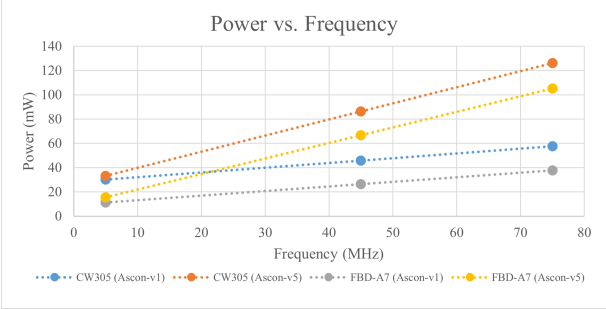

Figure 13: Comparison of Power Consumption at Different Frequencies for CW305 and FBD-A7

As expected, the smaller component has substantially lower power consumption. This difference is primarily due to the lower static power, which is only approximately 9 mW and 27 mW for the FBD-A7 and CW305 respectively. The dynamic power of both boards varies very little, which is expected since the number of state transitions is independent of the target and both are the same family of FPGA. We can also observe slight differences in the trace of the voltage across the CSM shown in Figure 11. While there are no capacitors in the V_{core} power circuitry of either board, the chip does have some internal capacitance. We can see that the CSM trace of the smaller chip used in the FBD-A7 board is less affected by this internal capacitance. We can also represent the estimated power as a function of the clock frequency as shown below. When comparing

the estimated results to the experimentally measured results, we can see the limitations of this power estimation tool. First, the static power is significantly higher than what was observed in our measurements and even shows the smaller device having higher static power consumption. We can also observe that the dynamic power estimate accuracy decreases with design complexity. The dynamic power for the smaller design is close to the experimental results, but for the larger design, it is substantially higher.

$$p_{cw305\_v1\_est} = 0.33 \cdot freq + 60.6\ mW$$
$$p_{cw305\_v5\_est} = 2.71 \cdot freq + 81.44\ mW$$
$$p_{fbd-a7\_v1\_est} = 0.36 \cdot freq + 85.41\ mW$$
$$p_{fbd-a7\_v5\_est} = 2.34 \cdot freq + 111.26\ mW$$

These results match the results of several previous works comparing estimates and measurements for older versions of Xilinx's and Intel's power estimation tools. One previous work had performed similar experiments with common cryptographic implementations such as AES and DES and found the error of the estimate varied between 17% and 200% [13]. Another work compared estimates and measurements for Spartan-6 FPGAs in which they compared results for several different types of algorithms, including AES, a Fast Fourier Transform, and multiple different implementations of a 32×32-bit multiplier. The estimation's error varied between -5% and 377% depending on the design and implementation language [16]. A similar work that compared estimates and measurements for various implementations of a 32×32 bit multiplier on a Cyclone-III FPGA found that the estimates were much more accurate, varying between -11.7% and 13% error [17].

6 CONCLUSIONS

This paper presents an open-source, multi-user platform for side-channel analysis and benchmarking. We introduced the FOBOS Shield, a measurement board, which in conjunction with a PYNQ-Z1 forms FOBOS Control and is capable of synchronized SCA measurement as well as power and energy consumption measurements. Additionally, we introduce a new low-cost FPGA target board and evaluate its performance. Our two case studies highlight the features of FOBOS 3 and its application to the evaluation and benchmarking of the finalists in the NIST LWC standardization process. All files required for building and running FOBOS 3 are provided under Apache 2 License (allowing commercial usage, requiring attribution) at https://github.com/GMUCERG/fobos. This includes HDL files (not including Xilinx IP) and Python for control, schematics and board design files for FOBOS Shield and FBD-A7, etc. Future work on this platform will be the support of the NewAE CW308 UFO board for targeting microcontrollers and the development of target boards for other popular FPGA families.

ACKNOWLEDGMENTS

Partially supported by the US Department of Commerce (NIST) (Grant no. 70NANB18H219).

REFERENCES

[1] Abubakr Abdulgadir, William Diehl, and Jens-Peter Kaps. 2019. An Open-Source Platform for Evaluation of Hardware Implementations of Lightweight Authenticated Ciphers. In *2019 International Conference on Reconfigurable Computing and FPGAs, ReConFig 2019, Cancun, Mexico*. https://doi.org/10.1109/ReConFig48160.2019.8994788

[2] Abubakr Abdulgadir, Jens-Peter Kaps, and Ahmad Salman. 2022. Enhancing Information Security Courses With Remotely Accessible Side-Channel Analysis Setup. In *Proceedings of the 2022 on Great Lakes Symposium on VLSI*. ACM, Irvine, CA. https://doi.org/10.1145/3526241.3530347

[3] George Mason University Cryptographic Engineering Research Group. 2022. Assignments, Commitments, and Reports of the LWC Side-Channel Security Evaluation Labs Targeting Hardware Implementations. https://cryptography.gmu.edu/athena/LWC/Lab_Implementation_Matching_HW.html.

[4] Christophe De Cannière and Bart Preneel. 2008. *Trivium*. Springer Berlin Heidelberg, Berlin, Heidelberg, 244–266. https://doi.org/10.1007/978-3-540-68351-3_18

[5] Gilbert Goodwill, Benjamin Jun, Josh Jaffe, and Pankaj Rohatgi. 2011. A Testing Methodology for Side-Channel Resistance Validation. In *NIST Non-invasive Attack Testing Workshop*. Nara, Japan.

[6] Hendra Guntur, Jun Ishii, and Akashi Satoh. 2014. Side-Channel AttacK User Reference Architecture Board SAKURA-G. In *2014 IEEE 3rd Global Conference on Consumer Electronics (GCCE)*. IEEE, Tokyo, Japan, 271–274.

[7] Ekawat Homsirikamol, William Diehl, Ahmed Ferozpuri, Farnoud Farahmand, Panasayya Yalla, Jens-Peter Kaps, and Kris Gaj. 2016. CAESAR Hardware API. Cryptology ePrint Archive, Report 2016/626. https://eprint.iacr.org/2016/626.

[8] Kengo Iokibe, Tomonubo Kan, and Yoshitaka Toyota. 2020. A Study on Evaluation Board Requirements for Assessing Vulnerability of Cryptographic Modules to Side-Channel Attacks. In *International Symposium on Electromagnetic Compatibility & Signal/Power Integrity (EMCSI)*. Reno, NV, USA.

[9] Jens-Peter Kaps, William Diehl, Michael Tempelmeier, Ekawat Homsirikamol, and Kris Gaj. 2022. Hardware API for Lightweight Cryptography v1.1 (with support for SCA-protected implementations). https://cryptography.gmu.edu/athena/LWC/LWC_HW_API_v1_1.pdf.

[10] Toshihiro Katashita, Akashi Satoh, Katsuya Kikuchi, Hiroshi Nakagawa, and Masahiro Aoyagi. 2010. Evaluation of DPA charactersitics of SASEBO for board level simulations. In *First International Workshop on constructive side channel analysis an secure design*. COSADE.

[11] David Knichel, Amir Moradi, Nicolai Müller, and Pascal Sasdrich. 2022. Automated Generation of Masked Hardware. *IACR Transactions on Cryptographic Hardware and Embedded Systems* 2022, 1 (2022), 589–629. https://doi.org/10.46586/tches.v2022.i1.589-629

[12] Paul Kocher, Joshua Jaffe, Benjamin Jun, and Pankaj Rohatgi. 2011. Introduction to Differential Power Analysis. *Journal of Cryptographic Engineering* 1, 1 (April 2011), 5–27. https://doi.org/10.1007/s13389-011-0006-y

[13] Dimitrios Meidanis, Konstantinos Georgopoulos, and Ioannis Papaefstathiou. 2011. FPGA Power Consumption Measurements and Estimations Under Different Implementation Parameters. In *2011 International Conference on Field-Programmable Technology*. https://doi.org/10.1109/FPT.2011.6132694

[14] NewAE. 2014. OpenADC Product Datasheet. http://newae.com/files/openadc-datasheet.pdf.

[15] Colin O'Flynn. 2017. *A Framework for Embedded Hardware Security Analysis*. Ph.D. Dissertation. Dalhousie University, Halifax, Nova Scotia.

[16] Juan P. Oliver, Julio Pérez Acle, and Eduardo Boemo. 2014. Power Estimations vs. Power Measurements in Spartan-6 Devices. In *2014 IX Southern Conference on Programmable Logic (SPL)*. https://doi.org/10.1109/SPL.2014.7002214

[17] Juan P. Oliver and Eduardo Boemo. 2011. Power Estimations vs. Power Measurements in Cyclone III Devices. In *2011 VII Southern Conference on Programmable Logic (SPL)*. IEEE, Cordoba, Argentina. https://doi.org/10.1109/SPL.2011.5782630

[18] Robert Primas. 2020. NIST LWC Hardware Reference Implementation of Ascon v1.2. https://github.com/ascon/ascon-hardware

[19] Rambus. 2019. DPA Workstation Analysis Platform - Rambus. https://www.rambus.com/security/dpa-countermeasures/dpa-workstation-platform/.

[20] Riscure. 2019. Side Channel Analysis Security Tools. https://www.riscure.com/security-tools/inspector-sca/.

[21] Ruhr-University Bochum. [n.d.]. Chair for Security Engineering GitHub. https://github.com/Chair-for-Security-Engineering/LWC-Masking.

[22] Tobias Schneider and Amir Moradi. 2015. *Leakage Assessment Methodology - a Clear Roadmap for Side-Channel Evaluations*. Cryptology ePrint Archive 2015/207.

[23] T. Steinbauer, R. Nagpal, R. Primas, and S. Mangard. 2022. *TVLA On Selected NIST LWC Finalists*. Technical Report.

[24] The Xoodyak Team. 2021. Unprotected Hardware Xoodyak Implantation. https://github.com/KeccakTeam/Xoodoo.

[25] Rajesh Velegalati and Jens-Peter Kaps. 2012. Introducing FOBOS: Flexible Open-source BOard for Side-channel analysis. Work in Progress (WiP), Third International Workshop on Constructive Side-Channel Analysis and Secure Design, COSADE 2012.

[26] C. Zhao, H. Zhao, W. Yang, W. Zhu, and L. Liu. 2022. *Leakage Assessment Report for Xoodyak R3 First Order*. Technical Report.