Cost effective techniques for chip delayering and in-situ depackaging

Philippe Loubet-Moundi

Gemalto, Security Hardware Lab, France
philippe.loubet-moundi@gemalto.com
http://www.gemalto.com

Abstract. *Invasive or semi-invasive attacks require, of course, because of their nature, the removal of metal layers or at least the package decapsulation of the chip. For many people - not expert in those sample preparation techniques - the simple access to the die surface and the observation of the chip structure after metal layers removal are the first obstacles to conduct an attack. In another direction, the development of embedded secure devices, sometime with very dense and complex assembly process, adds a new difficulty for an attacker to get a physical access to the silicon without intensive use of advanced soldering capabilities. This paper will deal with those two challenges: the first one is to provide an in-situ depackaging solution with limited ressources and then, the second one consists in finding the minimum mandatory tools required to perform chip delayering before metal layers imaging - or reverse engineering.*

Keywords: package removal, delayering, decapsulation, depackaging, reverse engineering, invasive attacks, semi-invasive attacks, sample preparation

1 Introduction

In recent years, the development of localized attack techniques is a fact [**?**]. As the global counter-measure strategies against global power analysis or optical glitch injections with bulb lamp [**?**] are less efficient the trend in the attacks development is to focus on the local leakage source or on the sensitive element of the circuit. In fault injection area, laser [**?**], UV light [**?**] [**?**], temperature [**?**], EM [**?**] [**?**] [**?**], or more recently FBI [**?**] require to have a direct access to the silicon surface. In side channel domain, photoemission techniques cannot be done without packaging removal [**?**]. Whereas EMA signal could be significantly better if the probe is directly in contact with the die [**?**].
In parallel, a prior knowledge of the chip layout is very useful to place the probe or the laser spot at the right location. Unfortunately, for current technology nodes, the dense metal routing on the top metal levels prevents any direct observation of the functional blocks with a microscope.
Because of this, chip delayering and package opening are the first steps in a localized attack realization.

Samples preparation techniques that are currently used by evaluation labs come directly from failure analysis domain. As a single defect in a semiconductor production line can have dramatic impact on chip manufacturers revenues. The tools developed must be reliable to find the smallest possible defect in an integrated circuit. Their price are usually in line with the return on investment that they could generate. In other words, the failure analysis equipments could reach a cost level not affordable for single hackers, accademic research security teams or even small evaluation labs.
In addition, the skills needed to perform sample preparation are a key point. Handling dangerous chemical acids or using very specific equipments is reserved to experts or at least to well trained operators.
Consequently, mathematicians, security software researchers - or hackers - that desire performing e.g. laser or EM investigations could hesitate before using acids on their new FPGA board or on the latest smartphone they are studying. They could also simply don't want or don't be able to waste money on costly dedicated equipments for their labs. This paper will provide some tips ans tricks that can be used as workaround because at the end, invasive or semi-invasive attacks are only performed after chip or package modifications.

Organization of the paper: After a short comparison of the basic techniques used in semiconductor samples preparation this paper will propose in the first part a low cost but reliable technique for in-situ depackaging, which deals with package decapsulation on assembled device. Then, in a second part a trial to achieve the minimum possible cost for chip delayering will be reported. Finally, the limitations compared to existing techniques and the impact on product security characterizations will be discussed in the conclusion.

2 Currently used techniques

2.1 Decapsulation and depackaging

The chip packaging domain proposes a huge number of package form factors and related options. For each application the right chip package must be used. Depending of the size of the die, the floorplan available, the temperature range, the soldering process, the price, the reliability, etc, the packaging of the same chip can be different. However, for decapsulation techniques, three main categories can be distinguished: the first one is metal package where the cover is side brazed or glued, the second one is the ceramic where the package is molded around the chip and the third one is the plastic where the package is also molded around the chip but with a different cheaper process. Exposing the chip in metal package requires removing the cover. This is done by applying mechanical constraints on the cover depending on the package size but this does not present any specific issues even for not specifically trained people. For ceramic and plastic package, the decapsulation techniques are most of the case based on chemical wet etching. Removing the entire package of a smart card, for example, can be done by dipping the entire package into fuming nitric acid (HNO_3)alone or mixed with

sulfuric acid (H_2SO_4). However, re-bonding the chip would be mandatory and decreases significantly the reliability of the sample preparation. So, in the rest of this paper, the decapsulation technique described will let the chip functional and in place within its package but also within the complete system board.
As described in some papers [**?**], manual opening techniques with hot fuming nitric acid or other etchants [**?**], [**?**] could be used efficiently on standalone plastic packages. The acid is dropped onto the compound until the chip surface is reached. Then, the chip is washed with acetone and isopropanol or ethanol. The fuming nitric acid does not directly attack neither the metal pads nor the gold or aluminum wires but corrosion can appear in wet environment. For more robust packages like ceramic or hard resins, a mix of nitric and sulfuric acids is prefered. In that case, semi-automated equipments minimize hazard for the operator. The acid or the acids mixture is pre-heated and send under pressure through a metallic cache window into the package during the required time. With such equipments [**?**], very high success rate close to 100% could be achieved.

However, for chips mounted in a high density printed circuit board, a complex desoldering operation is necessary to extract the chip from the board and puting it into the cabinet of the chemical etching equipment. In addition, for Ball Grid Array (BGA) packages, the re-balling operation with re-balling kits is always a risky and "funny" step. Finally, soldering the re-balled BGA chip with the open package in the previous board is not easy and not reliable without custom modifications of specialized soldering machine [**?**]. The in-situ decapsulating technique proposed hereafter in chapter **??** could help to open various package not only without compromising the integrity of the package itself but also without damaging the surrounding elements of the associated system board.

2.2 Chip delayering techniques

Accessing buried metal layers of a device can be mandatory for getting hidden useful information of the layout prior performing localized semi-invasive attacks. This is the first step for starting a partial or a full reverse engineering of the studied component [**?**]. Basic techniques for die deprocessing are usually classified in three main categories.
The first one is the wet etching methodology, several mixture of acids are used to selectively remove each layer of the chip process [**?**]. As silicon nitride or passivation are difficult to etch smoothly, very aggressive and hazardous acids must be used like Hydrofluoric (HF) acid. For selective metal etching or insulation materials different etchant preparations can be used depending of the chemical nature of the material. The main drawback of wet etching is the isotropy that makes the result of the removal of several metal layers very unpredictable. For acid and solvent manipulations and storage, an extraction hood, ventilated acids storage and individual safety protections are mandatory to avoid severe accident. But, for people who don't need to respect health and safety constraints - or who are not aware of the risk of chemical injuries -, the main advantage of this technique is the relatively low cost of the acids needed - some tens of Euros.

The second techniques used are the dry etching. The most commonly used is the plasma etching and more precisely the Reactive Ion Etching (RIE) [**?**]. The interaction between the ions present in the plasma and the surface of the targetted device is accelerated by an electric polarization of the vacuum chamber where the targeted device is exposed. As different gases are usually available, the user can select the right mix to obtain a selective etching of materials. With focused plasma, the results obtained can be very good but artifacts (fig: **??**) or over etching of the same layer (fig: **??**) appear often for deep buried layers. The main drawbacks of this equipment are the usage of dangerous gases (chlorine or fluorine based), the cost which is high (more than $200K$ Euros) and the relatively low speed of the preparation.

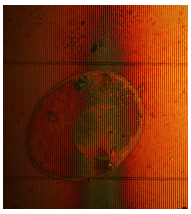

Fig. 1. RIE artefact over NVM after several metal layers etching, *1000x, optical microscope*

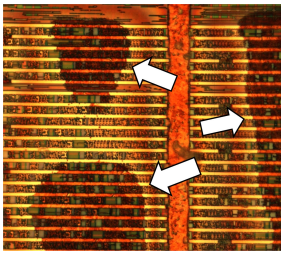

Fig. 2. plasma over etching on Metal 1, *500x, optical microscope*

The third technique is the parallel polishing technique. In that case, a mechanical etching is done by performing the grinding and lapping of the metal layers with rotating plates. A set of abrasive papers of micro sand grains of different sizes or diamond paste on glass plate are needed for the die polishing. The results obtained with this technique are strongly linked to the skill of the person who is performing the operation. The parallelism must be controlled and adjusted frequently. However, very good sample preparations could be obtained if it is performed properly and slowly.

Obviously, someone who master all the techniques above could efficiently combines them to achieve best in class delayered die samples.

Technique	Wet etching	Dry etching	Mechanical Polishing
Equipment cost	−	++	+
Duration	small	long	long
Operator dependency	reduced	reduced	high

Table 1. Summary of techniques vs cost, duration and operator dependency

3 Low cost techniques improvements

3.1 In-situ decapsulating

Purpose The usual depackaging techniques presented in the previous sections can be applied to packaged chips which are not soldered onto the application circuit board. If smart cards are exceptions, most of the other security chips are usually assembled in more complex and expensive devices. For BGA or tiny packages, the removal of the circuit from the printed circuit - prior to perform the decapsulation - could be very risky. The success rate of the attack could be dramatically reduced due to desoldering and re-soldering steps. The chip itself, the surrounding components or the printed circuit board could be damaged during these operations.
For that reasons, it should be more interesting to prepare the circuit directly in-situ without removing it from the board. So, the following work proposes one reliable methodology where the secure element remains soldered to the board during the depackaging. As the cost of the attacked device could be significant, or, as the number of samples could restricted, the technique must be very reliable to prevent any alteration of the complete system.
Depending on the orientation of the chip within the package, the backside or the top surface can be exposed.

Receipe The following methodology have been applied to the typical electronic devices shown in the pictures below.

- **1st step**
  The surface of the identified chip package is mechanically grinding (sharp knife, local polishing) to increase the roughness of the material and to increase the contact surface with the acid.

- **2nd step**
  The entire board is protected against acid exposure. A metallic adhesive tape is applied very carefully around the printed circuit board. This is definitively the most critical step and the junctions between the different pieces of adhesive tapes must be sealed perfectly. Several layers could be applied to avoid any bad surprise.
- **3rd step**
  The adhesive shield previously build is opened with a sharp knife to expose the area that will be etched. The opened window must be smaller than the package itself - ideally just smaller than the the size of the die.
- **4rd step**
  The acid (i.e. hot fuming nitric acid) is then dispensed by drops above the window. It should be possible to moderately pre-heating - without damaging the card - the targeted device with hot plate or hot air to accelerate the chemical reaction and minimize the exposure time with the aggressive acid. The acid temperature is chosen depending of the package material density. Hard resin needs high temperature around 70^oC whereas standard plastic devices could be etched efficiently with an acid at 40^oC.
- **5th step**
  The device is rinsed correctly to stop the acid reaction. Water is recommended but acetone followed by alcoholic solutions could be used. In both cases a generous final water rinse is strongly recommended.
- **6th step**
  The adhesive tape is then removed cautiously by taking a specific care of the chip opened and of the exposed wire bonding. If leakages were present during acid exposure or if solvent contaminations are visible after adhesive removing a new rinse with water could be done. In that case drying the entire board with pulsed nitrogen or clean air must be done during several minutes to avoid any further corrosion due to remaining moisture.

Decapsulation Results The methodology presented was applied on several electronic devices like USB keys, SD and micro SD cards, secure tokens, mobile phones.
The **first** (fig: **??**) and the **second** (fig: **??**) trials were performed on USB storage devices where a soldered element controls the access to the flash memory. After successful etching, the surface of the chip can be exposed and, if needed, further attacks could be done.

The **third** (fig: **??**) example shows a more complex device with several microcontrollers. Two of them have been exposed. For that test achievement, two windows were opened in the metallic tape and the acid was dropped on the two chips during the same experiment as shielding the device a second time could be more tricky to avoid damaging the wire bonding of the chip opened in first.

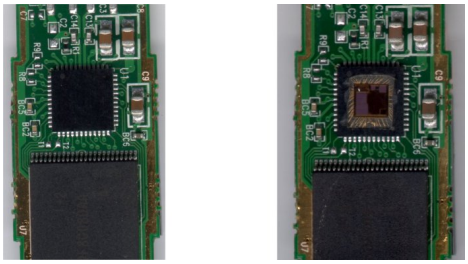

Fig. 3. 1st USB key before and after etching

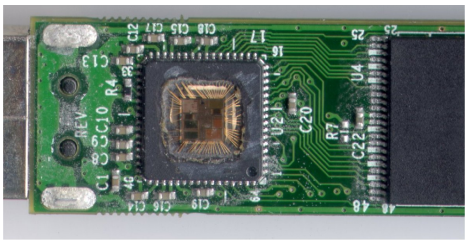

Fig. 4. 2nd USB device opened

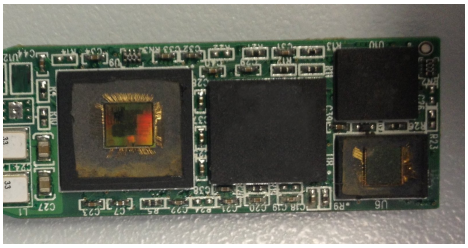

Fig. 5. Several packages opening on the same device

The goal of the **fourth** example (fig: **??**) is to validate the different process steps on another form factor. The pictures show how it is possible to perform exactly the same operation on a mobile phone circuit board.

Limitations and discussion Even if the trials were not performed on a significant numbers of the same device - mainly due to the assumptions: costly or limited number of samples - the reliability was very good as no device was destroyed during those tests.
However, the natural end of life of the opened circuit will be reduced by this decapsulating technique. Hot nitric fuming acid is difficult to rinse perfectly and corrosion appears frequently after a certain time on exposed bonding pads. That's why the best trade off between acid temperature and etching time must be found. For very brittle devices, it is recommended to heat the device around

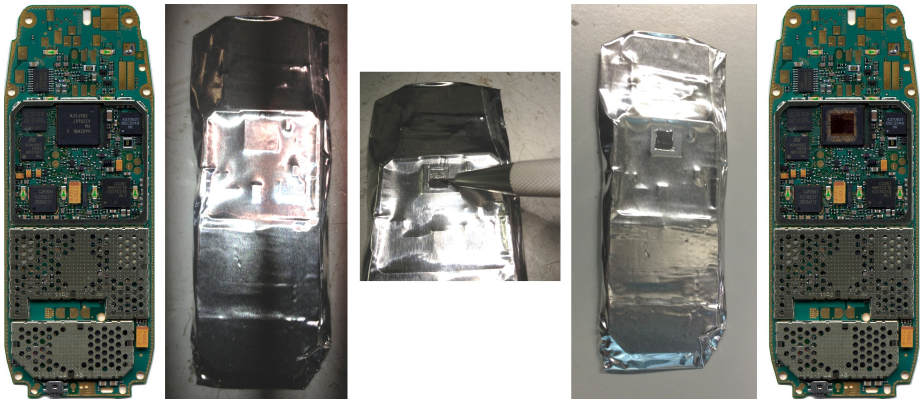

Fig. 6. in-situ depackaging applied on mobile phone circuit board

the acid temperature or to use ambient temperature for the etchant to avoid mechanical constraint due to thermal gradient during the operation.
Minimizing the etching time could also be realized by performing chip package thining with mechanical tools or ablation lasers.
The presented solution can be used for many devices without dimensions and shapes limitations if the attacked package can be dissolved by standard acids (nitric, sulfuric).

3.2 Ultra low cost chip delayering

Purpose Once the chip has been decapsulated, deeper semi-invasive investigations can be done in side-channel analysis or in fault attack injections. Delayering techniques could be used, as well, for chip reverse engineering or chip functional blocks analysis. The purpose is to have access to buried layers that are covered by security active shields or power routing shapes. As describe in the first part of this paper, those techniques usually requires dedicated expensive equipments or at least hazardous acids. The goal of the following tests is to measure and evaluate the capability to retreive hidden chip information with the minimum required equipment.

Methodology description This technique is inspired from parallel polishing with rotating plates with diamonds paste. The main difference is that the rotating plate machine which cost around $20K$ Euros is replaced by manual movements. A massive flat granite stone - from an anti-vibration stage - is used as polishing material. The die surface is placed in contact with the stone and the polishing movement was generated by applying smooth pressure with a finger on the backside and by drawing some small circles. This manual polishing was

monitored frequently using an optical microscope. The results are presented in the pictures below.

Results The trials were performed on a 130nm secure chip previously depackaged with nitric acid. A security protection against probing - active shield - hides the deeper metal layers from observation. Without removing the top metal layers, the localization of the different functional blocks like the memories, the logic part or the analogue circuitry is impossible to achieve.

The pictures below were taken regularly for progression monitoring. The time elapsed between the first and the last picture is around 1 hour.

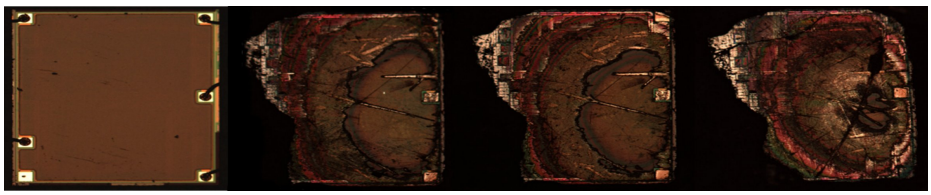

Fig. 7. Results of 1 hour of manual polishing, *low magnification*

At low magnification, the results seems awful compared to others methodologies. The parallelism is not preserved and the edges of the chip are etched faster than the middle. Small silicon particles fall from the edge of the die during the process and damaged the rest of the surface with scratches. However, if higher magnifications of the microscope are used, the result looks better.

The magnification level used for inspection is selected to have the smaller but the better view of local parts of the die. Three areas were reported here after: the logic part, the NVM memory and the RAM memory.

- **Logic area**
  The microscope observation of local areas of the logic part of the circuit is surprisingly good. For the buried metal layers like M2 and M1, the quality could be sufficient to perform cell layout analysis of some small portions of the circuit.
- **NVM area**
  In the sample studied, the NVM represents a huge part of the chip area. This is quite interesting, as the polishing is not perfectly parallel, different metal layers can be observed at the same time. By consequence, the main drawback is that there is no chance to take a clean picture of the entire NVM grinded at the same metal level.

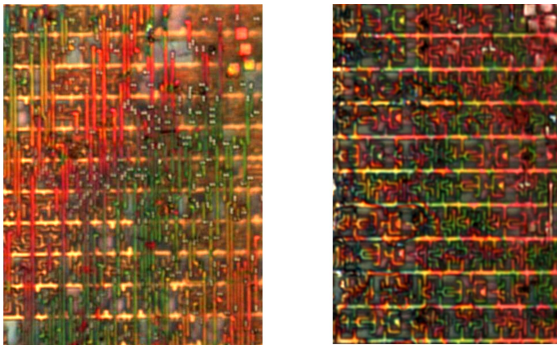

Fig. 8. Metal 2 (left) and Metal 1 (right) on the logic block of the chip

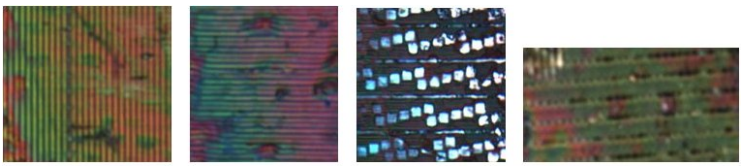

Fig. 9. Metal 4, Metal 3, Metal 2 and polysilicon levels on NVM area

– **RAM area**

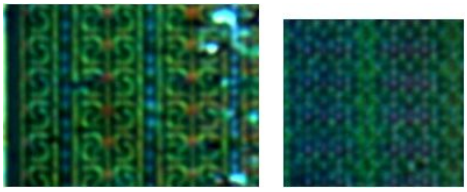

Fig. 10. Metal 1 and polysilicon level on RAM array

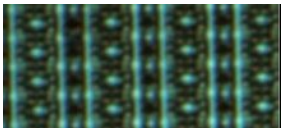

Fig. 11. RAM cell at polysilicon level with after wet etching process (HF)

The last levels reached in delayering are the polysilicon or the substrate implantation levels. The result obtained with the technique described in this paper (fig: **??**) is compared with sample preparation performed with Hydrofluoric acid (HF) wet etching (fig: **??**) on RAM area. The quality between the both techniques is similar but the low cost mechanical sample preparation allows to have also access to metal 1 (M1) which is not possible with basic HF layers removal.

Discussion The previous technique is the ultimate low cost technique as only a flat hard surface is needed. Of course, the results have a kind of randomness but they are sufficient to extract useful information like the kind of memories used, the location of the logic or the analogue areas... High magnification pictures - where planarity of the entire chip is not mandatory - are in some case better than wet etching. Obviously, without any improvement, this technique will not be used by people who have better capabilities. But, this feasibility study could really be considered as an alternative for people who do not have skills or equipments in failure analysis and who want to start chip investigations with the minimum of budget.

4 Conclusion

In the trend of localized attacks, the present contribution shows that - even with limited ressources - chip surface or hidden data in the die could be accessible. This paper has introduced a useful in-situ decapsulation technique for accessing chip surface without performing desoldering operations. By consequence, soldered secure elements should not be considered more secure or more complicated to attack compared to standlone package. In addition, this paper has also demonstrated that low cost approaches could be developped for chip delayering even by people that do not have a background in failure analysis for the semiconductors.

Fortunately, for the secure device provider - or unfortunately for the attacker - even with the results presented in this paper, performing semi-invasive or invasive attacks without expensive equipment will add more risks to damage the sample and the results obtained will remains far away from the state of the art.

However, the difficulty to perform satisfactory samples preparation will continue to be underestimated or not really taken into account during attack quotations. This can be explained easily, because most of evaluation labs can easily use expensive failure analysis equipments. In addition, for those experts who are performing daily sample preparations, accessing the chip surface, or performing reliable metal layer removal is an obvious and easy task. The massive introduction of copper and low-k insulation materials for advanced semiconductor technology nodes or the development of 3D integration process will probably provide them more challenging perspectives. Then, sample preparation techniques will

may be better considered during attack quotations. But for the time being, the security of the product will be the same, whatever the package used.

Remark: The related presentation of this paper was done during the short-tracks session of COSADE 2013.

References

1. https://berlin.ccc.de/wiki/Experiment:_IC-Entkapselung_mit_Kolophonium
2. http://www.ultratecusa.com/decapsulation
3. http://www.metcal.com/aprseries/
4. http://www.oxford-instruments.com/products/etching-deposition-and-growth/processes/etching-processes/organics/pi-etch
5. Anderson, R., Kuhn, M.: Low cost attacks on tamper-resistant devices. In: Security Protocols 5th International Workshop. pp. 125–136 (1997)
6. Beck, F.: Integrated circuit failure analysis: a guide to preparation techniques. Wiley series in quality and reliability engineering, Wiley (1998)
7. Dehbaoui, A., Dutertre, J.M., Robisson, B., Tria, A.: Electromagnetic transient faults injection on a hardware and software implementation of aes. In: FDTC proceedings. pp. 7–15. IEEE-CPS (2012)
8. Fournier, J., Loubet-Moundi, P.: Memory address scrambling reveals using fault attacks. In: Fault Diagnostic and Tolerance in Cryptography. pp. 30–36. FDTC '10 (2010)
9. Gandolfi, K., Mourtel, C., Olivier, F.: Electromagnetic analysis: Concrete results. In: CHES proceedings. Springer-Verlag (2001)
10. Laackmann, P., Janke, M.: Uncaging microchips. In: 31C3 (2013)
11. Maurine, P.: Electromagnetic transient faults injection on a hardware and software implementation of aes. In: FDTC proceedings. pp. 7–15. IEEE-CPS (2012)
12. Nohl, K., Evans, D., Starbug, Plotz, H.: Reverse-engineering a cryptographic rfid tag. In: USENIX Security Symposium (2008)
13. Poucheret, F., K.Tobich, Lisart, M., B.Robisson, Chusseau, L., Maurine, P.: Local and direct electromagnetic injection of power into cmos integrated circuits. In: FDTC'11 proceedings. pp. 100–104 (2011)
14. Schloesser, A., Nedospasov, D., Kraemer, J., Orlic, S., Seifert, J.P.: Simple photonic emission analysis of aes. In: CHES proceedings. pp. 41–57. Springer (2012)
15. Schmidt, J.M., Hutter, M., Plos, T.: Optical fault attacks on aes: A threat in violet. In: Naccache, D., Oswald, E. (eds.) 6th Workshop on Fault Diagnosis and Tolerance in Cryptography - FDTC 2009. pp. 13 – 22. IEEE-CS Press (2009)
16. Skorobogatov, S.: Semi-invasive attacks - a new approach to hardware security analysis,' university of cambridge, technical report 630. Tech. rep. (2005)
17. Skorobogatov, S.: Flash memory 'bumping' attacks. In: Proceedings of the 12th international conference on Cryptographic hardware and embedded systems. pp. 158–172. CHES'10, Springer-Verlag, Berlin, Heidelberg (2010)
18. Skorobogatov, S.P., Anderson, R.J.: Optical fault induction attacks. In: CHES proceedings. pp. 2–12. Springer-Verlag (2002)
19. Skorobogatovi, S.: Local heating attacks on flash memory devices. In: Proceedings of the 2nd IEEE International Workshop on Hardware-Oriented Security and Trust. HOST 2009 (2009)
20. Tobich, K., Maurine, P., Liardet, P.Y., Ordas, T.: Yet another fault injection technique by forward bias injection. Tech. rep. (2012)