2019 22nd Euromicro Conference on Digital System Design (DSD)

CMOS Illumination Discloses Processed Data

Jan Bělohoubek, Petr Fišer, Jan Schmidt
Faculty of Information Technology
Czech Technical University in Prague
Prague, Czech Republic
{jan.belohoubek, petr.fiser, jan.schmidt}@fit.cvut.cz

***Abstract*—As digital devices penetrate to many areas important for the present society, it is important to analyze even potential threats to mitigate vulnerabilities during their lifetime. In this paper, we analyze the data dependency of the photocurrent induced by a laser beam in the illuminated CMOS circuit. The data dependency may introduce potential threat(s) originating in the nature of the CMOS technology. The data dependency can be potentially misused to compromise the data processed by an embedded device. We show that also the devices employing dual-rail encoding to hide data-dependency are not safe.**

I. Introduction

In past years, the research of the security aspect of digital design and manufacturing processes itself attracted many researchers leading to a substantial scientific effort [1], [2], [3]. The research of weaknesses embedded into the digital design during the design and manufacturing phase is indeed very important. As digital devices continue to provide more vital services to our society [4], hidden security pitfalls may lead to significant damages [2], [3]. Especially, embedded device(s) security influences property and privacy protection of any being – a part of the present society [2], [5].

Attacks to embedded devices are often classified as *non-invasive*, *semi-invasive*, and *invasive* [6]. Any attack class considers different attacker(-strength) model, knowledge level, and available resources. A non-invasive attack, like *differential power analysis (DPA)* [7], [8], [9] requires only the knowledge of the algorithm, while invasive attacks may require even the detailed circuit layout knowledge [10].

Designers fighting with potential security threats often employ so-called *attack countermeasures* to make potential attacks more difficult. When concentrating on invasive *fault injection* attacks, the technology node itself may be understood as an attack countermeasure [11]. Behind this approach, there is a simple idea: "small parts (e.g. flip-flops) are hard to target" – in other words: it is hard, or even impossible, to induce a non-random fault into a specific location, if this location is too small [11].

Skorobogatov has shown, that reading individual bits from digital devices is possible when transistor sizes are large enough to allow precise laser beam localization on a single transistor [10], [11]. In the same paper, Skorobogatov argues that using deep-submicron technology, with very small transistors, is de-facto a countermeasure.

In [12], we described how to overcome the mentioned size-limitation in a special case: mounting an attack to a *combinational circuit*, namely the *majority voter*, which serves as an amplifier of a single bit to the power consumption side-channel. Such side-channel emission may lead to disclosing that bit. In this paper, we continue beyond to more general principles behind consequences described in [12].

In this paper, we describe data dependency of static current. Normally, it is very small and therefore hidden in noise. Yet, [11] showed that it can be modulated or amplified by light, in our case, a laser beam targeted to the combinational part of the CMOS circuit. In particular, the power consumption imprint of the logic gate illuminated by a laser beam depends on laser beam parameters and *the values at the gate inputs*.

The data dependency of the static current leads to a general method, which can be used to (partially) disclose input values in any combinational circuit, which is a part of a bigger system. This can be an issue for systems operating with sensitive data, e.g. crypto-units or security enclaves. We use SPICE simulation to reallistically demonstrate that combinational logic can be exploited to significantly decrease the *entropy* of the value processed by the *circuit under attack*.

The presented method, which can potentially compromise the security of digital circuits, is actually a *combined attack* [13]: an *optical attack* is combined with *(simple) power analysis* [6]. A precise control over a laser beam location is required: if the laser beam is targeted on the combinational logic, it induces a data-dependent current flow. The requirement of a precise laser beam location control may appear strong, however, there is a long history of using lasers for diagnostic purposes in digital design [14]. When the logic selected as the target of illumination is large enough, the targeting is completely possible and proved [1], [10], [15], [16], [17].

The rest of the paper is structured as follows: in Section II, the data dependency of the static current in common CMOS gates is described; in Section III, the photoelectric laser simulation is briefly described; in Section IV, the data-dependent behavior of CMOS gates illuminated by a laser beam is described; in Section V, we explain, why dual-rail encoding cannot be used as a vital countermeasure; Section VI describes how the presented method reduces entropy of data processed by embedded device; in Section VII, the discussion is provided and Section VIII summarizes the paper. Appendix A contains notes related to the simulation and experiment replicability.

II. CMOS Gate Power Analysis

A *static CMOS logic gate* in general has an *NMOS pull-down transistor network* (N) and a *PMOS pull-up transistor*

978-1-7281-2862-7/19/$31.00 ©2019 IEEE
DOI 10.1109/DSD.2019.00062

network (P) [18], as illustrated in Figure 1. The P/N parts arrangement ensures, that only one of both parts is ON and the other is OFF for any combination of input values.

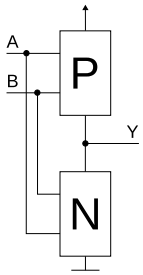

Fig. 1. Generalized (2-input) CMOS gate structure

The static CMOS gate requires a significant amount of energy to change its state – it is called *dynamic power*. The dynamic power is expressed by the (integral) equation [18]:

$$P_d = C \cdot V_{dd}^2 \cdot f, \tag{1}$$

where V_{dd} is the *supply voltage*, f is the *switching frequency*, and C is the *load capacitance* being charged/discharged. The load capacitance of a single gate is only charged/discharged when the output of the gate changes from $0 \rightarrow 1$ or $1 \rightarrow 0$. The dynamic power of the whole circuit, is, together with statistics employed in passive attack schemes like DPA, used to disclose a secret value: the widely used Hamming distance model [7], [9] in fact reflects the switched load capacitance.

The other component of CMOS gate power consumption, called *static power*, was addressed by several works since 2007 [19], [20]. The *Leakage Power Analysis attacks* (LPA) were introduced, but the *static power* security effects still remain on the edge of interest up today. The static power can be expressed by the following (integral) equation [18]:

$$P_s = I_s \cdot V_{dd}, \tag{2}$$

where V_{dd} is the *supply voltage* and I_s is the *static current*. It is not surprising, that I_s for a particular gate depends on many variables including manufacturing *process parameters* and variability, logic gate *geometry* (parallel vs. serial connection of transistors), and size.

One of the important parameters influencing the static current are logic values at the gate inputs – the *gate input pattern*. We simulated the static current for different input patterns for three standard CMOS gates [18] (2-input `NAND`, `NOR`, and `XOR` gates), namely the `NAND2X1`, `NOR2X1` and `XOR2X1` standard cell netlists were simulated – see the cell layouts in Figure 2. The results are shown in Figure 3.

Although the absolute difference of min/max currents for a single gate is very small, namely the maximum observed difference is below 7 nA for all simulated two-input gates, **data dependency is observable**.

To disclose the influence of the cell geometry and NMOS/PMOS transistors properties, we performed an additional simulation of transistor structures, which are not met in

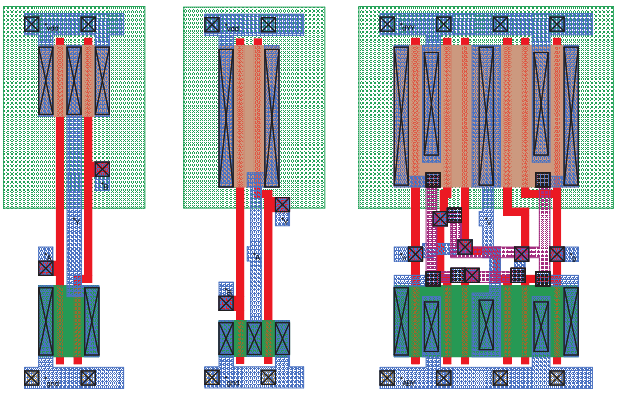

Fig. 2. Layout of `NAND2X1` (4 x 10.8 μm), `NOR2X1` (4 x 10.8 μm) and `XOR2X1` (7.2 x 10.8 μm) cells in 180nm TSMC technology

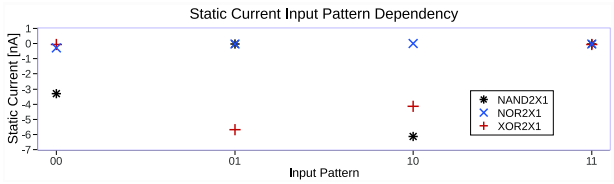

Fig. 3. The data dependency of the static current on the input pattern for three standard cell SPICE models – namely `NAND2X1`, `NOR2X1`, and `XOR2X1` – in 180nm TSMC technology

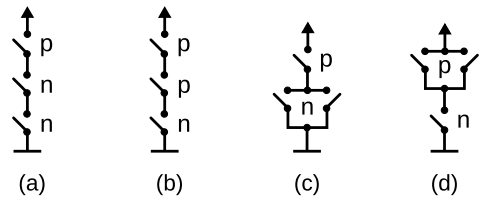

Fig. 4. Simulated transistor structures displayed as serial/parallel switches. Note, that these structures are not included in the standard cell library: SPICE models were derived from standard cells with equal geometry of NMOS/PMOS parts

static CMOS gates[1] The *structures of interest* are parallel/serial combinations of NMOS (n) and PMOS (p) transistors. The simulated structures are shown in Figure 4. The static current for all non-short "switch" states was obtained by simulation.

Simulations of the *switch* structures in Figure 4 have shown, that:

(i) in structures (a) and (c), the static current is influenced (mainly) by the state of the PMOS transistor – the dependency on any combination of NMOSes is distinctly less significant;

(ii) for NMOSes, it holds, that the serial structure (a) introduces a stronger data dependency than the parallel structure (c). Naturally, there is a difference in the static current when opening (only) the *top* or (only) the *bottom* transistor because of induced drain/source voltage differences;

(iii) for structures (b) and (d), we observed very little (almost none) data dependency on the (single) NMOS transistor state;

[1]Note, that in published SPICE models, the "switch" models are denoted as dynamic gates, as their layouts correspond to dynamic gates.

(iv) for PMOSes, in contrast to NMOSes, the parallel structure (d) introduces (a bit) stronger data dependency than the serial structure (b). This is apparently caused by lower hole mobility in PMOSes.

As static CMOS gates combine parallel/serial structures exercised in the previous paragraph, the resulting behavior is a "cocktail" of behaviors described above.

Taking the behavior simulations into account, the results provided in Figure 3 may be interpreted as follows:

(1) the `NAND2X1` cell is the most asymmetric one, which is caused by the NMOS serial and PMOS parallel arrangement;

(2) the `XOR2X1` cell power imprint allows a clear distinction of the XOR output state, which demonstrates the symmetry of the XOR gate;

(3) the `NOR2X1` power imprint is very narrow compared to the other two gates, which is apparently caused by the serial arrangement of PMOSes (introducing low conductivity when one of them is closed).

As the static current differences are very small, **evaluation of the static power is** very **impractical** in reality: when taking any complex digital circuit as an example, the static current can only be measured for a group of hundreds or even thousands of gates at once. This *cocktail effect* in connection with measurement inaccuracy and manufacturing variances (whose were not included in SPICE simulations and may change the static currents [21]) results in a situation, where finding a correlation between the circuit static power and data at any real and useful circuit inputs is almost impossible.

III. CMOS Photoelectric Laser Stimulation

In the next sections, we demonstrate how the static power consumption of a specified chip area can be modulated by a laser beam to increase the visibility of the useful information in the side-channel. First of all, the photoelectric Laser Stimulation will be briefly described.

The principle behind the *Photoelectric Laser Stimulation (PLS)* of a specified silicon chip area is based on the *photoelectric effect*. The laser beam passing through silicon creates, as a result of energy absorption, electron-hole pairs along its path. In *Space Charge Regions (SCR)* of PN junctions, the generated electron-hole pairs are separated by the internal electric field, and thus an *Optical Beam Induced Current (OBIC)* is generated [15], [22].

To perform an accurate electrical simulation of this process, accurate models of the transistor under laser stimulation are required. Sarafianos et al. published a series of papers related to Photoelectric Laser Stimulation (PLS), incrementally describing the electrical model of the pulsed photoelectric laser stimulation of an NMOS and PMOS respectively, e.g. [15], [16], [17]. Based on their work, we compiled the SPICE models, which are mounted on a public technology node (see Section A). The models are described in [12] and published online for better experiments replicability[2].

[2]http://ddd.fit.cvut.cz/prj/CMOS-PLS

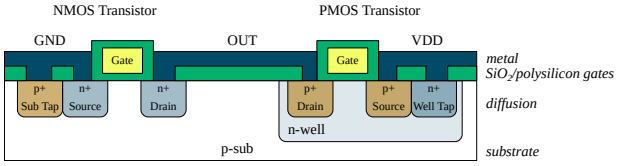

Fig. 5. CMOS cross-section showing the modeled PN junctions

For better idea about the PN junction-related photocurrent in CMOS, the modeled PN junctions – *p+/n-well*, *n+/p-sub* and *p-sub/n-well* – are shown in Figure 5 (for CMOS technology details refer to, e.g. [18]).

IV. Data-Dependent Power Imprint of CMOS Gates Under PLS

In Section II, we have shown, that the static power of a single gate is data-dependent. However, this dependency cannot be used to obtain useful data from the circuit, because of low measurement resolution and the *cocktail effect*, which mixes static currents of thousands of gates in one circuit together.

In this section, we demonstrate how the static power consumption of a single gate can be modulated by a laser beam to obtain useful information. This might be a security pitfall hidden in many designs.

The idea behind the static power analysis of the illuminated gate is as follows: if a single gate (standard cell) is illuminated, its data-dependent static power consumption is amplified, and thus made visible in the power trace of the whole circuit (refer to *simple power analysis – SPA*).

Figure 6 shows the static current consumption of three standard cells under PLS for different input patterns, as simulated in SPICE with 50mW laser power.

Note, that in our experiments, we omitted n-well/p-sub PN-junction currents, which are not data dependent (but only structure dependent). Thus, the currents presented in this paper are in fact lower than currents observable in reality (when illuminating a real CMOS circuit). Only data-dependent currents were simulated, thus the data-dependent differences remain equal.

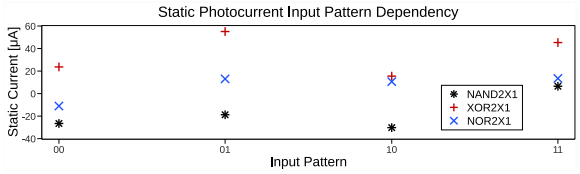

Fig. 6. The dependency of photocurrent, induced by 50mW laser beam, on input patterns for three standard cell SPICE models– namely `NAND2X1`, `NOR2X1` and `XOR2X1` – in 180nm TSMC technology

Compared to Figure 3, Figure 6 shows a significant difference in data-dependent power consumption: the laser beam provides the static current amplification, moving current differences from nanoamps in Figure 3 to microamps in Figure 6.

The `XOR2X1` cell is bigger than the other two gates causing a higher current in the area of the cell.

For both – NOR2X1 and NAND2X1 cells, it is possible to distinguish gate output values from the power traces (for NAND2X1: 00, 01, 10 $\rightarrow$ 1 and 11 $\rightarrow$ 0; and for NOR2X1: 01, 10, 11 $\rightarrow$ 0 and 00 $\rightarrow$ 1). Additionally, for XOR2X1 and NAND2X1, the input patterns can be distinguished clearly – the differences are above 8 μA.

The difference will be more significant for higher laser powers, as demonstrated in Figure 7 for the NAND2X1 cell and in Figure 8 for the NOR2X1 cell. For XOR2X1, the *cocktail effect* leads to a bit different situation, as displayed in Figure 9, however the dependency on the laser power is evident. For completeness, the data dependency of the single-input inverter (INVX1) is show in Figure 10. The data dependency of the inverter shows the analogy with the NAND2X1 gate. Bigger negative current component of NAND2X1 for 00 (compared to 0 in INV2X1) is given by the parallel composition of PMOSes in NAND2X1.

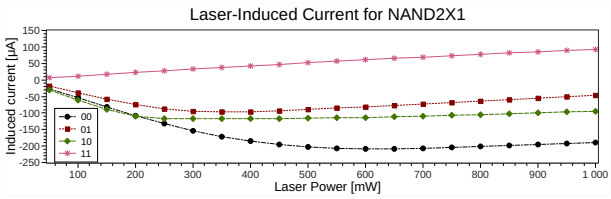

Fig. 7. The photocurrent for NAND2X1 for different input patters and increasing laser power. The 00 and 11 input patterns are easy to distinguish; patterns 01 and 10 cause similar currents, although the 20μA difference (for 100mW and above) is still distinguishable

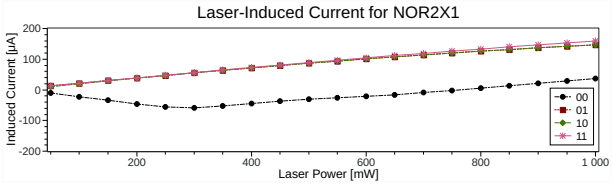

Fig. 8. The photocurrent for NOR2X1 for different input patters and increasing laser power. The (00) and (11, 01, 10) input pattern subsets are easy to distinguish

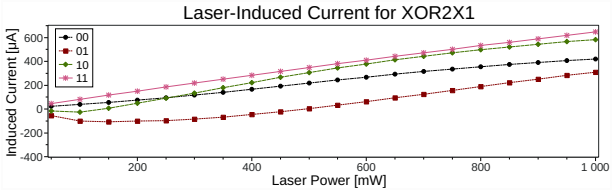

Fig. 9. The photocurrent for XOR2X1 for different input patters and increasing laser power. The 00 and 11, 01 and 10 input patterns can be distinguished

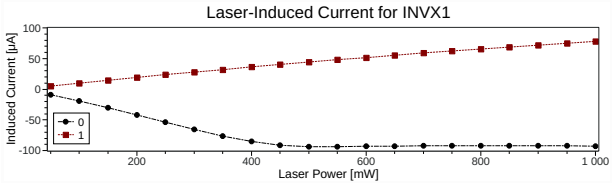

Fig. 10. The photocurrent for INVX1 for different input patters and increasing laser power. The 0 and 1 input patterns are easy to distinguish

Previous paragraphs summarized how the static current of a single gate can be modulated by a laser to disclose the state of a single gate. Let us expect that such a gate is a part of a bigger circuit and the task is to obtain the gate input pattern. To accomplish this task, several requirements must be met:

(i) the gate must be big enough to allow precise laser beam targeting to the gate area only;

(ii) it will be helpful, if the other circuit activity will be inhibited, e.g. by holding its clock signal stable;

If both requirements are met, the (static) consumption of a single gate under (laser beam) illumination will be distinguishable in the power trace of the whole circuit.

The advantage of the described method is that the induced photocurrent value can be modulated by the laser power: increasing the laser power increases the current – see Figure 7 for results related to the NAND2X1 cell. The current difference saturation for laser powers above 500mW (and 300mW respectively) is well distinguishable in Figures 7 and 8 and 9.

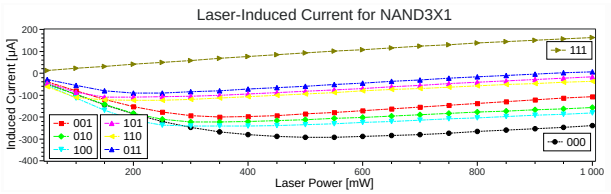

Fig. 11. The photocurrent for NAND3X1 for different input patterns and increasing laser power. The four sets of input patterns are easy to distinguish: these sets of input patterns distinguished by the Hamming Weight (HW): 000 with HW(0); 001, 010 and 100 with HW(1); 011, 101 and 110 with HW(2) and 111 with HW(3)

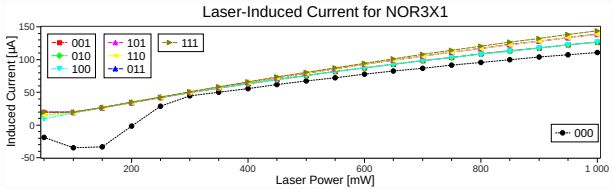

Fig. 12. The photocurrent for NOR3X1 for different input patterns and increasing laser power. It is simple to distinguish the 000 input pattern. Additionally, is possible to distinguish four sets of input patterns: these sets of input patterns distinguished by Hamming Weight (HW): 000 with HW(0); 001, 010 and 100 with HW(1); 011, 101 and 110 with HW(2) and 111 with HW(3)

For 3-input gates, the situation is slightly more complicated: the *cocktail effect* causes that just Hamming Weights of the processed data are easy to distinguish. See Figures 11 and 12.

We identified, that the serial arrangement of NMOS transistors (NAND2X1 and NAND3X1) implies the strongest and simple to identify data dependency. Additionally, complete opening/closing of the pull-up (PMOS-based) part of the transistor – see Figure 1 – is also well observable. This allows to clearly distinguish corner input patterns, whose open transistors are connected in series (e.g. 00 for 2-input NOR and 11 for 2-input NAND). These conclusions are in relation to the results presented in Section II.

V. Dual-Rail is Not Safe

As the presented method allows to reveal the input patterns of the combinational circuit, one could argue that the so-called *uniform power consumption* countermeasures will successfully prevent any probing attack. Nevertheless, such methods, frequently based on dual-rail encoding [23], were designed to balance the *dynamic power*. However, the presented method exploits the differences in the geometry and the data dependency of the *static power* modulated by a laser beam.

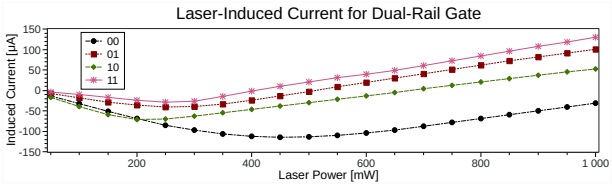

Fig. 13. The photocurrent for the conventional WDDL `NAND` gate composed of `NAND2X1` and `NOR2X1` gates for different input patterns and increasing laser power. The 00 input pattern (logical inputs) is easy to distinguish; it is possible to distinguish also the other patterns: 11, 01 and 10

Dual-rail circuits are composed such a way, that every bit value is encoded using two wires and every gate is replaced by a pair of complementary gates to increase the robustness or introduce uniform power consumption.

Even when the circuit is designed as a dual-rail circuit and complementary gates are placed such a way, so that it is impossible to target a single gate without affecting its complement, the method is still able to disclose the dual-rail gate input patterns, or at least decrease the entropy of the input pattern – to find the most probable input patterns.

In Figure 13, the data-dependency of the dual-rail WDDL-type gate composed of two standard cells is provided. The conventional WDDL (*Wave Dynamic Differential Logic*) [24] approach was used to construct the mentioned gate – the gate is composed of two standard cells, namely `NAND2X1` and `NOR2X1`, thus the power consumption is the sum of consumption of both gates, whose were balanced with equal load capacitance.

According to the best of our knowledge, many dual-rail circuit design styles (and actual designs) will suffer from this behavior, because of differences in gate geometries, whose cannot be removed [23], [24], [25].

VI. Using PLS to Decrease the Processed Data Entropy

In Sections IV and V, we have shown, how it is possible to distinguish patterns at a single (or a pair of) CMOS gate inputs (or outputs) in the case, where the gate is embedded into a larger circuit.

This method can help to disclose bit-values in the circuit, where CMOS gates are large enough to be targeted by a laser beam. However, this is not the case for deep-submicron technologies. In this section, we provide brief instructions on how to mount the photocurrent method to such circuits.

Given that the power imprint (current-data dependence model for given laser power) of all cells used in the circuit under illumination is available, it is possible to obtain the most probable pattern at the combinational circuit input. This is possible to realize for reasonably small combinational circuits only, where the summary data-dependent power is the *cocktail* of gate powers.

Let us consider a combinational sub-circuit, which is a part of a bigger (sequential) circuit. This can be any digital circuit, e.g. AES coprocessor. The selected combinational sub-circuit is large enough to allow illumination of its area only. The number of the sub-circuit inputs is n, thus there are 2^n input patterns. The proposed procedure to decrease the entropy – selecting the subset of 2^n input patterns containing just the inputs corresponding to the photocurrent power imprint – of the input patterns set – is as follows:

(i) make the clock signal stable at the moment, when values that need to be disclosed (e.g. secret key bits) are at the sub-circuit input;

(ii) illuminate the sub-circuit by using a defined laser power and perform the power measurement;

(iii) based on the gate power models (or sub-circuit power models in general), compile the power model of the sub-circuit under illumination by evaluating the expected power for all 2^n input vectors (patterns);

(iv) in the power model, find the vectors with the smallest difference from the measured power.

To be more illustrative, let us have a look at a particular example in Figure 14. The Figure illustrates two data dependent power models of the C17 circuit (mapped to `NAND2X1` cells), which is widely used for illustration purposes [26].

If the 5-input C17 circuit is understand as a black-box with no side-channels, the entropy of its input pattern would be:

$$S_{C17} = log_2 2^5 = 5\ bits$$

Based on the knowledge of the illuminated circuit schematic, the current consumption model for all 2^5 input patterns can be compiled, when power models of all building blocks (cells) are known – in the case of C17 circuit, only the `NAND2X1` cell is used. The column *Modeled photocurrent* in Figure 14 shows the modeled currents. Only 20 different data-dependent current values occur for this model – some of the values repeat, as the gate state combinations repeat. As shown in Figure 14, the same current is generated by three inputs at maximum. As a result, the entropy, based on this model, is:

$$S_{C17^M} = \lceil log_2 3 \rceil = 2\ bits$$

The schematic-based model however hides some of the circuit properties influencing the resulting data-dependent current consumption. The column *Simulated photocurrent* in Figure 14 shows the results for the whole netlist SPICE simulation. One can observe significant differences compared to the *Modeled photocurrent* column – the schematic-based model is misleading!

The positive outcome of the complete circuit simulation is, that 2^5 unique data-dependent current values occur, thus the

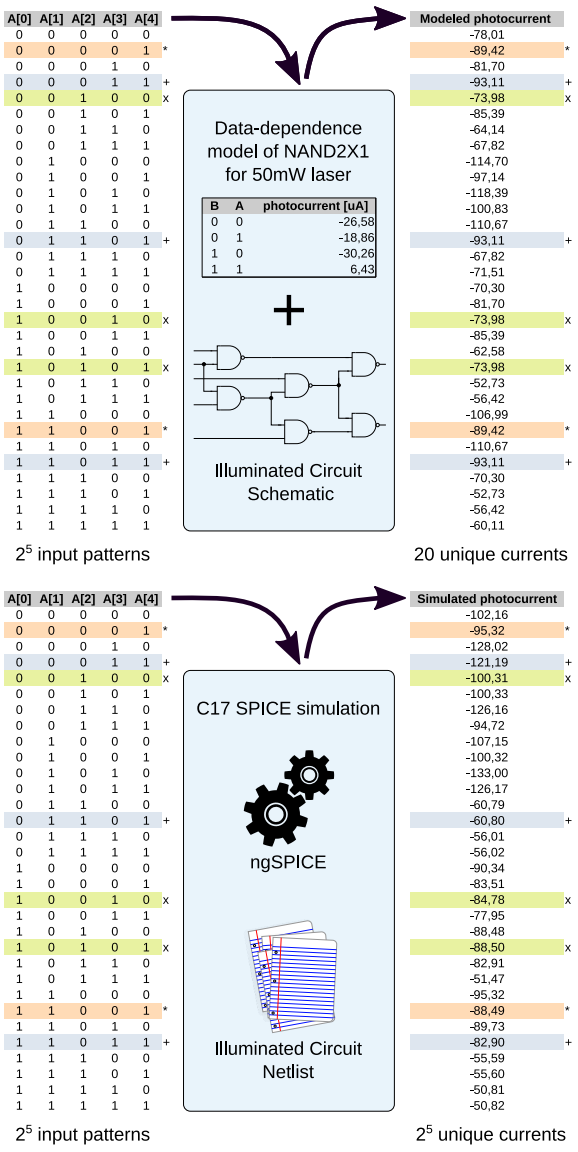

Fig. 14. Example of the data dependent power modeling for the C17 circuit; the C17 circuit is composed of 6 NAND2X1 standard cells; the *Modeled photocurrent* column shows the resulting currents for the composition of gate models and the *Simulated photocurrent* column shows the currents coming from the C17 netlist simulation

entropy based on this model, which reflects the reality closer than the previous one is:

$$S_{C17^S} = \lceil log_2 1 \rceil = 0 \; bits$$

The entropy of 0 bits represents the theoretical limit, however, bad news are also significant: (a) the minimal differences between simulated currents are in the order of nanoamps, which is hard to measure for bigger circuits in reality; and (b) the correspondence with the simpler model (*Modeled photocurrent*) is not maintained.

To be more realistic, the simulated data were rounded to tens of microamps, which created 8 groups of values. There were 1 to 6 data dependent currents in each group, representing the corresponding input patterns. Such an approach leads to entropy of 3 bits:

$$S_{C17^{S'}} = \lceil log_2 6 \rceil = 3 \; bits,$$

what is 2 bits less than for the C17 circuit as a black box.

One can wonder, why there is a difference when both modeled photocurrents are based on SPICE simulation under the same conditions. The reason is, that **the model based on gate powers was intentionally simplified**: the `NAND2X1` standard cell *data-dependence model* (lookup-table) has only 4 rows: `00`, `01`, `10` and `11`, but a precise model would require to be more detailed. As the laser beam induces the current flow through illuminated transistors, it also **causes the voltage drops** at gate outputs. Thus, logic `0` used in the simplified model may be `0.1V` or `0.3V` in reality (depending on the laser power), but the model was compiled with inputs equal to `VDD` for logic `1` and `VSS` for logic `0` respectively. The output voltage differences are reflected by the second model.

The unmodelled voltage drops limits the simpler model scalability. It follows that it is advantageous to **keep the laser power as low as possible to keep the model scalable and as simple (or as accurate) as possible at the same time** because higher laser power will cause more significant voltage drops at gate outputs.

In reality, the inaccuracy described for the simple model will not be the only source of total inaccuracy; at least the manufacturing process variability will make things more complicated; moreover: obtaining a precise current model for logic cells (or subcircuits in general) can be very challenging. Naturally, the alternative to SPICE simulations are physical measurements, but it is not always possible to perform them.

Nevertheless, we believe that a procedure similar to the described one, may lead to a decrease of the input pattern entropy for any combinational (sub-)circuit. This fact potentially influences the security of a wide class of CMOS devices.

VII. Discussion

The presented method, in theory, allows to use a laser beam to obtain data processed by a combinational sub-circuit in a complex digital device. Although the simulations show that the current differences, for the selected technology, are in the order of micro amps, real measurements must be performed to confirm the simulations.

The feasibility of the measurement may be limited in practice due to the presence of additional sources of photocurrent coming from surrounding logic or simply by noise. However, we believe that due to the significant difference in simulated currents, the presented method requires attention.

The advantage of the presented method is that the laser power can be increased in order to increase the significance of the response. Additionally, when the circuit activity is suppressed (the clock signal is stable), the photo-induced current is not pulsed but continuous, which simplifies the measurement.

The requirement of having control over the clock signal may be strong or weak depending on the implementation of the specific circuit. In the case where the circuit uses an external clock source, it is simple to meet this requirement.

The disadvantage of increasing the laser power is, that gate outputs may change: this may cause "partially open" or "partially closed" transistors instead of "open/closed" transistors. This leads to a limited scalability of the modeling: composition of simpler building block models into a complex structure model leads to inaccuracy or more complicated model composition (based on analog, not logic simulation).

If the laser power is low enough (and laser is targeted properly), no permanent fault is introduced into the device by the presented method, as only combinational logic is illuminated by the laser. Thus, fault-detection or intrusion detection mechanisms used as countermeasures may be overcome [1].

Naturally, there is a limit for the laser power, conditioned by devastating heating of the chip area.

The advantage of the presented method targeting combinational logic (instead of flip-flops or single transistors) is that the area affected by a laser beam may be significantly bigger. Thus, the presented method is usable for deep submicron devices.

Multiple metal layers in present devices may lead to the method efficiency reduction. This may lead to the need for illumination of the back-side of the die, causing higher data-independent currents [11].

The method presented in Section VI has strong requirements: it is not only necessary to know the circuit layout, but also the current imprints of the circuit (or the circuit's building blocks) under illumination. It may be very challenging to obtain accurate current models: the simulation can be far away from reality; data obtained from specially manufactured samples can be burdened by measurement errors, and differences in the manufacturing process and *profiling* [27] may be difficult.

Skorobogatov and Anderson have shown that it is possible to perform optical attacks with very cheap equipment [10]. However, environmental stability and replicability of the experiment (which is required by the presented entropy decreasing method) may require a relative costly equipment.

Although we have shown, that dual-rail based design styles may not provide sufficient level of security, very symmetric (both gate and transistor level) approaches employing dual-rail logic, e.g. carefully placed Seclib [28], may provide solid attack resistance.

VIII. Conclusions and Future Work

In this paper, we have shown how the static power of a selected CMOS sub-circuit, which is data dependent, can be modulated by a laser beam. This may lead to a visibility of the processed data in the device's power trace.

Although the simulation have shown that the current differences are in the order of microamps, real measurements must be performed to confirm the simulations.

The method allowing to use a laser beam to obtain potentially sensitive data processed by a combinational sub-circuit in a complex digital device has been provided. Although the method has strong requirements, we believe that the research of its potential requires attention.

We have identified that the arrangement of transistors – in `NAND2X1` or `NAND3X1` – implies the strongest and simple to identify data dependency. This increases the severity of the potential threat, as `NAND` gates are preferred due to smaller size and delay in digital design (optimization).

Additionally, we have shown, that dual-rail approaches cannot be simply used to fight against the potential attacks based on the presented method.

Acknowledgment

The authors acknowledge the support of the OP VVV MEYS funded project CZ.02.1.01/0.0/0.0/16_019/0000765 "Research Center for Informatics" and the CTU grant SGS17/213/OHK3/3T/18.

References

[1] D. Karaklajić, J. Schmidt, and I. Verbauwhede, "Hardware Designer's Guide to Fault Attacks," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 21, no. 12, pp. 2295–2306, Dec 2013.

[2] S. Ravi, A. Raghunathan, P. Kocher, and S. Hattangady, "Security in Embedded Systems: Design Challenges," *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 3, no. 3, pp. 461–491, 2004.

[3] K. Tiri, "Side-channel attack pitfalls," in *2007 44th ACM/IEEE Design Automation Conference*. IEEE, 2007, pp. 15–20.

[4] T. Snyder and G. Byrd, "The Internet of Everything," *Computer*, vol. 50, no. 6, pp. 8–9, 2017. [Online]. Available: doi.ieeecomputersociety.org/10.1109/MC.2017.179

[5] P. Kocher, R. Lee, G. McGraw, A. Raghunathan, S. Moderator-Ravi, and S. Moderator-Ravi, "Security as a New Dimension in Embedded System Design," in *Proceedings of the 41st annual Design Automation Conference*. ACM, 2004, pp. 753–760.

[6] Y. Li, M. Chen, and J. Wang, "Introduction to side-channel attacks and fault attacks," in *2016 Asia-Pacific International Symposium on Electromagnetic Compatibility (APEMC)*, vol. 01, May 2016, pp. 573–575.

[7] P. Kocher, J. Jaffe and B. Jun, "Introduction to Differential Power Analysis and Related Attacks," 1998, technical report. [Online]. Available: http://www.cryptography.com/dpa/

[8] F. Regazzoni, L. Breveglieri, P. Ienne, and I. Koren, *Interaction Between Fault Attack Countermeasures and the Resistance Against Power Analysis Attacks*. Springer Berlin / Heidelberg, 2012, ch. Interaction Between Fault Attack Countermeasures and the Resistance Against Power Analysis Attacks, pp. 257–272.

[9] V. Miškovský, H. Kubátová, and M. Novotný, "Influence of passive hardware redundancy on differential power analysis resistance of AES cipher implemented in FPGA," *Microprocessors and Microsystems*, vol. 51, pp. 220–226, 2017.

[10] S. P. Skorobogatov and R. J. Anderson, "Optical fault induction attacks," in *International workshop on cryptographic hardware and embedded systems*. Springer, 2002, pp. 2–12.

[11] S. Skorobogatov, "Optically enhanced position-locked power analysis," in *Workshop on Cryptographic Hardware and Embedded Systems*. Springer, 2006, pp. 61–75.

[12] J. Bělohoubek and P. Fišer and J. Schmidt, "Using Voters May Lead to Secret Leakage," in *IEEE 22nd International Symposium on Design and Diagnostics of Electronic Circuits and Systems (DDECS 2019)*, April 2019, pp. 1–4.

[13] F. Amiel, K. Villegas, B. Feix, and L. Marcel, "Passive and active combined attacks: Combining fault attacks and side channel analysis," in *Fault Diagnosis and Tolerance in Cryptography, 2007. FDTC 2007. Workshop on*, Sept 2007, pp. 92–102.

[14] D. H. Habing, "The use of lasers to simulate radiation-induced transients in semiconductor devices and circuits," *IEEE Transactions on Nuclear Science*, vol. 12, no. 5, pp. 91–100, 1965.

[15] A. Sarafianos, O. Gagliano, V. Serradeil, M. Lisart, J.-M. Dutertre, and A. Tria, "Building the electrical model of the pulsed photoelectric laser stimulation of an NMOS transistor in 90nm technology," in *IEEE International Reliability Physics Symposium (IRPS), 2013*. IEEE, 2013, pp. 5B–5.

[16] A. Sarafianos, R. Llido, O. Gagliano, V. Serradeil, M. Lisart *et al.*, "Building the electrical model of the pulsed photoelectric laser stimulation of an NMOS transistor in 90nm technology," in *38th International Symposium for Testing and Failure Analysis, (ISTFA) 2012*, 2012, pp. 5B–5.

[17] A. Sarafianos, O. Gagliano, M. Lisart, V. Serradeil, J. Dutertre, and A. Tria, "Building the electrical model of the pulsed photoelectric laser stimulation of a PMOS transistor in 90nm technology," in *Proceedings of the 20th IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, July 2013, pp. 22–27.

[18] N. Weste and D. Harris, *CMOS VLSI Design: A Circuits and Systems Perspective*, 4th ed. USA: Addison-Wesley Publishing Company, 2010.

[19] J. Giorgetti, G. Scotti, A. Simonetti, and A. Trifiletti, "Analysis of data dependence of leakage current in CMOS cryptographic hardware," in *Proceedings of the 17th ACM Great Lakes symposium on VLSI*. ACM, 2007, pp. 78–83.

[20] M. Alioto, L. Giancane, G. Scotti, and A. Trifiletti, "Leakage Power Analysis attacks: Well-defined procedure and first experimental results," in *2009 International Conference on Microelectronics - ICM*, Dec 2009, pp. 46–49.

[21] A. Srivastava, R. Bai, D. Blaauw, and D. Sylvester, "Modeling and Analysis of Leakage Power Considering Within-Die Process Variations," in *Proceedings of the International Symposium on Low Power Electronics and Design*. IEEE, 2002, pp. 64–67.

[22] R. Llido, A. Sarafianos, O. Gagliano, V. Serradeil, V. Goubier, M. Lisart, G. Haller, V. Pouget, D. Lewis, J.-M. Dutertre *et al.*, "Characterization and TCAD simulation of 90 nm technology transistors under continous photoelectric laser stimulation for failure analysis improvement," in *19th IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA) 2012*. IEEE, 2012, pp. 1–6.

[23] J. Sparsø and S. Furber, *Principles of Asynchronous Circuit Design: A Systems Perspective*, 1st ed. Kluwer Academic Publishers, Boston, 2001.

[24] K. Tiri and I. Verbauwhede, "A logic level design methodology for a secure DPA resistant ASIC or FPGA implementation," in *Proceedings Design, Automation and Test in Europe Conference and Exhibition*, vol. 1. IEEE, 2004, pp. 246–251.

[25] S. Bhasin, J. L. Danger, F. Flament, T. Graba, S. Guilley, Y. Mathieu, M. Nassar, L. Sauvage, and N. Selmane, "Combined SCA and DFA Countermeasures Integrable in a FPGA Design Flow," in *2009 International Conference on Reconfigurable Computing and FPGAs*, Dec 2009, pp. 213–218.

[26] F. Brglez and H. Fujiwara, "A Neutral Netlist of 10 Combinational Benchmark Circuits and a Target Translator in Fortran," in *Proceedings of IEEE International Symposium Circuits and Systems (ISCAS 85)*. IEEE Press, Piscataway, N.J., 1985, pp. 677–692.

[27] S. Chari, J. R. Rao, and P. Rohatgi, "Template attacks," in *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer, 2002, pp. 13–28.

[28] S. Guilley, F. Flament, Y. Mathieu, and R. Pacalet, "Security evaluation of a balanced quasi-delay insensitive library (seclib)," in *Conference on Design of Circuits and Integrated Systems*, 2008, pp. 6–pages.

[29] C. Roscian, A. Sarafianos, J. 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*, Aug 2013, pp. 89–98.

[30] B. University of California, "Berkeley logic interchange format (BLIF)," 2005.

APPENDIX A
NOTES TO RESOURCES AND REPLICABILITY

Sarafianos et al. [15], [16], [17] used the STM 90nm technology for their experiments. As the STM's technology details (SPICE models, cell libraries), are not publicly available, we decided to mount their models to publicly available technology node to increase the experiment replicability.

A. Technology node

For simulations, we used publicly available TSMC transistor models for the 180nm technology simulation. The TSMC 180nm technology advantage is the availability of open-source standard cell library and SPICE models provided by Oklahoma State University (OSU)[3]. Thanks to the availability of SPICE models and the standard cell library, it is possible to perform the simulation of a manufacturable circuit layout.

The 180nm technology does not represent the latest technology node, but it is still relevant for manufacturing devices like *smart-cards* or *key-fobs*, which may be compromised by the presented vulnerability. As the manufacturing in technologies like 180nm or 350nm is relatively cheap, the evaluation of the vulnerability at this technology node makes sense.

With the mentioned model, we repeated simulations presented by Sarafianos et al., while conserving other parameters from the original papers, including the laser power and transistor sizes. As a result, we achieved simulation outputs comparable with Sarafianos et al. and thus we showed, that using different transistor models does not lead to unrealistic results.

For real layout simulation, shrinking transistor sizes in the model is necessary. For scaling to lower transistor dimensions, we followed Roscian and Sarafianos et al. [29]. We used PN junction sizes coming from the actual layout under simulation.

B. Layout Synthesis

For the experiment replicability reasons, we have chosen a completely open toolchain to synthesize the layouts for experiments. The used available open *digital synthesis flow* is called *Qflow*[4]. Qflow incorporates well known open-source tools for different stages of synthesis including *Yosys*[5] for RTL Verilog synthesis, *Berkeley ABC*[6] for logic synthesis, *QRouter*[7] and *GrayWolf*[8] for place&route and *Magic*[9] as a VLSI layout tool.

All layouts were synthesized by using the TSMC 180nm technology provided by Oklahoma State University (OSU), which is distributed with Qflow.

The synthesis procedure was as follows: the "top" tools in Qflow were skipped and instead these steps were performed manually. The design started from a schematic, the circuit was translated to a BLIF [30] representation expressing the gate mapping, which is accepted by the place&route stage of Qflow. Qflow was then used to produce the layouts and also the SPICE netlists.

The resulting SPICE netlists were modified by hand – the transistors were replaced by transistors under PLS models – and used for simulation.

[3] https://vlsiarch.ecen.okstate.edu/flows/MOSIS_SCMOS
[4] http://opencircuitdesign.com/
[5] http://www.clifford.at/yosys/
[6] https://people.eecs.berkeley.edu/∼alanmi/abc/
[7] http://opencircuitdesign.com/qrouter/index.html
[8] https://github.com/rubund/graywolf
[9] http://opencircuitdesign.com/magic/