Side-Channel Leakage in Masked Circuits Caused by Higher-Order Circuit Effects

Zhimin Chen, Syed Haider, and Patrick Schaumont

Virginia Tech, Blacksburg, VA 24061, USA
{chenzm,syedh,schaum}@vt.edu

Abstract. Hardware masking is a well-known countermeasure against Side-Channel Attacks (SCA). Like many other countermeasures, the side-channel resistance of masked circuits is susceptible to low-level circuit effects. However, no detailed analysis is available that explains *how*, and *to what extent*, these low-level circuit effects are causing side-channel leakage. Our first contribution is a unified and consistent analysis to explain *how* glitches and inter-wire capacitance cause side-channel leakage on masked hardware. Our second contribution is to show that inter-wire capacitance and glitches are causing side-channel leakage of *comparable* magnitude according to HSPICE simulations. Our third contribution is to confirm our analysis with a successful DPA-attack on a 90nm COMS FPGA implementation of a glitch-free masked AES S-Box. According to existing literature, this circuit would be side-channel resistant, while according to our analysis and measurement, it shows side-channel leakage. Our conclusion is that circuit-level effects, not only glitches, present a practical concern for masking schemes.

1 Introduction

Side-channel power analysis on cryptographic circuits has been investigated since well over a decade. Differential Power Analysis (DPA) [1], and the related attacks present great concern for cryptographic hardware. As one of the mainstream countermeasures, masking [2, 3] provides protection by randomizing the intermediate circuit values.

However, like many other countermeasures, hardware masking faces problems at the circuit implementation phase. Glitches are a well-known circuit effect that deteriorates the effect of masking [4]. As a solution, dual-rail pre-charge circuit is employed.

In this paper, we take a close look at the circuit properties that depend on more than one value in the circuit. We call them higher-order circuit effects. We first derive two conditions for secure masking of circuits: 1) the random masks need to be unbiased, and 2) the power consumption needs to be independent of the unmasked data. In logic simulations, for example in simulations based on toggle counts, these two conditions can be satisfied. In a real circuit, however, we find the second condition cannot be met because of higher-order circuit effects. While higher-order effects have been mentioned as a possible source of side-channel leakage [5], no detailed analysis has been presented.

J.H. Park et al. (Eds.): ISA 2009, LNCS 5576, pp. 327–336, 2009.
© Springer-Verlag Berlin Heidelberg 2009

Our contribution includes detailed analysis of two common higher-order circuit effects, including glitches and inter-wire capacitance. Our analysis explains, in a consistent manner, why all higher-order circuit effects can cause side-channel leakage in masked circuits. As far as we are aware, no such analysis has been performed before. Our second contribution is to demonstrate and to quantify the impact of circuit effects on side-channel leakage. We present HSPICE simulations on an algorithmic masked $GF(2^2)$ multiplier. In our experiment, we conclude that the leakage caused by each of the above-mentioned higher-order circuit effects are comparable. Our third contribution is to demonstrate our argument with a successful DPA-attack on a 90nm CMOS FPGA implementation of a glitch-free masked AES S-Box.

The rest of the paper is organized as follows. Section 2 briefly introduces previous work. In Section 3, we present a detailed analysis on masked circuits. Section 4 provides experimental results. In Section 5, we conclude our work.

2 Previous Work

A circuit is called *perfectly masked* if there is no dependency, in a statistical sense, between the power signature and the unmasked circuit inputs [6]. Unfortunately, glitches have been indicated as a source of side-channel leakage in *perfectly masked* circuits [4]. Several dual-rail technologies [7] have since been proposed to address these issues . This makes glitch-free, perfectly-masked circuits a state-of-the-art solution for side-channel resistant hardware implementations. In this paper, we investigate the possibility of performing a *first-order* side-channel attack using higher-order circuit effects commonly found in deep-submicron implementations.

3 Analysis on Masked Circuits

In this section we present an analysis of masked circuits. We first derive two conditions to implement perfect masking. Next, we show that higher-order circuit effects (such as glitches and inter-wire capacitance) may break these conditions.

3.1 Two Conditions for Perfect Masking

The perfect masking condition requires that a logic-0 and a logic-1 appears with the same probability on all intermediate circuit nodes. Let's consider such an intermediate plaintext node a, which is masked using a mask m by means of Boolean masking (XOR) ($a_m = a \oplus m$).

Let F_0 be the probability that $m = 0$, and F_1 the probability that $m = 1$. Obviously $F_0 + F_1 = 1$. We can express the expectation of power consumption P with respect to the unmasked signal a as follows.

$$\begin{cases} P(a=0) = F_0 P(a_m = 0, m = 0) + F_1 P(a_m = 1, m = 1) \\ P(a=1) = F_0 P(a_m = 1, m = 0) + F_1 P(a_m = 0, m = 1) \end{cases} \tag{1}$$

In order to implement the perfect masking condition, the power consumption of the circuit needs to be independent from a in a statistical sense. Therefore, the first and second formula in Eq. (1) should have the same expectation. There are many ways to fulfill this condition. However, the most common approach is to require the following.

$$F_0 = F_1 \tag{2}$$

$$\begin{aligned} & P(a_m = 0, m = 0) + P(a_m = 1, m = 1) \\ = & P(a_m = 1, m = 0) + P(a_m = 0, m = 1) \end{aligned} \tag{3}$$

We consider Eq. (2) and (3) as two general conditions for secure masked circuits. Eq. (2) shows that the mask signal needs to be unbiased, while Eq. (3) shows that the circuit must consume the same power for each value of the unmasked input in a statistical sense.

The above conditions can be easily expanded to more general masking arrangements. Given a circuit block with masked data input a_m, and mask m, where each of these can be words, then we can write the power consumption in terms of the unmasked data a as follows.

$$P(a) = \sum F_m P(a_m, m)$$

F_m is the probability distribution of the mask, while $P(a_m, m)$ is the power consumption of the masked circuit for each possible combination of mask and masked value. Accordingly, we can define two general conditions for secure masked circuits as follows:

$$F_i = F_j |_{i \neq j} \tag{4}$$

$$\sum_m P(a_m = i \oplus m, m) = \sum_m P(a_m = j \oplus m, m) \ |_{i \neq j} \tag{5}$$

The first condition (4) requires the mask to have a uniform distribution. The second condition (5) requires the circuit to consume the same power for each possible value of the unmasked input in a statistical sense.

Thus, conditions (2) and (3), as well as their generalizations (4) and (5), express when perfect masking is achieved by a masked circuit implementation. First-order attacks on masked circuits are enabled by violation of either of these conditions. For example, it is known that a bias in the mask causes first-order side-channel leakage [8]. Indeed, bias represents a violation of Eq. (2) or Eq. (4). In order to understand the implications of the condition Eq. (3) or Eq. (5), we need a better understanding of the power consumption P. In a digital circuit, the dynamic power consumption is given by the following equations [9].

$$P_{avg} = \alpha \cdot f_c \cdot V^2 \cdot C \tag{6}$$

$$P = I \cdot V \tag{7}$$

In Eq. (6), P_{avg} is the average power consumption. α is the switching factor, f_c is the clock frequency, V is the voltage of power supply and C is the effective

capacitance. In Eq. (7), P is the instant power consumption. I is the instant current and V is the instant voltage. For a real attack, the average power consumption is the average value of all the power samples in one cycle. The instant power consumption corresponds to only one power sample. Both of them can be used for the side-channel analysis. We need to examine every term in Eq. (6) and Eq. (7) for a possible dependency to the value of the unmasked signal a. If such a dependency is found, then the condition of Eq. (3) or Eq. (5) no longer holds and a first-order side-channel leakage appears.

3.2 Higher-Order Circuit Effects Causing Side-Channel Leakage

As it turns out, real circuits have a large number of higher-order effects that can cause the terms of Eq. (3) and Eq. (5) to become dependent on the circuit state. We will describe two of them: glitches and inter-wire capacitance. There may be other higher-order effects in circuits, but it is not our intention to be exhaustive. Instead, our objective is to show *how* and *why* higher-order circuit effects can be used to relate different intermediate circuit values. When these values contain both the information of the mask and the corresponding masked value (not necessary to be exactly m and a_m), the previous 'perfect' masking is not perfect anymore.

Glitches. First, consider again the effect of glitches. A glitch results in additional net transitions in a circuit, so that the switching factor α in Eq. (6) appears to be increasing. Glitches are also state-dependent. For example, a glitch may appear when $a_m = 1$ and $m = 1 (a = 0)$ but not in any other combination. Hence, glitches may cause an imbalance in Eq. (3), which in turn results in a violation of the perfect masking condition.

Inter-wire Capacitance. Second, consider the effect of the capacitance C on the average power consumption. The total capacitance of a circuit has two components: gate capacitance and wire capacitance. In deep submicron technology, the inter-wire capacitance accounts for more than 50% of the total capacitance for narrow wires [10]. Modeling of the wire capacitance in an integrated circuit is a non-trivial task. For simplicity, the following discussion is on the basis of the simplified model as shown in Fig. 1(a). This circuit shows two wires $w1$ and $w2$,

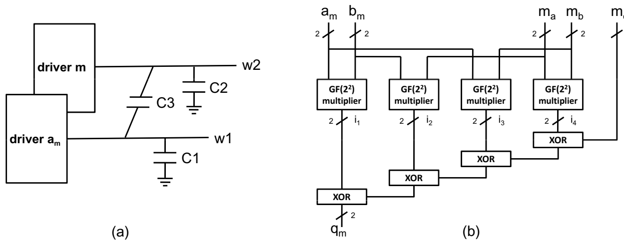

Fig. 1. (a) Model for inter-wire capacitance; (b) Circuit for HSPICE simulation

each with a different driver. There are two parts for the capacitance seen by each wire, one part between the wire and the semiconductor substrate ($C1$ and $C2$), and the other part between adjacent wires ($C3$). Wire $w1$ sees a single, fixed capacitor $C1$ and a second floating capacitor $C3$. The effective value of $C3$ as seen from wire $w1$ changes in function of $w2$'s voltage level. For example, when $w1$ carries a logic-1 while $w2$ carries a logic-0, then the effective capacitance on $w1$ is $C1 + C3$. However, when both $w1$ and $w2$ carry a logic-1, then the effective capacitance on $w1$ is only $C1$, since there is no current required to charge $C3$.

Now assume that wire $w1$ is driven by the masked signal a_m and wire $w2$ is driven by the mask m. Evaluating Eq. (3), and assuming that logic-0 does not consume power, we find for the left side of Eq. (3):

$$P(a_m = 0, m = 0) + P(a_m = 1, m = 1) = 0 + f_c \cdot V^2 \cdot (C1 + C2)$$

While the right side evaluates to:

$$P(a_m = 1, m = 0) + P(a_m = 0, m = 1) = f_c \cdot V^2 \cdot (C2 + C3) + f_c \cdot V^2 \cdot (C1 + C3)$$

Clearly, the right hand side is a factor $2 \cdot f_c \cdot V^2 \cdot C3$ bigger than the left hand side. This factor is caused by taking inter-wire capacitance $C3$ into account. As chip feature sizes continue to shrink, inter-wire capacitance becomes more significant. This implies that the possible asymmetry of Eq. (3) is likely to deteriorate further with shrinking feature size.

There are many other higher-order circuit effects, for example IR Drop, that can break Eq. (3) and Eq. (5). We do not discuss them in detail here because of lack of space.

According to the above analysis, higher-order circuit effects make the correlation between different intermediate circuit values a common phenomenon in a real circuit. This presents a risk for the perfect masking scheme.

4 Experimental Results

In this section, we describe a series of experiments that we did to test the above analysis. We will show that higher-order circuit effects, not only glitches, can cause measurable side-channel leakage. Our experiments are divided into two parts. In the first part, HSPICE simulation shows that the amount of leakage caused by glitches and inter-wire capacitance can be comparable. Given that the glitches can lead to successful SCA [4], there is no reason to believe that inter-wire capacitance is secure. In the second part, we move from simulation to real circuits. The purpose is to see, after ruling out the effect of glitches, the leakage caused by the other higher-order circuit effects is measurable.

4.1 HSPICE Simulation

To prove the analysis in Section 3, we did the simulations with HSPICE. Compared with experiments on real circuits, HSPICE simulation can easily test the

influence of individual higher-order circuit effects. Our approach is to explore the difference between the circuits with and without higher-order circuit effects.

The circuit under simulation is shown in Fig. 1(b). It is a masked Galois Field multiplier: a critical part of a masked AES S-Box. The circuit has 5 inputs: a_m, b_m, m_a, m_b, and m_q and 1 output: q_m (2 bit for each). a_m and b_m are masked values of a and b ($a_m = a$ XOR m_a, $b_m = b$ XOR m_b); q_m is the masked value of q ($q_m = q$ XOR m_q). The circuit consists of 4 $GF(2^2)$ multipliers and a set of XOR gates (36 standard gates in total).

Glitches. We perform two simulations to test glitches: the first with the entire circuit in Fig. 1(b); the second only with the $GF(2^2)$ multipliers. According to [4], only the XOR gates leak side-channel information through glitches. Therefore, we expect unmasked-data dependent power in the first experiment while no leakage in the second. In each simulation, we take the following steps: **1)** Switch the input from 0 to every possible value n. Accordingly, we obtain 1024 average current values for the entire circuit, proportional to the average power. **2)** Every average current value is mapped to a set of unmasked inputs a and b. We group the 1024 average current values in terms of the hamming weight (from 0 to 4) of the unmasked inputs. By averaging each group, we obtain the mean power for each hamming weight as the experimental result.

The result is shown in Fig 3(a). As we can see, for the first experiment, the majority momentum of mean power in the first simulation is increasing as the hamming weight of the unmasked inputs increases. In contrast, mean power in the second simulation almost remains the same. Hence, we can attribute the leakage to the glitches in the XOR gates.

What should be mentioned is that glitches in the XOR gates are related to the arrival time of the inputs which is decided by the layout of the circuit and other factors. In our experiment, we performed the first simulation several times with different arrival sequence of the inputs. The results turn out to be similar as the one shown in Fig. 3(a).

Inter-Wire Capacitance. Inter-wire capacitance can exist between the outputs of the $GF(2^2)$ multipliers. We can also find inter-wire capacitance between nets in different $GF(2^2)$ multipliers, if they are placed close to each other. Furthermore, inter-wire capacitance is influenced by many factors, for example the layout of the circuit. In our simulations, we made a reasonable assumption for inter-wire capacitance: comparable to the gate capacitance [10]. Because it is really hard to eliminate glitches from the XOR gates, to see the individual

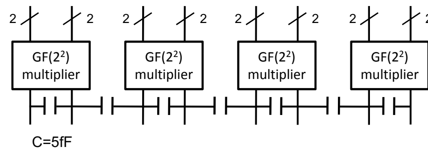

Fig. 2. Inter-wire capacitance between and inside $GF(2^2)$ multiplier

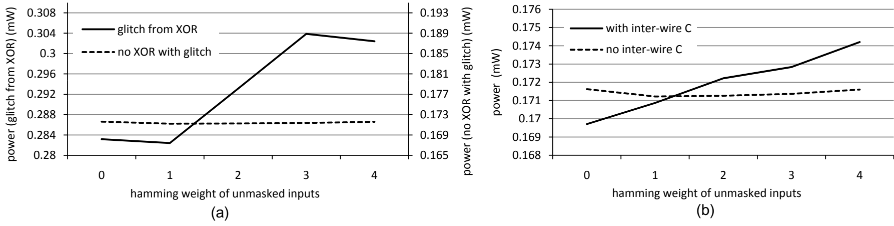

Fig. 3. (a) Result on glitches; (b) Result on inter-wire capacitance

influence of the inter-wire capacitance, we rule out the XOR gates and perform simulation just on the $GF(2^2)$ multipliers.

The first simulation is performed with inter-wire capacitances added between the outputs of the $GF(2^2)$ multipliers, shown in Fig. 2. The second one is done without inter-wire capacitance, which is exactly the same as the second simulation for glitches. In each simulation, we take the same steps mentioned in the previous subsection.

The results are presented in Fig. 3(b), where we can see the average power increases as the hamming weight goes from 0 to 4 in the first simulation but changes little in the second one.

According to perfect masking conditions, the power should be independent on the unmasked inputs. Clearly, it should also be independent on the hamming weight of them. From the above experimental results, we can see the two higher-order circuit effects discussed in Section 3 introduce dependence between the power and the unmasked input. Therefore, they are indeed sources of side-channel leakage.

We also quantify the relative side-channel leakage for each circuit effect. It is not easy to define the magnitude of a leakage. Usually, this is partially dependent on the attack method. Here, for the convenience, we base our analysis on the attack with hamming weight. From Fig. 3, we can find the maximum power variations for glitches and inter-wire capacitance are 0.025 mW, 0,0045 mW respectively. The ratio is 5.6 : 1. Clearly, their leakage are comparable.

4.2 Experiments on FPGA

In this section, we will show that, besides glitches, other higher-order circuit effects can also cause measurable side-channel leakage. All the other higher-order effects are considered together here, not individually. On the one hand, it is really hard to focus on one of them by disabling all the others; on the other hand, there is no need to do so because the purpose of this paper is to show that all the other higher-order effects as a whole present a practical concern for masking. The circuit under test is a glitch-free masked AES S-Box [3] in FPGA. Since the circuit is 'perfectly' masked and designed without glitches, we can then exclude the first-order circuit effects and the timing issues. In other words, if side-channel leakage is found, we can attribute it to the higher-order circuit effects other than glitches.

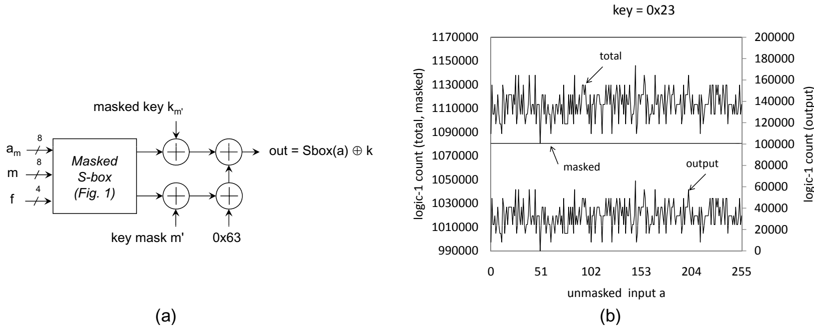

Fig. 4. (a) Design under test in FPGA; (b) Results from the first experiment - At the logic level, the masked S-Box works.

For the purpose of attack, we add a constant key addition (key = 0x23) to the S-Box as shown in Fig. 4(a). Only the output of the overall circuit is observed in unmasked form. The test setup also ensured that all mask values are unique and independently generated.

In the first experiment, we perform a logic simulation of the FPGA netlist and count the number of logic-1s in the whole design during simulation. This count is an estimate for the power consumption. Our objective is to demonstrate that the masked AES S-Box is correctly implemented and immune to a first-order attack on the logic level.

The testbench used for Fig. 4(a) enumerates every combination of the input a_m, mask m, and freshmask f to the masked S-Box. The logic-1 count from the FPGA netlist is obtained using logic simulation. We then group the resulting counts according to the value of the unmasked input a and accumulate each group. This results in the upper graph shown in Fig. 4(b), labeled 'total'. We can clarify the shape of this graph as follows. The sum of logic-1 counts for each a consists of two parts. The first part is from the masked section in Fig. 4(a), which should be independent from a. The second part comes from the unmasked output in Fig. 4(a), and that count should be proportional to the hamming weight of the unmasked output. The lower graph in Fig. 4(b) shows $hammingweight(Sbox(a) \oplus k)$, the hamming weight of each unmasked output in function of the unmasked input a. Clearly, the variations of 'total' logic-1 count and 'output' logic-1 count are identical. This means that the logic-1 count of the masked S-Box must be constant and independent of a, as shown by the middle line in Fig. 4(b), labeled 'masked'. Therefore, we conclude from the first experiment that the masking methodology works on the logic level, and that the masked AES S-Box is designed and implemented correctly.

In the second experiment, we implement a DPA attack on the implementation of the circuit from Fig. 4(a) in an FPGA. To make the circuit glitch-free, we transform the previous FPGA netlist to a pre-charged complementary logic,

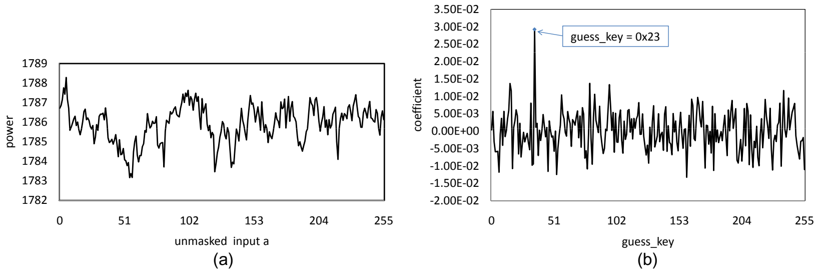

Fig. 5. (a) Results in the second experiment - At the circuit level, the glitch-free, masked S-Box shows first-order leakage; (b) A first-order attack on the glitch-free masked S-Box succeeds

based on WDDL [7] form. What's more, we also ensure timing issues cannot cause any side-channel leakage. Further, we generate the post place-and-route timing model. Using simulation, we verify that the circuit contains no glitches. An important fact is that the leakage caused by imbalanced dual-rail capacitances in WDDL does not affect our design, because the circuit is already masked by an algorithmic masking scheme.

In order to collect a large number of power traces, we created an automatic measurement system, which consists of three parts: the FPGA board, a Personal Computer (PC), and an oscilloscope. The FPGA board contains a Spartan-3E 1600 FPGA (90nm CMOS). Five identical WDDL based masked AES S-Boxes are implemented in the FPGA. A PicoBlaze is used to control the stimuli for the S-Box and the communication with PC through RS-232. During operation, each S-Box takes the same action. The oscilloscope is an Agilent DSO5032A. It samples the current running into the FPGA core through a current probe and also averages each trace 1024 times to eliminate noise. The PC automatically collects signal traces from the oscilloscope through a USB cable. The test program exhaustively enumerates all mask values m $(0 \ldots 255)$ for all possible masked inputs a_m $(0 \ldots 255)$. The freshmask f is randomly generated internally. Considering the 1024-time average, f is also exhausted and unbiased. This way, each data collection phase obtains 64K traces. Average power of these traces can be analyzed in the same manner as in the first experiment, by grouping the traces according to the value of the unmasked input a. The resulting power consumption is captured in Fig. 5(a) which is very different from the variation in Fig. 4(b). This means that the power consumption of the S-Boxes still depends on the value of the unmasked input, and therefore that the traces can be used for a first-order DPA attack. Moreover, we do the first-order DPA based on correlation of the unmasked input hamming weight and the corresponding average power. Fig. 5(b) shows the results of a successful first-order attack on the glitch-free, masked S-Box for a sample key k = 0x23. We also replaced the average power with instant power, the attack is still successful.

5 Conclusion

This paper introduced an analysis of masked circuits from the circuit-level perspective. This analysis is summarized with two general conditions for secure masking. Both of these conditions can be easily achieved at the logic-level, which abstracts voltage into discrete logic-levels and which abstracts time into clock cycles. Our results confirmed that a logic-level simulation of a masked circuit indeed remains side-channel-free. However, the conditions for secure masking are much harder to achieve in real circuits, in which we have to allow for various electrical and analog effects. We showed that glitches are not the only cause of side-channel leakage in masked circuits. As an example, the effect of inter-wire capacitance is elaborated. We evaluated our analysis using HSPICE simulations and measurements on a glitch-free masked AES S-Box in the FPGA. In HSPICE simulations, we found comparable side-channel leakage for glitches and inter-wire capacitance. We also successfully mounted first-order attacks on this FPGA. Our conclusion is that higher-order circuit effects, not only glitches, present a practical concern for masking schemes.

References

1. Kocher, P., Jaffe, J., Jun, B.: Differential Power Analysis. In: Wiener, M. (ed.) CRYPTO 1999. LNCS, vol. 1666, pp. 388–397. Springer, Heidelberg (1999)
2. Chari, S., Jutla, C.S., Rohatgi, P.: Towards Sound Approaches to Counteract Power-Analysis Attacks. In: Wiener, M. (ed.) CRYPTO 1999. LNCS, vol. 1666, pp. 398–412. Springer, Heidelberg (1999)
3. Oswald, E., Mangard, S., Pramstaller, N., Rijmen, V.: A Side-Channel Analysis Resistant Description of the AES S-Box. In: Gilbert, H., Handschuh, H. (eds.) FSE 2005. LNCS, vol. 3557, pp. 413–423. Springer, Heidelberg (2005)
4. Mangard, S., Schramm, K.: Pinpointing the Side-channel Leakage of Masked AES Hardware Implementation. In: Goubin, L., Matsui, M. (eds.) CHES 2006. LNCS, vol. 4249, pp. 76–90. Springer, Heidelberg (2006)
5. Mangard, S., Oswald, E., Popp, T.: Power Analysis Attacks - Revealing the Secrets of Smart Cards. Springer, Heidelberg (2007)
6. Blömer, J., Guajardo, J., Krummel, V.: Provably Secure Masking of AES. In: Handschuh, H., Hasan, M.A. (eds.) SAC 2004. LNCS, vol. 3357, pp. 69–83. Springer, Heidelberg (2004)
7. Tiri, K., Verbauwhede, I.: A Logic Level Design Methodology for a Secure DPA Resistant ASIC or FPGA Implementation. In: Proc. of DATE 2004, pp. 246–251 (2004)
8. Gierlilchs, B.: DPA-resistance without routing constraints? In: Paillier, P., Verbauwhede, I. (eds.) CHES 2007. LNCS, vol. 4727, pp. 107–120. Springer, Heidelberg (2007)
9. Rabaey, J.M., Chanadrakasan, A., Nikolic, B.: Digital Integrated Circuits: A Design Perspective, 2nd edn. Prentice Hall, Englewood Cliffs (2003)
10. Weste, N.H.E., Harris, D.: CMOS VLSI Design: A Circuits and Systems Perspective, 3rd edn. (2005) ISBN: 0-321-14901-7