Layout Dependent Phenomena A New Side-channel Power Model

Geir Olav Dyrkolbotn
Norwegian Information Security Laboratory, NISlab,
Gjøvik University College, Norway,
Email: geirolav.dyrkolbotn@gmail.com

Knut Wold and Einar Snekkenes
Norwegian Information Security Laboratory, NISlab,
Gjøvik University College, Norway,
Email: {knut.wold,einar.snekkenes}@hig.no

Abstract— **The energy dissipation associated with switching in CMOS logic gates can be used to classify the microprocessor's activity. In VLSI design, layout dependent phenomena, such as capacitive crosstalk, become a major contributor to the power consumption and delays of on-chip busses as transistor technology get smaller. These effects may be known to the security community but have received little attention.**

In a recent paper we presented a new power model, taking into consideration capacitive crosstalk. We have shown that capacitive crosstalk has a significant effect on gate energy dissipation. Our results confirm that the dissipated energy from CMOS switching gates depends not only on the hamming distance (HD), but also on the direction of switching activity on nearby data lines. We show that for an 8 bit data bus, crosstalk may improve detection performance from 2.5 bits (HD based detector) to a theoretical 5.7 bits and simulated 5.0 bits (crosstalk based detector) of information per sample.

In this paper we elaborate on the theory and simulations of layout dependent phenomena and how they must be considered when analyzing security implications of power and electromagnetic side-channels. We have also added a small case study, i.e. the electromagnetic side-channel of a smart card, that supports our simulations/theoretical results.

Index Terms— **Crosstalk, Power model, Switching CMOS, Side-channels, Classification, Entropy**

I. Introduction

When a microprocessor executes its program, power consumption (or resulting electromagnetic emanation) can be used to reveal the contents of program and/or data memory of the microprocessor. The correlation between power consumption and microprocessor activity has found many uses: to recover cryptographic keys [2], [3], [10]–[12], to reveal hidden hardware faults (trojans) on integrated circuits [1], to control the emanation through subversive software in the Wireless Covert Channel Attack [7] and to reverse engineer the code executed by microprocessors [14].

In side-channel attacks, a common power model used to simulate the power consumption is the Hamming Distance (HD) model, as it is simple and generic [12]. The model assumes the power consumption to be proportional to the number of transitions taking place. If this assumption is appropriate, signals transmitted on a parallel bus (e.g. intermediate values of the cryptographic algorithm) with the same HD should have equal power consumption and therefore be indistinguishable. This is not always the case, e.g. if Bayes classifier is used, as suggested by the template attack [3]. It has also been demonstrated in [8] that signals with the same number of transitions can be classified using a modified template attack.

The phenomena behind this may be known in the security community, but has received little attention. One paper by Z. Chen, S. Haider and P. Schaumont [4], investigates the effect of the coupling capacitance on masking schemes without a detailed examination of the phenomena. In their book ”Power Analysis Attacks”, S. Mangard, E. Oswald and T. Popp [12] mention power simulation at analog level as ”the most precise way to simulate the power consumption of digital circuits...”. Parasitic elements, such as capacitances between the wires and unwanted capacitances in the transistors are mentioned, however, it is also stated that it is very common to make simplifications by lumping together extrinsic and intrinsic capacitances into a single capacitance to ground. This will, in fact, make the model incapable of explaining the results we are addressing in this paper.

Parasitic couplings, and the coupling capacitance in particular, are however, a great concern within sub-micron VLSI design [5], [13], [15]. CMOS technology is currently being pushed into deep sub-micron range. As the number of transistors increase, the need for on-chip wiring increases as well and must be scaled accordingly. Parasitic couplings between interconnects, such as on-chip busses, must be taken seriously as they influence both the power consumption and maximum obtainable speed [5]. F. Moll, M. Roca and E. Isern [13] did a detailed analysis of the energy dissipation from two metal lines running close together. The lines were driven by CMOS inverters and transitions in one or two wires were studied. The effect of coupling capacitance between the two lines on the power consumption was shown analytically and simulated

doi:10.4304/jcp.7.4.827-837

in HSPICE. The main result was that if two bus lines have transitions in the same or opposite direction at the same time, the total energy is either lower or higher than if the two transitions are treated independently. This is due to the coupling capacitance. C. Duan, V.H.C. Calle and S.P. Khatri [5] focus on crosstalk avoidance codes that aim to reduce the effect of the coupling capacitances by avoiding specific data transition patterns. Their model considers coupling capacitance, C_C, between three adjacent lines. They show that 3 bit transition patterns can be divided into 5 crosstalk classes based on the influence of the coupling capacitances, C_C. The energy consumption therefore depends on which crosstalk class the transition pattern belong to, as seen in Table I reprinted from [5].

TABLE I.
CLASSES OF CROSSTALK FROM [5]. C_{eff} IS THE EFFICIENT CAPACITANCE, C_L THE LOAD CAPACITANCE AND $\lambda = C_C/C_L$

Class	C_{eff}	Transition pattern
0C	C_L	000 → 111
1C	$C_L(1+\lambda)$	011 → 000
2C	$C_L(1+2\lambda)$	010 → 000
3C	$C_L(1+3\lambda)$	010 → 100
4C	$C_L(1+4\lambda)$	010 → 101

The focus within VLSI design, such as [5] and [13] is on power consumption and delays caused by the coupling capacitance. They do not consider security implications, such as the ability to use the variation in energy consumption to classify transition patterns. However, correlations between data and energy consumption are exactly what side-channel attacks, such as DPA and Template attack, rely upon.

In this paper we elaborate on the hypothesis put forward in [9] that layout dependent phenomena, such as capacitive coupling between wires, can explain why it sometimes is possible to distinguish transition patterns with the same HD. We extend the theory and simulations of how the new power model, which takes into account capacitive crosstalk, affect our ability to classify activity in a microprocessor.

We look at the total dissipated energy from a parallel data bus driven by CMOS inverters. Our model is a generalization of [13], with inverters consisting of two MOSFET transistors, a load capacitance C_L connected to each inverter output and a coupling capacitance C_C connected between each bus line. Our model is generalized to n lines and simulations in PSPICE are done with eight bus lines.

The purpose of our simulation is to show that when the dissipated energy depends on the direction of change of nearby data lines, and not only the number of transitions taking place, the number of possible energy levels dissipating from the bus will increase, thus allowing classification of a larger number of transition patterns. Our hypothesis is that this can be used to explain why some signal with the same HD can be distinguished. Our model can easily take into consideration other layout dependent phenomena, potentially offering an explanation to classification of an even larger set of transition patterns. We will use entropy as our classifier performance indicator and show that a detector capable of detecting energy levels due to crosstalk can extract more information than a detector based on HD only.

Finally, in order to probe the practicality of our theory and simulations, we have included a small case study, in which the objective is to see if analysis of electromagnetic side-channel information also supports the division into crosstalk energy levels.

This paper is organized as follows: Section II presents the hypothesis of layout dependent phenomena. Section III presents our model and necessary theory to calculate the energy dissipation. Section IV is an analytic analysis of security implications. Section V presents simulation results. Section VI presents a case study and future work. Finally, a conclusion is drawn in Section VII.

II. LAYOUT DEPENDENT PHENOMENA

In a physical implementation of any circuit (e.g. CMOS based microprocessor) a number of phenomena will influence the energy dissipation and the resulting radiated electromagnetic field. These phenomena include inductance and capacitance of conductors, inductance and capacitance between conductors, wireless transmission characteristics (i.e. antenna properties) of conductors and other circuit elements and complex combinations of these phenomena. These phenomena apply to any transistors and wires in a circuit, but we choose to look at a portion of wires running parallel, as we expect them to be relatively good antennas and therefore a good source for side-channel information. This is illustrated in the model of a parallel bus, driven by CMOS inverters, seen in Fig. 1.

A. Inductance and Capacitance of Conductors

Any conductor, W_j, carrying an electric current will have an associated distributed resistance R_j, inductance L_j, conductance G_j and capacitance C_j, expressed as a characteristic impedance, Z_{0j}. The characteristic impedance is often modeled as an infinite series of lumped components. The inductance L_j and capacitance C_j will both block high frequency signals and act as a low pass filter. Small variations in the length and width of conductors result in small variations in the inductance. Small variations in the area and distance to ground plane result in small variations in the capacitance. There will therefore be small variations in how signals on different conductors (e.g. bus lines) are filtered.

B. Inductance and Capacitance between Conductors

Crosstalk can be defined as the coupling of energy between two conductors. Inductive coupling is caused by mutual inductance, $L_{j,j+1}$, (i.e. magnetic field) and capacitive coupling is caused by mutual capacitance, $C_{j,j+1}$, (i.e. electric field) between wire j and $j+1$. These couplings occur along the entire length of the conductor, but are also modeled as lumped components (Fig. 1). The

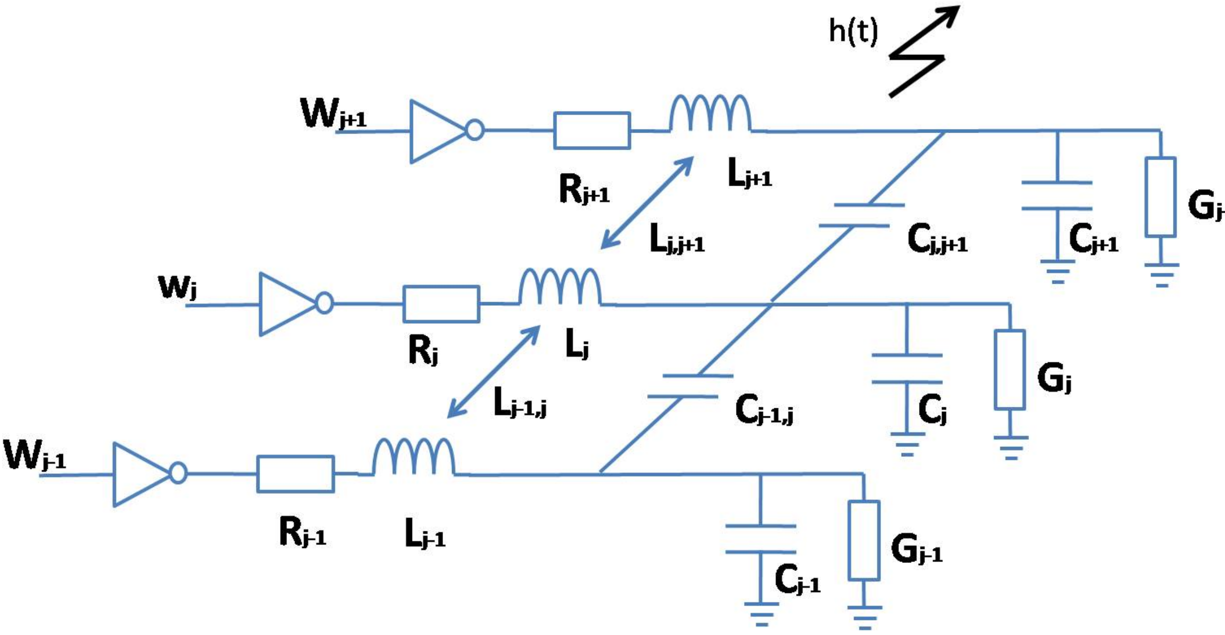

Figure 1. Model of layout dependent phenomena

interaction of magnetic and electric fields will effectively change the characteristic impedance, Z_{0j}, associated with the conductor. This interaction is layout dependent (e.g. distance and length of wires) and will effect both delays and energy dissipation. An important property of crosstalk is its dependency on the activity on the wires. F. Moll, M. Roca and E. Isern [13] state that, ”coupling capacitance is very different from the capacitance to ground because it depends on the switching activity... ”. If two lines are low and rise at the same time, the mutual capacitance coupling, $C_{j,j+1}$, does not have to be charged. However, if one line remains low and the other rises, $C_{j,j+1}$ must be charged, resulting in increased rise time and power consumption.

C. Wireless Transmission Characteristics

Any circuit element in the microprocessor, conducting electric current, can be considered an antenna. An antenna is a transducer converting electric current into electromagnetic waves, characterized by properties such as: resonant frequency, gain, radiation pattern, impedance, efficiency, bandwidth and polarization. These properties depend on factors such as: amount of current, length/shape and material of the circuit element. In addition, the electromagnetic waves will be influenced by filtering, reflection and interference from surrounding material and circuit elements. The relationship between the current (i.e. power consumption) and the electromagnetic wave can be expressed by a transfer function $h(t)$ (Fig. 1). Predicting $h(t)$ is not trivial, if possible at all, as most physical systems are non linear by nature. This is left for future work, but it is a fair assumption that relatively long bus lines are good antennas.

D. Complex Combinations of Factors

Finally, complex combinations of layout dependent phenomena may be the key to identify minute differences in microprocessor activity, e.g. the radiation efficiency of bus lines combined with data and layout dependencies of the line characteristics due to crosstalk suggest that the emanation detected will have data and layout dependent variations in power consumption and delay. In the following, we will assume that the coupling capacitance is the dominating factor, and show how this can explain why some signals with the same HD can be distinguished. This will show the potential effect of layout dependent phenomena on classifying microprocessor activity. Our work can easily be extended by including other layout dependent phenomena if a more precise model is needed.

III. Theoretical Considerations

By limiting the model to only coupling and load capacitances, the model in Fig. 1 can be simplified as seen in Fig. 2. This is a generalization of the model for two lines used in [13] and includes a model of the CMOS inverter.

In order to run simulations in PSPICE, we need an expression for the total energy dissipation, E_T. The energy dissipation for wire j in the p and n type transistor can be expressed as:

$$E_{pj} = \int (V_{DD} - V_j) i_{pj} dt \tag{1}$$

$$E_{nj} = \int V_j i_{nj} dt \tag{2}$$

The overall energy dissipation for an n wire bus is then

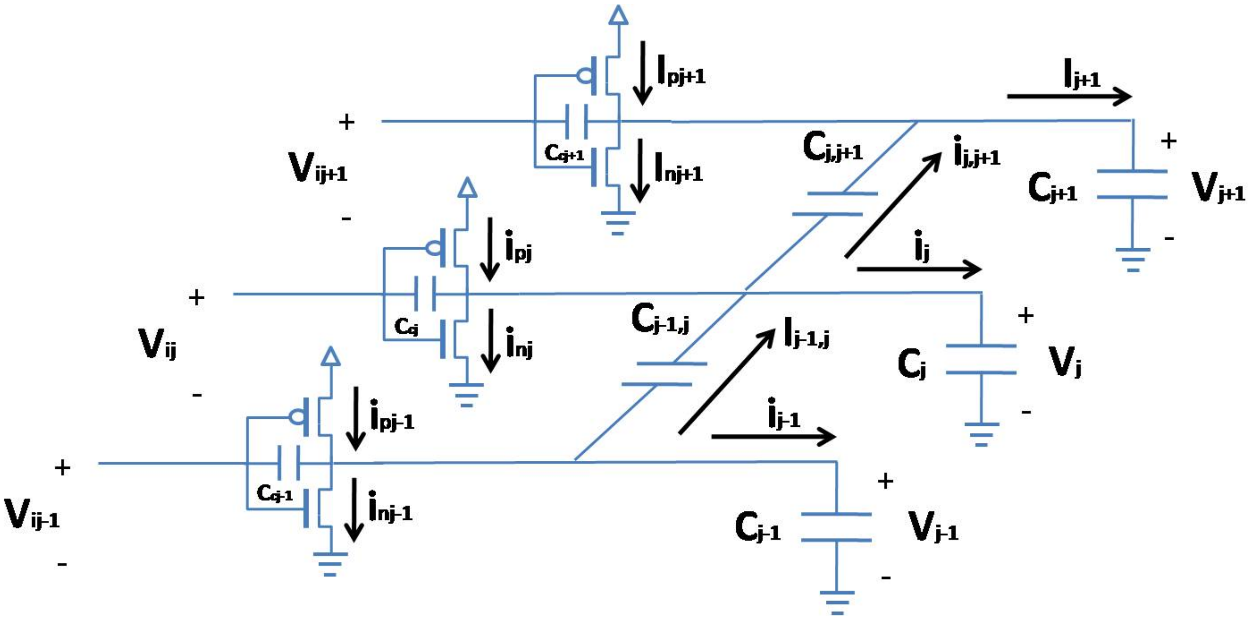

Figure 2. Simplified model, assuming load and coupling capacitances to be dominant

given by:

$$E_T = \sum_{j=1}^{n} (E_{pj} + E_{nj}) \tag{3}$$

Combining and rearranging (1), (2) and (3) the overall energy dissipation can be written as:

$$E_T = \sum_{j=1}^{n} V_{DD} \int i_{pj} dt - \sum_{j=1}^{n} \int V_j (i_{pj} - i_{nj}) dt \tag{4}$$

Using Kirchhoff's circuit laws and the current voltage relationship $i(t) = C\frac{dV(t)}{dt}$, the terms $(i_{pj} - i_{nj})$ can be written as:

$$\begin{aligned} i_{pj} - i_{nj} &= (C_j + C_{j,j+1} + C_{j-1,j} + C_{cj})\frac{dV_j}{dt} \\ &-C_{cj}\frac{dV_{ij}}{dt} - C_{j,j+1}\frac{dV_{j+1}}{dt} - C_{j-1,j}\frac{dV_{j-1}}{dt} \end{aligned} \tag{5}$$

Notice that the results in [13] are easily found from (4) and (5) by setting $n = 2$ (two adjacent lines). Equation (4) is used in PSPICE to simulate the total energy dissipation, $\hat{E}_T$, with the following assumptions:

- The transitions on the data bus are concurrent in time. It has been shown [13] that the effect of the coupling capacitance is maximum when transitions occur simultaneously on all bus lines.
- The load capacitances for data bus lines are identical ($C_j = C_L$ for $j = \{1, 2, \cdots, n\}$)
- Coupling capacitances are only found between adjacent line and are identical ($C_{j,j+1} = C_C$ for $j = \{1, 2, \cdots, n-1\}$)

These assumptions are not unrealistic in real bus architecture on a device. If, however, the transitions are shifted in time with more than the rise time of signal, the effect of the coupling capacitance is reduced and the transitions can be regarded as single transitions [13].

In order to compare the simulated energy dissipation ($\hat{E}_T$) with analytic values (E_T), different expressions than (4) and (5) are needed.

It is only when the individual line has a transition, that it is subject to capacitive crosstalk. Quantifying this crosstalk influence has to take into consideration voltage changes on the line itself and one (edges) or two adjacent lines. Let $\delta_j \in \{0, \pm 1\}$ be the normalized voltage change on line j, then the voltage change between two lines j and k is $\delta_{j,k} = \delta_j - \delta_k$. The crosstalk influence α_j on line j can then be defined as:

$$\alpha_j = \begin{cases} 0 & \textit{no transition line } j \\ |\delta_{j,j-1} + \delta_{j,j+1}| & \textit{otherwise} \end{cases} \tag{6}$$

It can be shown that $\alpha_j = \{0, 1, 2\}$ for lines with only one adjacent line (edges), and $\alpha_j = \{0, 1, 2, 3, 4\}$ for lines with two adjacent lines. Let the total crosstalk influence for an n line bus be called a crosstalk index α, defined as the sum of the crosstalk influence of each line:

$$\alpha = \sum_{j=1}^{n} \alpha_j \tag{7}$$

If the contributions from the load (C_L) and coupling capacitance (C_C) are dominant to the dissipated energy, then E_T can be expressed by the following power model [9]:

$$E_T = \frac{1}{2} C_L V_{DD}^2 (k + \alpha\lambda) = E_0(k + \alpha\lambda) \tag{8}$$

where $E_0 = \frac{1}{2} C_L V_{DD}^2$, V_{DD} is the power supply voltage, k is the number of transitions on the data bus, $\lambda = C_C/C_L$ and α is the crosstalk index of (7) indicating the coupling capacitance induced crosstalk, similar to the crosstalk classes in [5].

In the next section we will use (8) to analyze which transition patterns can be distinguished.

IV. Security Implications

The relationship between energy dissipation, number of transitions, crosstalk index, load capacitance and coupling capacitance in (8) can be used to analyze delays and energy dissipation of sub-micron VLSI design [5], [13], [15]. However, we are interested in the security implications of layout dependent phenomena, and in this paper the coupling capacitance in particular. How will a power model (8) that includes coupling capacitance affect our ability to predict the energy dissipation of activity in a microprocessor, such as data transfer on a parallel bus?

Let T be the set of possible transitions on an n bit parallel bus. Since "no transition" can be both $0 \rightarrow 0$ and $1 \rightarrow 1$ there are $|T| = 4^n$ possible transition patterns for an n-bit bus. Assuming that each transition pattern's energy dissipation is unique, a model should ideally predict a total of $|T| = 4^n$ energy levels. This may not be possible if physical properties are such that multiple transition patterns indeed use the same amount of energy.

Classification by energy dissipation can only distinguish transition patterns by the distinct energy levels explained by the model. A model that assumes energy dissipation proportional to the number of transition, can therefore only distinguish transition pattern into subsets T_k, $k = \{0, \cdots, n\}$ being subsets of T that has k transitions. The number of transition patterns in each subset is given by: $|T_k| = 2^n \binom{n}{k}$. The total number of possible transitions on an 8 wire bus ($|T| = 65536$) can be divided into 9 subsets, $T_0, T_1, \cdots, T_8$ based on the number of transitions, k. The energy dissipation, E_T (using (8) with $\alpha = 0$), associated with each subset $|T_k|$ can be seen in Table II. A model that assumes energy dissipation proportional to the number of transition, can only classify transition pattern by the energy level of these 9 subsets. For example, in Table II there are 14336 transition patterns with energy level $3E_0$ that are indistinguishable by the number of transitions alone.

Using the new power model (8), taking into consideration the coupling capacitor, each subset T_k can be split into a number of new energy levels. This gives a number of smaller subsets T_k^α, $|T_k| > |T_k^\alpha|$ and $\sum_{\forall \alpha \in q_k} |T_k^\alpha| = |T_k|$, where α is the crosstalk index of (7) and q_k is the set of possible values of α for k transitions.

Computing $|T_k^\alpha|$ for a fixed number of bus lines n can be done by constructing a table of $(2^k)^2$ elements corresponding to all possible transition patterns. For each of these, first compute the crosstalk index α (7), then the energy dissipation E_T (8). $|T_k^\alpha|$ can then be computed by counting the table entries for each tuple $\{k, \alpha\}$. Notice that for a finite n, there are restrictions on the sets q_k of possible values of α. As the number of transitions increase, all energy levels are not possible. This applies to 6,7 and 8 transitions for an 8 bit bus.

The results for an 8 bit bus can be seen in Table II. The results show that taking into consideration the coupling capacitance increases the number of energy levels from 9 in the HD model to 93 in the crosstalk model, e.g. the 14336 transition patterns with 3 transitions previously indistinguishable can now be split into 10 energy levels. The largest increase in energy levels is found for 6 transitions with 21 new energy levels. Note that energy level $\alpha = 20$ does not exist.

Also notice that given an ideal classifier, there is no confusion between subsets of the same k as they all have unique energy levels. There may, however, be confusion between subsets of different k. The extent of this confusion is architecture dependent, expressed by λ, e.g. subset T_2^6 has the same energy level as T_3^2 if $\lambda = 1/4$, in case they should be treated as one subset. It is easy to show that confusion between transition A (energy E_{TA}, k_A transitions and crosstalk index α_A) and B (energy E_{TB}, k_B transitions and crosstalk index α_B) happens when:

$$\lambda_{AB} = \frac{k_B - k_A}{\alpha_A - \alpha_B} \tag{9}$$

λ_{AB} values that are close to the real $\lambda = C_C/C_L$ indicate subsets that will be difficult to distinguish.

Finally, we have only shown how to split the subset T_k into smaller subsets T_k^α by considering the effect of the coupling capacitance (i.e. α). This idea can easily be generalized, such that T_k is split into subsets T_k^β, where $|T_k| > |T_k^\beta|$, and β is the influence of other layout dependent phenomena. Examples of phenomena for future work include: variations in coupling and load capacitance, coupling capacitance between line j and $j+2$, inductance, effect of bends in circuit paths and multi-layer capacitance (3-dimensional). We believe that the key to identify minute differences in microprocessor activity is to combine several layout dependent phenomena, $\beta_1, \cdots, \beta_m$, such that:

$$|T_k| > |T_k^{\beta_1}| > |T_k^{\beta_1+\beta_2}| > \cdots > |T_k^{\sum_{i=1}^m \beta_i}| \tag{10}$$

A. Classification Performance

Table II shows that, taking into consideration the coupling capacitance, we are able to increase the number of subsets (or energy levels) T_k to T_k^α. For the purpose of comparing alternative detectors we will assume uniform random transition. Thus for an 8 bit bus we would like the detector to extract 16 bits of information, i.e. high or low (2 bits of information) for each of the 8 wires. We will use entropy as our classifier performance indicator. The entropy (i.e. bits of information) for a detector, when there are r energy levels, can be calculated using:

$$H(x) = -\sum_{i=1}^{r} p(x_i) log\, p(x_i) \tag{11}$$

In the following, we have assumed an 8 bit bus width, thus there are $4^8 = 65536$ possible transitions. Call the detector that can extract 16 bits of information a level detector. If we assume that one only has bus activity when initial and final state are different, and that $0 \rightarrow 1$ and $1 \rightarrow 0$ can be distinguished, an observation will give us the following entropy: $-(1/2 log 1/2 + 1/4 log 1/4 + 1/4 log 1/4) = 3/2$ bits as we cannot distinguish $0 \rightarrow 0$

TABLE II.
THE TABLE SHOWS THE NUMBER OF TRANSITION PATTERNS, WITHOUT ($|T_k|$) AND WITH ($|T_k^\alpha|$) CROSSTALK INFLUENCE, BELONGING TO A CERTAIN ENERGY LEVEL, E_T. k IS THE NUMBER OF TRANSITIONS (HAMMING DISTANCE) AND α IS THE CROSSTALK INDEX

k	E_T [pJ]	$\|T_k\|$	α	E_T [pJ]	$\|T_k^\alpha\|$
0	0	256	0	0	256
1	E_0	2048	1	$E_0(1+\lambda)$	512
			2	$E_0(1+2\lambda)$	1536
2	$2E_0$	7168	1	$E_0(2+1\lambda)$	256
			2	$E_0(2+2\lambda)$	896
			3	$E_0(2+3\lambda)$	2560
			4	$E_0(2+4\lambda)$	2560
			5	$E_0(2+5\lambda)$	256
			6	$E_0(2+6\lambda)$	640
3	$3E_0$	14336	1	$E_0(3+1\lambda)$	128
			2	$E_0(3+2\lambda)$	512
			3	$E_0(3+3\lambda)$	2048
			4	$E_0(3+4\lambda)$	2560
			5	$E_0(3+5\lambda)$	3328
			6	$E_0(3+6\lambda)$	1792
			7	$E_0(3+7\lambda)$	2048
			8	$E_0(3+8\lambda)$	1536
			9	$E_0(3+9\lambda)$	128
			10	$E_0(3+10\lambda)$	256
4	$4E_0$	17920	1	$E_0(4+\lambda)$	64
			2	$E_0(4+2\lambda)$	288
			3	$E_0(4+3\lambda)$	1152
			4	$E_0(4+4\lambda)$	1728
			5	$E_0(4+5\lambda)$	2496
			6	$E_0(4+6\lambda)$	1824
			7	$E_0(4+7\lambda)$	2816
			8	$E_0(4+8\lambda)$	2304
			9	$E_0(4+9\lambda)$	2496
			10	$E_0(4+10\lambda)$	864
			11	$E_0(4+11\lambda)$	1152
			12	$E_0(4+12\lambda)$	576
			13	$E_0(4+13\lambda)$	64
			14	$E_0(4+14\lambda)$	96
5	$5E_0$	14336	1	$E_0(5+\lambda)$	32
			2	$E_0(5+2\lambda)$	160
			3	$E_0(5+3\lambda)$	512
			4	$E_0(5+4\lambda)$	896
			5	$E_0(5+5\lambda)$	896
			6	$E_0(5+6\lambda)$	1024
			7	$E_0(5+7\lambda)$	1536
			8	$E_0(5+8\lambda)$	1920
			9	$E_0(5+9\lambda)$	1728
			10	$E_0(5+10\lambda)$	1088
			11	$E_0(5+11\lambda)$	1536
			12	$E_0(5+12\lambda)$	1152
			13	$E_0(5+13\lambda)$	896
			14	$E_0(5+14\lambda)$	256
			15	$E_0(5+15\lambda)$	512
			16	$E_0(5+16\lambda)$	128
			17	$E_0(5+17\lambda)$	32
			18	$E_0(5+18\lambda)$	32
6	$6E_0$	7168	1	$E_0(6+\lambda)$	16
			2	$E_0(6+2\lambda)$	88
			3	$E_0(6+3\lambda)$	160
			4	$E_0(6+4\lambda)$	320
			5	$E_0(6+5\lambda)$	80
			6	$E_0(6+6\lambda)$	360
			7	$E_0(6+7\lambda)$	640
			8	$E_0(6+8\lambda)$	960
			9	$E_0(6+9\lambda)$	160
			10	$E_0(6+10\lambda)$	560
			11	$E_0(6+11\lambda)$	960
			12	$E_0(6+12\lambda)$	960
			13	$E_0(6+13\lambda)$	160
			14	$E_0(6+14\lambda)$	400
			15	$E_0(6+15\lambda)$	640
			16	$E_0(6+16\lambda)$	320
			17	$E_0(6+17\lambda)$	80
			18	$E_0(6+18\lambda)$	120
			19	$E_0(6+19\lambda)$	160
			21	$E_0(6+21\lambda)$	16
			22	$E_0(6+22\lambda)$	8
7	$7E_0$	2048	1	$E_0(7+\lambda)$	8
			2	$E_0(7+2\lambda)$	48
			5	$E_0(7+5\lambda)$	48
			6	$E_0(7+6\lambda)$	240
			9	$E_0(7+9\lambda)$	120
			10	$E_0(7+10\lambda)$	480
			13	$E_0(7+13\lambda)$	160
			14	$E_0(7+14\lambda)$	480
			17	$E_0(7+17\lambda)$	120
			18	$E_0(7+18\lambda)$	240
			21	$E_0(7+21\lambda)$	48
			22	$E_0(7+22\lambda)$	48
			25	$E_0(7+25\lambda)$	8
8	$8E_0$	256	0	$E_0(8)$	2
			4	$E_0(8+4\lambda)$	14
			8	$E_0(8+8\lambda)$	42
			12	$E_0(8+12\lambda)$	70
			16	$E_0(8+16\lambda)$	70
			20	$E_0(8+20\lambda)$	42
			24	$E_0(8+24\lambda)$	14
			28	$E_0(8+28\lambda)$	2

from $1 \rightarrow 1$, but $0 \rightarrow 0$, $0 \rightarrow 1$, $1 \rightarrow 0$ can be distinguished. Thus, each observation will give us $3/2$ bits per line. The theoretical optimum for an 8 bit bus with a 'transition detector' would be $8 \cdot 3/2\ bits = 12\ bits$, assuming all observable transitions are distinguishable. In other words, by observing transitions rather than levels, we loose 4 bits ($1/2$ bit per line) compared to the setting where we would observe the states.

Using the results of Table II, we can now calculate the entropy extracted by a detector that can distinguish HD only (T_k) and a detector that can distinguish energy levels due to crosstalk (T_k^α).

The entropy extracted by a HD detector is found using (11) with 9 energy levels ($r = 9$) and $p(x_i) = |T_{i-1}|/65536$ ($|T_{i-1}|$ from column 3 and 9 Table II) giving an entropy of 2.5 bits. The entropy extracted by a crosstalk detector is found using (11) with 93 energy levels ($r = 93$) and $p(x_i) = |T_{i-1}^\alpha|/65536$ ($|T_{i-1}^\alpha|$ from column 6 and 12 Table II) giving an entropy of 5.7 bits.

The difference between the ideal value of a level detector and the entropy extracted by other detectors, represent the amount of guessing needed for classifying an observation. By considering the coupling capacitance and not only HD, we extract more information out of each observation, therefore reducing the amount of "guessing" needed for classification. In the next section we present simulations validating the effect of the coupling capacitance.

V. SIMULATIONS

The simulations are performed in PSPICE with $C_L = 400fF$, $C_C = 250fF$, $V_{dd} = 3V$ and a rise- and fall-time of $200ps$ of the input voltages (same as [13]). The inverter drivers are equal and balanced. Equation (4) is used in PSPICE to find the simulated energy dissipation $\hat{E}_T$.

TABLE III.
DISSIPATED ENERGY WHEN CONSIDERING CROSSTALK FOR 2 ADJACENT WIRES

Transition pattern	**Transitions k**	**Crosstalk** α	**Theoretical** E_T **[pJ]**	**Simulated** $\hat{E}_T$ **[pJ]**
00 → 01	1	1	2.9	2.7
00 → 10	1	1	2.9	2.7
00 → 11	2	0	3.6	3.5
01 → 10	2	4	8.1	8.0

TABLE IV.
DISSIPATED ENERGY WHEN CONSIDERING CROSSTALK FOR BUS WITH 3 LINES

Transition pattern	**Transitions k**	**Crosstalk** α	**Theoretical** E_T **[pJ]**	**Simulated** $\hat{E}_T$ **[pJ]**
000 → 111	3	0	5,4	5.4
000 → 011	2	1	4,7	4.4
000 → 010	1	2	4,1	3.9
010 → 100	2	5	9,2	9.5
010 → 101	3	8	14,4	14.2

A. *Model Validation*

Simulations were initially carried out and compared with the results of [5], [13] as a model validation. The results are shown in Table III and IV. Transition pattern refers to transitions in the output voltage V_j (Fig. 2) and also shows the number of bus lines used. Column 2 is the number of transitions k followed by the crosstalk index α. Theoretical energy, E_T, is calculated from (8) and simulated energy, $\hat{E}_T$ is from PSPICE simulations.

The simulations for two lines are consistent with [13]. For two wires, as seen in Table III, it is clear that the energy dissipation for two simultaneous transitions is either lower or higher than if treated as two single transitions, depending on the direction of the transitions, as expected. This means that introducing the coupling capacitance it is possible to explain a difference in the energy dissipation for transition patterns $00 \leftrightarrow 11$ from $01 \leftrightarrow 10$. Without this difference in energy dissipation the two transition patterns should not be distinguishable.

Simulations of three lines confirms the difference in energy dissipation of the 5 crosstalk classes (Table I) introduced in [5]. Notice that only the transition pattern with the same number of transitions (first and last, second and fourth) can be used to evaluate the effect of the coupling capacitance.

The small differences between analytic and simulated energy dissipation can be explained by simplifications in deriving (8) (e.g. omitting leakage currents, such as short-circuit and sub-threshold currents). Having validated our model, all the following simulations are done on an 8 bit bus.

B. *Results and Discussion*

Simulation results for 8 lines are shown in Table V. The table is not exhaustive, but includes results for all possible subsets T_k^α.

The simulated energy levels $\hat{E}_T$ are similar to the analytic values E_T. The results confirm that energy consumption is proportional to the number of transitions and the crosstalk index, α. The crosstalk index depends on switching activity on adjacent lines and position, edge (one adjacent wire) or middle (two adjacent wires). As seen in Table V, the results also confirm that there is no confusion between energy levels for subsets of an equal number of transitions. However, there may be some confusion between some of the 93 subgroups, e.g. the energy dissipation of subset T_2^6 and T_3^4 are almost equal. This is expected as $\lambda_{AB} = 0,5$ is close to $\lambda = 0,63$ used in this experiment. Other examples can be found and this reduces the number of subsets depending on how accurate our detector is. A theoretical crosstalk detector capable of separating all 93 energy levels can extract 5.7 bits of information. It is therefore expected that a practical crosstalk detector will extract less information, due to some subset having almost equal energy levels. Which of the simulated energy levels that should be considered indistinguishable will depend on the accuracy of the detector and the number of observations available. A random loss of 20% of the subsets will still, on average, have an entropy of 5.0. Even with this loss due to similar energy levels, the information gain is still 2.5 bits compared to the HD detector. The performance of the detectors is summarized in Table VI.

TABLE VI.
COMPARING THE ABILITY TO EXTRACT INFORMATION OF DIFFERENT DETECTORS FOR AN 8 WIRE BUS

Type of detector	Entropy (information) [bits]
Level detector	16,0
Optimum transition detector	12,0
Crosstalk detector (theoretical)	5,7
Crosstalk detector (simulated)	5,0
HD detector	2,5

VI. CASE STUDY AND FUTURE WORK

In order to probe the practicality of our theory and simulations, we collected a small set of experimental data. The objective was to see if analysis of electromagnetic side-channel information also supports the division into crosstalk energy levels of Table II and V.

When classifying two transition patterns by their energy dissipation, we expect a lower probability of error (P_e) when the difference in energy level is large and higher P_e as the difference in energy levels decreases. When the energy dissipation of two transition patterns are equal, we don't expect to be able to do any better than flipping a coin. Transition patterns with an unequal number of transitions have a relatively large difference in energy dissipation and are therefore fairly easy to distinguish, as shown in [8].

Consider two transition patterns A (crosstalk index α^A) and B (crosstalk index α^B) of an equal number of transitions. Let α-distance, $\Delta\alpha = |\alpha^A - \alpha^B|$, be the difference in crosstalk index between transition patterns A and B. According to our model (8), patterns with $\Delta\alpha = 0$ dissipate the same amount of energy and

TABLE V.
ANALYTIC (E_T) AND SIMULATED ($\hat{E}_T$) DISSIPATED ENERGY WHEN CONSIDERING CROSSTALK FOR BUS WITH 8 LINES. K IS THE NUMBER OF TRANSITIONS (HAMMING DISTANCE) AND α IS THE CROSSTALK INDEX

Transition pattern	k	α	E_T [pJ]	$\hat{E}_T$ [pJ]	Transition pattern	k	α	E_T [pJ]	$\hat{E}_T$ [pJ]
0000 0000 → 0000 0001	1	1	2,9	2.9	0000 0000 → 0011 1111	6	1	11,9	12.3
0000 0000 → 0000 0010	1	2	4,1	4.1	0000 0000 → 1001 1111	6	2	13,1	13.4
0000 0000 → 0000 0011	2	1	4,7	4.8	0000 0000 → 0101 1111	6	3	14,2	14.5
0000 0000 → 1000 0001	2	2	5,9	5.9	0000 0000 → 1010 1111	6	4	15,3	15.7
0000 0000 → 0000 0101	2	3	7,0	7.0	0000 0001 → 0011 1110	6	5	16,4	16.7
0000 0000 → 0000 1010	2	4	8,1	8.1	0000 0001 → 1001 1110	6	6	17,6	17.9
0000 0010 → 0000 0001	2	5	9,2	9.3	0000 0001 → 0101 1110	6	7	18,7	19.0
0000 0100 → 0000 0010	2	6	10,4	10.1	0000 0001 → 1010 1110	6	8	19,8	20.0
0000 0000 → 0000 0111	3	1	6,5	6.7	0000 0010 → 0011 1101	6	9	20,9	20.8
0000 0000 → 1000 0011	3	2	7,7	7.8	0000 0010 → 1001 1101	6	10	22,1	21.9
0000 0000 → 0000 1011	3	3	8,8	8.9	0000 0010 → 0101 1101	6	11	23,2	23.0
0000 0000 → 1000 1001	3	4	9,9	10.0	0000 0010 → 1010 1101	6	12	24,3	24.1
0000 0000 → 0001 0101	3	5	11,0	11.1	0000 0101 → 0011 1010	6	13	25,4	25.5
0000 0000 → 0010 1010	3	6	12,2	12.3	0000 0101 → 1001 1010	6	14	26,6	26.5
0000 0010 → 0000 1001	3	7	13,3	13.3	0000 0101 → 0101 1010	6	15	27,7	27.7
0000 0100 → 0001 0010	3	8	14,4	14.4	0000 0101 → 1010 1010	6	16	28,8	28.9
0000 0010 → 0000 0101	3	9	15,5	15.4	0000 1010 → 0011 0101	6	17	29,9	29.7
0000 0100 → 0000 1010	3	10	16,7	16.6	0000 1010 → 1001 0101	6	18	31,1	30.9
0000 0000 → 0000 1111	4	1	8,3	8.6	0000 1010 → 0101 0101	6	19	32,2	32.0
0000 0000 → 1000 0111	4	2	9,5	9.6	0010 1010 → 0001 0101	6	21	34,4	34.0
0000 0000 → 0001 0111	4	3	10,6	10.8	0101 0100 → 0010 1010	6	22	35,6	35.2
0000 0000 → 1000 1011	4	4	11,7	11.9	0000 0000 → 0111 1111	7	1	13,7	14.2
0000 0000 → 0010 1011	4	5	12,8	13.0	0000 0000 → 1011 1111	7	2	14,9	15.3
0000 0000 → 1001 0101	4	6	14,0	14.2	0000 0001 → 0111 1110	7	5	18,2	18.5
0000 0000 → 0101 0101	4	7	15,1	15.3	0000 0001 → 1011 1110	7	6	19,4	19.7
0000 0010 → 1001 0001	4	8	16,2	16.0	0000 0010 → 0111 1101	7	9	22,7	22.7
0000 0100 → 1001 0010	4	9	17,3	17.5	0000 0010 → 1011 1101	7	10	23,9	23.8
0000 0010 → 1000 0101	4	10	18,5	18.3	0000 0101 → 0111 1010	7	13	27,2	27.4
0000 0010 → 0100 0101	4	11	19,6	19.4	0000 0101 → 1011 1010	7	14	28,4	28.5
0000 0100 → 0100 1010	4	12	20,7	20.7	0000 1010 → 0111 0101	7	17	31,7	31.6
0000 1010 → 0000 0101	4	13	21,8	21.5	0000 1010 → 1011 0101	7	18	32,9	32.6
0001 0100 → 0000 1010	4	14	23,0	22.7	0001 0101 → 0110 1010	7	21	36,2	36.2
0000 0000 → 0001 1111	5	1	10,1	10.4	0001 0101 → 1010 1010	7	22	37,4	37.3
0000 0000 → 1000 1111	5	2	11,3	11.5	0010 1010 → 0101 0101	7	25	40,7	40.5
0000 0000 → 0010 1111	5	3	12,4	12.6	0000 0000 → 1111 1111	8	0	14,4	14.9
0000 0000 → 1001 0111	5	4	13,5	13.8	0000 0001 → 1111 1110	8	4	18,9	19.3
0000 0000 → 0101 0111	5	5	14,6	14.9	0000 0010 → 1111 1101	8	8	23,4	23.4
0000 0000 → 1010 1011	5	6	15,8	16.0	0000 0101 → 1111 1010	8	12	27,9	28.1
0000 0010 → 0111 0001	5	7	16,9	16.8	0000 1010 → 1111 0101	8	16	32,4	32.3
0000 0010 → 1011 0001	5	8	18,0	17.9	0001 0101 → 1110 1010	8	20	36,9	36.9
0000 0010 → 0101 1001	5	9	19,1	19.3	0010 1010 → 1101 0101	8	24	41,4	41.1
0000 0010 → 1010 1001	5	10	20,2	20.4	0101 0101 → 1010 1010	8	28	45,9	45.6
0000 0010 → 0110 0101	5	11	21,4	21.3					
0000 0010 → 1010 0101	5	12	22,5	22.5					
0000 0010 → 0101 0101	5	13	23,6	23.5					
0000 1010 → 1000 0101	5	14	24,8	24.5					
0000 1010 → 0100 0101	5	15	25,9	25.6					
0001 0100 → 0100 1010	5	16	27,0	27.0					
0000 1010 → 0001 0101	5	17	28,1	27.9					
0001 0100 → 0010 1010	5	18	29,3	29.1					

are therefore assumed to be indistinguishable with an expected classification error, $P_e = 0,5$ (guessing), e.g. $00000000 \rightarrow 00011111$ and $00000000 \rightarrow 11111000$ both belonging to T_5^1 (Table II). Patterns with $\Delta\alpha > 0$ are assumed to be distinguishable with $P_e < 0,5$ and P_e is expected to decrease as $\Delta\alpha$ increases, e.g. it is expected to be easier (lower P_e) to classify $00000000 \rightarrow 00011111$ from $00010100 \rightarrow 00101010$ ($\Delta\alpha = 17$), than $00000000 \rightarrow 00011111$ from $00000000 \rightarrow 01010111$ ($\Delta\alpha = 4$) (Table V), simply because $\Delta\alpha = 17$ indicate a larger difference in energy levels than $\Delta\alpha = 4$.

An experiment was designed to validate the expected relationship between difference in energy levels, $\Delta\alpha$, and classification error, P_e. The experiment consisted of three steps: (i) Measure the electromagnetic emanation from a set of transition patterns. (ii) Evaluate the performance (P_e) of a classifier trained by the recorded data. (iii) Look at average P_e as a function of $\Delta\alpha$. Do we see the expected relationship?

An 8 bit internal data bus on a smart card (i.e. PIC 16F84A microprocessor) was chosen as the source for the electromagnetic radiation. All 18 possible crosstalk indexes for transition patterns with 5 transitions were studied ($T_5^\alpha, \alpha = 1, \cdots, 18$ in Table II). A total of 1000 traces (observations) of the electromagnetic emanation, for each of the 18 transition patterns, were collected. A 10 Gs/s oscilloscope with a broadband E near-field probe was used. The probe was positioned as close to the microprocessor as possible, without any decapsulation, see Fig. 3.

The challenge of this experiment is to manipulate the microprocessor, such that the appropriate transition patterns are generated on the internal data bus. It is also essential that the power consumption (i.e. electromagnetic radiation) is correlated with this bus activity and not dominated by noise (e.g. other irrelevant microprocessor activities).

Our approach was to combine detailed knowledge of

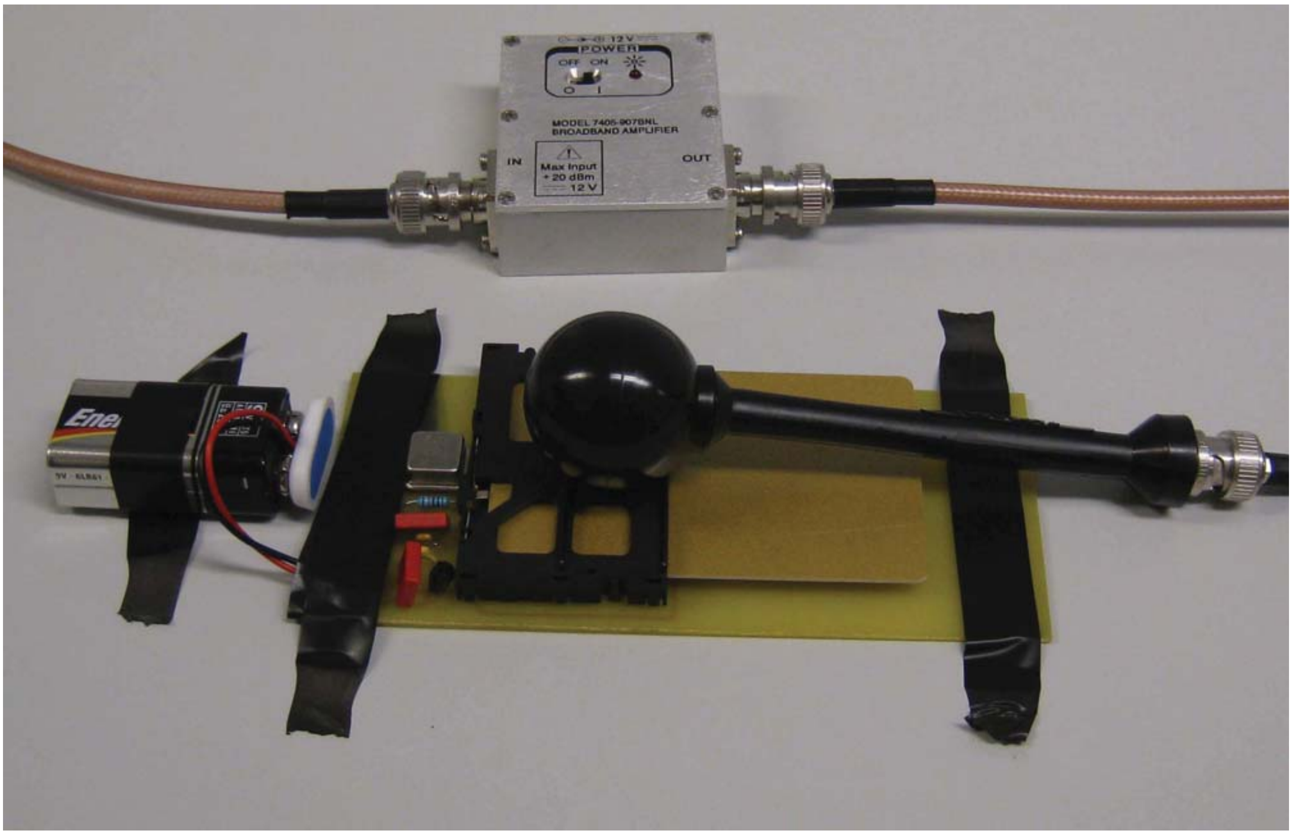

Figure 3. Measurement setup: Smart card (PIC 16F84A) executing a code in a customized terminal. A broadband E near-field probe, with amplification, is positioned as close as possible.

the execution sequence of the microprocessor with careful assembly programming. The code was written off-line, using vendors development kits, and loaded to the smart card with a standard smart card terminal. To execute the code, a customized smart card terminal was used to provide power and clock signals only. This was to limit noise from external circuitry. The microprocessor automatically executed from the beginning of the program and no I/O communication was required, with the exception of an initial trigger signal.

The objective of the code is to create a transition between value a and value b on the microprocessors internal 8 bit data bus and minimizing irrelevant activity. PIC 16F84A has a 2-stage pipeline architecture. Each instruction is executed during four clock cycles (Q1-Q4). The transition between value a and b should ideally take place in consecutive clock cycles or ”near” consecutive such that no data bus activity takes place between handling the two values. In addition, parallel activity, due to pipelining, must be avoided or kept constant for all transitions. This is possible to achieve through careful choice of instructions as shown in [8]. The code used in this experiment is seen in Table VII.

Code lines $1-4$ toggles the smart cards I/O and provides a trigger-point for the oscilloscope. The following 10 NOP’s create a buffer between electromagnetic disturbances caused by the relatively strong I/O toggle and the rest of the program. Code lines $15-18$ are used to

```
        Test Code
        ;Main program
        Start
        ;Trigger Turn I/O ON and OFF
 1      movlw 80h          ; Turn I/O ON
 2      movwf PORTB        ; by moving the value 80h onto port B
 3      movlw 00h          ; Turn I/O OFF
 4      movwf PORTB        ; by moving the value 00h onto port B
        ;10 NOP’s to create buffer from I/O disturbances
 5      nop
 .
 .
 .
14      nop
        ; Transition: a:0000 0000 - b: 0001 1111
15      movlw 00h          ; a into W register
16      movwf DATA1        ; mov a from W to DATA1 register
17      movlw 1Fh          ; (b-a) into W register
18      addwf DATA1,1      ; Q2 read a, Q4 write b=(a+(b-a))
        ; Transition: a:0000 0000 - b: 1000 1111
19      movlw 00h          ; a into W register
20      movwf DATA1        ; mov a from W to DATA1 register
21      movlw 8Fh          ; (b-a) into W register
22      addwf DATA1,1      ; Q2 read a, Q4 write b=(a+(b-a))
        ; Continue for all 18 transition patterns
 .
 .
 .
        ; Back to the start of the program
23      goto Start
```

TABLE VII.
CODE USED TO GENERATE 18 DIFFERENT TRANSITIONS PATTERNS ON THE INTERNAL DATA BUS OF MICROPROCESSOR PIC 16F84A

create a transition from bit pattern 00000000 to 00011111 (T_5^1). Code lines $15-17$ are initialization, making sure value a is available in DATA1 register and value $(b-a)$ is found in the working register. Transition between value a and b is then made possible by the ADDWF instruction of line 18. In clock cycle 2 (Q2) the value a is read over the data bus, in clock cycle 3 (Q3), a is added to $(b-a)$ found in the working register. The result, b, is written back over the databus in clock cycle 4 (Q4), creating the desired transition without unwanted data bus activity. The process can now be repeated for all other values of a and b, as shown in code lines 19-22. Finally code line 23 repeats the program indefinitely.

Classification between all pairs of transition patterns was done according to the Modified Template Attack [8]. This includes feature selection, training and evaluating the performance of a quadratic Bayes classifier (for details refer to [8]). The probability of error, P_e, was found from the confusion matrix [6]. Since the classification accuracy depends on how the observations are split, the average of 100 random permutations of 200 training observations and 800 test observations was used. Finally, the average classification error as a function of α distance ($\Delta\alpha$) was calculated and plotted in Fig. 4.

The results (Fig. 4) show that transition patterns belonging to equal energy levels ($\Delta\alpha = 0$) have $P_e = 0,5$. This is equal to guessing as expected. When the difference in energy level increase (larger $\Delta\alpha$) the results suggest that the average classification error decrease. This supports our simulation/theoretical results. We hypothesize that the discrepancy between our simulation/theoretical results for alpha distance 4 is a consequence of statistical uncertainty/noise in the experimental data. Currently, we cannot offer any explanation for why classification error seems to increase for high values of the alpha distance.

The classification results are a result of measuring electromagnetic emanation from a real system. The assumptions made for analytic (E_T) and simulated ($\hat{E}_T$) energy dissipation are therefore not necessarily valid. A real system will be subject to all layout dependent phenomena of section II, and not limited to coupling capacitance and the assumptions that load and coupling capacitances are equal. To evaluate these assumptions and how the power model can be revised if the assumption do not hold is subject of future work. Future work also include calculating confidence intervals for the results in Fig. 4. This is not trivial, due to non-uniform distribution of energy levels (see Table II). When noise causes energy levels to overlap, the results looks to be a function of which energy levels merge, and must therefore be modeled carefully.

Finally, to validate the impact of the new power model on security, we encourage the power model to be applied in a DPA attack and compared to the performance of other power models (e.g. HD/HW). Such comparison could also be done with extended versions of our power model, e.g. adding other layout dependent phenomena or removing some of the simplifying assumptions made (equal coupling and load capacitance).

VII. Conclusion

It is known that one can distinguish bus activity generated from signal transitions having different HD. In this paper we elaborate on the theory and simulations on the hypothesis from [9] that layout dependent phenomena, such as inductance and capacitance in and between conductors and radiation properties of circuit elements, can explain why it sometimes is possible to distinguish transition patterns with the same HD. Our simulations show that capacitive crosstalk has a significant effect on gate energy dissipation, and confirm that the dissipated energy from CMOS switching gates depend not only on the HD, but also on the direction of switching activity on nearby data lines. For an 8 bit bus, this increases the number of possible energy levels from 9 (HD) to 93 (crosstalk), and therefore allows us to explain why signals with the same HD sometimes can be distinguished. Where as an HD based detector can provide about 2.5 bits of information per sample, a crosstalk based detector will yield about 5.7 bits (theoretical) or 5.0 bits (simulated) of information per sample - in all cases for an 8 bit bus.

In this paper we have also shown that experimental data, i.e. the electromagnetic side-channel of a smart card, suggest that the average classification error gets reduced as α distance (i.e. difference in energy level due to capacitive crosstalk) increases. This supports our simulations/theoretical results that layout specific phenomena (e.g. capacitance) must be considered when analyzing security implications of electromagnetic side-channels.

References

[1] D. Agrawal, S. Baktir, D. Karakoyunlu, P. Rohatgi, and B. Sunar. Trojan detection using ic fingerprinting. In *IEEE Symposium on Security and Privacy*, pages 296 –310, may 2007.

[2] Dakshi Agrawal, Bruce Archambeault, Josyula Rao, and Pankaj Rohatgi. The em side-channel(s). In *Cryptographic Hardware and Embedded Systems - CHES*, volume 2523 of *Lecture Notes in Computer Science*, pages 29–45. Springer Berlin / Heidelberg, 2003.

[3] Suresh Chari, Josyula Rao, and Pankaj Rohatgi. Template attacks. In *Cryptographic Hardware and Embedded Systems - CHES*, volume 2523 of *Lecture Notes in Computer Science*, pages 51–62. Springer Berlin / Heidelberg, 2003.

[4] Zhimin Chen, Syed Haider, and Patrick Schaumont. Side-channel leakage in masked circuits caused by higher-order circuit effects. In *Advances in Information Security and Assurance*, volume 5576 of *Lecture Notes in Computer Science*, pages 327–336. Springer Berlin / Heidelberg, 2009.

[5] Chunjie Duan, V.H.C. Calle, and S.P. Khatri. Efficient on-chip crosstalk avoidance codec design. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 17(4):551 –560, april 2009.

[6] R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern Classification*. John Wiley and Sons, Inc, 2001.

[7] Geir Olav Dyrkolbotn and Einar Snekkenes. A wireless covert channel on smart cards (short paper). In *Information and Communications Security - ICICS*, volume 4307 of *Lecture Notes in Computer Science*, pages 249–259. Springer Berlin / Heidelberg, 2006.

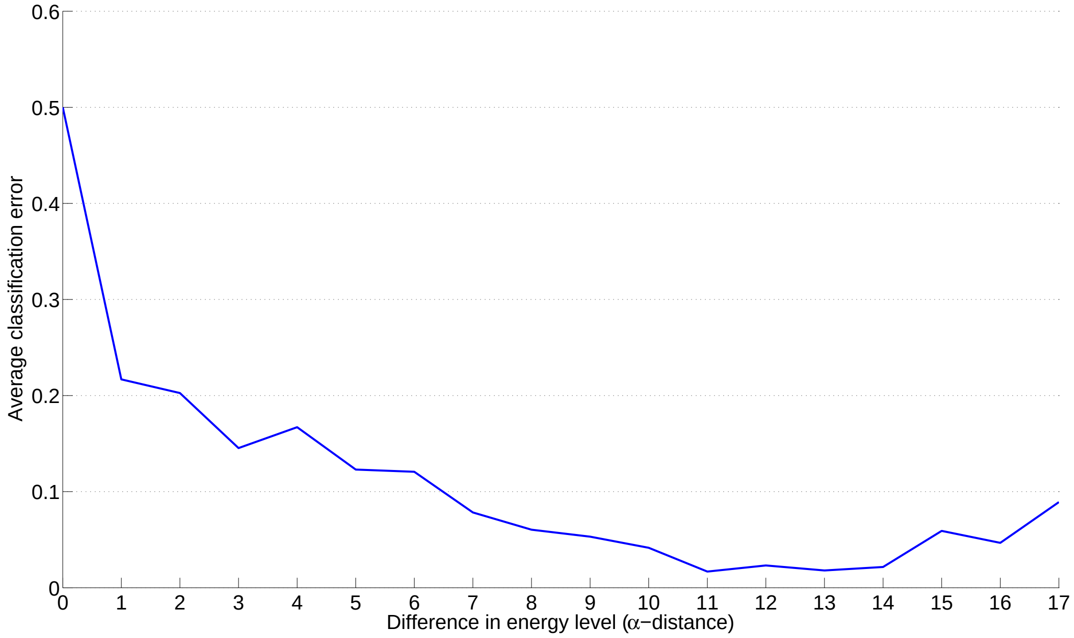

Figure 4. Average classification error as a function of difference in energy level (expressed as α distance, $\Delta\alpha$)

[8] Geir Olav Dyrkolbotn and Einar Snekkenes. Modified template attack: Detecting address bus signals of equal hamming weight. In *Annual Norwegian Information Security Conference - NISK*, pages 43–56. Tapir akademisk forlag, 2009.

[9] Geir Olav Dyrkolbotn, Knut Wold, and Einar Snekkenes. Security implications of crosstalk in switching cmos gates. In *Information Security Conference - ISC*, volume 6531 of *Lecture Notes in Computer Science*, pages 269–275. Springer Berlin / Heidelberg, 2010.

[10] Karine Gandolfi, Christophe Mourtel, and Francis Olivier. Electromagnetic analysis: Concrete results. In *Cryptographic Hardware and Embedded Systems - CHES*, volume 2162 of *Lecture Notes in Computer Science*, pages 251–261. Springer Berlin / Heidelberg, 2001.

[11] Paul Kocher, Joshua Jaffe, and Benjamin Jun. Differential power analysis. In *Advances in Cryptology*, volume 1666 of *Lecture Notes in Computer Science*, pages 388–397. Springer Berlin / Heidelberg, 1999.

[12] S. Mangard, E. Oswald, and T. Popp. *Power Analysis Attack - Revealing the Secret of Smart Cards*. Springer, 2007.

[13] Francesc Moll, Miquel Roca, and Eugeni Isern. Analysis of dissipation energy of switching digital cmos gates with coupled outputs. *Microelectronics Journal*, 34(9):833 – 842, 2003.

[14] Jean-Jacques Quisquater and David Samyde. Automatic code recognition for smart cards using a kohonen neural network. In *Conference on Smart Card Research and Advanced Application Conference*, Berkeley, CA, USA, 2002. USENIX Association.

[15] P.P. Sotiriadis and A. Chandrakasan. Low power bus coding techniques considering inter-wire capacitances. In *Custom Integrated Circuits Conference, 2000. CICC. Proceedings of the IEEE 2000*, pages 507 –510, 2000.