Spectral Coherence Analysis
- First Experimental Results -

Amine Dehbaoui[1], Sébastien Tiran[1], Philippe Maurine[1],
François-Xavier Standaert[2], Nicolas Veyrat-Charvillon[2].

[1] Université de Montpellier II, LIRMM, France.
[2] Université catholique de Louvain, Crypto Group, Belgium.

Abstract. This paper introduces a new family of distinguishers for side-channel analysis, based on the spectral coherence between leakage traces. Its main goal is to allow adversaries and evaluators of cryptographic devices to take advantage of both time domain and frequency domain intuitions, while also allowing to keep a generic attack in case such intuitions are not available. Compared to previous side-channel analysis tools working in the frequency domain, Spectral Coherence Analysis has the significant advantage to directly capture the degree of similarity between different time domain traces, rather than comparing them with an hypothetical (e.g. Hamming distance) leakage model. In other words, we exploit leakage models to build partitions of the leakage, but not to correlate with an estimated spectrum. As a result, we obtain a more generic and remarkably robust distinguisher. First experiments performed against an unprotected DES implementation suggest that we also gain an improved efficiency in certain meaningful application contexts.

1 Introduction

In modern cryptography, security evaluations and proofs of algorithms and protocols are typically obtained under computational assumptions. For example, breaking the Advanced Encryption Standard is not supposed to be impossible (in an information theoretic sense), but to require a computational power that is likely to remain out of reach for decades [4]. While this setting has allowed for a large number of positive results, concerns about the relevance of such a computational-only model have also appeared in the late 1990s. For example, Kocher, Jaffe and Jun showed that, by monitoring the power consumption of a cryptographic implementation, it is sometimes possible to completely recover the cryptographic keys used, e.g. for encryption [9]. Following this seminal paper, the investigation of so-called side-channel attacks has become an important topic, both for academic research and industrial developments. Other types of side-channels have been discovered, e.g. the electromagnetic one [6, 14], and many different ways to exploit this physical information have been introduced, denoted as side-channel distinguishers. As usual in cryptography, the development and understanding of new attacks are closely connected with the one of sound countermeasures. In fact, present security evaluations of embedded devices are

performed in laboratories that typically apply a battery of tests, in order to detect possible weaknesses. In this respect, the development of strong attacks, as we envision in this paper, is a necessary step for securing implementations. But as an exhaustive application of all possible attacks is intensive, it also leads to the question of what are the most relevant distinguishers.

Because of their easy connection with engineering intuition, many side-channel attacks proposed in the literature directly work with time domain traces. They usually proceed according to a divide-and-conquer strategy (i.e. they recover cryptographic keys byte per byte) and proceed in three main steps. First, the adversary computes a prediction of the physical leakage, e.g. for the 256 values of a target key byte. Second, he measures the physical leakages for a device manipulating the secret key to be recovered. Finally, he compares the key dependent predictions with the physical leakages. In successful attacks, the key byte giving rise to the best prediction is the correct one. Following this description, the relevance of a side-channel distinguisher can be rated along two different axes. First, distinguishers should be efficient, i.e. output the correct key with a minimum number of measurements (and offline computations). Second, they should be generic, i.e. allow to recover the key for most (if not all) devices. This tradeoff between efficiency and genericity mainly relates to the leakage model used by the adversary in his prediction step. That is, taking advantage of precise leakage assumptions improves efficiency, but implies less genericity, as the distinguisher may fail in case these assumptions are not respected. Correlation Power Analysis (CPA), analyzed in [2], is an example of efficient but specific distinguisher. Mutual Information Analysis, introduced in [8], is an example of less efficient but more generic one.

By contrast, and perhaps surprisingly, much less attention has been paid to side-channel analysis performed in the frequency domain. To the best of the authors' knowledge, Gebotys, Ho and Tiu first proposed to apply a differential type of attack after the application of a Fast Fourier Transform (FFT) [7]. This work was then extended towards CPA-like attacks in [11, 12]. Various similar approaches were described at InsCrypt 2010 [13]. Compared to attacks performed in the time domain, these different works share a number of significant advantages. First, side-channel attacks in the time domain always need to detect so-called samples of interest, where the secret information lies. Those samples are usually found by an exhaustive analysis, inspecting all the samples in a trace, which can be computationally intensive. Also, for certain attacks such as CPA, combining several leakage samples is not trivial (as it typically implies to assign"weights" to the contributions of these samples). By contrast, moving to the frequency domain allows a type of dimensionality reduction, summarizing the information of all samples into a few components with rich engineering intuition, as they correspond to frequencies. Second, these attacks are inherently more robust against trace misalignment, that typically increases the difficulty of distinguishing in the time domain [3].

On the negative side, previous approaches for side-channel attacks in the frequency domain are essentially based on a statistical relation between the spectrum of some leakage traces and an adversary-chosen leakage model. But the selection of these (e.g. Hamming weight) leakage models was first motivated by time domain intuitions. So, although these proposals did lead to successful key recoveries in different practical scenarios, it is not always clear if and why they are efficient. Hence, one could wonder if it would not be possible to take advantage of the best of two worlds. That is, can we use time domain intuitions to partition the leakage traces, and frequency domain intuitions to select the bandwidth where most information lies, without explicitly using the leakage models when applying a statistical test in the frequency domain?

In this paper, we answer this question positively, by introducing a new class of side-channel distinguishers, exploiting the spectral coherence between different leakage traces. Rather than computing a correlation between a leakage model and some estimated spectrum, the main idea of Spectral Coherence ANalysis (SCAN) is to directly estimate the correlation between the spectrum of two (or more) leakage traces. Interestingly, this proposal relies on a strongly established theory and can take advantage of well known tools in signal processing, such as the Magnitude Squared Coherence (MSC). Given two leakage traces l_1 and l_2, the MSC corresponds to the ratio between the squared cross-spectral density between l_1 and l_2 and the product of their auto-spectral densities. For a given frequency f, the MSC provides a value between 0 and 1 that captures the degree of similarity between the two time domain traces. Hence, SCAN is expected to take advantage of time domain intuitions, when partitioning the leakage traces, and of frequency domain intuitions, when selecting the bandwidths where the information lies, in a principled and separated manner. As a result, it provides more genericity than previous frequency-based distinguishers. First experimental results provided in this paper indicate that, at least in certain meaningful scenarios, this increased genericity does not come at the cost of reduced efficiency. Note that previous works already pointed out that MSC could be a relevant tool in side-channel analysis. For example, [5] used it to perform a "weighted difference-of-means test". In the following, we take a different approach and directly use the spectral coherence as the distinguishing tool of our attacks, allowing us to fully exploit its powerful discriminating features.

2 Side-channel attacks

Side-channel attacks generally follow the three steps mentioned in introduction (i.e. prediction, measurement and comparison). In order to keep our descriptions as simple as possible, we follow the formalism of standard Differential Power Analysis (DPA) attacks introduced in [10] and intuitively recalled in Figure 1. In the rest of this section, we briefly introduce these physical attacks with a simple example.

Say one wants to recover the secret key byte s implied in a (6-bit to 4-bit) S-box computation $z_i = \mathsf{S\text{-}box}(x_i \oplus s)$, exploiting the leakage trace l_i obtained by

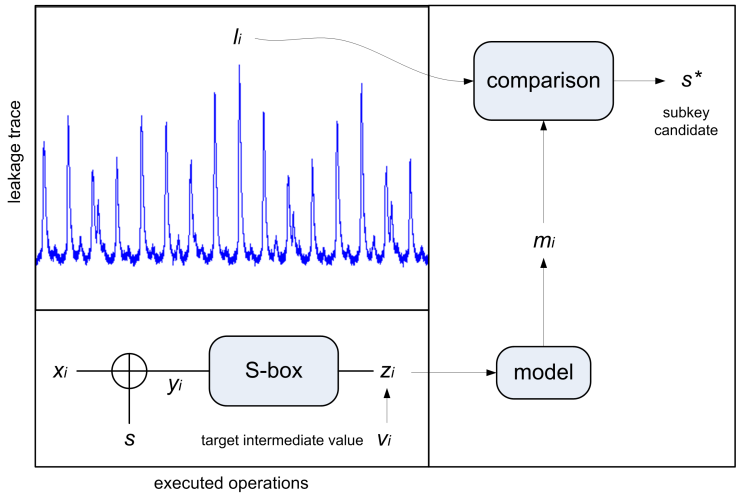

Fig. 1. Illustrative side-channel attack.

measuring a device performing this computation. For this purpose, the adversary first predicts the 64 possible values at the output of this S-box, denoted as $v_i^{s^*}$ for input plaintext x_i and key candidate s^*. Next, he translates these values in so-called models $m_i^{s^*}$. In our (non-profiled) attack scenario, we focus on discrete models that can be used to partition the leakage traces, as illustrated in Table 1. For example, in the left part of the table, we consider a 1-bit model leading to 2-column partitions. That is, for each key candidate s^*, one separates the measurements in two sets: set $p^1_{s^*}$ grouping the leakages such that one bit of $v_i^{s^*}$ is set to zero; set $p^2_{s^*}$ grouping the leakages such that the same bit of $v_i^{s^*}$ is set to one. Finally, the adversary uses a statistical test to check the relevance of these partitions. For example, in Kocher's original DPA [9], we simply compute:

$$\Delta_{s^*} = \left(\hat{\mathbf{E}}_{l_i \in p^1_{s^*}} l_i \right) - \left(\hat{\mathbf{E}}_{l_i \in p^2_{s^*}} l_i \right), \tag{1}$$

where $\hat{\mathbf{E}}$ denotes the sample mean operator. If the attack is successful, there is at least one leakage sample in the traces, for which this difference-of-means will be significant, leading to a so-called DPA peak. Hence, the adversary just has to select the key candidate that maximizes this peak. Most side-channel distinguishers naturally extend to partitions based on more elaborate models, of which the goal is to increase the signal to noise ratio of the attack, at the cost of more specific assumptions. For example, the middle and right parts of Table 1 illustrate partitions obtained from a 2-bit and a Hamming weight leakage model, respectively.

$p^1_{s^*}$	$p^2_{s^*}$
l_1	l_2
l_3	l_5
l_4	l_7
l_6	l_8
l_{10}	l_9
l_{12}	l_{11}
l_{14}	l_{13}
l_{15}	l_{16}

$p^1_{s^*}$	$p^2_{s^*}$	$p^3_{s^*}$	$p^4_{s^*}$
l_3	l_1	l_5	l_6
l_4	l_2	l_9	l_7
l_{11}	l_{10}	l_8	l_{12}
l_{15}	l_{14}	l_{16}	l_{13}

$p^1_{s^*}$	$p^2_{s^*}$	$p^3_{s^*}$	$p^4_{s^*}$	$p^5_{s^*}$
l_5	l_2	l_1	l_3	l_{14}
	l_7	l_4	l_6	
	l_9	l_8	l_{12}	
	l_{16}	l_{10}	l_{13}	
		l_{11}		
		l_{15}		

Table 1. Examples of 1-bit, 2-bit and Hamming weight partitions.

3 Spectral Coherence Analysis

Taking the example of 1-bit DPA in the previous section, it is quite simple to explain how a 1-bit SCAN would proceed. In fact, the prediction step and partitions would be exactly the same, as well as the measurement of the leakage traces l_i. The only difference appears in the statistical tool used for evaluating the partitions. For this purpose, we first need to define the MSC between two leakage traces l_1 and l_2 as follows:

$$\mathsf{MSC}_{l_1,l_2}(f) = \frac{|\mathsf{PSD}_{l_1,l_2}(f)|^2}{\mathsf{PSD}_{l_1,l_1}(f) \cdot \mathsf{PSD}_{l_2,l_2}(f)}, \tag{2}$$

where PSD_{l_1,l_2} and PSD_{l_i,l_i} denote the cross-spectral density between l_1 and l_2 and the auto-spectral density of l_i, respectively. Compared to standard DPA attacks, performing a SCAN additionally requires to estimate the PSD functions. Interestingly, one can take advantage of different tools for this purpose, allowing an effective reduction of the noise. For example, a usual solution is to use Welch's method[1] [16], that is frequently embedded in mathematical toolboxes. In other words, SCAN is not a fully non-parametric distinguishers but the selection of parameters follows standard strategies of signal processing textbooks.

To illustrate the previous definitions, 5 electromagnetic traces were acquired, corresponding to the encryption of 5 different plaintexts, with a hardware (FPGA) DES implementation running at 50MHz. These curves, illustrated in Figure 2, have been collected with a 0.5 mm diameter probe placed above the DES module. Next, Figure 3 contains the FFT of these different traces. Finally, Figure 4 plots the evolution of their MSC in function of the frequency.

Given that one can efficiently estimate the MSC, a simple variant of 1-bit SCAN would then work as follows. First, for each key candidate s^*, the adversary averages the leakage traces according to the partitions, i.e. he computes two average traces $\mathbf{l}^1_{s^*}$ and $\mathbf{l}^2_{s^*}$ as follows:

[1] Which works by dividing the the traces in several overlapping segments.

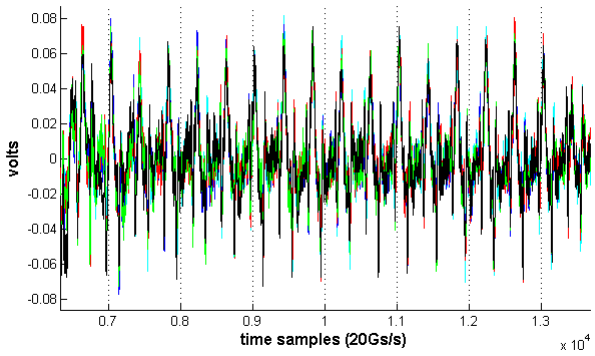

Fig. 2. Five exemplary EM side-channel traces.

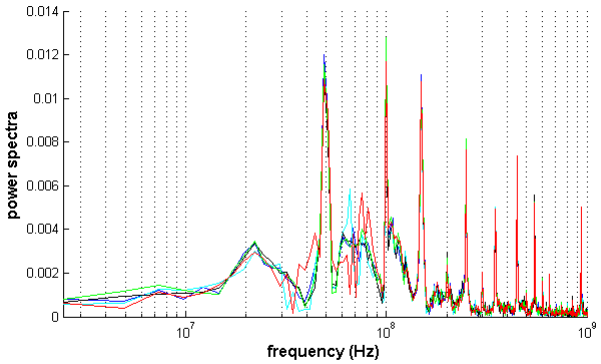

Fig. 3. FFT of five exemplary EM side-channel traces.

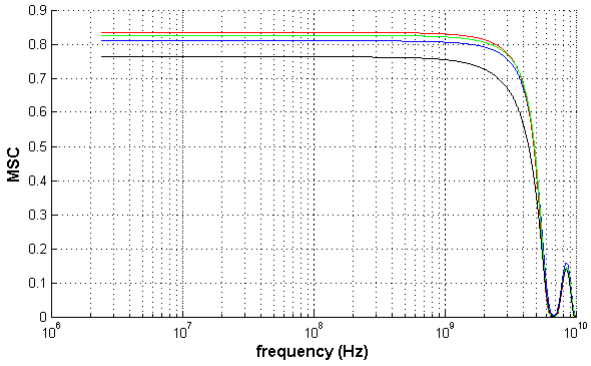

Fig. 4. MSC between one and four exemplary EM side-channel traces.

$$\mathbf{l}^1_{s^*} = \underset{l_i \in p^1_{s^*}}{\hat{\mathbf{E}}} l_i, \tag{3}$$

$$\mathbf{l}^2_{s^*} = \underset{l_i \in p^2_{s^*}}{\hat{\mathbf{E}}} l_i. \tag{4}$$

Next, for each key candidate s^* and frequency f, he estimates $\mathsf{MSC}_{\mathbf{l}^1_{s^*}, \mathbf{l}^2_{s^*}}(f)$. Finally, he determines a bandwidth B from which he computes the distinguisher value (corresponding to a global coherence scalar value):

$$\mathsf{C}_{s^*} = 1 - \underset{f \in B}{\hat{\mathbf{E}}} \, \mathsf{MSC}_{\mathbf{l}^1_{s^*}, \mathbf{l}^2_{s^*}}(f). \tag{5}$$

As mentioned in introduction, the MSC is between 0 and 1 and, for each frequency f, measures the degree of similarity between the traces. Hence, a good partition, corresponding to the correct key candidate, should lead to a larger value for the distinguisher C. Note that it is typically in the selection of the bandwidth that engineering intuition can be exploited. But in case no such information is available to adversaries, it is of course possible to average over all frequencies.

Quite naturally, single-bit attacks are not optimal in terms of signal-to-noise ratio. Various solutions exist in the literature, allowing to turn a single-bit time domain distinguisher into a multi-bit one. Most of these solutions can be applied in our new setting. In order to keep the distinguisher as generic as possible, our following experiments focus on a proposal of Bevan and Knudsen, where one just sums the values of the statistics obtained for different single-bit DPA [1]. More precisely, we focused on the following multi-bit distinguisher: let $C_{s^*}(i)$ denote the result of a single-bit SCAN for bit i, we define the multi-bit SCAN distinguisher as:

$$\mathsf{MC}_{s^*} = \sum_i \mathsf{C}_{s^*}(i). \tag{6}$$

4 First experimental results

In order to confirm the relevance of our new approach, this section reports on first experiments of different attacks against the same unprotected DES implementation as in the previous section. For this purpose, we tried to select a list of distinguishers that are representative of the state-of-the-art.

In addition to multi-bit SCAN (MB-SCAN), and as a reference, we started our analysis with a CPA exploiting a Hamming distance leakage model. This test is usually the first one considered in the literature for evaluating a leaking device. In the present context, it allowed us to confirm experimentally that a Hamming distance leakage model was a reasonable abstraction to predict ou electromagnetic radiation traces. Second, we applied the multi-bit DPA of Bevan and Knudsen in the time domain (MB-DPA). This allowed us to have a

straightforward comparison of a time domain vs. SCAN attack, using exactly the same assumptions and partitions of the leakage. Third, we applied the Correlation Power Frequency Analysis (CPFA) described in [13] and further analyzed in [11], in order to provide a comparison with previous attempts to exploit frequency-based attacks. Finally, we implemented two types of MIA attacks, one estimating the probability density functions with histograms, as advocated in [8], and one based on a Gaussian assumption that turned out to be a quite accurate approximation in our experiments.

Our evaluations followed the framework in [15] and, in particular, we computed both a global success rate and the minimum and maximum guessing entropies, taken over the 8 DES subkeys. These metrics have been estimated for up to 5000 traces per attack, and sampled over 50 independent attacks. Figures 5, 6 and 7 clearly illustrate the promising features of SCAN. As detailed in the previous section, the new distinguisher can be considered as a quite generic one, as it does not rely on any specific assumption. Still, it is by far the most efficient in our experiments against an unprotected device. The reasons of this increased efficiency is assumably related to the ability to capture information scattered over different time samples with MSC, and is a scope for further research. In this respect, a final and intriguing observation is that SCAN tolerates certain arbitrary deterioration of the leakage models much better than other distinguishers. For example, it was experimentally confirmed that the leakage of our target device was reasonably correlated with the Hamming distance of the manipulated data, which is typical of hardware implementations. But we additionally investigated the performances of attacks using weight-based models. As illustrated in Figures 8, 9 and 10, only the new distinguisher could efficiently deal with such a context. Of course, this is an partially artificial situation (why to use a wrong model when there is a better one available?). But we believe it clearly emphasizes the type of robustness against incorrect assumptions that is allowed by exploiting the spectral coherence in the analysis of leaking devices.

5 Conclusions and open problems

We presented a new class of side-channel distinguishers, based on the spectral coherence between leakage traces. First attack results obtained against an unprotected DES implementation highlight excellent performances, both in terms of efficiency and genericity. For illustration purposes, our experiments focused on a simple multi-bit version of the distinguisher. But different variants of SCAN could be considered and would require further investigation. In particular, alternative versions of the distinguisher, maximizing the coherence between clusters of curves, and obtained from more complex partitions (e.g. based on a Hamming distance leakage model), could lead to more powerful attacks. In this respect, different strategies could be considered. As a starting point, the present work exploited the coherence between the average curves obtained from different partitions of the measurements, leading to limited time complexities for the attacks. But an alternative could be to compute the coherence directly between single

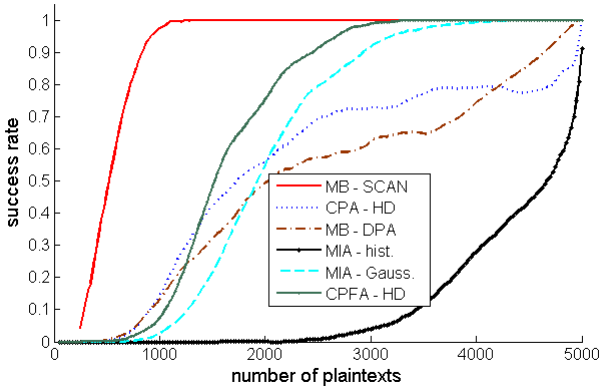

Fig. 5. Global success rate of our experiments, distance-based models.

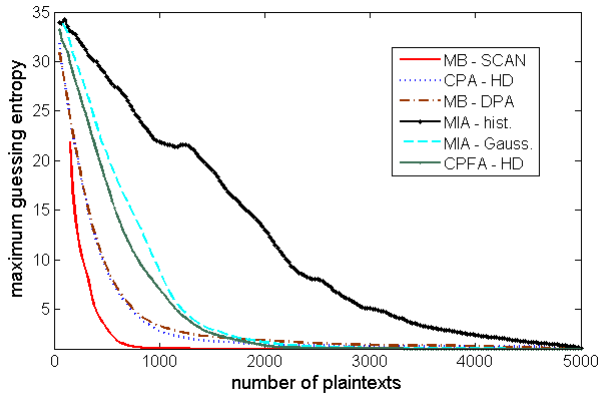

Fig. 6. Maximum guessing entropy of our experiments, distance-based models.

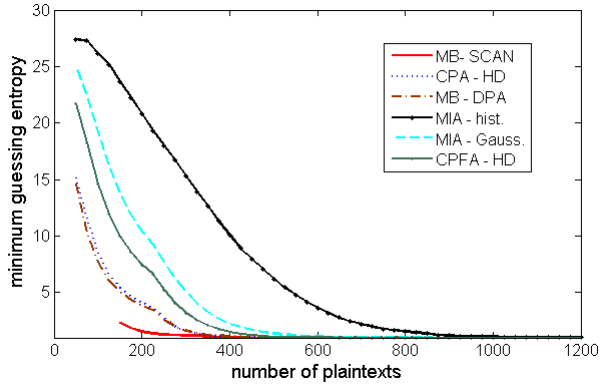

Fig. 7. Minimum guessing entropy of our experiments, distance-based models.

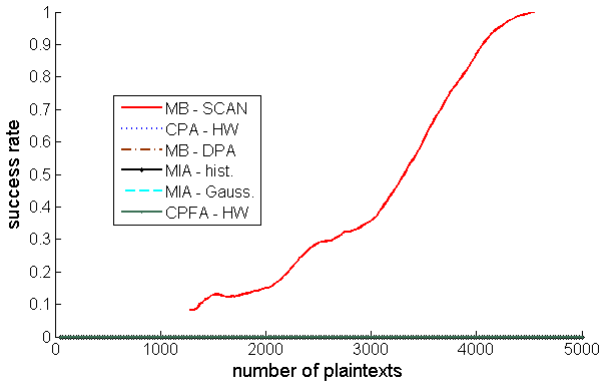

Fig. 8. Global success rate of our experiments, weight-based models.

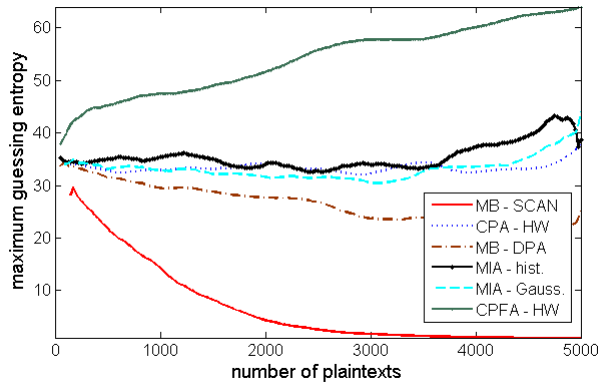

Fig. 9. Maximum guessing entropy of our experiments, weight-based models.

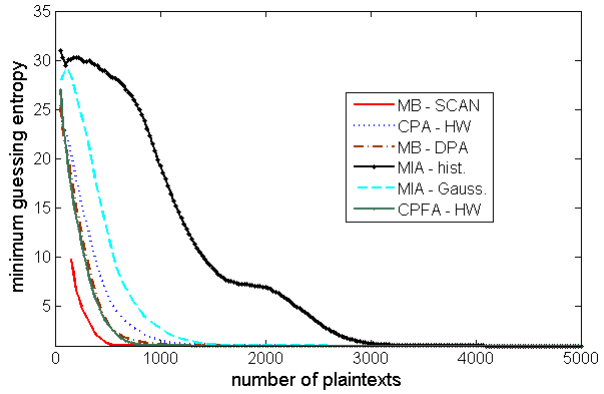

Fig. 10. Minimum guessing entropy of our experiments, weight-based models.

(non average) curves and to perform the statistical treatment afterwards. Such a proposal would increase the time complexity of the attacks, but potentially apply much better in case of implementations where no information leakage can be found in the mean traces (i.e. secure against first-order DPA). In this respect, a particularly interesting direction is to investigate the effectiveness of SCAN against implementations protected with masking or other countermeasures, taking advantage of the dimensionality reduction, and constructive sample combination, allowed by frequency analysis. Preliminary experiments suggest that the direct application of coherence-based attack could be used to defeat certain countermeasures.

Acknowledgements. Amine Dehbaoui, Sébastien Tiran and Philippe Maurine have been supported by the French National Research Agency (ANR-10-SEGI-005). François-Xavier Standaert is an associate researcher of the Belgian Fund for Scientific Research. Nicolas Veyrat-Charvillon is postdoctoral researcher funded by the Walloon region SCEPTIC project.

References

1. Régis Bevan and Erik Knudsen. Ways to enhance differential power analysis. In Pil Joong Lee and Chae Hoon Lim, editors, *ICISC*, volume 2587 of *Lecture Notes in Computer Science*, pages 327–342. Springer, 2002.
2. Eric Brier, Christophe Clavier, and Francis Olivier. Correlation power analysis with a leakage model. In Marc Joye and Jean-Jacques Quisquater, editors, *CHES*, volume 3156 of *Lecture Notes in Computer Science*, pages 16–29. Springer, 2004.
3. Christophe Clavier, Jean-Sébastien Coron, and Nora Dabbous. Differential power analysis in the presence of hardware countermeasures. In Çetin Kaya Koç and Christof Paar, editors, *CHES*, volume 1965 of *Lecture Notes in Computer Science*, pages 252–263. Springer, 2000.
4. Joan Daemen and Vincent Rijmen. *The Design of Rijndael*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2002.
5. Amine Dehbaoui, Victor Lomné, Thomas Ordas, Lionel Torres, Michel Robert, and Philippe Maurine. Enhancing electromagnetic analysis using magnitude squared incoherence. *IEEE Trans. Very Large Scale Integrated Circuits*, xx:yyy–zzz, to appear.
6. Karine Gandolfi, Christophe Mourtel, and Francis Olivier. Electromagnetic analysis: Concrete results. In Çetin Kaya Koç, David Naccache, and Christof Paar, editors, *CHES*, volume 2162 of *Lecture Notes in Computer Science*, pages 251–261. Springer, 2001.
7. Catherine H. Gebotys, Simon Ho, and C. C. Tiu. Em analysis of rijndael and ecc on a wireless java-based pda. In Josyula R. Rao and Berk Sunar, editors, *CHES*, volume 3659 of *Lecture Notes in Computer Science*, pages 250–264. Springer, 2005.
8. Benedikt Gierlichs, Lejla Batina, Pim Tuyls, and Bart Preneel. Mutual information analysis. In Elisabeth Oswald and Pankaj Rohatgi, editors, *CHES*, volume 5154 of *Lecture Notes in Computer Science*, pages 426–442. Springer, 2008.
9. Paul C. Kocher, Joshua Jaffe, and Benjamin Jun. Differential power analysis. In Michael J. Wiener, editor, *CRYPTO*, volume 1666 of *Lecture Notes in Computer Science*, pages 388–397. Springer, 1999.

10. Stefan Mangard, Elisabeth Oswald, and Francois-Xavier Standaert. One for all - all for one: Unifying standard dpa attacks. Cryptology ePrint Archive, Report 2009/449 - to appear in IET Information Security, 2009. `http://eprint.iacr.org/`.
11. Edgar Mateos and Catherine H. Gebotys. A new correlation frequency analysis of the side-channel. In *Proceedings of WESS*, Scottsdale, Arizona, USA, October 2010.
12. E. Bohl J. Hayek O. Schimmel, P. Duplys and W. Rosenstiel. Correlation power analysis in frequency domain. In *Proceedings of COSADE 2010*, Darmstadt, Germany, February 2010.
13. Sylvain Guilley F. Flament Jean-Luc Danger et Frédéric Valette Olivier Meynard, Denis Réal. Characterization of the electromagnetic side channel in frequency domain. In *Proceedings of INSCRYPT*, Shangai, China, October 2010.
14. Jean-Jacques Quisquater and David Samyde. Electromagnetic analysis (ema): Measures and counter-measures for smart cards. In Isabelle Attali and Thomas P. Jensen, editors, *E-smart*, volume 2140 of *Lecture Notes in Computer Science*, pages 200–210. Springer, 2001.
15. François-Xavier Standaert, Tal Malkin, and Moti Yung. A unified framework for the analysis of side-channel key recovery attacks. In Antoine Joux, editor, *EUROCRYPT*, volume 5479 of *Lecture Notes in Computer Science*, pages 443–461. Springer, 2009.
16. P.D. Welch. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short. *IEEE Trans. Audio Electroacoustics*, 15:70–73, 1967.