EMpowering Side–Channel Attacks

Josyula R. Rao* Pankaj Rohatgi*

May 11, 2001

Abstract

In this paper, we report preliminary results obtained as a result of a systematic investigation of leakage of compromising information via EM emanations from chipcards and other devices. Our findings show that the EM side–channel is more powerful than other side–channels such as timing and power analysis. Specifically, in some cases, one can obtain much more compromising information about computations and one can use this information to defeat the protection provided by countermeasures to the other side–channel attacks.

1 Introduction

Side–channel cryptanalysis, *i.e.*, cryptanalysis using information leaked during the computation of cryptographic primitives has been used successfully in extracting cryptographic material such as the secret keys of block ciphers stored on a device [3, 4, 5]. Most of the publicly available literature on side–channels deals with attacks based on timing or power. While it is rumored that there is a large body of classified literature on exploiting leakages due to electromagnetic (EM) emanations, there is scant information about this in the public domain.

Recently, with de–classification of some portions of the Tempest documents by the US government [2] and some preliminary claims by researchers such as Jean–Jacques Quisquater, an awareness of the potential and power of the EM side–channel is developing. However, it is imperative to conduct a full and thorough investigation into this matter as a number of unanswered questions remain. For instance, how does information leaked via EM emanations compare to information obtained from the other side–channels? Do the countermeasures that have been developed and deployed to provide effective protection against other side–channel attacks suffice to counteract exposure from EM as well?

With questions like these in mind, we began a systematic investigation of EM side–channel leakage from chipcards. Although our investigation is still ongoing, we have important results to report.

Our investigations show that although the EM side–channel superficially resembles the power side–channel in the nature of information revealed, there are instances and situations where the EM side–channel can carry much more useful information. In particular, there exist classes of "bad" instructions on some platforms which leak much more information in the EM side–channel as compared to the power side–channel. Thus, the EM side–channel

*I.B.M. T.J.Watson Research Center P.O.Box 704, Yorktown Heights, NY 10598, U.S.A. Email: {jrrao,rohatgi}@watson.ibm.com

could be used to reduce the effectiveness of existing countermeasures against power analysis. In this paper, we show how, on certain platforms, a secret–sharing scheme effective against power analysis, can be rendered ineffective using EM emanations since it uses a “bad” instruction.

In view of our findings, we strongly believe that a careful reconsideration of the deployed countermeasures to side–channel attacks is in order.

2 Simple and Differential Electromagnetic Attacks

The terms Simple and Differential Electromagnetic Attacks, abbreviated as SEMA and DEMA, were introduced by Jean-Jacques Quisquater at numerous rump session talks at Eurocrypt ’00, Crypto ’00 and CHES ’00. In this section, we describe the equipment used to monitor electromagnetic emanation from chipcards and show the information that is available in the signals.

2.1 Obtaining EM Signals

The setup for monitoring electromagnetic emanation makes use of information available in the recently de-classified and publicly available documents on Tempest [5]. Specifically, Chapter 1 of NACSIM 5000, paragraphs 1–2.b, states “...the harmonics are radiated fields which occur at some multiple of the frequency of the originating signal and represent, in effect, a great many compromising signals. These signals can be acquired by not only being tuned to the fundamental frequency but also at any of the harmonic frequencies...” In chipcards, the fundamental frequency is usually the clock frequency. Information about the computation on the chipcard modulates these frequencies as shown in Figure 1.3 of the NACSIM document. These radiated fields can be captured by placing an appropriate antenna in the vicinity of the device being monitored and information can be extracted by demodulation using an EM receiver tuned to either the fundamental frequency or better still, to one of its harmonics. The rest of the set–up is identical to the equipment used for power analysis: software to operate and control the chipcard being attacked, DSP hardware and software for capturing and analyzing the samples.

2.2 Information in EM Signals

Just as in power analysis, the EM signal contains information about the computation done on the chipcard at various levels of granularity. For instance, at a macroscopic level, one can see the structure of the computation, including loops and similar code. Figures 1 shows 16 rounds of DES. The end of a round is demarcated by a sharp negative peak. Figure 2 shows two rounds of DES. At a microscopic level, one can see emanations at the clock cycle level (see Figure 3).

2.3 SEMA

In a SEMA attack, an adversary is able to extract compromising information from a single EM sample. If a computation makes use of conditional branches based on secret information, then sometimes this can be observed as relative shifts in the distances between major computational structures. In some cases, these shifts may be sufficient to reveal the branch taken, which in turn confirms the value of the secret information. This is analogous to what

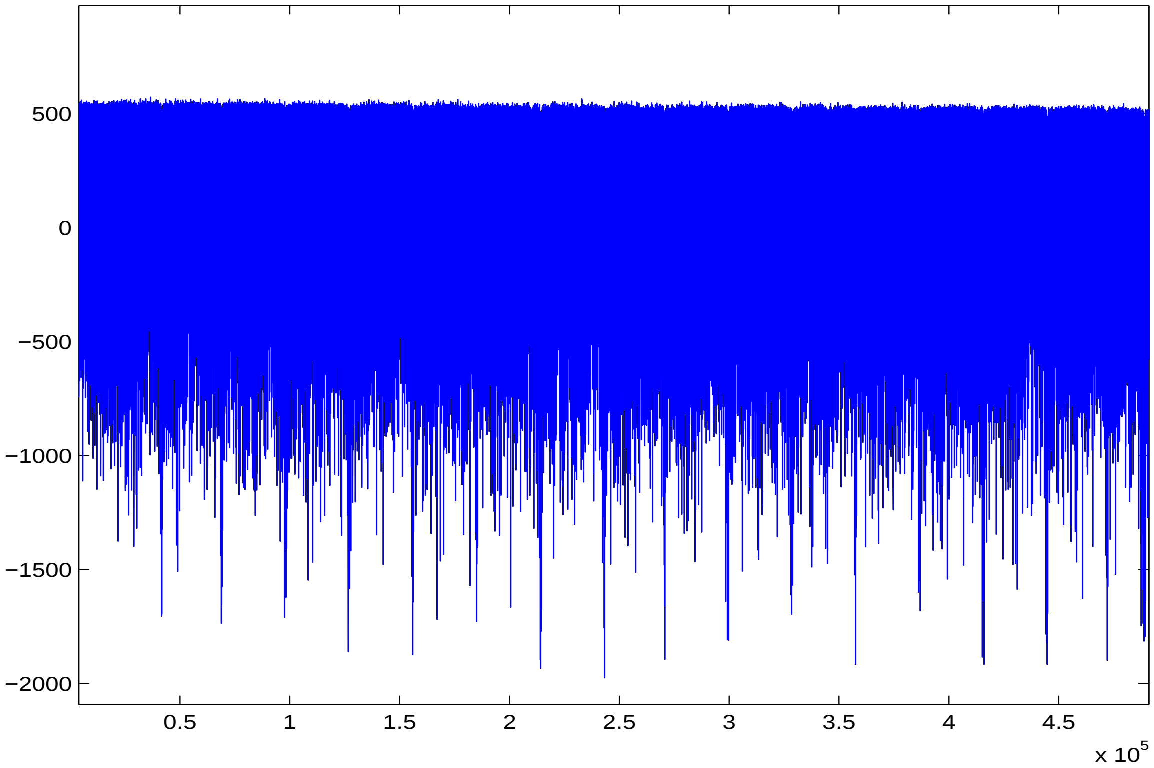

Figure 1: EM Signal from a chipcard executing DES

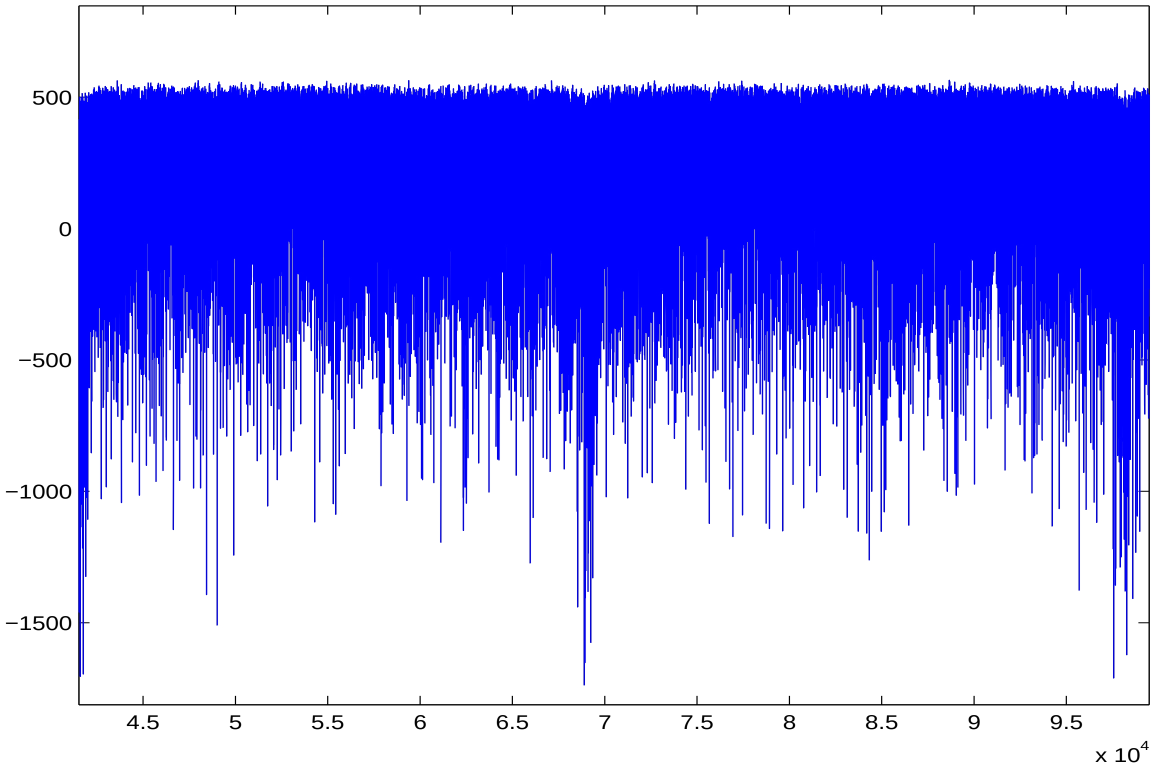

Figure 2: EM Signal of two rounds of DES

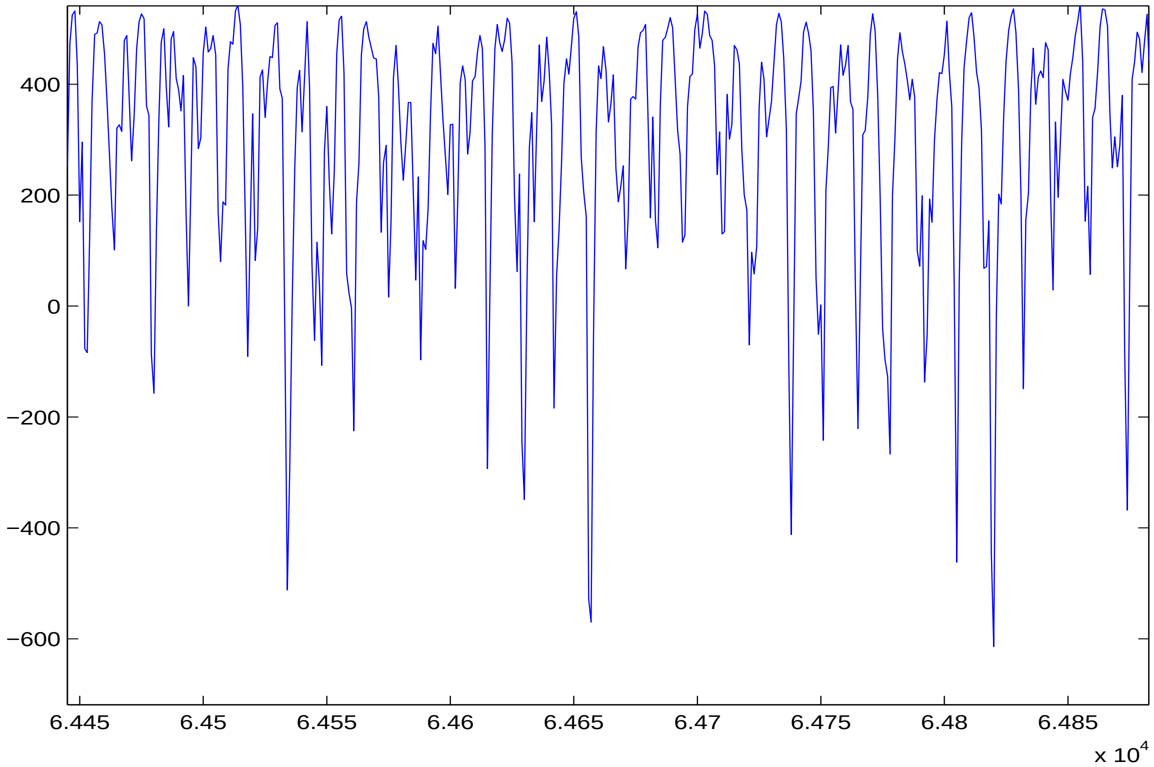

Figure 3: EM Signal showing cycle level information

has already been demonstrated for power samples [4]. Thus conditional statements in the code could provide valuable opportunities for both SPA and SEMA.

In our opinion, the interesting case is where SEMA attacks are successful in extracting information whereas SPA attacks fail. This is possible if the EM side–channel leaks more information than the power side–channel. The following experiment confirms this possibility.

In the following set of figures, we considered a chipcard in which the internal noise generators had been turned off. In such a setting, we observed that an instruction that tests a bit of a byte in memory leaks information from a *single* signal about the value of the bit in the EM channel but not in the power channel.

Figure 4 shows two EM signals in which the bits tested are both 0. This is seen as a low value in both the signals at the point 18915. Figure 5 shows two EM signals in which one of the bits tested is 0 and the other is 1. This is seen as a low value in one of the signals and a high value in the other at the point of interest which in this case is 18780.

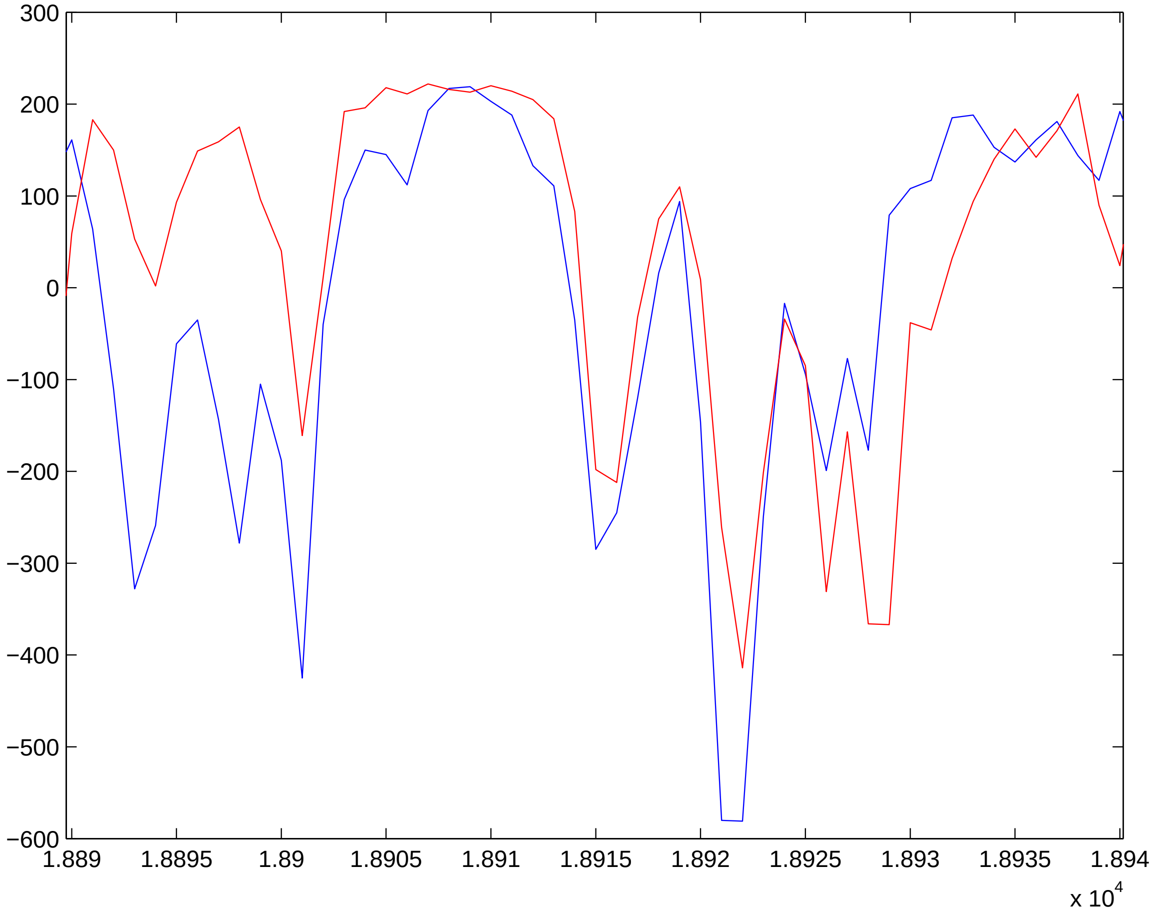

Figure 4: Two EM Signals where tested bits are 0 (seen as low values at 18915)

The same experiment when repeated for the power side–channel does not reveal this information. The corresponding figures are shown in Figures 6 (at point 19260) and 7 (at point 18990) respectively. The power signal levels, at the corresponding points where the

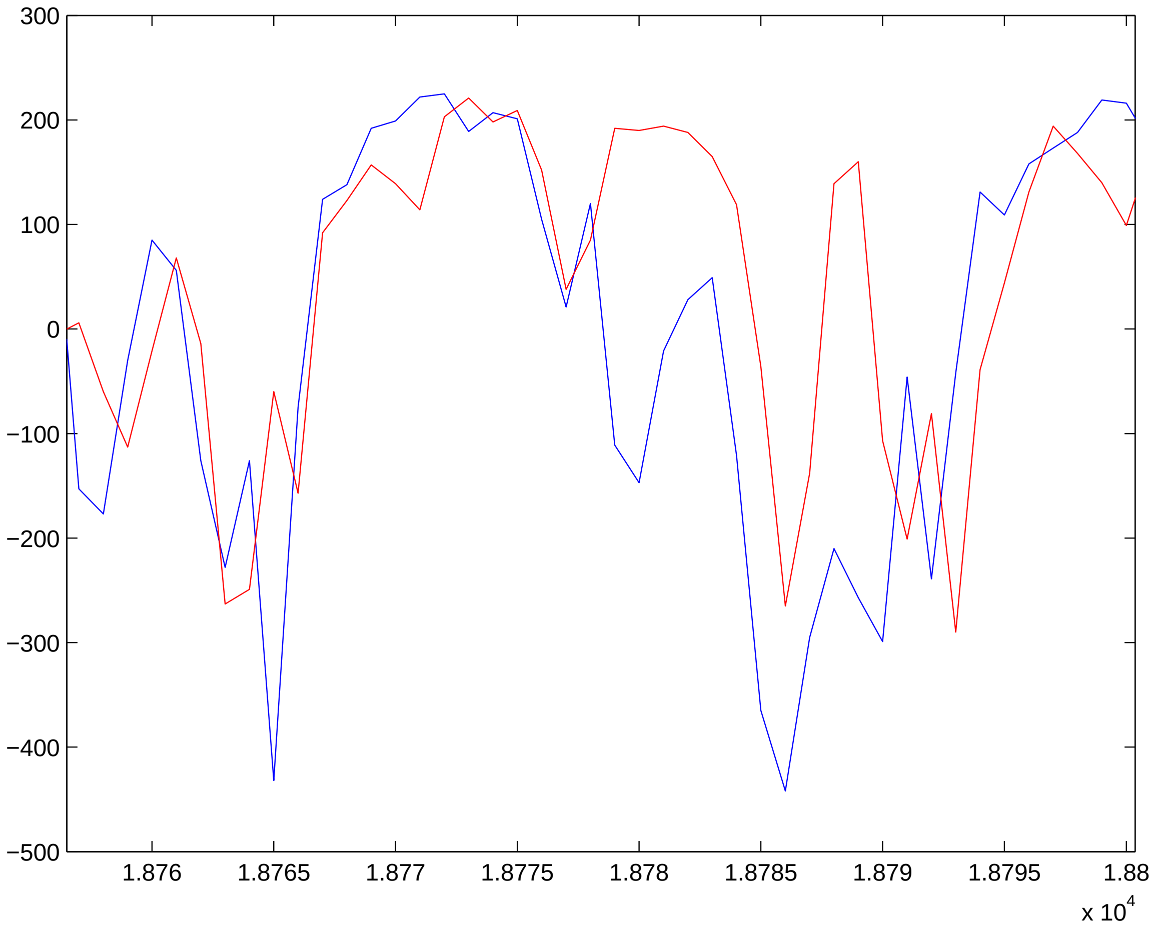

Figure 5: Two EM Signals where tested bits are 0 and 1 (seen as low and high values at 18780)

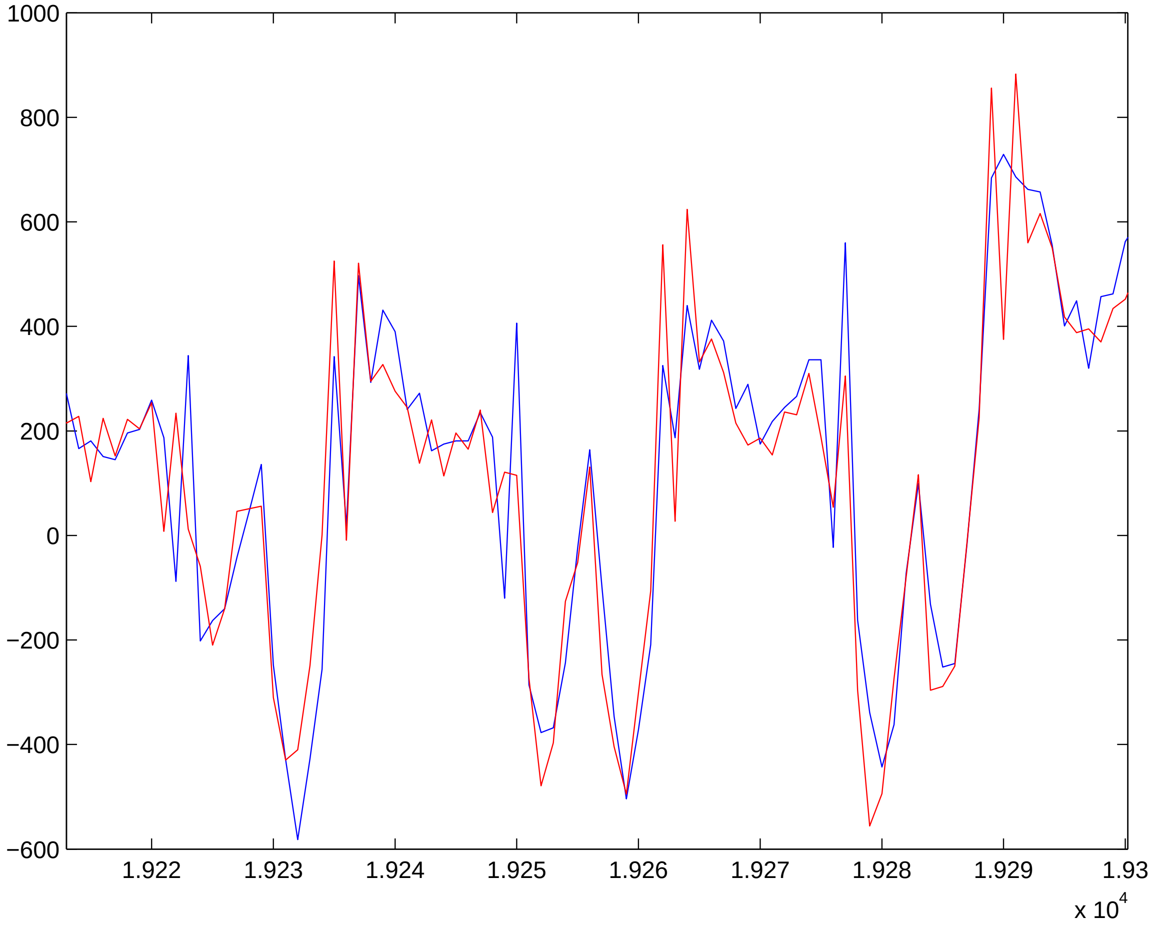

Figure 6: Two Power Signals where tested bits are 0 at 19260

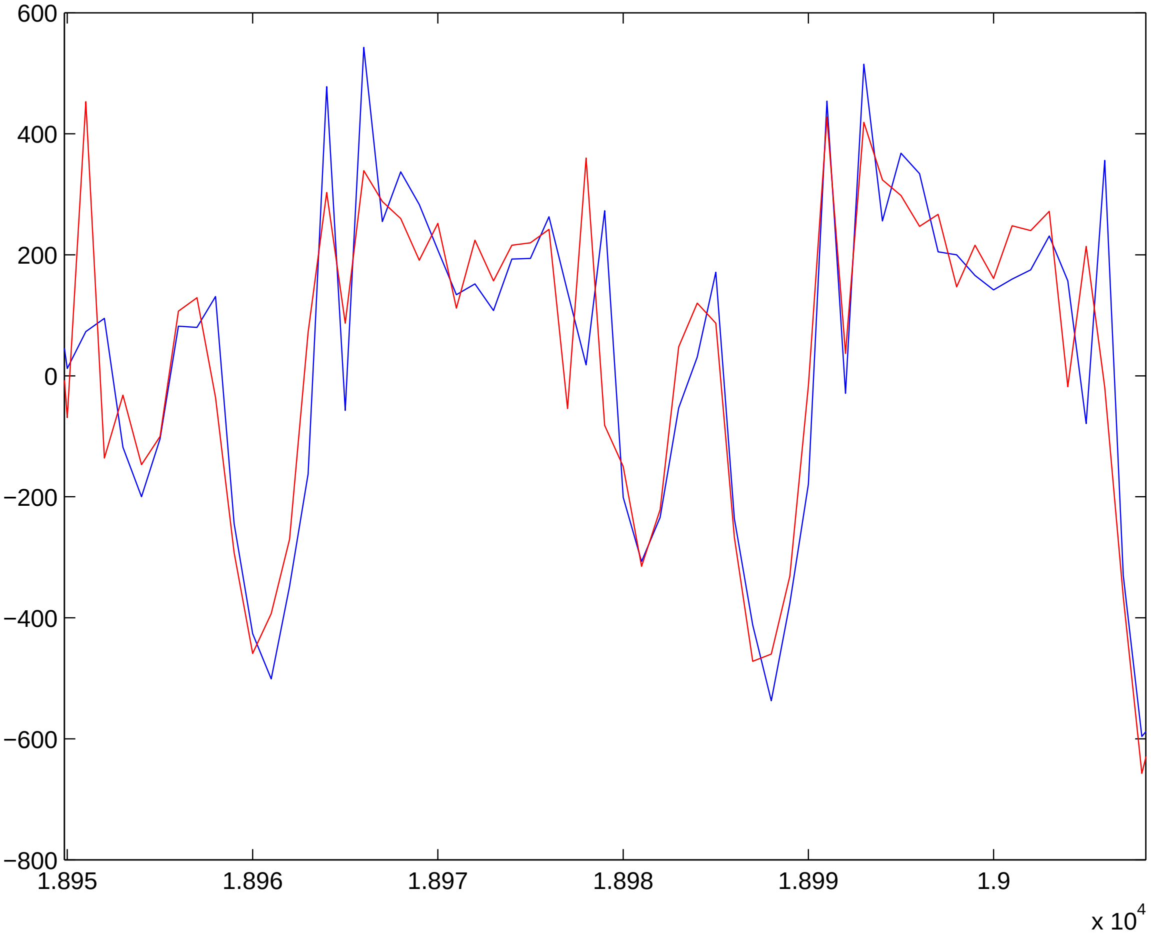

Figure 7: Two Power Signals where tested bits are 0 and 1 at 18990

EM emanations differed widely, are very close. This was also verified by taking averages of 500 power samples. The experiment once again confirmed that the averaged signal at the point of interest was identical for the 0 and 1 bit.

2.4 DEMA

The analogy for differential power analysis (DPA) is DEMA. DEMA can be used to attack DES in a manner akin to DPA. The following two figures (Figures 8 and 9) show the correlation peaks for the right hypothesis for an output bit of an S–box in the first round of DES in both power and EM.

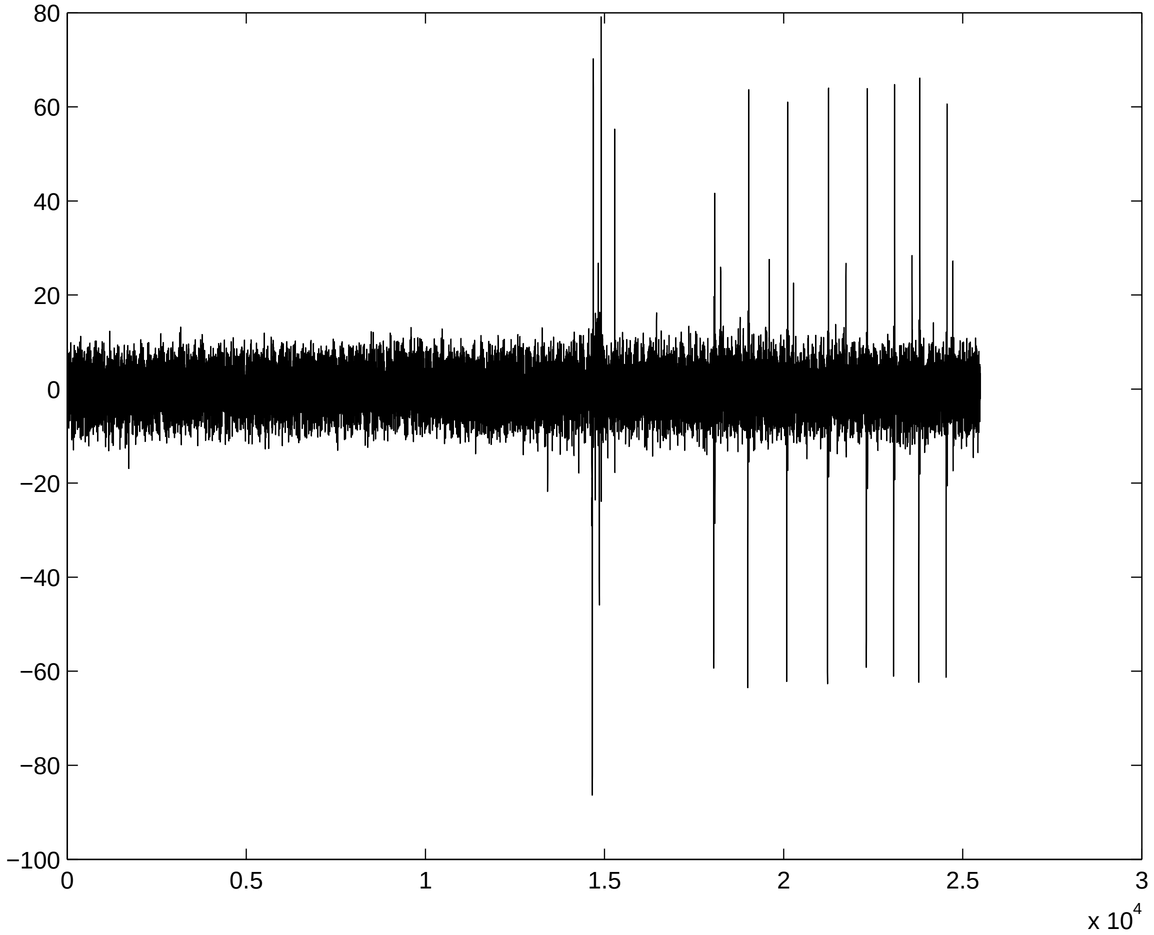

Figure 8: Correlation peaks for DPA attack on DES

2.5 Summary

Thus one can see that the information available in the EM signal is at least as useful as power but some cases even more information is available.

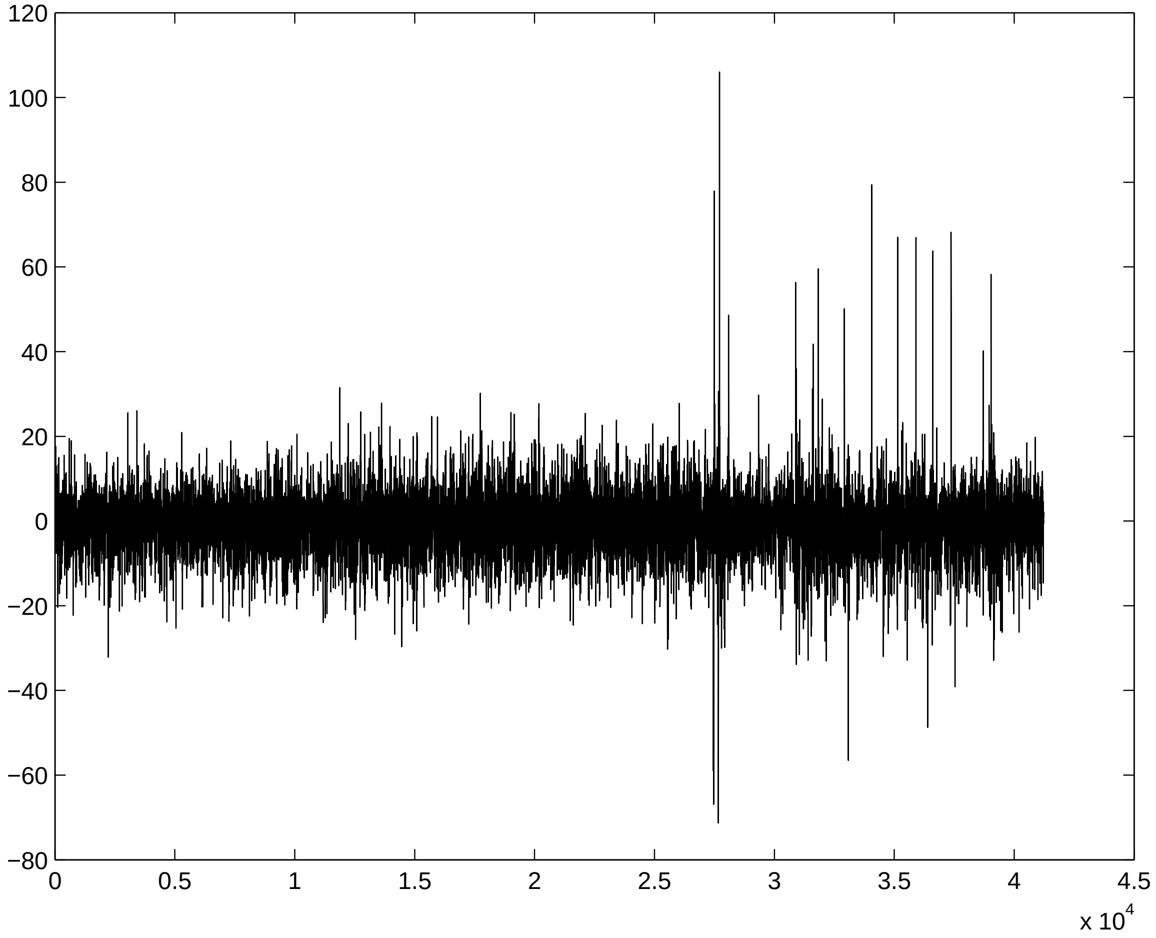

Figure 9: Correlation peaks for DEMA attack on DES

3 Comparing side–channels: EMF vs. Power

As shown in the previous section, the information in the EM signal is at least as useful as the power signal and in some cases, it is more. This opens the possibility that implementations that are secure against power analysis could be vulnerable to EM analysis.

3.1 Using SEMA to defeat DPA countermeasures

In [4], a suggested countermeasure to power analysis is to code using only those instructions where the leakage is not excessive and the key is refreshed at every invocation of DES using a non–linear key update. This means that an adversary is forced to figure out a significant portion of the key from a single sample. Otherwise any information gleaned from a single sample becomes useless for subsequent samples.

In [1], a secret sharing scheme is proposed as a countermeasure. The basic idea is that uncertainty about the information contained at each share is exponentially magnified in proportion to the number of shares.

Both of these countermeasures assume that there is no excessive leakage that will enable SPA on a DPA–protected implementation.

Any countermeasure that assumes that the amount of leakage of information about a particular instruction is not excessive and is within limits based on empirical observation on the power signal becomes vulnerable if that instruction leaks excessively in the EM signal. For example, the bit test instruction described in Section 2 is one such instruction for the chipcard that we examined. Based on its leakage characteristics in power, it can be potentially used in DPA protected DES implementations for performing bit–level permutations such as those used in key expansion (to compute round keys) and the P–permutations etc. However, its leakage characteristics in the EM signal are so egregious that both the countermeasures described above can be rendered useless if this instruction is excessively used in protecting implementations.

In the chipcard that we examined, once the noise generators are enabled, the SEMA attacks described above will not be successful as the noise can cause incorrect classification of the the bit value. However, one can use statistical techniques on multiple samples to develop workarounds against some countermeasures based on secret sharing ([1] etc. but probably not against [4]).

3.2 Using statistical techniques to defeat some DPA countermeasures

DPA countermeasures based on secret sharing schemes (such as those proposed in [1] and other places) rely on choosing an appropriate value for the number of shares based on the leakage characteristics and the desired level of resistance against DPA attacks in terms of the number of samples required to break the implementation. In the case where EM leaks more information than power, the assumptions about the leakage made in a share–based DPA resistant implementation do not hold for the EM channel. Hence, the implementation becomes vulnerable to EM attacks using a fewer number of samples.

To test this hypothesis, we implemented a two–way XOR–based secret sharing scheme for bits on the chipcard where the bit test instruction leaks more information in the EM channel. This sample code split the input bits into pairs of shares and tested the values of the shares using the bit test instruction. This was done at two distinct points in the computation corresponding to the points where the shares were being manipulated. As a sanity check, we confirmed that DPA and DEMA did not work, i.e., no single point in the power/EM signal correlated with any of the input bits.

We took 500 EM signals and subjected them to a second order differential EM analysis. Specifically, we defined a statistical measure on the signal at the two shares. We noticed that there was significant difference in the measure for the case where a zero bit was shared as opposed to where a one bit was shared. In fact, the difference was observable with just a few hundred samples. This difference was not observed in the power samples. We will illustrate these results in the next section which deals with a more general case.

3.3 Dealing with unknown code

In the previous section, it may seem that an adversary would need to know the points in the computation where the shares are being manipulated in order to perform the attack. This may suggest that knowledge of the code may be necessary. We now describe a method to circumvent this problem.

We would like to note that the techniques that we describe work under the assumption that the chipcard does not implement hardware countermeasures to complicate alignment of signals. In practice, many of these hardware countermeasures can be removed by signal processing and the techniques that we have described are still applicable.

Let us say that we are given a chipcard containing unknown k–way secret-sharing based DPA protected code for a known algorithm. Further assume that the chipcard hardware has already been analyzed for EM and power leakage, and it is known that there are vulnerable instructions that leak more information through the EM channel than the power channel. Further assume that some of these instructions have been used to manipulate shares. These, of course, are necessary conditions for EM attacks to be more effective than power attacks.

The value of k is usually small and for the time being, let us assume that k is 2. The procedure outlined below can be generalized for slightly larger values of k as well.

Fix, a priori, a reasonable limit L on the number of EM samples that we are willing to collect. The idea is that if k is small and if with knowledge of the code we could have broken the protected code using L samples, then with the procedure outlined below, we should be able to break the unknown protected code with $O(L)$ samples.

For the case of two–way split, we would like to locate the places where the shares of some algorithmic quantity are being manipulated using vulnerable instructions. Given that we know the algorithm, we can provide two different inputs such that the algorithmic quantity is different for these inputs and most of the other algorithmic quantities are the same within the window of interest. We take L EM samples for each of these two different inputs. If we know the exact locations where the shares of the algorithmic quantity are being manipulated, then we know that there is second order statistic, S, for the two points which can distinguish between the two different inputs, thus enabling hypothesis testing.

In the absence of location information, one can only assume that the distances between the two points where the shares are being manipulated is an integral number, D, of clock cycles. So the strategy is to compute, for all reasonable values of D, the statistic for each point on the signal with respect to a corresponding point that is D cycles away. This exercise is done for both sets of inputs.

If the shares of the algorithmic quantity are not manipulated at distance D, then the values of the statistic S at all points will be similar for the two inputs. However, for the right value of D, i.e., if the shares of the algorithmic quantity are indeed manipulated D cycles apart, then there will be a significant difference in the computed statistic S exactly at the point where the first share is manipulated. In practice, one would choose the two inputs so that a few (more than one) algorithmic quantities are different and this exercise will yield candidate locations where the shares of these quantities are manipulated. Once these locations are identified, then one can perform the second order attacks just as though the code was known.

We illustrate this exercise for an attack based on the bit test instruction mentioned in the previous section. In our code, the shares of one of the input bits were tested 40 cycles apart. In Section 2, when the bit is 1, the signal value at the bit test instruction is high and when the bit is 0, the signal value is low. For the case where the bit is split using an XOR-scheme, if the algorithmic bit (the input bit) is 0 then the shares would either be (0, 0) or (1, 1) with equal probability. When the algorithmic bit is 1 then the shares would be (0, 1) or (1, 0) with equal probability. This suggests that a good statistic would be the correlation coefficient between the corresponding signal points where the shares are being tested. This statistic would clearly be positive when the bit is 0 and negative when the bit is 1. The following figures show the results.

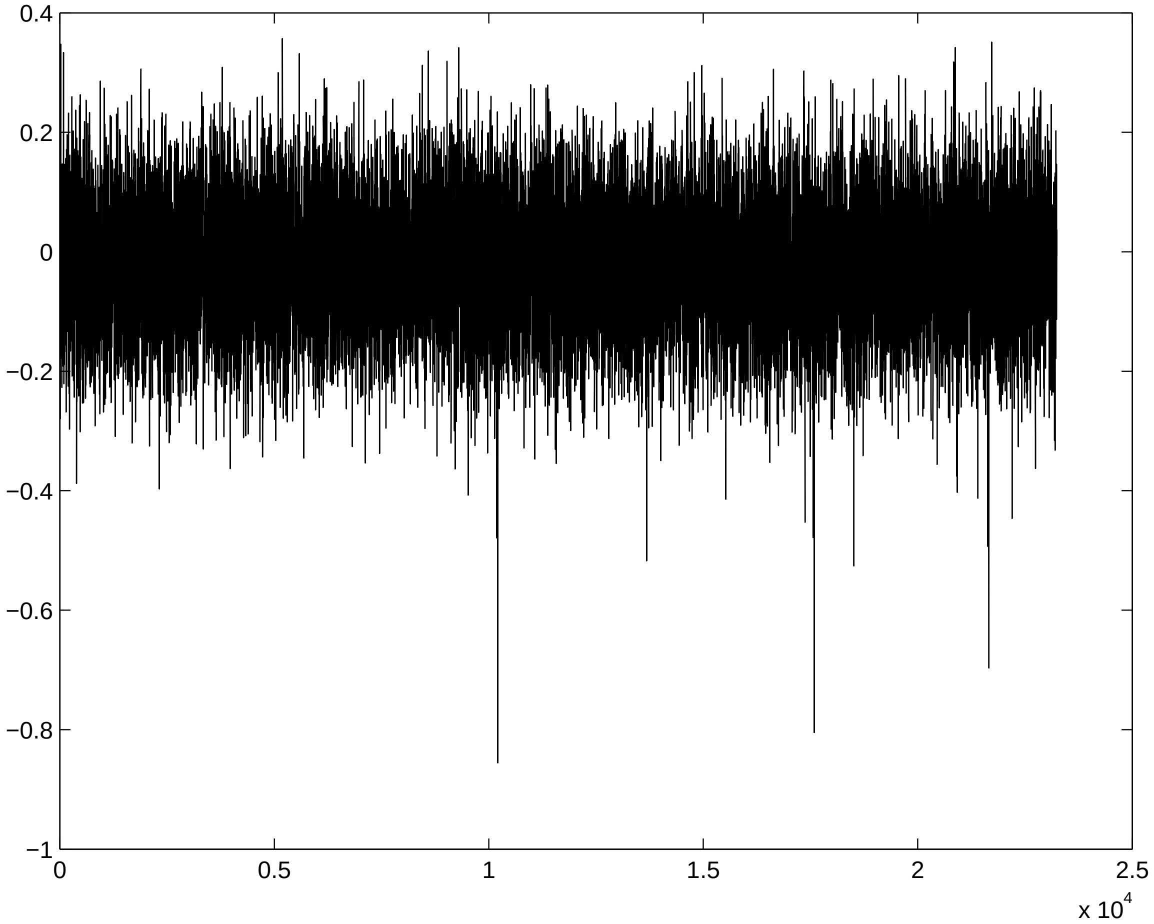

Figure 10: Difference in correlation statistics for $D = 40$, $L = 500$

We experimented with $L = 500$, for two different inputs, which differed in exactly three

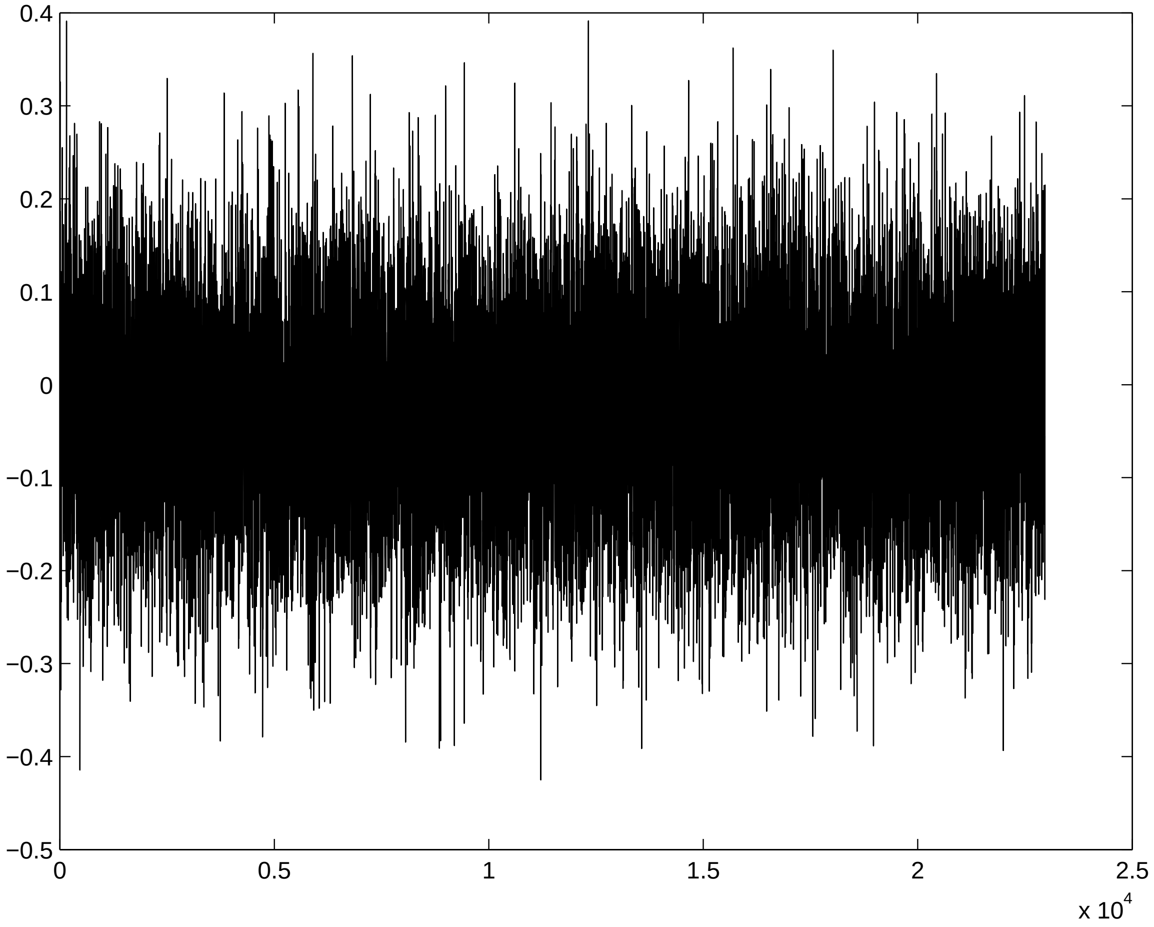

Figure 11: Difference in correlation statistics for $D = 50$, $L = 500$

bits. Figure 10 shows the difference in the correlations when the distance D is 40 for the two sets of inputs. We confirmed that the three significant negative peaks are at exactly the points where the first shares of the three bits (that differ) are being manipulated.

For illustrative purposes, Figure 11, shows the difference in correlations when the distance D is 50, which does not correspond to the distance between the shares of any of the three algorithmic bits. As can be seen, there is no significant difference at any point.

Our findings also show that $L = 200$ is adequate enough for this EM attack.

4 Ongoing Work

As stated in the Introduction, this work is part of a systematic effort to fully understand the nature of the EM Side-Channel. We are currently investigating several exciting avenues of research in this area.

One of our observations is that different harmonics of the fundamental frequency carry somewhat different information about the computation. One exciting consequence is that tools from multivariate statistical analysis can be used to combine information from multiple frequencies, to get much more information about the computation than is available from any single frequency. This will enable one to focus on specific aspects of the computation such as loads from memory. A related research area is to develop tools to determine the best carrier frequency to launch a specific side-channel attack. Another ongoing area of research is the development of a methodology to assess and protect chipcards and other devices from EM Attacks.

5 Acknowledgments

First and foremost, we would like to thank Bruce Archambeault for his invaluable help and guidance in assembling the experimental setup. We also owe a debt of gratitude to Helmut Scherzer for key components that enabled us to explore EM attacks further, in particular, help with the experimental setup, data collection and analysis tools and with smart card programming. This work has benefited greatly from interactions with Emmanuel Yashchin for his suggestions on the use of various statistics and statistical tools. Last but not the least, we would like to thank Elaine and Charles Palmer for several useful comments and their constant encouragement.

References

[1] Suresh Chari, Charanjit S. Jutla, Josyula R. Rao and Pankaj Rohatgi. Towards Sound Countermeasures to Counteract Power–Analysis Attacks. Advances in Cryptology — Proceedings of Crypto '99, Springer–Verlag, LNCS 1666, August 1999, pages 398–412.

[2] NSA Tempest Series `http://cryptome.org/#NSA--TS`.

[3] P. Kocher. Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS and Other Systems. Advances in Cryptology-Crypto '96, Lecture Notes in Computer Science # 1109, pp 104–113.

[4] P. Kocher, J. Jaffe and B. Jun. Differential Power Analysis: Leaking Secrets. Advances in Cryptology — Proceedings of Crypto '99, Springer Verlag, LNCS 1666, pages 388–397. One version of the paper is available online at `http://www.cryptography.com/dpa/technical/index.html`.

[5] The complete unofficial TEMPEST web page. Available at `http://www.eskimo.com/ joelm/tempest.html`.