Amplifying Side-Channel Attacks with Techniques from Block Cipher Cryptanalysis

Raphael C.-W. Phan[1] and Sung-Ming Yen[2⋆]

[1] Information Security Research (iSECURES) Lab,
Swinburne University of Technology (Sarawak Campus), 93576 Kuching, Malaysia
E-mail: rphan@swinburne.edu.my
[2] Laboratory of Cryptography and Information Security (LCIS)
Dept of Computer Science and Information Engineering
National Central University, Chung-Li, Taiwan 320, R.O.C.
E-mail: yensm@csie.ncu.edu.tw
http://www.csie.ncu.edu.tw/~yensm/

Abstract. We introduce the notion of *amplified side-channel attacks*, i.e. the application of block cipher cryptanalysis techniques to amplify effects exploitable by side-channel attacks. Such an approach is advantageous since it fully exploits the special characteristics of each technique in situations where each thrives the most. As an example, we consider the integration of block cipher cryptanalysis techniques into a particular type of side-channel attack, the differential fault attack (DFA). In more detail, we apply the DFA on the AES key schedule or on intermediate states within the AES and then exploit distinguishers based on Square attacks and impossible differential cryptanalysis to cover the remaining rounds. The use of techniques from conventional differential cryptanalysis in DFAs is not new; however, to the best of our knowledge, more advanced differential-like attack techniques have so far not been applied in collaboration with DFA. Further, while previous DFA attacks can only be mounted if faults are induced in the last or first (but with more restrictions) few rounds, our attacks alternatively show that even when faults are induced into some middle rounds, the DFA attacks still work, complementing existing results in literature; and thus showing that DFA attacks work regardless of where faults are induced. This is of importance because redundancy is a costly countermeasure against DFA and thus it is vital to study which rounds have to be protected. We hope that this completes the picture on the applicability of DFAs to block ciphers, and motivates thoughts into applying other advanced block cipher cryptanalysis techniques into other types of side-channel attacks.

Keywords: Attacks and countermeasures in hardware and software, side-channel attacks, cryptanalysis, fault attacks, Advanced Encryption Standard.

⋆ S.-M. Yen's research in this work was supported in part by the National Science Council of the Republic of China under contract NSC 94-2213-E-008-009 and also the University IT Research Center Project.

1 Introduction

Since the introduction of side-channel attacks [27, 11, 28], the importance of designing block ciphers resistant to these attacks has been in the limelight, and this resistance is now part of a cipher's design criteria. Security against side-channel attacks is especially significant in situations where ciphers are implemented in hardware like smart cards or tamper-resistant devices, where secrets are meant to be closely guarded and thus no room for compromise via leakage of secrets.

The study of block cipher cryptanalysis has developed tremendously in recent years. Attacks on block ciphers can typically be grouped into two major types, namely block cipher cryptanalysis which attacks the block cipher's design, and the side-channel attacks (also known as physical cryptanalysis [39]) which attack the block cipher's implementation. Each of these two categories of cryptanalysis has its own cryptanalytic assumption and theoretical foundation. Block-cipher cryptanalysis has been considered extensively in literature, while the physical cryptanalysis is still a somewhat new branch of research in applied cryptography.

Block cipher cryptanalysis refers to attacks that exploit the intrinsic weaknesses of block cipher components, e.g. differential cryptanalysis (DC) [7], linear cryptanalysis [30], slide attacks [9] and rectangle attacks [5]. Meanwhile, side-channel attacks are those that exploit potential physical properties or signal leakages, or bugs that occur when these block ciphers are implemented in particular situations and devices. Such attacks include timing analysis [27], power analysis [28], electromagnetic (EM) analysis [1], and fault analysis [11, 8].

Different though they be, both block cipher cryptanalysis and side-channel attacks have their own individual advantages over each other. More recently, block cipher cryptanalysis has considered more realistic attack models by fully exploiting advances of every kind of computing machine, e.g., developing specific cryptanalysis hardware based on reconfigurable devices e.g. field programmable gate arrays (FPGAs). Block cipher cryptanalysis has hence gone beyond pure theoretical work and now also considers practical issues. On the other hand, side-channel attacks in recent years have gone much further than merely exploiting implementation bugs. They sometimes fully exploit fundamental characteristics of the underlying cipher or related algorithms used to implement the cipher, and at times these attacks might reveal a possible vulnerability of the cipher. Side-channel attacks have hence moved towards theoretical design aspects. As both types of attacks gradually progress towards each other, it seems feasible therefore to consider bridging the gap by directly integrating attacks of both types so that we could exploit and use either type to suit the situation in which they thrive the most. This is the intent of this paper.

We discuss attempts to integrate side-channel attacks, including those in [2, 33, 38]. We then proceed with the first main contribution of this paper, i.e. to introduce the notion of *amplified side-channel attacks* which refers to the integration of one or more block cipher cryptanalysis techniques into side-channel attacks. In particular, we show that in situations where the effects due to conventional side-channel attacks on their own may be lost after some rounds of

the cipher, we can apply techniques from block cipher cryptanalysis to amplify these effects so that they cover more rounds and become more distinguishable.

To illustrate this, we consider the particular integration of the Square attack and the impossible differential attack into the differential fault analysis (DFA) [8], a type of side-channel attack. We apply this to the AES [13].

The second main contribution of this paper is that our approach of integrating block cipher cryptanalysis techniques into the DFA makes a much *weaker* assumption on the fault location in that it does not restrict the fault location to be within the last (or sometimes first[3]) few rounds only, as is the case with previous DFAs [10, 19, 15, 12, 32]. This leads to a more reasonable attack from the view point of fault attacks, and a less restricted attack model.

We therefore see that the advantages of our amplified approach is twofold. One, it allows the individual power of block cipher cryptanalysis techniques to be fully exploited by side-channel attacks. Side-channel attacks on their own would not be able to cover as many rounds of a cipher. Two, it allows DFA attacks to be mounted with a more flexible attack model, that faults could be induced even in rounds where previous DFAs are inapplicable. This study is important because redundancy is a costly countermeasure against DFA, thus one should ascertain exactly which rounds need to be protected.

Our attacks do not improve on previous work in situations where previous attacks are applicable, but our contribution is in showing that situations previously not susceptible to DFAs can now be attacked. Our work here therefore complement previous work; and together they show the universality of DFAs and how important it is to guard against them.

In the process, our discussions also provide an insight into the link between side-channel attacks and techniques from block cipher cryptanalysis

1.1 Attack Models: Block Cipher Cryptanalysis vs Side Channels

Block cipher cryptanalysis assumes an attacker has access to or control over input plaintexts and corresponding output ciphertexts − and even secret key relationships in the case of related-key (RK) attacks. He has no access to or control over what happens within the cipher's encryption process but knows the internal structure of the block cipher and exploits this to his advantage.

In contrast, side-channel attacks assume an attacker has much more access or control, not only over the inputs and outputs but also able to induce differences into intermediate rounds (via DFA) and/or predicting behaviour in these intermediate rounds (via timing, power or EM traces). Similarly, he also exploits his knowledge of the internal block cipher structure.

Therefore, the attack model used in side-channel attacks is much more powerful compared to that used in block cipher cryptanalysis. In fact, the former can be considered a superset of the latter.

[3] But with a higher text complexity or stricter text requirements.

1.2 Outline of This Paper

We describe the AES in Section 2. We recall in Section 3 previous attempts to integrate different side-channel attacks, and propose the notion of amplified side-channel attacks. In Section 4, we review past work on the DFA of the AES. We also comment on limitations of three DFA countermeasures proposed in [12] and argue that they would still allow for DFA to work. This observation recalls the importance of the DFA and its applicability to the proposed attacks considered in this paper. We then show in Section 5 how we could exploit techniques from the Square attack and impossible differential cryptanalysis to cause DFAs to work in situations where they were previously inapplicable. We conclude in Section 6.

2 The AES

The AES is a 128-bit block cipher which uses a 128-, 192- or 256-bit secret key, where the number of rounds are then 10, 12 and 14, respectively. For the rest of this paper, we will use AES to refer to the much-analysed 128-bit secret key version, unless otherwise stated. The 128-bit data block of the AES can be represented as a matrix of 4×4 bytes.

The input 128-bit block is passed through a round function, ρ iterated R times, hence R is the number of rounds. Simultaneously, the secret key, K is input to a key schedule to produce round keys RK_i ($i \in \{0 \dots R\}$) for use in each round. Each round function consists of four components applied in sequence:

- `SubBytes,SB`: a non-linear byte substitution.
- `ShiftRows,SR`: a cyclic shift of each row by different byte amounts.
- `MixColumns,MC`: a linear combination of all 4 bytes in the same column.
- `AddRoundKey,AR`$_i$: an exclusive-OR of data block with round key, RK_i.

Each round is identical except that an additional `AddRoundKey` is added before the first round and `MixColumn` is excluded from the last round.

3 Amplified Side-Channel Attacks

In this section, we first discuss previous attempts to integrate side-channel attacks, and then introduce the notion of *amplified side-channel attacks*.

In side-channel attacks, the attacker derives the embedded secret key by collecting and analyzing the obtained side-channel signals, or abnormal behavior and response in the case of fault-based attacks. Integrating several side-channel attacks means that the attacker collects multiple side-channel signals simultaneously and tries to obtain more information than what would be achieved from each attack if only applied individually. This enables the attacker to deduce possible secrets from each side-channel, then either intersect the results from individual side-channels to obtain the secret key, or deduce the secret key from the union of all collected side-channel signals. Furthermore, the secret key

can sometimes be deduced from some useful relationship between different side-channel signals. To summarize, the purpose of integrating side-channel attacks is to optimize the information retrieved from the limited amount of individual side-channel information.

Agrawal et al. [2] proposed a formal *multi-channel attack* framework for integrating multiple side-channel attacks, in particular by simultaneously collecting the power and electromagnetic signals. They demonstrated that integrating such multiple side-channel signals in the scenario considered in their work will lead to a two- to three-fold reduction in the requirement of samples needed for a traditional *differential power analysis* (DPA) attack [28]. In [38], another combined side-channel attack was developed by Walter and Thompson which employs previous techniques for timing attacks in order to exploit useful timing information from power signals. Note that this combined side-channel attack is applicable to a pure timing-attack-resistant and pure power-attack-resistant device. Later on, the efficiency of this integrated attack was enhanced by a factor of five and generalized considerably by Schindler [33].

3.1 Integrating Block Cipher Cryptanalysis with Side Channels

Though most side-channel attacks apply to full rounds of the cipher, they also have restrictions. For example, the differential fault analysis (DFA) requires that the attacker induce faults into some final rounds of a cipher. Faults induced earlier cannot be exploited by conventional DFA attacks. It is therefore reasonable to consider integrating block cipher cryptanalysis techniques into side-channel attacks to cover more rounds of the attacked cipher.

Referring to our example of the DFA, its limitation of requiring faults to be induced in the final rounds of the cipher can be overcome by allowing faults to be induced much earlier, and then applying block cipher cryptanalysis techniques to the rounds after where the fault was induced. Later in Section 5, we will show two examples of such *amplified side-channel attacks* on the AES, namely the Square-DFA and Impossible-DFA attacks.

Also in [35, 34], Schramm et al. proposed to overcome limitations of collision attacks on cipher implementations by using techniques from either the power analysis [28] or electromagnetic (EM) analysis [1], both of which are side-channel attacks. In more detail, collision attacks had so far been applied successfully to hash functions [14] and are essentially variants of the differential cryptanalysis in that they study the propagation of a collision − which is a non-difference − between a pair through some internal rounds. Nevertheless, collisions eventually disappear as the rounds increase, due to the diffusing nature of round functions, and hence cannot be directly observed at the output. Schramm et al. overcame this limitation by measuring the power or EM traces of the cipher implementation in the second round in order to predict whether collisions had occurred in the first round. To trigger such collisions, they collected sufficiently many chosen plaintext pairs with certain differences for input to the cipher.

Therefore, the collision side-channel attack proposed by Schramm et al. can in fact be viewed as the combination of differential cryptanalysis techniques

with the power or EM attack. We remark that this attack also falls into our amplified side-channel attack framework, though in direct complement to our Square/Impossible-DFA in Section 5. Whereas our Square/Impossible-DFA uses block cipher cryptanalysis techniques to enhance the effects of side-channel attacks, Schramm et al.'s collision side-channel uses side-channel techniques to enhance the effects of block cipher cryptanalysis.

We consider the unique advantages of each of the relevant block cipher cryptanalysis techniques or side-channel attacks:

- KP attack: allows attacker to obtain random plaintexts.
- CP attack: allows attacker to choose plaintexts with specific differences.
- RK attack: allows attacker to know or choose relationships (differences) between two or more unknown secret keys.
- DC attack: studies the propagation of differences between pairs through rounds of a cipher, and checks for corresponding differences at cipher output.
- DFA: allows attacker to induce differences into an intermediate round of a cipher.
- Timing/Power/EM attack: allows attacker to predict the behaviour (eg. difference or non-difference/collision) in some intermediate round of a cipher.

With this, we formalize the notion of the amplified side-channel attack:

Definition 1. *The amplified side-channel attack integrates block cipher cryptanalysis techniques with side-channel attacks, and consists of the following steps:*

1. (a) Use KP attack to collect some random plaintexts and/or RK attack to control relationship between two or more secret keys, OR
   (b) Use CP attack to control input plaintexts and/or RK attack to control relationship between two or more secret keys, OR
   (c) Use FA (fault attack) to induce differences into intermediate rounds of the cipher.
2. (a) Use DC attack to study propagation of differences through rounds and further use observed output to guess secret key bits, OR
   (b) Use timing, power or EM attack to predict difference or non-difference behaviour in intermediate rounds to guess secret key bits.

We can now express block cipher cryptanalysis, side-channel attacks or their combination under this amplified side-channel framework. e.g. differential cryptanalysis is simply the sequence of steps <1(b),2(a)>, timing/power/EM attack is <1(a),2(b)> or at times simply <2(b)>, DFA is <1(c),2(a)>, collision side-channel is <1(b),2(b)> and Square/Impossible-DFA (Section 5) is <1(c),2(a)>.

4 Previous DFAs on the AES and Countermeasures

In this section, we review past work on the DFA of AES. We also comment on the limitations of three DFA countermeasures proposed in [12] and argue that they would still allow for the DFA to work. All this recalls the importance of the DFA and the difficulty of guarding against it. This will motivate the choice of integrating the block cipher cryptanalysis techniques into the DFA in Section 5.

4.1 Previous DFAs on the AES

Blömer and Seifert [10] first considered the DFA on AES but worked with a restricted fault model. Their first attack required that a certain chosen bit of the intermediate state just after $\texttt{AR}_0$ be forced to 0, and required 128 faulty ciphertexts in order to determine the full key. Their second attack is implementation-dependent, and requires 256 faulty ciphertexts to obtain the full key.

This was followed by two attacks on the AES by Giraud [19]. The first attack also required to induce a bit fault at the beginning of the last round, R, and required 50 faulty ciphertexts. The second attack required 250 faulty ciphertexts and the faults had to be induced on a byte of the round keys, RK_{R-2}, and RK_{R-1}, and on the intermediate state before the second to last round, $R-1$.

Later, Dusart, Letourneux and Vivolo [15] presented another attack that required a fault to be induced on a byte before `MC` in the second to last round, $R-1$ and required about 50 faulty ciphertexts.

Chen and Yen [12] improved on Giraud's second attack to require about 30 faulty ciphertexts. Their attack similarly needed several byte faults to be induced in the last few rounds, but all on the round keys and none on intermediate states. In particular, faults had to be induced one at a time on one of four bytes of RK_{R-1}, followed by faults one at a time on each of 7 bytes of RK_{R-2}. Their attack model is efficient on AES key schedules that are generated on the fly.

Piret and Quisquater [32] presented two attacks on the AES. Their first attack required 8 faulty ciphertexts and that a byte fault be induced on the intermediate state between `MC` in round $R-2$ and `MC` in round $R-1$. Their second attack requires 2 faulty ciphertexts and that a byte fault be induced on the intermediate state between `MC` in round $R-3$ and `MC` in round $R-2$.

4.2 Comments on Countermeasures Against DFA

In [12], Chen and Yen presented a DFA on the AES key schedule based on three stages. The first stage involves inducing a fault in a byte of the 9th round key, RK_9. The next stage involves inducing a fault in a byte of the 8th round key, RK_8. Finally, the last stage involves inducing another fault in a different byte of the 8th round key, RK_8. All in all, the attack requires less than 30 faulty ciphertexts. Their attack depended on a fault being induced in the middle of the key schedule, as the round keys are generated on the fly, and hence relies on an induced fault in a round key inducing further faults on subsequent round keys and propagating the faults all the way to the ciphertext output.

Therefore, such an attack would have to occur during key accesses, during which faults are induced as the round keys are generated. Besides this limitation of their fault model, Chen and Yen also suggested some countermeasures [12].

Their first countermeasure suggests that in order to prevent DFA on the AES key schedule, round keys should not be generated on the fly, but should be pre-generated and then stored in memory. This eliminates the need for a key schedule, and also prevents the DFA attack described in [12].

We agree that such a countermeasure prevents the DFA attack on the AES key schedule described in [12]. However, even though round keys have been pre-generated and stored in memory, it is still possible to induce faults into them. In fact, it is at times even more desirable since faults induced in a round key would not cause any further faults in other subsequent round keys. This allows the attacker to have more control over the position of the faults that will be induced. Also, this removes the limitation that the attacker *must* induce the faults *during* key accesses when the round keys are generated. Since now the round keys are residing in memory all the time, the attacker could induce the faults at any time convenient to him, and hence is able to attack under a less restricted time duration. Therefore, it appears that this first countermeasure does not entirely prevent DFA attacks on the key schedule. On the contrary, it gives the attacker more control of the location and propagation of the faults induced, and less restrictions on when to induce the faults. This suggests that permanently storing the round key may not be sufficient to prevent DFA attacks. In Section 5, we will describe DFA attacks that work *especially* with this countermeasure.

The second countermeasure suggests to generate the round keys once whenever there is a need for an update. But again, for the round keys to be used, they would need to be stored somewhere in memory. Therefore, though this prevents the DFA attacks in [12], it falls to the same problem as the first countermeasure.

The third countermeasure suggests to apply a two-dimensional parity check on the round keys that are generated. Nevertheless, we point out that such an error check would inherit the limitations of conventional two-dimensional parity checks, that 4-bit errors or in this case faults would be undetectable. Therefore, this countermeasure will not prevent DFA attacks on the AES key schedule that involve inducing faults into 4 specific bits of the round keys. Though it may be argued that it is hard to induce 4 bits into exactly specified positions, this is not at all impossible with the optical fault induction attack that requires just US$30 worth of equipment bought at a second-hand camera shop [36].

5 Amplified Differential Fault Attacks on the AES

We describe two special cases of amplified side-channel attacks by exploiting techniques of block cipher cryptanalysis to enhance the DFA. These serve solely to illustrate the idea behind the notion of amplified side-channel attacks. Sections 5.1 and 5.2 respectively discuss how to integrate the Square attack and impossible differential cryptanalysis into the DFA.

5.1 Square-DFA on the AES

To mount a Square attack [13] on the AES requires us to use a Square distinguisher that works for three rounds of the AES. Suppose we have a group of 256 plaintexts that are totally identical to each other except for one byte in which they would have entirely different values. Then the Square distinguisher specifies that after encryption by 3 rounds of the AES, the 256 texts would have

the property that the XOR of all the 256 ciphertexts would result in a zero for all byte positions. This is a very interesting property and has been previously exploited to attack the AES up to 7 rounds [16, 18, 29].

Consider if we use equipment similar to that described in [36] but replaced with a suitable laser to increase precision, to induce a bit of fault in a byte of the 6th round key, RK_6, and repeating for 255 times, each time inducing one or more bits of fault into that same byte of RK_6 such that it would have all 256 (one correct and 255 faulty) values. These faults will not affect any of the other round keys. However, they will affect the AES encryption starting from the 6th round onwards. Therefore at the end of round 6, the 1 correct encryption and 255 faulty encryptions under these RK_6 values would be identical except for that one byte in which they would all have different values. By the Square distinguisher, this would propagate through the next three rounds until the end of round 9 when the XOR of all these 256 texts would result in a zero in all byte positions. What we have basically done is using the DFA to induce faults into RK_6 so that we can apply a 3-round Square distinguisher from rounds 7 to 9.

We can now guess all possible values of any byte of RK_{10} and partially decrypt these 256 (one correct and 255 faulty) ciphertexts by one round up to the output of round 9, and then check if their XOR gives a zero. The correct byte value of RK_{10} will always satisfy this, while a wrong value would only satisfy this with a very low probability, so it is almost guaranteed that only the right byte value remains. In the same way, move on to guess all possible values of another byte of RK_{10}. Repeat this for all 16 bytes of RK_{10}.

In summary, we need 1 correct ciphertext and 255 faulty ciphertexts, which can be reused for guessing all 16 bytes of RK_{10}. To guess each byte of RK_{10}, we make 256 guesses of the key byte and do 256 single-round AES encryptions, so in total $256 \times 256 \times 16 = 2^{20}$ single-round AES encryptions or $2^{20}/10 \approx 2^{16.5}$ AES encryptions for this DFA-induced Square attack.

Generalizations. Our attack considered inducing faults on one byte of RK_6. It equally applies when faults are induced on the *intermediate state* between `MC`s in rounds 6 and 7, or more generally between the `MC`s in rounds $R-4$ and $R-3$. In order to generalize this further, we recall that our attack outlined above induces the byte faults between the `MC`s in rounds $R-4$ and $R-3$, and applies a 3-round Square distinguisher in the rounds $R-3$ to $R-1$. In fact, we could also induce the byte faults a bit deeper into the middle of the AES, in particular between the `MC`s in rounds $R-5$ and $R-4$, in either the intermediate state or the corresponding round key, and again apply the 3-round Square distinguisher to the rounds $R-4$ and $R-2$. Then, to attack the last two rounds, we guess any column of RK_9 and the corresponding 4 bytes of RK_{10}, partially decrypt our ciphertexts by those last two rounds up to just before round 9 and check if the XOR is zero in any byte of the column corresponding to that column of RK_9. Repeating this four times, we obtain the entire RK_9 and RK_{10} with the same number of faulty ciphertexts.

Alternatively, we could also induce the byte faults between the MCs in rounds $R-3$ and $R-2$, in either the intermediate state or the corresponding round key, and hence apply the first 2 rounds of the 3-round Square distinguisher to the rounds $R-2$ and $R-1$. In this case, we are guaranteed that after round $R-1$ we would always have all 256 unique values in each byte of the correct and the faulty encryptions. This allows one to consider each byte of the last round key, RK_{10} at a time and performing an attack similar to the above with the same number of faulty ciphertexts, except that instead of computing the resultant XOR value, one would further have to check that all 256 unique values exist.

Finally, we can induce the faults between the MCs in rounds $R-2$ and $R-1$ and apply the first round of our 3-round Square distinguisher to the round $R-1$. This states that after round $R-1$ we would always have all 256 unique values in the column in which the fault was induced. We guess at a time each of the 4 bytes of RK_{10} that correspond to that column, each time reusing the same faulty ciphertexts. We repeat this four times to obtain all 4 columns of the key, and hence requiring a total of 2^{10} faulty ciphertexts.

Discussion. Our attacks are the *only* DFA-style attacks that can be applied to the AES if faults can only be injected between the rounds $R-4$ and $R-3$, and between the rounds $R-5$ and $R-4$, which would be the case for AES implementations that incorporate countermeasures against standard previous DFAs. Previous DFAs do not work for these rounds at all, even with the entire code book! Our results therefore stress that one should guard against DFAs in any round of the AES, and not just the outer (first or last) few rounds.

5.2 Impossible-DFA on the AES

Before we proceed with a description of the attack, we briefly introduce a 3-round impossible differential of the AES, which is a variant of the 4-round impossible differential discussed in [6]. Specifically, our 3-round impossible differential states that given a pair of plaintexts equal in all bytes (called *passive bytes*) except one (*active*) byte in which the pair differs, then the ciphertexts after 3 rounds cannot be equal in any of the 16 bytes at the state just before MC in round 3. Note that only the ShiftRows and MixColumns operations affect the number and positions of the active bytes, and that MC and AR are invariant of each other [13].

We use this distinguisher for our attack. Consider that a fault is induced on any byte of the 6th round key, RK_6 that is stored in memory. This fault will not affect any of the other round keys. However, it will affect the AES encryption starting from the 6th round onwards. A correct and a faulty encryption would then differ in a byte prior to the 7th round. This difference will propagate to 4 bytes after round 7, and if we consider our 3-round impossible differential distinguisher previously discussed, this will suggest that after round 9 we would never have any equal byte between the correct and the faulty encryptions at the state just before MC in round 9. We will henceforth denote this state as X.

We have in essence used concepts from the DFA to induce a fault into any byte of RK_6, in order to cause a byte of difference between a correct and a

faulty encryption prior to the 7th round. We then apply the 3-round impossible differential from rounds 7 to 9 up to X, and with this in place, we guess all 2^{32} possible values of the four bytes of the last round key, RK_{10} that correspond to any column at X, say the first column, partially decrypt the correct and the faulty ciphertexts by one round up to X and check if we get any equal bytes in that column of X. If this is the case, then the guessed values of RK_{10} are wrong since they caused the impossible differential to occur. These values are removed from the list of 2^{32} possible values of RK_{10}. Doing this with one faulty encryption causes about $(1 - 2^{-6}) \times 2^{32}$ possible key values to remain[4] [6]. Repeating this with a sufficient number of faulty encryptions, in this case about 2^{11}, will leave $2^{32}(1 - 2^{-6})^{2^{11}} \approx 0$ wrong key values, so only the correct key value remains [6]. With this, we obtain 4 bytes of RK_{10} that correspond to that column of X. We can repeat the same steps for the bytes of RK_{10} that correspond to the other 3 columns of X, and hence obtain the entire RK_{10}.

To obtain each column of RK_{10}, the attack needs 1 correct ciphertext and 2^{11} faulty ciphertexts which can be reused. Also, to obtain each column of RK_{10}, we do 2^{32} single-round AES encryptions, so this makes it 2^{34} single-round AES encryptions or $2^{32}/10 \approx 2^{28.5}$ AES encryptions.

Generalizations. This can be generalized similarly to Section 5.1, hence the flexibility of inducing the byte fault in the round key or in the intermediate state between the `MC`s in rounds $R-4$ and $R-3$. However, in contrast to the case of the DFA and Square attacks, it is not possible to further generalize and make this attack work when the fault is induced at other locations simply because the first few rounds of the 3-round impossible differential are in fact probability-one differentials, so the propagation of the active and passive would always occur irrespective of the guessed key values, hence cannot be used for filtering wrong keys. For AES-192 (respectively AES-256), one could consider applying the 4-round (respectively 5-round) impossible differentials reported by Kim et al. [24].

Discussion. As was the case with our attacks in Section 5.1, our attacks in this section are the *only* DFA-style attacks that can be applied to the AES if faults can only be injected between the rounds $R-4$ and $R-3$.

6 Concluding Remarks

We have introduced the notion of amplified side-channel attack, and illustrated specifically with Square-DFA and impossible-DFA attacks on the AES. In Table 1, we compare between previous DFAs and our amplified DFA attacks on the AES. We have indicated in Table 1 the best DFAs based on the fault location. Clearly, Dusart, Letourneux and Vivolo's [15] attack is the best for faults induced in round $R-1$ while Piret and Quisquater's [32] attacks are the best for

[4] The probability of getting a passive byte is 2^{-8} so the probability of getting any passive byte in a column is 2^{-6}.

faults induced between the rounds $R-3$ and $R-1$. Our amplified DFA attacks are the best and only attacks that are applicable for faults induced between the rounds $R-5$ through to $R-3$. Therefore, we can think of all these attacks as complementing each other. Depending on where the faults can be induced, the cryptanalyst has the option to choose the best that is currently available. Our results also complete the picture of applying DFAs to the AES, and demonstrate that it is sometimes useful to apply techniques from block cipher cryptanalysis to amplify effects caused by side-channel attacks. The integration of two or more cryptanalysis techniques often results in a more powerful attack. This is due to the fact that since we are using more than one attack, we could selectively exploit the special features of each attack in situations or parts of the cipher where it thrives the most. Thus, we ensure the most suitable attack is applied to block cipher components most susceptible to it in order to get an optimum result.

Table 1. Comparison of DFAs on AES.

Attack type	Fault model	Fault location (Which round)	Faulty texts	Source	Best attack
DFA	Bit faults	1 (after $\texttt{AR}_0$)	128	[10]	
DFA	Impl-depend.	-	256	[10]	
DFA	Bit faults	$R-1$ (after $\texttt{AR}_{R-1}$)	50	[19]	
DFA	Byte faults	$R-1$ (after SR)	50	[15]	✓
DFA	Byte faults	$R-2$ and $R-1$ (RK_{R-2}, and RK_{R-1})	250	[19]	
DFA	Byte faults	$R-2$ and $R-1$ (RK_{R-2}, and RK_{R-1})	30	[12]	
DFA	Byte faults	Between MCs in $R-2$ and $R-1$	8	[32]	✓
Square-DFA	Byte faults	Between MCs in $R-2$ and $R-1$	2^{10}	This paper	
DFA	Byte faults	Between MCs in $R-3$ and $R-2$	2	[32]	✓
Square-DFA	Byte faults	Between MCs in $R-3$ and $R-2$	256	This paper	
Impossible-DFA	Byte faults	Between MCs in $R-4$ and $R-3$	2^{11}	This paper	
Square-DFA	Byte faults	Between MCs in $R-4$ and $R-3$	256	This paper	✓
Square-DFA	Byte faults	Between MCs in $R-5$ and $R-4$	256	This paper	✓

Note: Best attack is indicated based on various different fault locations.

References

1. D. Agrawal, B. Archambeault, J.R. Rao, P. Rohatgi, "The EM Side-Channel(s)," *CHES '02*, LNCS 2523, pp. 29–45, Springer-Verlag, 2002.
2. D. Agrawal, J.R. Rao, P. Rohatgi, "Multi-Channel Attacks," *CHES '03*, LNCS 2779, pp. 2–16, Springer-Verlag, 2003.
3. E. Biham, "New Types of Cryptanalytic Attacks using Related Keys," *Advances in Cryptology – EUROCRYPT '93*, LNCS 765, pp. 398–409, Springer-Verlag, 1994.
4. E. Biham, A. Biryukov, A. Shamir, "Miss in the Middle Attacks on IDEA, Khufu and Khafre," *Advances in Cryptology – EUROCRYPT '99*, LNCS 1636, pp. 124–138, Springer-Verlag, 1999.
5. E. Biham, O. Dunkelman, N. Keller, "The Rectangle Attack – Rectangling the Serpent," *Advances in Cryptology – EUROCRYPT '01*, LNCS 2045, pp. 340–357, Springer-Verlag, 2001.
6. E. Biham, N. Keller, "Cryptanalysis of Reduced Variants of Rijndael," Submitted to 3rd AES Conference, U.S., 2000.
7. E. Biham, A. Shamir, "Differential Cryptanalysis of the Data Encryption Standard," Springer-Verlag, 1993.
8. E. Biham, A. Shamir, "Differential Fault Analysis of Secret Key Cryptosystems," *Advances in Cryptology – CRYPTO '97*, LNCS 1294, pp. 513–525, Springer-Verlag, 1997.
9. A. Biryukov, D. Wagner, "Slide Attacks," *FSE '99*, LNCS 1636, pp. 245-259, Springer-Verlag, 1999.
10. J. Blömer, J.-P. Seifert, "Fault Based Cryptanalysis of the Advanced Encryption Standard," *Financial Cryptography '03*, LNCS 2742, pp. 162–181, Springer-Verlag, 2003.
11. D. Boneh, R.A. Demillo, R.J. Lipton, "On the Importance of Checking Cryptographic Protocols for Faults," *Advances in Cryptology – EUROCRYPT '97*, LNCS 1233, pp. 37–51, Springer-Verlag, 1997.
12. C.-N. Chen, S.-M. Yen, "Differential Fault Analysis on AES Key Schedule," *ACISP '03*, LNCS 2727, pp. 118–129, Springer-Verlag, 2003.
13. J. Daemen, V. Rijmen, "AES proposal: Rijndael (version 2)," Updated Documentation and Complete Specification, 1999.
14. H. Dobbertin, "Cryptanalysis of MD4," *Journal of Cryptology*, vol. 11, pp. 235–271, Springer-Verlag, 1998.
15. P. Dusart, G. Letourneux, O. Vivolo, "Differential Fault Analysis on A.E.S.," IACR Cryptology ePrint Archive, No. 010, 2003.
16. N. Ferguson, J. Kelsey, S. Lucks, B. Schneier, M. Stay, D. Wagner, D. Whiting, "Improved Cryptanalysis of Rijndael," *3rd AES Conference*, 2000.
17. S. Furuya, "Slide Attacks with a Known-Plaintext Cryptanalysis," *ICISC '01*, LNCS 2288, pp. 214–225, Springer-Verlag, 2002.
18. H. Gilbert, M. Minier, "A Collision Attack on 7 Rounds of Rijndael," *3rd AES Conference*, 2000.
19. C. Giraud, "DFA on AES," IACR Cryptology ePrint Archive, No. 008, 2003.
20. M. Hellman, S. Langford, "Differential-linear Cryptanalysis," *Advances in Cryptology – CRYPTO '94*, LNCS 839, pp. 17–26, Springer-Verlag, 1994.
21. T. Jakobsen, L.R. Knudsen, "The Interpolation Attack on Block Ciphers," *FSE '97*, LNCS 1267, pp. 28–40, Springer-Verlag, 1997.
22. G. Jakimoski, Y. Desmedt, "Related-Key Differential Cryptanalysis of 192-bit Key AES Variants," *SAC '03*, LNCS 3006, pp. 208–221, Springer-Verlag, 2004.

23. J. Kelsey, T. Kohno, B. Schneier, "Amplified Boomerang Attacks against Reduced-round MARS and Serpent," *FSE '00*, LNCS 1978, pp. 75–93, Springer-Verlag, 2001.
24. J. Kim, S. Hong, J. Sung, S. Lee, J. Lim, S. Sung, "Impossible Differential Cryptanalysis for Block Cipher Structures," *Progress in Cryptology – INDOCRYPT '03*, LNCS 2904, pp. 82–96, Springer-Verlag, 2003.
25. J. Kim, G. Kim, S. Hong, S. Lee, D. Hong, "The Related-Key Rectangle Attack – An Application to SHACAL-1," *ACISP '04*, LNCS 3108, pp. 123–136, Springer-Verlag, 2004.
26. L.R. Knudsen, D. Wagner, "Integral Cryptanalysis," *FSE '02*, LNCS 2365, pp. 112–127, Springer-Verlag, 2002.
27. P. Kocher, "Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and Other Systems," *Advances in Cryptology – CRYPTO '96*, LNCS 1109, pp. 104–113, Springer-Verlag, 1997.
28. P. Kocher, J. Jaffe, B. Jun, "Differential Power Analysis," *Advances in Cryptology – CRYPTO '99*, LNCS 1666, pp. 388–397, Springer-Verlag, 1999.
29. S. Lucks, "Attacking Seven Rounds of Rijndael under 192-bit and 256-bit Keys," *3rd AES Conference*, 2000.
30. M. Matsui, "Linear Cryptanalysis Method for DES Cipher," *Advances in Cryptology – EUROCRYPT '93*, LNCS 765, pp. 386–397, Springer-Verlag, 1994.
31. R.C.-W. Phan, S. Furuya, "Sliding Properties of the DES Key Schedule and Potential Extensions to the Slide Attacks," *ICISC '02*, LNCS 2587, pp. 138–148, Springer-Verlag, 2003.
32. G. Piret, J.-J. Quisquater, "A Differential Fault Attack Technique against SPN Structures, with Application to the AES and KHAZAD," *CHES '03*, LNCS 2779, pp. 77–88, Springer-Verlag, 2003.
33. W. Schindler, "A Combined Timing and Power Attack," *PKC '02*, LNCS 2274, pp. 263–279, Springer-Verlag, 2002.
34. K. Schramm, G. Leander, P. Felke, C. Paar, "A Collision-Attack on AES: Combining Side Channel- and Differential-Attack," *CHES '04*, LNCS 3156, pp. 163–175, Springer-Verlag, 2004.
35. K. Schramm, T. Wollinger, C. Paar, "A New Class of Collision Attacks and its Application to DES," *FSE '03*, LNCS 2887, pp. 206–222, Springer-Verlag, 2003.
36. S.P. Skorobogatov, R.J. Anderson, "Optical Fault Induction Attacks," *CHES '02*, LNCS 2523, pp. 2–12, Springer-Verlag, 2003.
37. D. Wagner, "The Boomerang Attack," *FSE '99*, LNCS 1636, pp. 156–170, Springer-Verlag, 1999.
38. C.D. Walter, S. Thompson, "Distinguishing Exponent Digits by Observing Modular Subtractions," *Topics in Cryptology – CT-RSA '01*, LNCS 2020, pp. 192–207, Springer-Verlag, 2001.
39. S.-M. Yen, S. Kim, S. Lim, S.-J. Moon, "A Countermeasure Against One Physical Cryptanalysis May Benefit Another Attack," *ICISC '01*, LNCS 2288, pp. 414–427, Springer-Verlag, 2001.

A Integrated Block Cipher Cryptanalysis

In this appendix, we summarize previous attempts to integrate block cipher cryptanalysis techniques. This is hoped to motivate more work in this direction. The first was notably the *differential-linear cryptanalysis* [20] in 1994, which combined differential cryptanalysis (DC) [7] with linear cryptanalysis [30]. Denote

the block cipher, $E(P) = E_2(E_1(P))$ as the composition of two halves[5], where E_1 (respectively E_2) denotes the earlier (respectively later) half of the cipher. Then the differential-linear cryptanalysis applies differential cryptanalysis to E_1 to enable linear cryptanalysis to be applied to E_2. Differential cryptanalysis is a *chosen-plaintext* (CP) attack where the attacker needs to obtain encryptions of plaintexts with a certain chosen difference between them. Meanwhile, linear cryptanalysis is a *known-plaintext* (KP) attack in that the attacker simply needs to be able to obtain some known plaintext values and their corresponding ciphertexts. CP attacks that are of the differential cryptanalysis naturally can be converted to KP attacks but with a considerably high increase in text complexity. In particular, suppose that we need m pairs of CPs with a certain difference between them. Then with $2^{n/2}\sqrt{2m}$ random KPs, we can form $2^n \times m$ pairs of KPs, of which the probability of getting a pair with a certain difference is 2^{-n}, and therefore we get m pairs of CPs with the desired difference [7].

In 2001, Furuya [17] considered combining the slide attacks [9] with KP attacks such as linear cryptanalysis. We consider that such attacks should rightly be called the *slide-KP attacks*. These apply the slide attacks to the entire cipher E to enable KP attacks to be applicable to some outer rounds of E. Slide attacks are generally KP attacks, but if chosen-plaintext queries are possible, the attacker could mount the slide attacks with a much reduced text complexity.

In 2002, the *integral-interpolation attacks* [26] where presented, which applies integral cryptanalysis [26] to E_1 to enable the interpolation attacks [21] on E_2. Integral cryptanalysis is a CP attack while interpolation attacks are KP attacks.

Finally, in cases where it is possible for the attacker to obtain the encryptions of plaintexts under two related keys, K and K', he could then mount related-key versions of any of these block cipher cryptanalysis attacks. Examples of such considerations include the related-key differential attacks [3], related-key slide attacks [3], related-key square attacks [16], related-key impossible differential cryptanalysis [22], and the related-key rectangle attack [25].

As an aside, we note that some attacks have been proposed that apply the same kind of attacks to both E_1 and E_2. In this respect, we consider such attacks as a special case of integrated block cipher cryptanalysis. For instance, the *boomerang attack* [37] uses chosen plaintexts to mount differential cryptanalysis to E_1 and then enables differential cryptanalysis on E_2 by making adaptively-chosen ciphertext queries from the other end of the cipher. Note that adaptively-chosen plaintext-ciphertext attacks are much harder to mount than CP or KP attacks. The *amplified boomerang attack* [23] and *rectangle attack* [5] are enhancements of the boomerang attack. They similarly apply differential cryptanalysis to E_1 but the number of chosen plaintext queries used is increased considerably such that enough texts with the desired chosen difference appear probabilistically after E_1 to allow differential cryptanalysis to be further mounted on E_2. The *inside-out attack* [37] obtains a high number of known plaintexts such that enough texts with the desired chosen difference appear probabilistically in the middle of the cipher so that the difference will propagate outwards in both direc-

[5] Not necessarily consisting of the same number of rounds.

tions through E_1 and E_2. The *miss-in-the-middle attack* [4] applies differential cryptanalysis to both E_1 and E_2 in such a way that the differences between the texts in the middle of the cipher contradict each other.

Our main observation is that one starts by first applying a CP attack or a KP attack on E_1, to enable a KP attack to be mountable on E_2. In some cases where it is possible to considerably increase the number of texts obtained, then one could also apply CP attacks to E_2.

Definition 2. *Integrated block cipher cryptanalysis applies different types of cryptanalysis attacks to the first and second halves of a cipher, E. In particular, CP or KP attacks are applied to E_1 to enable KP attacks on E_2.*

Fact 1 *CP attacks can be converted to KP by increasing the text complexity.*

Corollary 1. *In some cases one could also mount CP attacks on E_2 when it is possible to considerably increase the number of texts obtained.*

Corollary 2. *In cases where it is only possible to apply CP or KP attacks on one sequence of rounds (one half instead of two) within E, then this can be viewed as a special case of integrated block cipher cryptanalysis where the attack is applied to either E_1 or E_2.*

The notion of integrated block cipher cryptanalysis opens doors to numerous possible attacks where previous attacks on their own failed. In general, any integration of CP and KP attacks could be mounted on ciphers. Further, related-key versions of the aforementioned integrated attacks are also possible.

In Table 2, we consider previous integration of CP and KP attacks, where the rows and columns indicate attacks applied to E_1 and E_2, respectively: differential-differential attacks e.g. the boomerang [37], inside-out [37], amplified boomerang [23], rectangle [5], and miss-in-the-middle [4]; we also have differential-linear attacks [20] and integral-interpolation attacks [26]. The slide-linear attack [17] is just one of the ways one could mount his proposed slide-KP attacks, another variant he suggested being the slide-partitioning attacks [17]. On this note, we also remark that it would be possible to have slide-interpolation attacks. The slide-slide (double slide) attack [31] has also been considered.

Table 2. Previous integration of block cipher cryptanalysis attacks.

	Differential	Integral	Linear	Interpolation	Partitioning	Slide
Differential	[37, 23, 5, 4]		[20]			
Integral				[26]		
Linear						
Interpolation						
Partitioning						
Slide			[17]	New	[17]	[31]