Systematic Construction and Comprehensive Evaluation of Kolmogorov-Smirnov Test Based Side-Channel Distinguishers

Hui Zhao[1], Yongbin Zhou[1,⋆], François-Xavier Standaert[2], and Hailong Zhang[1]

[1] State Key Laboratory of Information Security,
Institute of Information Engineering, Chinese Academy of Sciences,
89A, Mingzhuang Rd, Beijing, 100093, P.R. China
{zhaohui,zhouyongbin,zhanghailong}@iie.ac.cn
[2] UCL Crypto Group, Université catholique de Louvain, Belgium
fstandae@uclouvain.be

Abstract. Generic side-channel distinguishers aim at revealing the correct key embedded in cryptographic modules even when few assumptions can be made about their physical leakages. In this context, Kolmogorov-Smirnov Analysis (KSA) and Partial Kolmogorov-Smirnov analysis (PKS) were proposed respectively. Although both KSA and PKS are based on Kolmogorov-Smirnov (KS) test, they really differ a lot from each other in terms of construction strategies. Inspired by this, we construct nine new variants by combining their strategies in a systematic way. Furthermore, we explore the effectiveness and efficiency of all these twelve KS test based distinguishers under various simulated scenarios in a univariate setting within a unified comparison framework, and also investigate how these distinguishers behave in practical scenarios. For these purposes, we perform a series of attacks against both simulated traces and real traces. Success Rate (SR) is used to measure the efficiency of key recovery attacks in our evaluation. Our experimental results not only show how to choose the most suitable KS test based distinguisher in a particular scenario, but also clarify the practical meaning of all these KS test based distinguishers in practice.

Keywords: Side-Channel Analysis, Distinguisher, Kolmogorov-Smirnov Test, Construction, Evaluation.

1 Introduction

Side-channel attack aims at identifying the secret information embedded in a cryptographic device from its physical leakages. One of the most famous side-channel attacks is Differential Power Analysis (DPA), which was proposed by Kocher in his seminal work [1]. Generally, DPA employs some type of statistics (also referred to as *distinguisher*) to reveal the correct key hypothesis about the secret key or part of it within a set of candidates. In side-channel attacks, the

⋆ Corresponding author.

R.H. Deng and T. Feng (Eds.): ISPEC 2013, LNCS 7863, pp. 336–352, 2013.

most famous two distinguishers known are distance-of-means [1], Pearson correlation coefficient in Correlation Power Analysis (CPA) [3]. Meanwhile, other variants of these two distinguishers, such as Multi-bit DPA [2] and Partitioning Power Analysis (PPA) [4], are also proposed to enhance the performance of DPA and CPA respectively. Concerning these distinguishers, recent works [5,6] has shown that DPA, CPA and even PPA are in fact asymptotically equivalent to each other, given that they are provided with the same a priori information about the leakages. Therefore, these distinguishers are collectively called CPA-like distinguishers. Essentially, all these CPA-like distinguishers exploit linear dependency between key-dependent hypothetical power consumptions and physical leakages.

Even though CPA-like distinguishers are well capable of measuring linear dependency between hypothetical power consumptions and physical leakages, they become less efficient when the dependency is not strictly linear [10]. In light of this, Mutual Information Analysis (MIA) was proposed by Gierlichs in [7] to measure total dependency (both linear and nonlinear) between the hypothetical power consumptions and the physical leakages. Consequently, MIA is considered to be generic because it is capable of dealing with the total dependency. Although MIA is generic, it also bears some technical challenges. For example, the Probability Density Function (PDF) estimation in MIA is widely accepted to be a difficult problem [8,9]. Experiments in [10,11] confirmed that the PDF estimation methods have a decisive impact on the performance of MIA. Therefore, the performance of MIA depends on the accuracy of the estimation methods. Considering the PDF of MIA is hard to estimate accurately, Kolmogorov-Smirnov Analysis (KSA) [10] and Partial Kolmogorov-Smirnov analysis (PKS) [13] were independently proposed. KSA and PKS use Cumulative Density Function (CDF) estimation, instead of PDF estimation, to avoid explicit PDF estimation. Both KSA and PKS sound like promising alternatives for MIA, but which one is a better alternative for MIA in key recovery attacks?

On the one hand, although both KSA and PKS are based on Kolmogorov-Smirnov (KS) test, they differ a lot from each other in terms of construction strategies. One natural yet important question is that whether or not we can construct more efficient distinguishers via combining different construction strategies by both KSA and PKS. For all these KS test based distinguishers, how can we choose the most suitable distinguisher in a certain scenario? For all these KS test based distinguishers, to what extent do they pose severe threats on the implementations of cryptographic modules in practice? In order to answer these questions above, we will investigate the efficiency of all these KS test based distinguishers in a comprehensive comparison framework. Since it seems difficult to study the relationship of all KS test based distinguishers theoretically, we will explore the advantages and limitations of KS test based distinguishers experimentally.

Note that we only compare KS test based distinguishers in a univariate setting, due to the fact that PKS does not have multivariate extensions.

1.1 Our Contributions

The contributions of this paper are threefold. First, we systematically construct nine new variants of KS test based distinguishers via combining different construction strategies by both KSA and PKS. Second, we consider the impacts of leakage function, noise level and power model to twelve KS test based distinguishers and MIA in a comprehensive comparison framework. Experimental results show that how to choose the most suitable distinguisher in a certain scenario. Third, we also demonstrate the practical meaning of all these KS test based distinguishers in practice.

2 Preliminaries

In this section, we will first introduce KS test, and then briefly recall KSA distinguisher and PKS distinguisher.

2.1 Kolmogorov-Smirnov Test

In statistics, KS test is a nonparametric test whose main target is to determine if two distributions differ significantly. Assume that the random variable X has n samples. Its empirical CDF is $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{A_i \leqslant x}$. $I_{A_i \leqslant x}$ is the indicator function, where its value is 1 when $A_i \leqslant x$; otherwise, it is 0. For a given CDF $F(x)$, formula (1) is used to test their similarity.

$$D_n = sup_x |F_n(x) - F(x)| \tag{1}$$

where sup_x is the supremum of the set of distances. Specifically, the largest distance between two distributions represents the similarity between them. On the other hand, p-value can also be used to measure the similarity of two distributions. The smaller of the p-value, the less similar between them.

2.2 KSA Distinguisher

KSA distinguisher is based on two-sample KS test. Its central idea is to measure the maximum distance between the global trace distribution L and the conditional trace distribution $L|M$, and then average the distances over the prediction space, where M denotes hypothetical power consumption model. Denote l the leakages, and m the hypothetical power consumption values. Denote Pr the probability. KSA is shown in the formula (2).

$$E_{m \epsilon M}(D_{KS}(Pr[L = l|M = m]||Pr[L = l])) \tag{2}$$

KSA can be extended to a normalized version (norm-KSA) that is shown in the formula (3).

$$E_{m \epsilon M}(\frac{1}{|L|M = m|} D_{KS}(Pr[L = l|M = m]||Pr[L = l]))) \tag{3}$$

Both KSA and norm-KSA will produce a large average difference when the key hypothesis is correct.

2.3 PKS Distinguisher

PKS distinguisher is based on single-sample KS test. Its central idea is to measure the p-value produced by comparing normal distribution and part of conditional trace distribution $L|M$. For convenience, leakages L and the hypothetical power consumptions M are usually processed by Z-score transformation in PKS. p is an empirical parameter in PKS from zero to one. N(0,1) represents standard normal distribution. PKS, a two-partial KS test distinguisher, is shown in the formula (6).

$$D_{KS_l} = P_{value}(D_{KS}(Pr[L = l|M \leq p]||N(0,1))) \tag{4}$$

$$D_{KS_r} = P_{value}(D_{KS}(Pr[L = l|M > p]||N(0,1))) \tag{5}$$

$$D_{PKS} = D_{KS_l} \times D_{KS_r} \tag{6}$$

PKS will return the smallest product of p-values when the key hypothesis is correct.

3 Systematic Construction of KS Test Based Side-Channel Distinguishers

From section 2, we learn that both KSA and PKS are based on KS test, and they are able to recover the correct key by partitioning the leakages correctly. However, KSA and PKS are really different from each other in terms of their construction strategies. Therefore, we will show how to construct other new variants of KS test based distinguishers by combining their different construction strategies in a systematic way. For this purpose, we will analyze the construction strategies using by KSA and PKS, and then we will present nine new variants of KS test based distinguishers.

3.1 Construction Strategies of KSA and PKS

In this subsection, we will compare the construction differences between KSA and PKS in four aspects: partition method, similarity measure used by KS test, assumption about leakages, and normalization.

Partition Method. In a partition attack [16], leakages are divided into several sets $p_k^1, p_k^2, ..., p_k^n$ according to each key hypothesis k. These sets are built according to a power model H. In this paper, partition method is classified as non-cumulative partition method and cumulative partition method. Examples of hypothetical leakages that can be used to partition 16-element leakages are shown in Table 1. Specifically, non-cumulative partition used by KSA is shown in the left part of Table 1, while cumulative partition used by PKS is shown in the right part of Table 1.

Similarity Measure Used by KS Test. Distance is used by KSA to measure the similarity of two distributions. In contrast, p-value is adopted in PKS to indicate whether or not partial leakages follow a normal distribution.

Table 1. Examples of non-cumulative partition (left) and cumulative partition (right)

partition	leakages
p_k^1	l_5
p_k^2	$l_2\ l_7\ l_9\ l_{16}$
p_k^3	$l_1\ l_4\ l_8\ l_{10}\ l_{11}\ l_{15}$
p_k^4	$l_3\ l_6\ l_{12}\ l_{13}$
p_k^5	l_{14}

partition	leakages
p_k^1	l_5
p_k^2	$l_5\ l_2\ l_7\ l_9\ l_{16}$
p_k^3	$l_5\ l_2\ l_7\ l_9\ l_{16}\ l_1\ l_4\ l_8\ l_{10}\ l_{11}\ l_{15}$
p_k^4	$l_5\ l_2\ l_7\ l_9\ l_{16}\ l_1\ l_4\ l_8\ l_{10}\ l_{11}\ l_{15}\ l_3\ l_6\ l_{12}\ l_{13}$
p_k^5	$l_5\ l_2\ l_7\ l_9\ l_{16}\ l_1\ l_4\ l_8\ l_{10}\ l_{11}\ l_{15}\ l_3\ l_6\ l_{12}\ l_{13}\ l_{14}$

Assumption about Leakages. PKS distinguisher considers that leakages follow a normal distribution, while KSA makes no assumption about leakages.

Normalization. [10] suggested that normalization could improve the performance of KSA. Our question is whether or not the normalization is always effective in some typical scenarios for KSA. We will also try to answer this question in this work.

3.2 Nine New Variants of KS Test Based Distinguishers

In subsection 3.1, we analyzed the construction strategies of both KSA and PKS. We find that KSA and PKS have different choices for a specific construction strategy. One natural yet pertinent question is that is it possible to construct other (more efficient) KS test based distinguisher by combining the construction methods of both KSA and PKS? To answer this question, we combine construction strategies using by both KSA and PKS to construct nine new variants of KS test based distinguishers, in a systematic way.

For convenience, we will label each strategy that was used by KSA and PKS. Denote A0 the non-cumulative partition, and A1 the cumulative partition. Denote B0 the expectation of distance as the similarity measure of KS test, and B1 the product of p-values as the similarity measure of KS test. Denote C0 the distinguisher that makes no assumption about leakage distribution, and C1 the distinguisher that assumes the leakage follows a normal distribution. Denote D0 that we perform normalization on a distinguisher, and D1 that we do not.

By combining these strategies systematically, one can, in total, construct sixteen ($16 = 2^4$) KS test based distinguishers. Among these sixteen distinguishers, three are existing and they are KSA (A0,B0,C0,D1), PKS (A1,B1,C1,D1) and norm-KSA (A0,B0,C0,D0). On the other hand, note that B1 and D0 conflict with each other, therefore four combinations (A1,C1,B1,D0; A1,C0,B1,D0; A0,C1,B1,D0; A0,C0,B1,D0) do not make any sense. Additionally, three combinations , which are (A0, B0, C1, D1), (A0, B0, C1, D0) and (A0,B1,C1,D1), fail to work in the key recovery attacks. We free the limitation of Z-score on hypothetical power consumptions of D-PKS (A1, B0, C1, D1), norm-D-PKS (A1, B0, C1, D0) and PKS (A1,B1,C1,D1) to form C-PKS (A1, B0, C1, D1), norm-C-PKS (A1, B0, C1, D0) and MPC-PKS (A1, B1, C1, D1). Finally, the remaining combinations are MP-KSA (A0, B1, C0, D1), C-KSA (A1, B0, C0, D1), norm-C-KSA (A1, B0, C0, D0) and MPC-KSA (A1, B1, C0, D1). Therefore, we only

Table 2. Nine new variants of KS test based distinguishers

Distinguisher	Equation
MP-KSA	$log_2(\prod_{m \epsilon M} P_{value}(D_{KS}(Pr[L = l \vert M = m] \vert\vert Pr[L = l])))$
C-KSA	$E_{m \epsilon M}(D_{KS}(Pr[L = l \vert M \leqslant m] \vert\vert Pr[L = l]))$
norm-C-KSA	$E_{m \epsilon M}(\frac{1}{\vert L \vert M = m \vert} D_{KS}(Pr[L = l \vert M \leqslant m] \vert\vert Pr[L = l]))$
MPC-KSA	$log_2(\prod_{m \epsilon M} P_{value}(D_{KS}(Pr[L = l \vert M \leqslant m] \vert\vert Pr[L = l])))$
D-PKS	$E(D_{KS}(Pr[L = l \vert M \leqslant p] \vert\vert N(0, 1)))$
norm-D-PKS	$E(\frac{1}{\vert L \vert M \leqslant p \vert} D_{KS}(Pr[L = l \vert M \leqslant p] \vert\vert N(0, 1)))$
C-PKS	$E_{m \epsilon M}(D_{KS}(Pr[L = l \vert M \leqslant m] \vert\vert N(0, 1)))$
norm-C-PKS	$E_{m \epsilon M}(\frac{1}{\vert L \vert M \leqslant m \vert} D_{KS}(Pr[L = l \vert M \leqslant m] \vert\vert N(0, 1)))$
MPC-PKS	$log_2(\prod_{m \epsilon M} P_{value}(\frac{1}{\vert L \vert M \leqslant m \vert} D_{KS}(Pr[L = l \vert M \leqslant m] \vert\vert N(0, 1))))$

construct nine ($9{=}2^4-4-3-3+3$) new variants. These nine new distinguishers are summarized in Table 2.

In these nine new variants of KS test based distinguishers, the distinguisher which contains B0 strategy will return the largest expected distance under the correct key hypothesis, while the distinguisher which contains B1 strategy will return the smallest product of p-values. Additionally, in order to avoid arithmetic underflow, one typically applies the logarithm to the distinguisher which contains B1 strategy.

4 A Comprehensive Evaluation of All Twelve KS Test Based Side-Channel Distinguishers

So far, we have constructed nine new variants of KS test based distinguishers. The performance of these distinguishers in a univariate setting has a huge impact on how to choose the most suitable distinguisher in a certain scenario. Consequently, we will evaluate the performance of all these distinguishers by amounting key recovery attacks, and analyze their effectiveness and efficiency by using Success Rate (SR) [15] in typical scenarios. On the one hand, we will evaluate the performance of these KS test based distinguishers in a unified comparison framework inspired by [12]. In this framework, we will evaluate the influence of different factors, such as leakage function, noise level and power model, on the

performance of each KS test based distinguisher. We will compare the attacking efficiency of these distinguishers in terms of SR. On the other hand, we will perform a series of attacks against the real traces from both OpenSCA and DPA Contest v2, respectively. With these practical attacks, we will demonstrate the practical meaning of all these KS test based distinguishers. Note that we do not compare the running cost for different distinguishers.

4.1 Simulated Experiments

In simulated scenarios, we choose the output of the first S-box of the first round AES operation as the target intermediate value. Three typical leakage functions, i.e. Hamming Weight (HW) leakage function, an Unevenly Weighted Sum of the Bits (UWSB) leakage function and highly nonlinear leakage function, are adopted to test the adaptability of KS test based distinguishers and MIA. Noise level in simulated leakages is measured by Signal-to-noise ratio (SNR). We particularly employ seven SNRs, i.e. 0.125, 1, 8, 16, 32, 64 and positive infinity, to test the influence of Gaussian noise on these distinguishers.

In each scenario, we perform key recovery attacks using all twelve KS test based distinguishers and MIA (MIA(HW,bins=9) and MIA(ID,bins=256)) as well. For each one of these sixteen kinds of attacks, we repeat it 300 times by uniformly choosing different plaintexts.

Our experiments are also carefully organized in order to make them understood more easily. Specifically, we divide the results of all these thirteen distinguishers into three groups, and denote these groups by A, B and C, respectively. Group A consists of four existing distinguishers and they are PKS, KSA, normKSA and MIA. For each scenario, we select the most efficient one from Group A, and the selected one is set to be a benchmark for this scenario. Next, the other new nine KS test based distinguishers are classified into two groups, according to their relative efficiency over the selected benchmark. Namely, for each scenario, those distinguishers that are more efficient than the benchmark are set into Group B, while the others that are less efficient than the benchmark are put into Group C.

Hamming Weight Leakage

In the following scenarios, we assume that the leakage of a cryptographic device consists of HW of target intermediate value and Gaussian noise. Under this reasonable assumption, we will investigate the performance of different distinguishers with two adversarial characterization abilities.

- An Adversary with a Perfect Power Model. Figure 1 shows the SR of twelve KS test based distinguishers and MIA using a HW model against HW leakage of the first AES S-box. When the SNR is 0.125, PKS(HW,p=0.618) in Figure 1(a) is used as the benchmark for Figure 1(d) and 1(g). Figure 1(d) shows that, C-KSA(HW), MPC-PKS(HW) and C-PKS(HW) are better than the benchmark, and C-KSA(HW) is the best distinguisher. Distinguishers in Figure 1(g) are less efficient than the benchmark, so we do not explain them in more

details. Due to fact that other SNRs can be analyzed in a similar way as that of SNR of 0.125, we do not explain them in more details.

In summary, C-KSA(HW) is the best choice in all twelve KS test based distinguishers when the SNRs are 0.125 and 1 respectively, while MP-KSA(HW) is the best choice when the SNR is 8 and positive infinity. Additionally, MPC-PKS(HW) is better than the benchmark when the SNRs are 0.125, 1 and 8 respectively. Another interesting observation is that norm-KSA(HW) is less efficient than KSA(HW).

- An Adversary with a Generic Power Model. Figure 2 shows the SR of twelve KS test based distinguishers and MIA using an Identity (ID) model against HW leakage of the first AES S-box. When the SNR is 0.125, PKS(ID,p=0.618) in Figure 2(a) is chosen as the benchmark in Figure 2(d) and 2(g). Figure 2(d) shows that C-KSA(ID), norm-C-KSA(ID), MPC-KSA(ID), C-PKS(ID), norm-C-PKS(ID) and MPC-PKS(ID) are more efficient than the benchmark, and they have similar performance. Distinguishers in Figure 2(g) are less efficient than the benchmark, so we do not explain them in more details. When the SNRs are 1, 8, and positive infinity, they can be analyzed in a similar way as that of SNR of 0.125.

In a word, although C-KSA(ID), norm-C-KSA(ID), MPC-KSA(ID), C-PKS(ID), norm-C-PKS(ID) and MPC-PKS(ID) are more efficient than the benchmark and they have similar performance under four noise levels, C-KSA(ID), norm-C-KSA(ID) and MPC-KSA(ID) are slightly more efficient than C-PKS(ID), norm-C-PKS(ID) and MPC-PKS(ID). Another interesting point is that KSA(ID), norm-KSA(ID) and MIA(ID) all fail to reveal the correct key, while both PKS(ID,p=0.25) and PKS(ID,p=0.618) succeeds to do that.

An Unevenly Weighted Sum of the Bits Leakage Scenario

In the following scenarios, we assume that the least significant bit dominates in the leakage function with a relative weight of 10 and other bits with a relative weight of 1, and we will investigate the performance of twelve KS test based distinguishers and MIA with two adversarial characterization abilities.

- An Adversary with an Imprecise Power Model. Figure 3 shows the SR of twelve KS test based distinguishers and MIA using a HW model against UWSB leakage of the first AES S-box. When the SNR is 0.125, PKS(HW,p=0.618) in Figure 3(a) will be chosen as the benchmark for Figure 3(d) and Figure 3(g). Figure 3(d) shows that C-KSA(HW) exhibits consistently better performance compared with the benchmark. Distinguishers in Figure 3(g) are less efficient than the benchmark, so we do not explain them in more details. When the SNRs are 1, 8, and positive infinity, they can be analyzed in a similar way as that of SNR of 0.125.

In summary, C-KSA(HW) is the best choice of all twelve KS test based distinguishers when the SNRs are 0.125, 1 and 8 respectively, while MIA(HW) is the best choice when the SNR goes into positive infinity. Additionally, MPC-KSA(HW) is no worse than the benchmark, and KSA(HW) is more efficient than norm-KSA(HW).

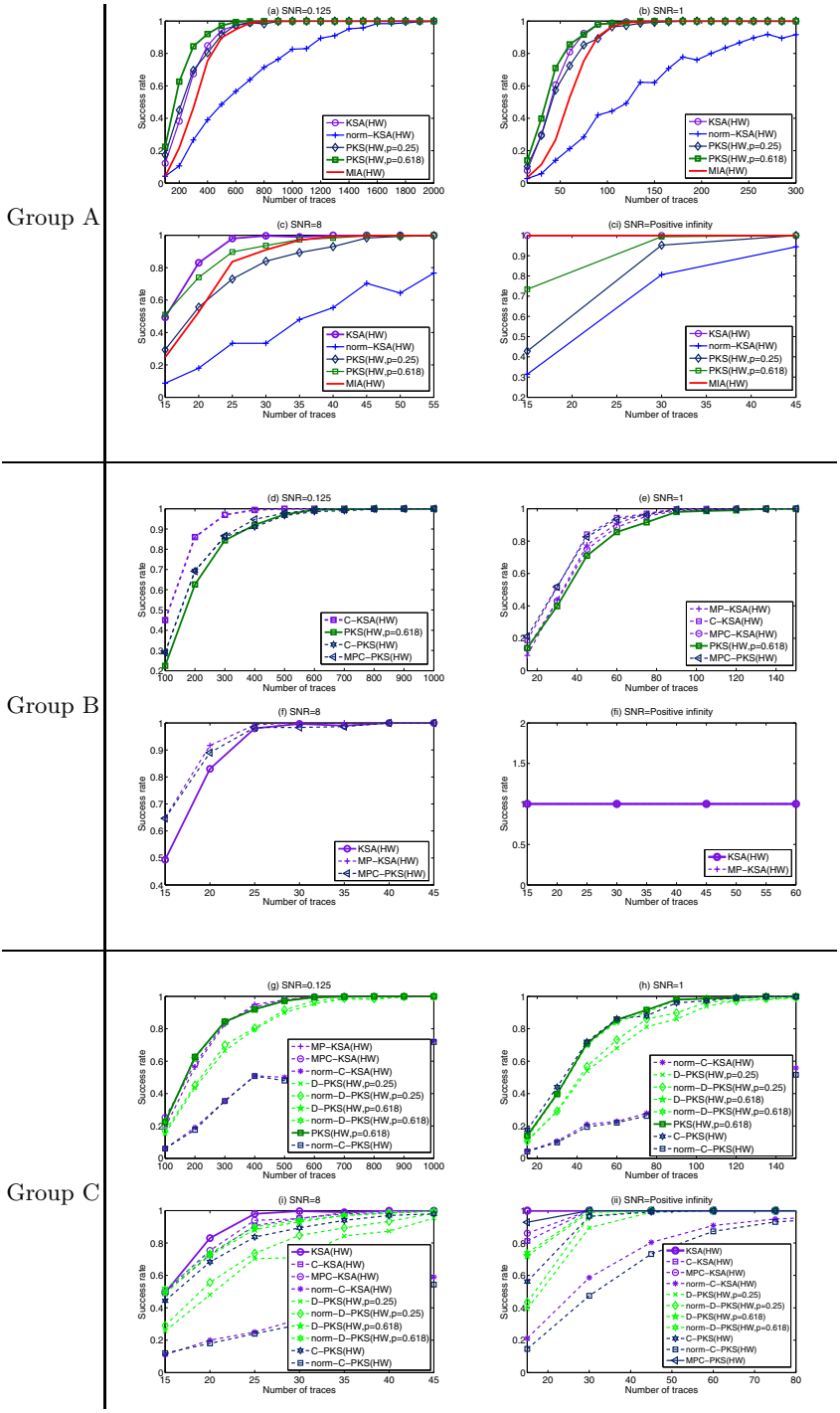

Fig. 1. SRs of different distinguishers against the first AES S-box in HW leakage

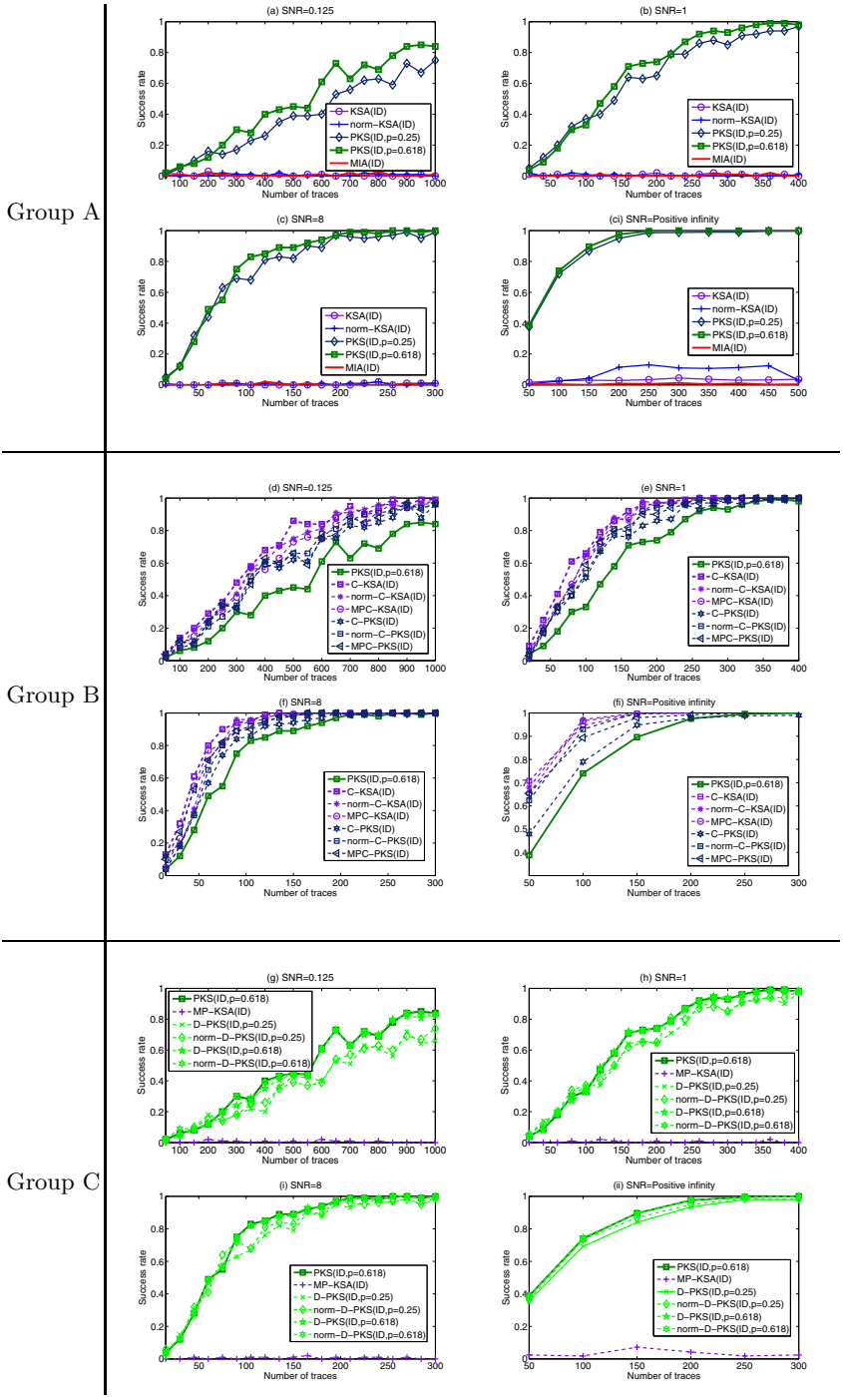

Fig. 2. SRs of different distinguishers against the first AES S-box in HW leakage

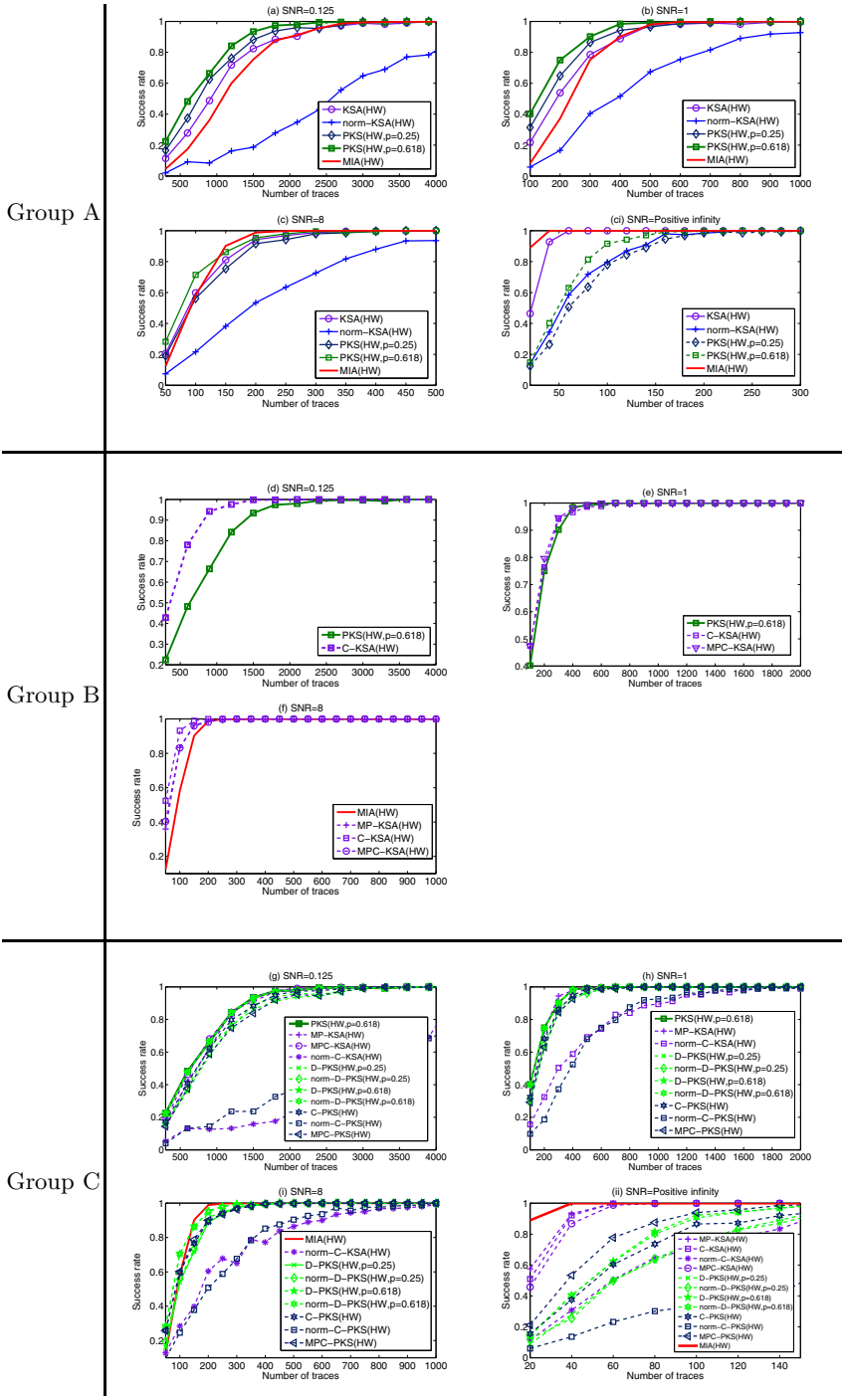

Fig. 3. SRs of different distinguishers against the first AES S-box in UWSB leakage

- An Adversary with a Generic Power Model. Due to the computation cost, we select the SNRs of 16, 32, 64 and positive infinity in this scenario. Figure 4 shows the SR of twelve KS test based distinguishers and MIA using an ID model against UWSB leakage of the first AES S-box. In Group A, Figure 4 shows that, KSA(ID), norm-KSA(ID), MIA(ID) and PKS(ID) all fail to recover the correct key with a relative small number of traces. Therefore, the benchmark for Group B and Group C is whether or not a distinguisher can recovery the correct key with a relative small number of traces. The distinguishers in Group B can recover the correct key with a trace number of 4,000, while the distinguishers in Group C fail to do that. For example, when the SNR is 16, C-KSA(ID), norm-C-KSA(ID) and MPC-KSA(ID) in Group B can recovery the correct key (see Figure 4(d)), while other new variants of KS test based distinguishers fail to do that with 4,000 power traces (see Figure 4(g)). When the SNRs are 32, 64, and positive infinity, they can be analyzed in a similar way as that of SNR of 16.

To sum up, C-KSA(ID), norm-C-KSA(ID) and MPC-KSA(ID) are more efficient than the benchmark, and C-KSA(ID) is the best choice when the SNRs are 16, 32, 64 and positive infinity.

Highly Nonlinear Leakage Scenario. In this scenario, the leakage function of cryptographic device is a highly nonlinear function. Without loss of generality, S-box is used in this leakage scenario [14]. Our experimental results show that twelve KS test based distinguishers and MIA all fail to recover the correct key in this scenario.

Note: When SNR goes into positive infinity, the performance of PKS with a fixed parameter may decrease with the increase of the trace number. This indicates that the parameter in PKS is critical to the performance of PKS, as is shown in [13].

4.2 Practical Experiments

In order to show how these twelve KS test based distinguishers behave in practical scenarios, we perform attacks against unprotected software AES implementation on 8-bit microcontroller (Case 1) and unprotected hardware AES implementation on Xilinx Vertex-5 FPGA (Case 2), respectively. These power traces are from OpenSCA and from DPA Contest v2, respectively.

In the view of an adversary, we will choose the power model according to our priori knowledge. Specifically, we will use HW model in Case 1, and Hamming distance (HD) model in Case 2. We will choose SR to evaluate the efficiency, by amounting key recovery attacks 300 times. In this part, the experiments are also organized exactly in the same way as that in our simulated experiments, except that we also perform CPA attacks. This means that we place CPA distinguisher in Group D. That is to say, in practical experiments, we will show the performance of traditional CPA distinguisher, which is widely believed to be well capable of characterizing linear leakages.

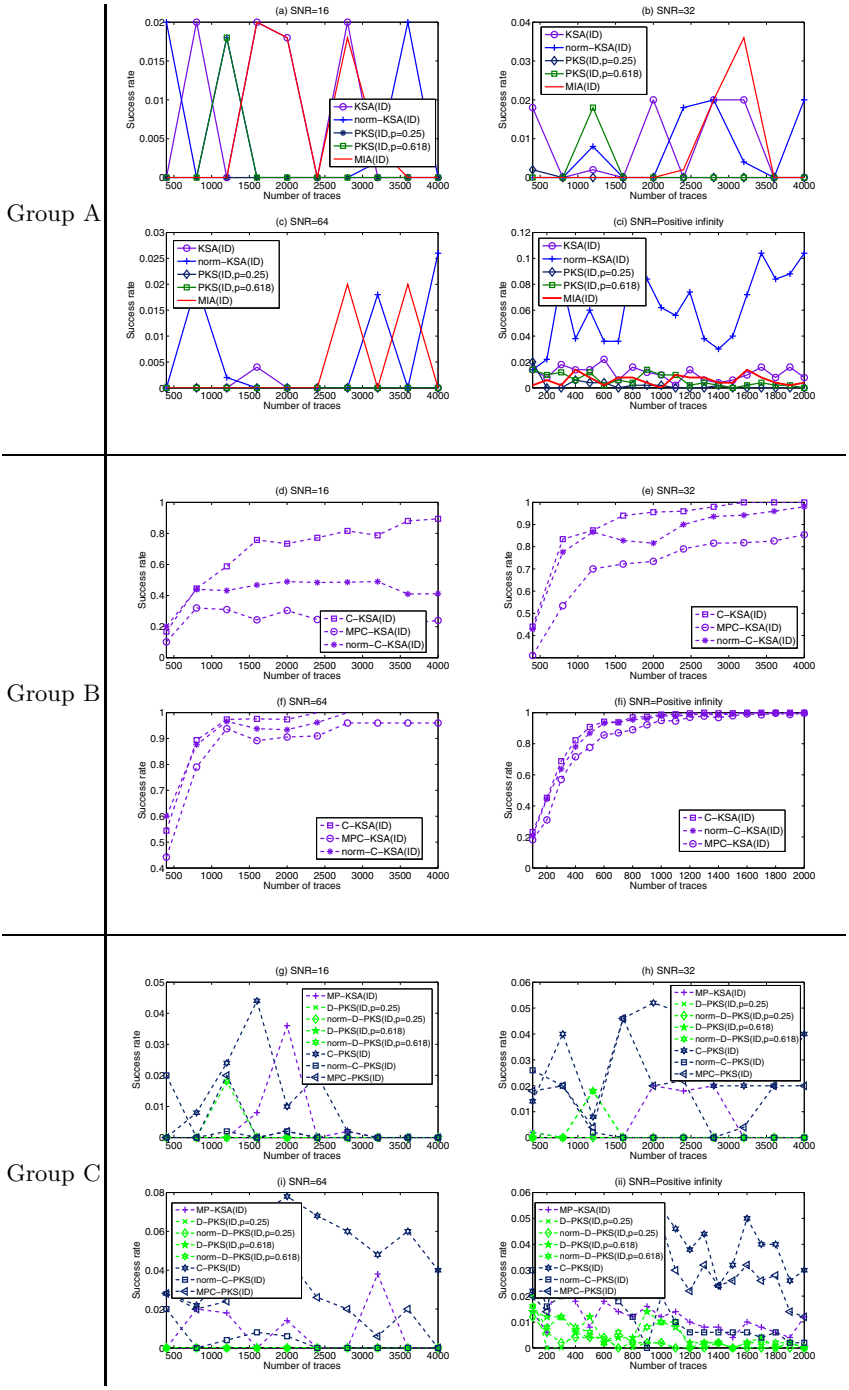

Fig. 4. SRs of different distinguishers against the first AES S-box in USWB leakage

Case 1: Unprotected Software AES Implementation Provided by OpenSCA

In this scenario, the output of the first S-box of the first round of AES operation is chosen as the target. In Group A, Figure 5 (a) shows that, KSA(HW) exhibits the best performance among three existing KS test based distinguishers, so KSA(HW) is used as the benchmark for Group B (see Figure 5(b)) and Group C (see Figure 5(c)). In Group B, Figure 5(b) shows that, MP-KSA(HW) is more efficient than the benchmark. In Group C, Figure 5(c) shows that, other new variants of KS test based distinguishers are less efficient than the benchmark. In Group D, Figure 5(d) shows that, MP-KSA(HW) is less efficient than CPA(HW). In summary, MP-KSA(HW) is the best choice in all these KS test

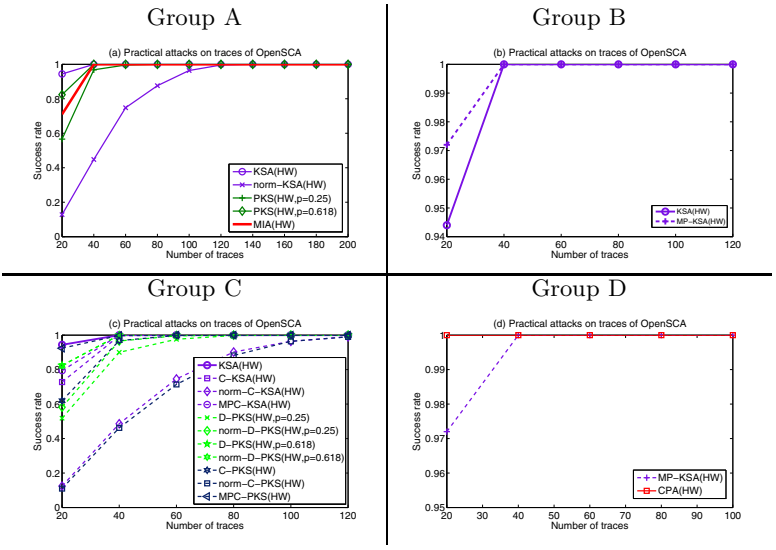

Fig. 5. SRs for twelve KS test based distinguishers, MIA and CPA with HW model in attacks against the first AES S-box

based distinguishers in this case. In the view of an adversary, CPA is an ideal distinguisher. This indicates that, when the leakage of a cryptographic device could be accurately characterized, CPA is the best choice compared with all KS test based distinguishers.

Case 2: Unprotected Hardware AES Implementation Provided by DPA Contest v2

In this scenario, the input of the first S-box of the last round of AES operation is chosen as the target. In Group A, Figure 6(a) shows that, both PKS(HD) and MIA(HD) (MIA(HD,bins=9)) can reveal the correct key, while KSA(HD) and norm-KSA(HD) fail to do that. The empirical parameter in PKS(HD) can largely improve the performance of PKS(HD). Therefore, PKS(HD,p=0.01) is

selected as the benchmark for finding the most promising variants in this case. In Group B, Figure 6(b) shows that, MPC-KSA(HD) and MPC-PKS(HD) outperform PKS(HD,p=0.01) in terms of achieving a partial success rate of 80%. In Group C, other KS test based distinguishers are less efficient than the benchmark, so we do not discuss them in more details. In Group D, Figure 6(d) shows that, MPC-KSA(HD) is even better than CPA(HD).

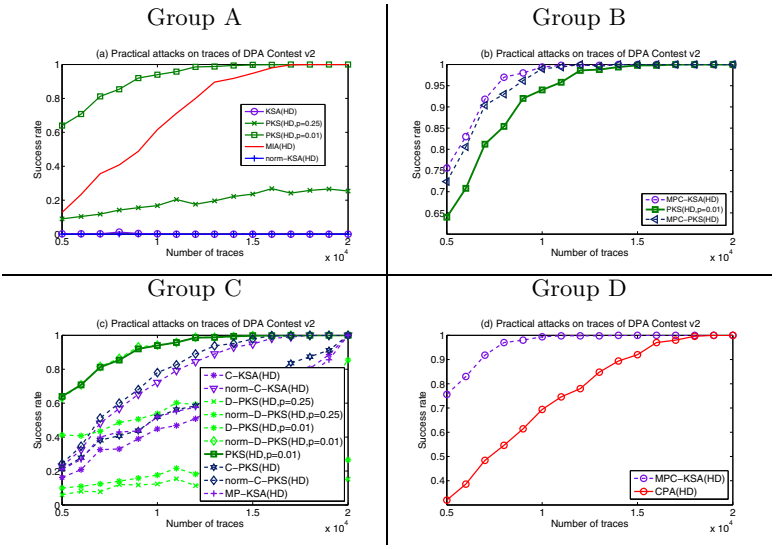

Fig. 6. SRs for twelve KS test based distinguishers, MIA and CPA with HD model in attacks against the first AES S-box of the last round

Table 3. Number of traces required to achieve partial SR of 80% on individual byte

	byte 1	byte 2	byte 3	byte 4	byte 5	byte 6	byte 7	byte 8
MPC-KSA	5,300	6,100	5,700	9,800	9,600	5,500	4,800	6,800
CPA	12,500	10,000	6,900	7,000	12,700	6,000	5,900	7,400
	byte 9	byte 10	byte 11	byte 12	byte 13	byte 14	byte 15	byte 16
MPC-KSA	4,500	5,200	9,200	3,500	4,100	14,500	6,000	5,500
CPA	6,800	3,600	10,000	3,000	6,600	16,900	15,000	5,100

In order to enhance the understanding of whether or not MPC-KSA is a reasonable alternative for CPA, we perform attacks on all sixteen bytes of AES encryption. Table 3 shows the number of traces required to achieve partial SR of 80% of attacks on individual bytes. Although CPA is more efficient than MPC-KSA on four bytes (byte 4, byte 10, byte 12, byte 16), it is less efficient on other twelve bytes. For example, for byte 15, the number of required traces for MPC-KSA to achieve partial SR of 80% is 6,000, while that of CPA is 15,000. However, for byte 4, the number of required traces for MPC-KSA to achieve

partial SR of 80% is 9,800, while that of CPA is 7,000. Although MPC-KSA does not perform consistently better than CPA, it performs better than CPA on 75% of sixteen bytes. As a whole, MPC-KSA is more efficient than CPA in terms of the required number of traces to achieve the global SR of 80%. In summary, MPC-KSA is the best choice in this case. This experimental result indicates that, when the leakages of a cryptographic device could not been accurately characterized, MPC-KSA exhibits better performance than CPA in terms of SR, as the former is capable of measuring the total dependency between hypothetical power consumptions and physical leakages.

5 Conclusions

Distinguishers play an vital role in exploiting physical leakages in side-channel attacks. Due to the capability of dealing with both linear and nonlinear dependencies, generic side-channel distinguishers are being increasingly popular. Among those are KS test based distinguishers, such as KSA and PKS. In this paper, we constructed nine new variants of KS test based distinguishers via combining different construction strategies of KSA and PKS, and then explored the effectiveness and efficiency of twelve KS test based distinguishers and MIA in typical simulated scenarios and practical scenarios.

In a whole, we experimentally investigated the performance of KS test based distinguishers, and provided some helpful guides on how to choose a suitable distinguisher. One of the most interesting observations is that MPC-KSA is more efficient than CPA against unprotected hardware AES implementation on Xilinx Vertex-5 FPGA in DPA Contest v2. However, we did not provide any theoretical analysis yet about why this happens, which could be part of our future work.

Acknowledgments. This work was supported in part by National Natural Science Foundation of China (No. 61272478, 61073178, 60970135 and 61170282), Beijing Natural Science Foundation (No. 4112064), Strategic Priority Research Program of the Chinese Academy of Sciences (No.XDA06010701), and IIE Cryptography Research Project (No. Y2Z0011102).

References

1. Kocher, P., Jaffe, J., Jun, B.: Differential Power Analysis. In: Wiener, M. (ed.) CRYPTO 1999. LNCS, vol. 1666, pp. 388–397. Springer, Heidelberg (1999)
2. Messerges, T.S., Dabbish, E.A., Sloan, R.H.: Examining Smart-Card Security under the Threat of Power Analysis Attacks. IEEE Trans. Comput. 51(5), 541–552 (2002)
3. Brier, E., Clavier, C., Olivier, F.: Correlation Power Analysis with a Leakage Model. In: Joye, M., Quisquater, J.-J. (eds.) CHES 2004. LNCS, vol. 3156, pp. 16–29. Springer, Heidelberg (2004)

4. Le, T.-H., Clédière, J., Canovas, C., Robisson, B., Servière, C., Lacoume, J.-L.: A Proposition for Correlation Power Analysis Enhancement. In: Goubin, L., Matsui, M. (eds.) CHES 2006. LNCS, vol. 4249, pp. 174–186. Springer, Heidelberg (2006)
5. Mangard, S., Oswald, E., Standaert, F.-X.: All for one-one for all: Unifying univariate DPA attacks. IET Information Security 5(2), 100–110 (2011)
6. Doget, J., Prouff, E., Rivain, M., Standaert, F.-X.: Univariate side channel attacks and leakage modeling. Journal of Cryptographic Engineering 1(2), 123–144 (2011)
7. Gierlichs, B., Batina, L., Tuyls, P., Preneel, B.: Mutual Information Analysis: A Generic Side-Channel Distinguisher. In: Oswald, E., Rohatgi, P. (eds.) CHES 2008. LNCS, vol. 5154, pp. 426–442. Springer, Heidelberg (2008)
8. Moon, Y.-I., Rajagopalan, B., Lall, U.: Estimation of Mutual Information using Kernal Density Estimators. Physical Review E 52, 2318–2321 (1995)
9. Walters-Williams, J., Li, Y.: Estimation of mutual information: A survey. In: Wen, P., Li, Y., Polkowski, L., Yao, Y., Tsumoto, S., Wang, G. (eds.) RSKT 2009. LNCS, vol. 5589, pp. 389–396. Springer, Heidelberg (2009)
10. Veyrat-Charvillon, N., Standaert, F.-X.: Mutual Information Analysis: How, When and Why? In: Clavier, C., Gaj, K. (eds.) CHES 2009. LNCS, vol. 5747, pp. 429–443. Springer, Heidelberg (2009)
11. Batina, L., Gierlichs, B., Prouff, E., Rivain, M., Standaert, F.-X., Veyrat-Charvillon, N.: Mutual Information Analysis: A Comprehensive Study. Journal of Cryptology 24(2), 269–291 (2011)
12. Whitnall, C., Oswald, E., Mather, L.: An Exploration of the Kolmogorov-Smirnov Test as Competitor to Mutual Information Analysis. In: Prouff, E. (ed.) CARDIS 2011. LNCS, vol. 7079, pp. 234–251. Springer, Heidelberg (2011)
13. Liu, J.-Y., Zhou, Y.-B., Yang, S.-G., Feng, D.-G.: Generic Side-Channel Distinguisher Based on Kolmogorov-Smirnov Test: Explicit Contruction and Practical Evaluation. Chinese Journal of Electronics 21(3), 547–553 (2012)
14. Whitnall, C., Oswald, E.: A Fair Evaluation Framework for Comparing Side-Channel Distinguishers. Journal of Cryptographic Engineering 1(2), 145–160 (2011)
15. Standaert, F.-X., Malkin, T.G., Yung, M.: A unified framework for the analysis of side-channel key recovery attacks. In: Joux, A. (ed.) EUROCRYPT 2009. LNCS, vol. 5479, pp. 443–461. Springer, Heidelberg (2009)
16. Standaert, F.-X., Gierlichs, B., Verbauwhede, I.: Partition *vs.* Comparison Side-Channel Distinguishers: An Empirical Evaluation of Statistical Tests for Univariate Side-Channel Attacks against Two Unprotected CMOS Devices. In: Lee, P.J., Cheon, J.H. (eds.) ICISC 2008. LNCS, vol. 5461, pp. 253–267. Springer, Heidelberg (2009)