Chapter 8

ATTACKS ON HIDING

The goal of hiding countermeasures is to make the power consumption of cryptographic devices independent of the performed operations and the processed values. However, in practice this goal can only be achieved to a certain degree, see Chapter 7. Attacks on protected devices are therefore still possible. In most cases though, these attacks require significantly more effort than attacks on unprotected devices.

In this chapter, we first discuss the effectiveness of hiding countermeasures in general. In particular, we analyze how the different types of hiding countermeasures increase the number of power traces that are needed for DPA attacks. Subsequently, we look at two specific countermeasures in detail. We discuss DPA attacks on countermeasures that destroy the alignment of power traces, and we look at the effectiveness of the DRP logic styles presented in Section 7.4.

8.1 General Description

In Chapter 7, we have introduced two types of hiding countermeasures. On the one hand, there are countermeasures that randomize the execution of cryptographic algorithms. On the other hand, there are countermeasures that reduce the SNR of the executed operations.

We now analyze the effectiveness of these two types of countermeasures against DPA attacks by determining their effect on $\rho_{ck,ct}$. Recall that this is the correlation between the hypothetical power consumption for the correct key hypothesis H_{ck} and the power consumption at the moment of time ct. This is the moment of time when the device processes the attacked intermediate result. As pointed out in Section 6.4, the correlation $\rho_{ck,ct}$ determines the number of power traces that are needed to perform DPA attacks. For our analysis, we model the power consumption of the cryptographic device as in (4.8). This

means that we use P_{total} to denote the power consumption at the moment ct. The correlation $\rho_{ck,ct}$ therefore corresponds to $\rho(H_{ck}, P_{total})$.

8.1.1 Time Dimension

The random insertion of dummy operations and shuffling change the execution of the cryptographic algorithm randomly. The attacked intermediate result is therefore processed at a different moment of time in each power trace, *i.e.* ct is randomly distributed. The statistical distribution of ct depends on the way the random insertion of dummy operations and the shuffling are implemented. If only shuffling is used, ct is typically uniformly distributed. The random insertion of dummy operations often leads to binomial or uniform distributions of ct. If the random insertion of dummy operations and shuffling are combined, the resulting distribution is the superposition of the corresponding distributions.

Independent of the shape of the distribution of ct, we denote the maximum of this distribution by $\hat{p}$. Furthermore, we denote the power consumption that is located at this position by $\hat{P}_{total}$. Hence, $\hat{P}_{total}$ has the following properties. With probability $\hat{p}$, $\hat{P}_{total}$ corresponds to the power consumption of the attacked intermediate result, *i.e.* with probability $\hat{p}$ it holds that $\hat{P}_{total} = P_{total}$. With probability $(1-\hat{p})$, $\hat{P}_{total}$ corresponds to the power consumption of some other operations. We refer to the power consumption of these operations as P_{other}. The covariance $Cov(H_{ck}, \hat{P}_{total})$ can therefore be calculated as follows:

$$Cov(H_{ck}, \hat{P}_{total}) = \hat{p} \cdot Cov(H_{ck}, P_{total}) + (1-\hat{p}) \cdot Cov(H_{ck}, P_{other})$$

Since $\hat{p}$ is the maximum probability, the correlation between H_{ck} and $\hat{P}_{total}$ leads to the highest correlation coefficient that occurs in a DPA attack on a protected device. The correlation $\rho(H_{ck}, \hat{P}_{total})$ hence determines the number of power traces that are needed for the attack. This correlation can be calculated based on $\rho(H_{ck}, P_{total})$ as shown in (8.1). The simplifications in this equation are possible because we assume that P_{other} and P_{total} are independent.

$$\begin{aligned}
\rho(H_{ck}, \hat{P}_{total}) &= \frac{\hat{p} \cdot Cov(H_{ck}, P_{total}) + (1-\hat{p}) \cdot Cov(H_{ck}, P_{other})}{\sqrt{Var(H_{ck}) \cdot Var(\hat{P}_{total})}} \\
&= \frac{\hat{p} \cdot Cov(H_{ck}, P_{total})}{\sqrt{Var(H_{ck}) \cdot Var(\hat{P}_{total})}} \\
&= \rho(H_{ck}, P_{total}) \cdot \hat{p} \cdot \sqrt{\frac{Var(P_{total})}{Var(\hat{P}_{total})}} \qquad (8.1)
\end{aligned}$$

The effect of shuffling and the random insertion of dummy operations mainly depends on $\hat{p}$. The probability $\hat{p}$ linearly reduces the correlation $\rho(H_{ck}, \hat{P}_{total})$. Halving $\hat{p}$, halves this correlation and hence quadruples the number of needed