Analysis of Communication Channels Related to Physical Unclonable Functions

Georg Maringer[1[0000−0002−2868−5131]], Marvin Xhemrishi[1], Sven Puchinger[1], Kathrin Garb[2], Hedongliang Liu[1[0000−0001−7512−0654]], Thomas Jerkovits[1,3], Ludwig Kürzinger, Matthias Hiller[2[0000−0003−1238−1114]], and Antonia Wachter-Zeh[1[0000−0002−5174−1947]] ⋆

[1] Technical University of Munich, Germany
{georg.maringer, marvin.xhemrishi, sven.puchinger, lia.liu, antonia.wachter-zeh}@tum.de, ludwig.kuerzinger@mytum.de

[2] Fraunhofer Institute for Applied and Integrated Security AISEC, Garching (near Munich), Germany
{matthias.hiller, kathrin.garb}@aisec.fraunhofer.de

[3] Deutsches Luft- und Raumfahrtzentrum, Germany
{thomas.jerkovits@dlr.de}

Abstract. Cryptographic algorithms rely on the secrecy of their corresponding keys. On embedded systems with standard CMOS chips, where secure permanent memory such as flash is not available as a key storage, the secret key can be derived from Physical Unclonable Functions (PUFs) that make use of minuscule manufacturing variations of, for instance, SRAM cells. Since PUFs are affected by environmental changes, the reliable reproduction of the PUF key requires error correction. For silicon PUFs with binary output, errors occur in the form of bitflips within the PUF response. Modeling the channel as a Binary Symmetric Channel (BSC) with fixed crossover probability p is only a first-order approximation of the real behavior of the PUF response. We propose a more realistic channel model, referred to as the Varying Binary Symmetric Channel (VBSC), which takes into account that the reliability of different PUF response bits may not be equal. We investigate its channel capacity for various scenarios which differ in the channel state information (CSI) present at encoder and decoder. We compare the capacity results for the VBSC for the different CSI cases with reference to the distribution of the bitflip probability according to a work by Maes et al.

Keywords: Physical Unclonable Functions · Channel Model · Channel Capacity · Channels with Random State

⋆ G. Maringer's work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) under Grant No. WA3907/4-1. The work of H. Liu has been supported by a German Israeli Project Cooperation (DIP) grant under grant no. PE2398/1-1 and KR3517/9-1. The work of S. Puchinger and A. Wachter-Zeh has been supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 801434) The work of K. Garb and M. Hiller has been supported by the DFG Priority Program SPP 2253 Nano Security (Project STAMPS).

1 Introduction

Secure keys are the foundation for secure cryptographic operations. As a wide range of integrated circuits does not have access to secure key storage, Physical Unclonable Functions (PUFs) evaluate physical properties of a circuit to derive a unique fingerprint and thus generate a device-specific cryptographic key. A PUF can be seen as the fingerprint of a physical object [7, 9], emanated from minuscule manufacturing variations that vary from object to object. This physical fingerprint can, hence, contribute to a hardware root of trust for security applications, such as authentication or identification of a device, and also the generation of a key for cryptographic primitives.

SRAM is widely available on embedded devices and a popular PUF primitive [6]. Upon power-up, the state of SRAM cells is not deterministically defined, however they will mostly start up in the same state. Due to manufacturing variations, the state of a certain SRAM cell will either be “0” or “1”. Multiple SRAM bits can therefore be grouped into a (randomly distributed) PUF-response. As characterized in [11] and later refined in [10], SRAM PUF cells vary significantly in their reliability. This reliability can be estimated with multiple measurements. In contrast, other PUFs such as the Ring-Oscillator PUF directly output analog or finely quantized discrete outputs, such that reliability information is directly available for each measurement. This also allows to estimate the reliability of a specific measurement as channel state information during decoding.

From these random SRAM bits, a key can be generated, which is exemplarily shown in Fig. 1 for the code-offset fuzzy extractor [4]. The PUF key is generated (once) during the manufacturing process of the device, which is referred to as *enrollment*. After that, the PUF response is repeatedly read out whenever needed, such that the key can be reproduced. However, since the PUF is subject to noise and also environmental effects such as temperature changes, certain SRAM bits can flip and impact the reliability of the key. Hence, an error correction algorithm was included into the key generation scheme.

Fig. 1 shows that the key generation phase and the reproduction phase contain encoding and decoding steps. During the reproduction phase, the PUF will be read out again and the PUF response will deviate from the original one in the enrollment phase due to changes in environmental conditions. This corresponds to transmitting the PUF response over a noisy channel, which in case of SRAM cells can be described by a Binary Symmetric Channel (BSC).

The reliability of the power-up state of an SRAM cell heavily depends on the mismatch between its transistors so that a BSC model with constant error probability discards parts of the information. Experimental work compared in [8] has shown that fuzzy extractors benefit from reliability information, but still lack work on fundamental limits.

In this paper, we extend the BSC model as used, e.g. in [2], by taking into account that the crossover probabilities may vary from cell to cell. This leads to the channel model discussed in Section 3 which we then analyze using information-theoretic methods in Section 4. We distinguish whether CSI is present at encoder and/or decoder. Furthermore within this section we show a capacity-achieving

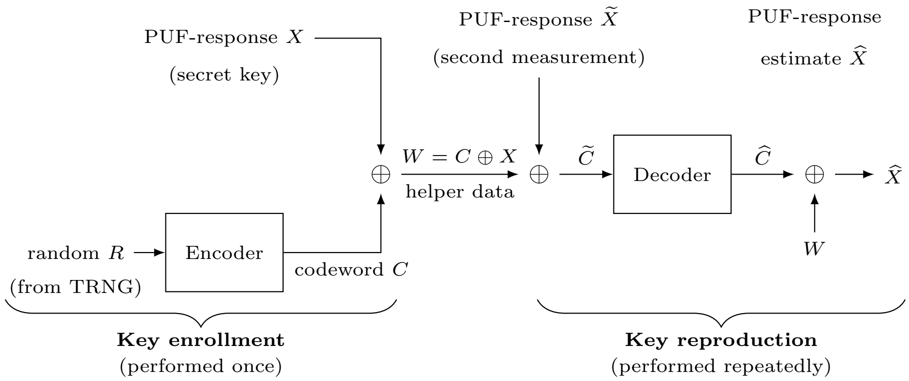

Fig. 1. Simplified schematic of a key generation scheme based on PUFs.

code construction using polar codes for CSI at encoder and decoder. Section 5 shows the results of our analysis when the probability density function (pdf) of the crossover probabilities is specified according to [10]. Section 6 concludes the paper.

2 Notation

We denote vectors by bold lowercase letters, e.g. $\boldsymbol{v}$, and its i-th component by v_i. We denote random variables by uppercase letters, e.g. X denotes a random variable and we denote its probability mass function by P_X if the random variable is discrete and its probability density function by $f_X(x)$ in case the random variable X is continuous. We denote the expected value of a random variable X by $\mathbb{E}\left[X\right]$. In this work logarithms are to the base 2 unless otherwise stated. We denote the entropy of a discrete random variable by $H(X)$ and the mutual information between two random variables by $I(X;Y)$. We denote the binary entropy function by $H_2(p) := -p \log p - (1-p)\log(1-p)$.

3 Channel Model

In Section 1, it has been motivated why reading out a PUF response corresponds to data transmission over a noisy channel. More specifically the channel corresponds to a binary symmetric channel (BSC) that changes its crossover probability p before each channel use according to a pdf $f_P(p)$. In the following, it is assumed that $f_P(p)$ is known to the transmitter and to the receiver. The pdf has to be estimated by the manufacturer because it may be highly dependent on the manufacturing process.

For a block transmission with block length n, we denote the crossover probability for the i-th channel use by p_i. The p_i can be given to the encoder and

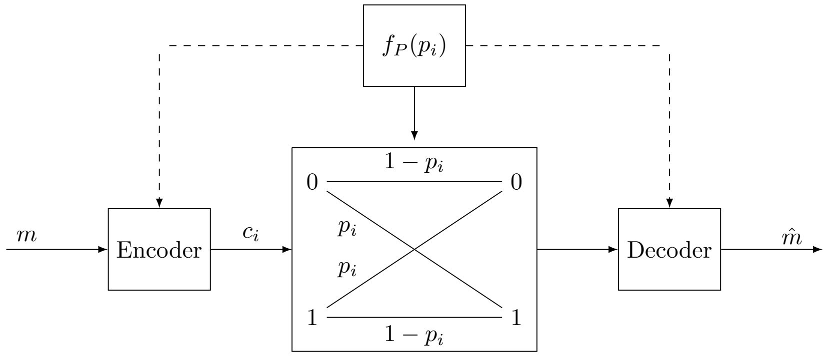

Fig. 2. Varying Binary Symmetric Channel (VBSC) for the i-th channel use without knowledge of p_i at encoder and decoder.

the decoder potentially resulting in larger capacities for the varying binary symmetric channel (VBSC). There are several cases to be distinguished. The p_i can be given to the encoder, the decoder, to both or to none of them. This is reflected in Fig. 2 by using dashed lines for the transmission of the p_i to encoder and decoder. Furthermore, we distinguish causal and non-causal channel state information at the encoder.

Definition 1. *The channel state information (CSI) at the encoder is* causal *if the encoder only has knowledge about* $p^i := p_1, \ldots, p_i$ *to determine the* i*-th codeword symbol* c_i*, i.e.,* $c_i := c_i(m, p^i)$.

Definition 2. *The channel state information (CSI) at the encoder is* non-causal *if the encoder has full knowledge about* $p^n := p_1, \ldots, p_n$ *during the entire encoding process, i.e.,* $c_i := c_i(m, p^n)$ *for all* i.

Notice that it is not necessary to distinguish causal and non-causal CSI at the decoder. For the causal case the decoder can simply wait until the entire block y^n and the complete CSI p^n has been received before it starts the decoding procedure. Therefore, causal CSI is equivalent to non-causal CSI at the decoder.

On the contrary, causal and non-causal CSI at the encoder have to be distinguished in general.

4 Information Theoretic Analysis of the channel

In this section, we aim to obtain the capacity for the VBSC. As introduced in Section 3 we distinguish several cases according to the availability of CSI at the encoder and decoder. We present the capacity results for all cases except non-causal CSI only at the encoder. The proofs in the extended version of this paper [12] or within this section are frequently only shown for a continuous random variable P describing the channel state. Practically, it may be more convenient to quantize the random variable P since its value has to be estimated. The proofs for a discrete random variable P describing the channel state follow analogously.

We state the following lemma which is standard textbook knowledge (e.g. [3]) as a reminder for the reader.

Lemma 1 (BSC Capacity, [3]). *The capacity of the binary symmetric channel (BSC) with crossover probability p is given by*

$$C_{BSC}(p) = 1 - H_2(p)$$

4.1 No Channel State Information at Encoder and Decoder

Drawing the p_i randomly for each channel use can be interpreted as part of the channel's noise. We denote the input sequence of the channel by $X_1, \dots, X_n$ and the corresponding outputs by $Y_1, \dots, Y_n$. Since the VBSC is a discrete memoryless channel (DMC), it holds that

$$C = \max_{P_X} I(X;Y) \ .$$

Theorem 1. *Let $f_P(p)$ denote the pdf of the random variable P from which the p_i are drawn for a VBSC. Let $f_P(p)$ be known at encoder and decoder and let the realizations of the p_i be unknown at the encoder and decoder. Then, the capacity of the VBSC is*

$$C_{VBSC} = C_{BSC}\left(\mathbb{E}\left[P\right]\right) \tag{1}$$

Proof. See extended version [12]. □

4.2 Channel State Information at Encoder and Decoder

Theorem 2 ([5, Chapter 7.4]). *Let a discrete memoryless channel (DMC) with state space $\mathcal{P} := \{1, \dots, |\mathcal{P}|\}$ be given and let the state for each channel use be sampled i.i.d. from a probability mass function (pmf) $P_P(p)$. The capacity of this channel for CSI at encoder and decoder is given by*

$$C_{VBSC-ED} = \max_{P_{X|P}} I(X;Y|P) \ .$$

Using Theorem 2 we determine the capacity of the VBSC for finite channel state space $\mathcal{P}$.

Theorem 3. *For the capacity of the VBSC with CSI at encoder and decoder and finite channel state space $\mathcal{P}$ it holds that*

$$C_{VBSC-ED} = \mathbb{E}\left[C_{BSC}(P)\right] \ .$$

Proof. See extended version [12]. □

Now, we show that for continuous pdf $f_P(p)$ describing the channel state we can approach the rate

$$\tilde{R} := \max_{P_{X|P}} I(X;Y|P) = 1 - \int_{\mathrm{supp}(f_P)} H_2(p) f_P(p)\, dp = \mathbb{E}\big[C_{BSC}(P)\big]$$

arbitrarily close.

Corollary 1. *Let $f_P(p)$ denote the pdf of a random variable P from which the crossover probability of the VBSC is drawn. Furthermore, let $f_P(p)$ be differentiable with a continuous derivative in $(0,1)$. If CSI is known at encoder and decoder the capacity of the channel is arbitrarily close to $\tilde{R}$.*

Proof. See extended version [12]. □

Theorem 1 in [11] states that publication of the CSI does on average not leak information about the expected response of the SRAM cell. Therefore, the requirement of having CSI at the decoder does not compromise the security of the SRAM-PUF and the channel model with CSI at the encoder and decoder is relevant in practice.

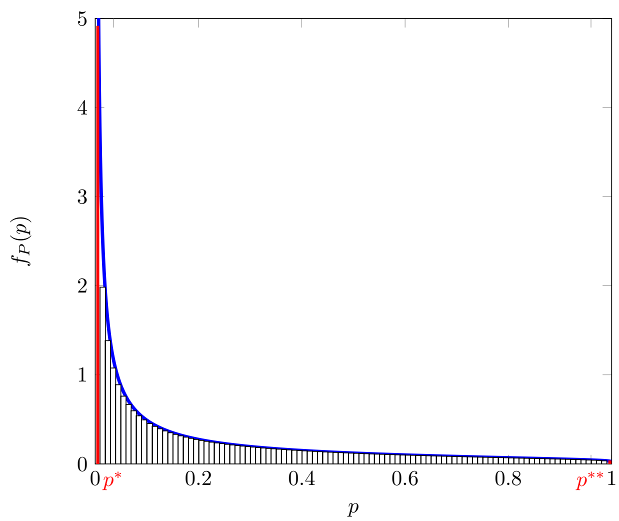

Fig. 3. The probability density function (pdf) of the error for the SRAM cell. The pdf is computed as shown in [10] with parameters $\lambda_1 = 0.1213$ and $\lambda_2 = 0.021$.

Achievability of Capacity using Polar Codes The proof of Corollary 1 presented in the extended version [12] also shows that by constructing a finite number of polar codes, it is possible to achieve the capacity of the VBSC with CSI at encoder and decoder. Basically for each of the intervals in Fig. 3, it is possible to construct a polar code for a BSC with crossover probability equal to the value of the minimal $C_{BSC}(p)$ (largest p if the interval is in $[0, 1/2]$ and smallest p if the interval is in $[1/2, 1]$). Polar codes are capacity achieving for BSCs [1]. Since every interval occurs with strictly positive probability, on average each of them occurs linearly in the block length n. Using the respective polar code for each interval achieves the capacity of the channel arbitrarily close.

4.3 Channel State Information only at the Decoder

As already mentioned in Section 3, it makes no difference whether channel state information is given to the decoder in a causal or non-causal manner.

Theorem 4. *The capacity of the VBSC with channel state information at the decoder is equal to*

$$C_{VBSC-D} = \mathbb{E}\left[C_{BSC}(P)\right]$$

Proof. See extended version [12]. □

Notice that CSI at the encoder (causally or non-causally) has no effect on the capacity of the VBSC as long as CSI is given to the decoder. This effect occurs because the input distribution has no effect on $H(Y|X,P)$ and the uniform distribution maximizes $H(Y|P)$ independently of P.

4.4 Causal Channel State Information at the Encoder

As discussed in Subsection 4.2, publishing CSI does not compromise the security of the PUF. However, CSI only at the encoder may still be relevant because including the CSI into the helper data of the PUF increases the channel capacity at the expense of also increasing the helper data size. In this subsection, we consider causal channel state information at the encoder.

We will show the capacity of the VBSC with causal CSI at the encoder. Furthermore, we will show that the capacity is higher (except for some special cases) than for the case without CSI at encoder and decoder. This statement is non-trivial due to the symmetry of the channel.

Theorem 5. *Let $f_P(p)$ denote the pdf from which the crossover probability is sampled i.i.d. for each channel use. Then the capacity of the VBSC with causal CSI is equal to*

$$C_{VBSC-E} = 1 - H_2\left(\int_0^{0.5} f_P(p)(1-p)\,dp + \int_{0.5}^{1} f_P(p)p\,dp\right) \quad . \tag{2}$$

In order to prove this theorem, we make use of a generic result on the capacity of channels with causal CSI at the encoder which is stated in [5].

Theorem 6. *The capacity of a DMC with CSI causally available at the encoder can be computed by*

$$C_{CSI-E} = \max_{P_U, x(U,P)} I(U;Y) \ , \tag{3}$$

where for the alphabet size $|\mathcal{U}|$ of the auxiliary variable U, it holds that $|\mathcal{U}| \leq \min\{(|\mathcal{X}|-1)|\mathcal{P}|+1, |\mathcal{Y}|\}$.

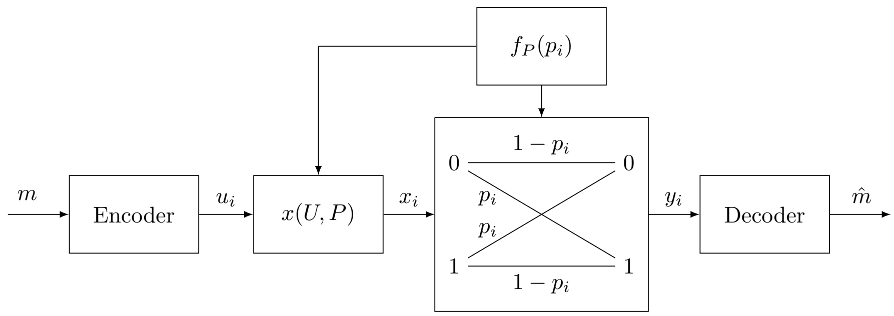

Fig. 4. Analysis of the VBSC with causal CSI at the encoder according to Theorem 6

In the theorem above, the auxiliary variable U can be interpreted as the output of an encoder which encodes an arbitrary message m independently of the channel state. U functions as the input to a mapping device that maps the input according to the channel state onto the channel's input alphabet. This resulting X is then transmitted over the channel resulting in the channel output Y. The transmission of the i-th symbol within the block is illustrated in Fig. 4. The proof of Theorem 5 is based on the following two lemmas.

Lemma 2. *The mapper*

$$x(u,p) = \begin{cases} u & \text{for } p \leq 0.5 \\ \overline{u} & \text{else} \end{cases} \tag{4}$$

minimizes the conditional entropy $H(Y|U)$. Furthermore, $H(Y|U)$ is independent of P_U and it holds

$$H(Y|U) = H_2\left(\int_0^1 f_P(p)\max\{p, 1-p\}\, dp\right) \ . \tag{5}$$

Proof. See extended version [12]. □

Lemma 3. *If the mapper is chosen as specified in Lemma 2, choosing P_U to be the uniform distribution leads to a uniform output distribution P_Y.*

Proof. See extended version [12]. □

Finally, we are ready to determine the capacity of the VBSC with causal CSI at the encoder.

Proof (Theorem 5). Expanding the mutual information in Theorem 6 we have

$$C_{VBSC-E} = \max_{P_U, x(U,P)} I(U;Y) = \max_{P_U, x(U,P)} H(Y) - H(Y|U) \ .$$

By Lemma 2 we have that the conditional entropy $H(Y|U)$ is minimized independent of the choice of P_U for the mapper specified in (4). Furthermore, according to Lemma 3, a uniform P_U leads to a uniform P_Y and consequently $H(Y) = 1$. Therefore, the choice we made for P_U and $x(U,P)$ maximizes $I(U;Y)$ and the channel capacity is obtained. Combining the previous results, we get

$$C_{VBSC-E} = 1 - H_2\left(\int_0^{0.5} f_P(p)(1-p)\,dp + \int_{0.5}^{1} f_P(p)p\,dp\right)$$

completing the proof. □

It holds that $C_{VBSC} \leq C_{VBSC-E}$ with equality if and only if $f_P(p) = 0,\ \forall p \in (1/2, 1]$ almost everywhere. For an elaborate discussion on this statement see the extended version of this paper [12].

4.5 Non-causal Channel State Information at the Encoder

Proposition 1. *For the capacity of the VBSC with non-causal CSI at the encoder denoted as $C_{VBSC-E/nc}$ it holds that*

$$C_{VBSC-E} \leq C_{VBSC-E/nc} \leq C_{VBSC-ED}$$

Proof. Both bounds are trivial. It is obvious that non-causal CSI at the encoder can only increase capacity compared to the case with causal CSI and clearly CSI at encoder and decoder can obviously only increase the capacity compared to the case for non-causal CSI at the encoder only. □

5 Results for a Fixed Crossover Probability Distribution

In this section, we take the model for the error probabilities introduced in [10] and compute the capacities for CSI at encoder and decoder as well as for the case without CSI at encoder and decoder.

In order to do so, we use Theorems 1, 3, 4 and 5 as well as Corollary 1. This results in the values presented in Table 1. We observe that CSI at encoder and decoder increases the capacity of the VBSC for the proposed crossover probability distribution by 0.179 bits per channel use (or by 25%) compared to the case without CSI at encoder and decoder. Table 1 furthermore shows that causal CSI at the encoder increases the channel capacity by 0.0688 bits per channel use again compared to the case without CSI.

Table 1. Numerical computation of the capacities for the pdf $f_P(p)$ from [10] which is depicted in Fig. 3.

C_{VBSC}	C_{VBSC-E}	C_{VBSC-D}	$C_{VBSC-ED}$
0.6961 bpcu	0.7649 bpcu	0.8751 bpcu	0.8751 bpcu

6 Conclusion

In this work, we have introduced the Varying Binary Symmetric Channel (VBSC) to model the difference of PUF responses between the key enrollment and the key reproduction phase. We have derived capacity results depending on the available channel state information (CSI) at encoder and decoder. To exemplify our results, we computed the channel capacities for the crossover probability model proposed in [10]. The results show that the capacity of the corresponding VBSC increases by about 25% if encoder and decoder have CSI compared to the case when both have no knowledge about the channel state. Furthermore, we argued that polar codes can be used to achieve capacity if CSI is available at encoder and decoder.

References

1. Arikan, E.: Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. IEEE Transactions on information Theory **55**(7), 3051–3073 (2009)
2. Chen, B., Ignatenko, T., Willems, F.M.J., Maes, R., van der Sluis, E., Selimis, G.: A robust SRAM-PUF key generation scheme based on polar codes. In: IEEE Global Communications Conference (GLOBECOM). pp. 1–6 (2017)
3. Cover, T.M.: Elements of information theory. John Wiley & Sons (1999)
4. Dodis, Y., Reyzin, L., Smith, A.: Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In: International conference on the theory and applications of cryptographic techniques. pp. 523–540. Springer (2004)
5. El Gamal, A., Kim, Y.H.: Network information theory. Cambridge university press (2011)
6. Guajardo, J., Kumar, S.S., Schrijen, G.J., Tuyls, P.: FPGA intrinsic PUFs and their use for IP protection. In: Paillier, P., Verbauwhede, I. (eds.) Workshop on Cryptographic Hardware and Embedded Systems (CHES). LNCS, vol. 4727, pp. 63–80. Springer Berlin / Heidelberg (2007)
7. Herder, C., Yu, M.D.M., Koushanfar, F., Devadas, S.: Physical unclonable functions and applications: A tutorial. Proceedings of the IEEE **102**(8), 1126–1141 (2014)
8. Hiller, M., Kürzinger, L., Sigl, G.: Review of error correction for PUFs and evaluation on state-of-the-art FPGAs. Journal of Cryptographic Engineering (2020)
9. Maes, R.: Physically Unclonable Functions: Constructions, Properties and Applications. Dissertation (2012)
10. Maes, R.: An accurate probabilistic reliability model for silicon PUFs. In: Bertoni, G., Coron, J.S. (eds.) Workshop on Cryptographic Hardware and Embedded Systems (CHES). pp. 73–89. LNCS, Springer, Berlin / Heidelberg (2013)
11. Maes, R., Tuyls, P., Verbauwhede, I.: A soft decision helper data algorithm for SRAM PUFs. In: IEEE International Symposium on Information Theory (ISIT). pp. 2101–2105 (2009)
12. Maringer, G., Xhemrishi, M., Puchinger, S., Garb, K., Liu, H., Jerkovits, T., Kürzinger, L., Hiller, M., Wachter-Zeh, A.: Analysis of communication channels related to physical unclonable functions (extended version of this paper). arXiv preprint (2021)