LOW-COMPLEXITY AND RELIABLE TRANSFORMS FOR PHYSICAL UNCLONABLE FUNCTIONS

Onur Günlü and Rafael F. Schaefer

Information Theory and Applications Chair, Technische Universität Berlin
{guenlue, rafael.schaefer}@tu-berlin.de

ABSTRACT

Noisy measurements of a physical unclonable function (PUF) are used to store secret keys with reliability, security, privacy, and complexity constraints. A new set of low-complexity and orthogonal transforms with no multiplication is proposed to obtain bit-error probability results significantly better than all methods previously proposed for key binding with PUFs. The uniqueness and security performance of a transform selected from the proposed set is shown to be close to optimal. An error-correction code with a low-complexity decoder and a high code rate is shown to provide a block-error probability significantly smaller than provided by previously proposed codes with the same or smaller code rates.

Index Terms— physical unclonable function (PUF), no multiplication transforms, secret key agreement, low complexity.

1. INTRODUCTION

Biometric identifiers such as fingerprints are useful to authenticate a user. Similarly, secret keys are traditionally stored in non-volatile memories (NVMs) to authenticate a physical device that contains the key. NVMs require hardware protection even when the device is turned off since an attacker can try to obtain the key at any time. A safe and cheap alternative to storing keys in NVMs is to use physical identifiers, e.g., fine variations of ring oscillator (RO) outputs, as a randomness source. Since invasive attacks to physical identifiers permanently change the identifier output, there is no need for continuous hardware protection for physical identifiers [1].

Physical unclonable functions (PUFs) are physical identifiers with reliable and high-entropy outputs [2,3]. PUF outputs are unique to each device, so they are used for safe and low-complexity key storage in digital devices. These keys can be used for private authentication, secure computation, and encryption. Replacing such identifiers is expensive, so key-storage methods should limit the information the public data leak about the identifier outputs. Moreover, the same device should be able to reconstruct a secret key generated from the noiseless outputs by using the noisy outputs and public information. The ultimate secret-key vs. privacy-leakage rate tradeoffs are given in [4–6]. The secret-key and privacy-leakage rate limits for a suboptimal chosen-secret (CS) model called *fuzzy commitment scheme* (FCS) [7] are given in [8]. We consider the FCS to compare different post-processing methods applied to PUFs. Asymptotically optimal CS model constructions are given in [9] and similar comparison results can be obtained by using these constructions.

Physical identifier outputs are highly correlated and noisy, which are the two main problems in using PUFs. If errors in the extracted sequences are not corrected, PUF reliability would be low. If correlations are not eliminated, machine learning algorithms can model the PUF outputs [10]. To solve the two problems, the discrete cosine transform (DCT) is used in [11] to generate a uniformly-distributed bit sequence from PUFs under varying environmental conditions. Similarly, the discrete Walsh-Hadamard transform (DWHT), discrete Haar transform (DHT), and Karhunen-Loève transform (KLT) are compared in [12] in terms of the maximum secret-key length, decorrelation efficiency, reliability, security, and hardware cost. The DCT, DWHT, and DHT provide good reliability and security results, and a hardware implementation of the DWHT in [12] shows that the DWHT requires a substantially smaller hardware area than other transforms. There are two main reasons why the DWHT can be implemented efficiently. Firstly, the matrix that represents the DWHT has elements 1 or -1, so there is no matrix multiplication. Secondly, an input-selection algorithm that is an extension of the algorithm in [13] allows to calculate two-dimensional (2D) DWHT recursively. Based on these observations, we propose a new set of transforms that preserve these properties and that significantly improve the reliability of the sequences extracted from PUFs.

The FCS requires error-correction codes (ECCs) to achieve the realistic block-error probability of $P_{\mathrm{B}} = 10^{-9}$ for RO PUFs. The ECCs proposed in [12] have better secret-key and privacy-leakage rates than previously proposed codes, but in some cases it is assumed that if multiple bits are extracted from each transform coefficient, each bit is affected by independent errors. This assumption is not valid in general. Thus, we extract only one bit from each transform coefficient. The contributions of this work are as follows.

- We propose a new set of 2D orthogonal transforms that have low-complexity hardware implementations and no matrix multiplications. The new set of transforms are shown to provide an average bit-error probability smaller than the most reliable transform considered in the PUF literature, i.e., DCT.
- Bit sequences extracted using a transform selected from the new set of transforms are shown to give good uniqueness and security results that are comparable to state-of-the-art results.
- We propose a joint transform-quantizer-code design method for the new set of transforms in combination with the FCS to achieve a block-error probability substantially smaller than the common value of 10^{-9} with perfect secrecy.

This paper is organized as follows. In Section 2, we review the FCS. The transform-coding algorithm to extract secure sequences from RO PUFs is explained in Section 3. A new set of orthogonal transforms that require a small hardware area and that result in bit-error probabilities smaller than previously considered transforms is proposed in Section 4. In Section 5, we compare the new transforms with previous methods and show that the proposed ECC provides a

O. Günlü and R. F. Schaefer were supported by the German Federal Ministry of Education and Research (BMBF) within the national initiative for "Post Shannon Communication (NewCom)" under Grant 16KIS1004.

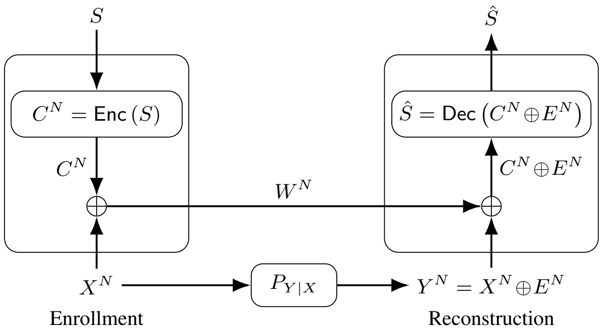

Fig. 1. The fuzzy commitment scheme (FCS).

block-error probability for the new selected transform (ST) that is smaller than for previously considered transforms.

2. REVIEW OF THE FUZZY COMMITMENT SCHEME

Fig. 1 shows the FCS, where an encoder $\mathsf{Enc}(\cdot)$ adds a codeword C^N, uniformly distributed over a set with cardinality $|\mathcal{S}|$, modulo-2 to the binary noiseless PUF-output sequence X^N during enrollment. We show in Section 3 that the sequence X^N and its noisy version Y^N can be obtained by applying the post-processing steps in Fig. 2 to RO outputs $\widetilde{X}^L$ and its noisy version $\widetilde{Y}^L$, respectively. The sum $W^N = C^N \oplus X^N$ is publicly sent through a noiseless and authenticated channel, and it is called *helper data*. The modulo-2 sum of W^N and the noisy PUF-output sequence $Y^N = X^N \oplus E^N$, where E^N is the binary error vector, gives the noisy codeword $C^N \oplus E^N$. Using the noisy codeword, a channel decoder $\mathsf{Dec}(\cdot)$ estimates the secret key S during reconstruction. A reliable secret-key agreement is possible by using X^N, Y^N, and W^N [14, 15].

One can achieve a (secret-key, privacy-leakage) rate pair (R_{s}, R_ℓ) using the FCS with perfect secrecy if, given any $\epsilon > 0$, there is some $N \geq 1$, and an encoder and a decoder for which $R_{\text{s}} = \dfrac{\log|\mathcal{S}|}{N}$ and

$$\Pr[S \neq \hat{S}] \leq \epsilon \qquad \text{(reliability)} \qquad (1)$$

$$I\big(S; W^N\big) = 0 \qquad \text{(perfect secrecy)} \qquad (2)$$

$$\frac{1}{N} I\big(X^N; W^N\big) \leq R_\ell + \epsilon. \qquad \text{(privacy)} \qquad (3)$$

Condition (2) ensures that the public side information W^N does not leak any information about the secret key, so one achieves perfect secrecy. The normalized information that W^N leaks about the PUF output sequence X^N is considered in (3). If one should asymptotically limit the unnormalized privacy leakage $I(X^N; W^N)$, private keys available during enrollment and reconstruction are necessary [4], which is not realistic or practical; see the discussions in [9].

Suppose the measurement channel $P_{Y|X}$ is a binary symmetric channel (BSC) with crossover probability p, and X is independent and identically distributed (i.i.d.) according to a uniform distribution. Define $H_b(p) = -p \log p - (1-p)\log(1-p)$ as the binary entropy function. The region $\mathcal{R}$ of all achievable (secret-key, privacy-leakage) rate pairs for the FCS with perfect secrecy is [8]

$$\mathcal{R} = \big\{ (R_{\text{s}}, R_\ell)\colon \quad 0 \leq R_{\text{s}} \leq 1 - H_b(p), \quad R_\ell \geq 1 - R_{\text{s}} \big\}. \qquad (4)$$

We plot this region in Section 5 to evaluate the secret-key and privacy-leakage rates achieved by the proposed ECC.

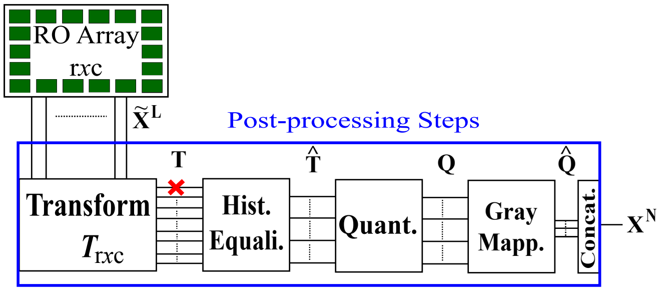

Fig. 2. The transform-coding steps.

The FCS is a particular realization of the CS model. The region $\mathcal{R}_{\text{cs}}$ of all achievable (secret-key, privacy-leakage) rate pairs for the CS model, where a generic encoder is used to confidentially transmit an embedded secret key to a decoder that observes Y^N and the helper data W^N, is given in [4, 5] as the union over all $P_{U|X}$ of the set of achievable rate pairs (R_{s}, R_ℓ) such that

$$\Big\{ 0 \leq R_{\text{s}} \leq I(U;Y), \qquad R_\ell \geq I(U;X) - I(U;Y) \Big\} \qquad (5)$$

where P_X is the probability distribution of X and the alphabet $\mathcal{U}$ of the auxiliary random variable U can be limited to have the size $|\mathcal{U}| \leq |\mathcal{X}| + 1$ as $U - X - Y$ forms a Markov chain. The FCS achieves a boundary point of $\mathcal{R}_{\text{cs}}$ for a BSC $P_{Y|X}$ only at the point $(R_{\text{s}}^*, R_\ell^*) = (1 - H_b(p), H_b(p))$. To achieve the other points on the rate-region boundary, one should use a nested code construction as in [9] or a binning based construction as in [16], both of which require careful polar code [17] designs. This is not necessary to illustrate the gains from the new set of transforms and it suffices to combine the new set with the FCS.

3. POST-PROCESSING STEPS

We consider a 2D array of $r \times c$ ROs. Denote the continuous-valued outputs of $L = r \times c$ ROs as the vector random variable $\widetilde{X}^L$, distributed according to $f_{\widetilde{X}^L}$. Suppose that the noise component $\widetilde{E}_j$ on the j-th RO output is Gaussian distributed with zero mean for all $j = 1, 2, \ldots, L$ and that the noise components are mutually independent. Denote the noisy RO outputs as $\widetilde{Y}^L = \widetilde{X}^L + \widetilde{E}^L$. We extract binary vectors X^N and Y^N from $\widetilde{X}^L$ and $\widetilde{Y}^L$, respectively, and define binary error variables $E_i = X_i \oplus Y_i$ for $i = 1, 2, \ldots, N$.

The post-processing steps used during the enrollment (and reconstruction) to extract a bit sequence X^N (and its noisy version Y^N) are depicted in Fig. 2. These steps are transformation, histogram equalization, quantization, Gray mapping, and concatenation. Since RO outputs $\widetilde{X}^L$ are correlated, we apply a transform $T_{r\times c}(\cdot)$ for decorrelation. We model all transform coefficients and noise components as random variables with Gaussian marginal distributions. A transform-coefficient output T that comes from a distribution with mean $\mu \neq 0$ and variance $\sigma^2 \neq 1$ is converted into a standard Gaussian random variable during histogram equalization, which reduces the hardware area when multiple bits are extracted. Independent bits can be extracted from transform coefficients by setting the quantization boundaries of a K-bit quantizer to

$$b_k = Q^{-1}\left(1 - \frac{k}{2^K}\right) \text{ for } k = 0, 1, \ldots, 2^K \qquad (6)$$

where $Q(\cdot)$ is the Q-function. Quantizing a coefficient $\hat{T}$ to k if $b_{k-1} < \hat{T} \leq b_k$ ensures that X^N is uniformly distributed, which is necessary to achieve the rate point where the FCS is optimal.

One can use scalar quantizers without a performance loss in security if the RO output statistics satisfy certain constraints [6]. We do not use the first transform coefficient, i.e., DC coefficient, for bit extraction since it corresponds to the average over the RO array, known by an attacker [6]. Furthermore, Gray mapping ensures that the neighboring quantization intervals result in only one bit flip. This is a good choice as the noise components E_i for all $i = 1, 2, \ldots, N$ have zero mean. The sequences extracted from transform coefficients are concatenated to obtain the sequence X^N (or Y^N).

4. NEW ORTHOGONAL TRANSFORMS

A useful metric to measure the complexity of a transform is the number of operations required for computations. Consider only RO arrays of sizes $r = c = 8$ and 16, which are powers of 2, so fast algorithms are available. In [6], the DWHT is suggested as the best candidate among the set of transforms {DCT, DHT, KLT, DWHT} for RO PUF applications with a low-complexity constraint such as internet of things (IoT) applications.

In [12], we extend an input-selection algorithm to compute the 2D 16×16 DWHT by applying a 2×2 matrix operation recursively to illustrate that the DWHT requires a small hardware area in a field programmable gate array (FPGA) since it does not require any multiplications. Following this observation, we propose a set of transforms that are orthogonal (to decorrelate the RO outputs better), that have matrix elements 1 or -1 (to eliminate multiplications), and that have size of 16×16 (to apply the input-selection algorithm given in [12] to further reduce complexity). We show in the next section that these transforms provide higher reliability than other transforms previously considered in the literature.

4.1. Orthogonal Transform Construction and Selection

Consider an orthogonal matrix A with elements 1 or -1 and of size $k \times k$, i.e., $AA^T = I$, where T is the matrix transpose and I is the identity matrix of size $k \times k$. It is straightforward to show that the following matrices are also orthogonal:

$$\begin{bmatrix} A & A \\ A & -A \end{bmatrix}, \begin{bmatrix} A & A \\ -A & A \end{bmatrix}, \begin{bmatrix} A & -A \\ A & A \end{bmatrix}, \begin{bmatrix} -A & A \\ A & A \end{bmatrix},$$
$$\begin{bmatrix} -A & -A \\ -A & A \end{bmatrix}, \begin{bmatrix} -A & -A \\ A & -A \end{bmatrix}, \begin{bmatrix} -A & A \\ -A & -A \end{bmatrix}, \begin{bmatrix} A & -A \\ -A & -A \end{bmatrix}. \quad (7)$$

Since 2^{k^2} possible matrices should be checked for orthogonality, we choose $k = 4$ to keep the complexity of the exhaustive search for orthogonal matrices low. The result of the exhaustive search is a set of orthogonal matrices A of size 4×4. By applying the matrix construction methods in (7) twice consecutively, we obtain 12288 unique orthogonal transforms of size 16×16 with elements 1 or -1.

We apply these orthogonal transforms, one of which is the DWHT, to an RO dataset to select the orthogonal transform whose maximum bit-error probability over the transform coefficients is minimum. This selection method provides reliability guarantees to every transform coefficient. An ECC that has a higher code dimension than it is achievable according to the Gilbert-Varshamov (GV) bound [18, 19] for the maximum error probability over the transform coefficients of the ST, is given in Section 5.3. This illustrates that our selection method is conservative and the block-error probability is substantially smaller than 10^{-9}.

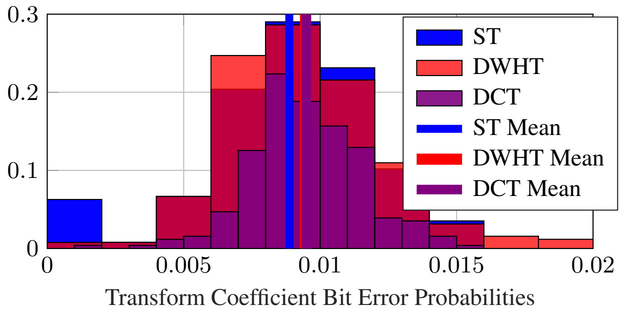

Fig. 3. The histograms and means of the bit-error probabilities of the transform coeeficients obtained from the DCT, DWHT, and the selected transform (ST) from the new set.

There are also other orthogonal transforms of size 16×16 but we illustrate in the next section that the new set suffices to significantly increase the reliability of the extracted bits as compared to previously considered transforms and previous RO PUF methods.

5. PERFORMANCE EVALUATIONS

We use RO arrays of size 16×16 from the RO dataset in [20] and apply the transform-coding steps in Fig. 2 to compare the previously considered transforms with the new set of transforms in terms of their reliability, uniqueness, and security. We illustrate that a Bose-Chaudhuri-Hocquenghem (BCH) code can be used for error correction in combination with the FCS to achieve a block-error probability smaller than the common value of 10^{-9}.

5.1. Transform Comparisons

We compare the orthogonal transform selected from the new set, i.e., the ST, with the DCT and DWHT in terms of the bit-error probabilities of the 255 transform coefficients obtained from the RO dataset in [20]. Fig. 3 illustrates the bit-error probabilities of the DCT, DWHT, and the ST. The mean of the ST is smaller than the means of the DCT and DWHT. Furthermore, the maximum bit-error probability of the DCT and ST are almost equal and are less than the maximum error probability of the DWHT. Most importantly, the ST has a large set of transform coefficients with bit-error probabilities close to zero, so an ECC design for the maximum or mean bit-error probability of the ST would give pessimistic rate results. We propose in the next section an ECC for the ST to achieve a smaller block-error probability than the block-error probability for the DCT.

5.2. Uniqueness and Security

A common measure to check the randomness of a bit sequence is uniqueness, i.e., the average fractional Hamming distance (HD) between the sequences extracted from different RO PUFs [21]. The rate region in (4) is valid if the extracted bit sequences are uniformly distributed, making the uniqueness a valid measure for the FCS.

Uniqueness results for the DCT, DWHT, KLT, and DHT have a mean HD of 0.5000 and HD variances of approximately 7×10^{-4} [12], which are close to optimal and better than previous RO PUF results. For the ST, we obtain a mean HD of 0.5001 and a HD variance of 2.69×10^{-2}. This suggests that the ST has good average uniqueness performance, but there might be a small set of RO PUFs from which slightly biased bit sequences are extracted. The latter can be avoided during manufacturing by considering uniqueness as a parameter in yield analysis of the chip that embodies the PUF. We apply the national institute of standards and technology (NIST) randomness tests [22] to check whether there is a detectable deviation from the uniform distribution in the sequences extracted by using the ST. The bit sequences generated with the ST pass most of the randomness tests, which is considered to be an acceptable result [22]. A correlation thresholding approach in [11] further improves security.

5.3. Code Selection

Consider the scenario where secret keys are used as an input to the advanced encryption standard (AES), a symmetric-key cryptosystem, with a key size of 128 bits, so the code dimension of the ECC should be at least 128 bits. The maximum error probability over the transform coefficients of the ST is $p_{\text{max}} = 0.0149$, as shown in Fig. 3. Furthermore, assume that we use an ECC with a bounded minimum distance decoder (BMDD) to keep the complexity low. A BMDD can correct all error patterns with up to $\lfloor\frac{d_{\text{min}}-1}{2}\rfloor$ errors, where d_{min} is the minimum distance of the code. It is straightforward to show that the ECC should have at least a minimum distance of $d_{\text{min}} = 41$ to achieve a block-error probability of $P_{\text{B}} \leq 10^{-9}$ if all transform coefficients are assumed to have a bit-error probability of p_{max}. None of binary BCH and Reed-Solomon (RS) codes, which have good minimum-distance properties, can satisfy these parameters. Similarly, the GV bound computed for p_{max} shows that there exists a linear binary ECC with code dimension 98. Consider the binary BCH code with the block length 255, code dimension 131 that is greater than the code dimension of 98 given by the GV bound, and minimum distance $d_{\text{min,BCH}} = 37$ that is close to the required value of $d_{\text{min}} = 41$. We illustrate in the next section that this BCH code provides a block-error probability significantly smaller than 10^{-9}.

5.4. Reliability, Privacy, and Secrecy Analysis of the Code

We now show that the proposed ECC satisfies the block-error probability constraint. The block-error probability P_{B} for the BCH$(255, 131, 37)$ code with a BMDD is equal to the probability of having more than 18 errors in the codeword, i.e., we have

$$P_{\text{B}} = \sum_{j=19}^{255} \left[\sum_{\mathcal{D}\in\mathcal{F}_j} \prod_{i\in\mathcal{D}} p_i \cdot \prod_{i\in\mathcal{D}^c} (1-p_i) \right] \tag{8}$$

where $p_i \leq p_{\text{max}}$ is the bit-error probability of the i-th transform coefficient, as in Fig. 3, for $i=2, 3, \ldots, 256$, $\mathcal{F}_j$ is the set of all size-j subsets of the set $\{2, 3, \ldots, 256\}$, and $\mathcal{D}^c$ denotes the complement of the set $\mathcal{D}$. The bit-error probabilities p_i represent probabilities of independent events due to the mutual independence assumption for transform coefficients and one-bit quantizers used.

The evaluation of (8) requires $\sum_{j=0}^{18} \binom{255}{j} \approx 1.90\times10^{27}$ different calculations, which is not practical. We therefore apply the discrete Fourier transform - characteristic function (DFT-CF) method [23] to (8) and obtain the result $P_{\text{B}} \approx 2.860 \times 10^{-12} < 10^{-9}$. This value is smaller than the block-error probabilitiy $P_{\text{B,DCT}} = 1.26 \times 10^{-11}$ obtained in [6] for the DCT with the same code. The block-error probability constraint is thus satisfied by using the BCH code although the conservative analysis suggests otherwise.

The rate regions given in (4) and (5) are asymptotic results, i.e., they assume $N \to \infty$. Since separate channel and secrecy coding is optimal for the FCS, we can use the finite length bounds for a BSC $P_{Y|X}$ with crossover probability $p = \frac{1}{L-1}\sum_{i=2}^{L} p_i \approx 0.0088$, i.e., the error probability averaged over all used coefficients. In [6], we show that the BCH$(255, 131, 37)$ code achieves $(R_{\text{s,BCH}}, R_{\ell,\text{BCH}}) \approx (0.514,\ 0.486)$ bits/source-bit, significantly better than previously proposed codes in the RO PUF literature, so it suffices to compare the proposed code with the best possible finite-length results for the FCS. We use Mrs. Gerber's lemma [24], giving the optimal auxiliary random variable U in (5), to compute all points in the region $\mathcal{R}_{\text{cs}}$. We plot all achievable rate pairs, the (secret-key, privacy-leakage) rate pair of the proposed BCH code, and a finite-length bound for the block length of $N = 255$ bits and $P_{\text{B}} = 10^{-9}$ in Fig. 4.

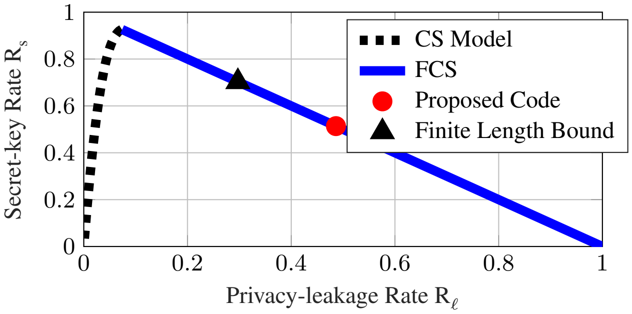

Fig. 4. Boundaries of asymptotically achievable rate regions for the CS model and the FCS, operation point of the proposed code, and a finite-length bound for $N = 255$ bits and $P_{\text{B}} = 10^{-9}$.

The maximum secret-key rate is $R_{\text{s}}^* \approx 0.9268$ bits/source-bit with a corresponding minimum privacy-leakage rate of $R_\ell^* \approx 0.0732$ bits/source-bit. The gap between the points $(R_{\text{s,BCH}}, R_{\ell,\text{BCH}})$ and $(R_{\text{s}}^*, R_\ell^*)$ can be partially explained by the short block length of the code and the small block-error probability. The finite-length bound given in [25, Theorem 52] shows that the rate pair $(R_{\text{s}}, R_\ell) = (0.7029, 0.2971)$ bits/source-bit is achievable by using the FCS, as depicted in Fig. 4. One can thus improve the rate pairs by using better codes and decoders with higher hardware complexity, which is undesirable for IoT applications. Fig. 4 also illustrates the fact that there are operation points of the region $\mathcal{R}_{\text{cs}}$ that cannot be achieved by using the FCS and, e.g., a nested polar code construction from [9] should be used to achieve all points in $\mathcal{R}_{\text{cs}}$.

6. CONCLUSION

We proposed a new set of transforms that are orthogonal (so that the decorrelation efficiency is high), that have elements 1 or -1 (so that the hardware complexity is low), and that have a size of $k \times k$ where k is a power of 2 (so that an input-selection algorithm can be applied to further decrease complexity). By using one-bit uniform quantizers for each transform coefficient obtained by applying the ST, we obtained bit-error probabilities that are on average smaller than the bit-error probabilities obtained from previously considered transforms. We proposed a BCH code as the ECC for RO PUFs in combination with the FCS. This code achieves the best rate pair in the RO PUF literature and it gives a block-error probability for the ST that is substantially smaller than for the DCT. We illustrated that the FCS cannot achieve all possible rate points. In future work, in combination with the new set of transforms, we will apply a joint vector quantization and error correction method by using nested polar codes to achieve rate pairs that cannot be achieved by the FCS.

7. REFERENCES

[1] C. Böhm and M. Hofer, *Physical Unclonable Functions in Theory and Practice*, Springer, New York, NY, Oct. 2012.

[2] B. Gassend, "Physical random functions," M.S. thesis, M.I.T., Cambridge, MA, Jan. 2003.

[3] R. Pappu, *Physical One-way Functions*, Ph.D. thesis, M.I.T., Cambridge, MA, Oct. 2001.

[4] T. Ignatenko and F. M. J. Willems, "Biometric systems: Privacy and secrecy aspects," *IEEE Trans. Inf. Forensics Security*, vol. 4, no. 4, pp. 956–973, Dec. 2009.

[5] L. Lai, S.W. Ho, and H. V. Poor, "Privacy-security trade-offs in biometric security systems - Part I: Single use case," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 1, pp. 122–139, Mar. 2011.

[6] O. Günlü, *Key Agreement with Physical Unclonable Functions and Biometric Identifiers*, Ph.D. thesis, TU Munich, Germany, Nov. 2018, published by Dr. Hut Verlag.

[7] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in *ACM Conf. Comp. Commun. Security*, New York, NY, Nov. 1999, pp. 28–36.

[8] T. Ignatenko and F. M. J. Willems, "Information leakage in fuzzy commitment schemes," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 337–348, Mar. 2010.

[9] O. Günlü, O. İşcan, V. Sidorenko, and G. Kramer, "Code constructions for physical unclonable functions and biometric secrecy systems," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 11, pp. 2848–2858, Nov. 2019.

[10] G. Hospodar, R. Maes, and I. Verbauwhede, "Machine learning attacks on 65nm Arbiter PUFs: Accurate modeling poses strict bounds on usability," in *IEEE Int. Workshop Inf. Forensics Security*, Tenerife, Spain, Dec. 2012, pp. 37–42.

[11] O. Günlü, O. İşcan, and G. Kramer, "Reliable secret key generation from physical unclonable functions under varying environmental conditions," in *IEEE Int. Workshop Inf. Forensics Security*, Rome, Italy, Nov. 2015, pp. 1–6.

[12] O. Günlü, T. Kernetzky, O. İşcan, V. Sidorenko, G. Kramer, and R. F. Schaefer, "Secure and reliable key agreement with physical unclonable functions," *Entropy*, vol. 20, no. 5, May 2018.

[13] K. Komatsu and K. Sezaki, "Lossless 2D discrete Walsh-Hadamard transform," in *IEEE Int. Conf. Acoustics, Speech Sign. Process.*, Salt Lake City, UT, May 2001, pp. 1917–1920.

[14] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography - Part I: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, July 1993.

[15] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 2733–742, May 1993.

[16] R. A. Chou, M. R. Bloch, and E. Abbe, "Polar coding for secret-key generation," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 6213–6237, Nov. 2015.

[17] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.

[18] E. N. Gilbert, "A comparison of signalling alphabets," *Bell Sys. Techn. J.*, vol. 31, no. 3, pp. 504–522, May 1952.

[19] R. R. Varshamov, "Estimate of the number of signals in error correcting codes," *Docklady Akad. Nauk SSSR*, vol. 117, pp. 739–741, 1957.

[20] A. Maiti, J. Casarona, L. McHale, and P. Schaumont, "A large scale characterization of RO-PUF," in *IEEE Int. Symp. Hardware-Oriented Security Trust*, Anaheim, CA, June 2010, pp. 94–99.

[21] O. Günlü and O. İşcan, "DCT based ring oscillator physical unclonable functions," in *IEEE Int. Conf. Acoustics, Speech Sign. Process.*, Florence, Italy, May 2014, pp. 8198–8201.

[22] A. Rukhin et al., "A statistical test suite for random and pseudorandom number generators for cryptographic applications," Tech. Rep., National Inst. Stand. Techno., 2001, Rev. in 2010.

[23] Y. Hong, "On computing the distribution function for the sum of independent and nonidentical random indicators," Tech. Rep., Dep. Stat., Virginia Tech., Blacksburg, VA, Apr. 2011.

[24] A. D. Wyner and J. Ziv, "A theorem on the entropy of certain binary sequences and applications: Part I," *IEEE Trans. Inf. Theory*, vol. 19, no. 6, pp. 769–772, Nov. 1973.

[25] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.