

# Key Agreement with Correlated Noise and Multiple Entities or Enrollments

Onur Günlü

Information Theory and Applications Chair

Technische Universität Berlin, Germany

Email: guenlue@tu-berlin.de

**Abstract**—We extend a basic key agreement model with a hidden identifier source to a multi-user model with joint secrecy and privacy constraints over all entities that do not trust each other. Different entities that use different measurements of the same remote source through broadcast channels (BCs) to agree on mutually-independent local secret keys are considered. This model is the proper multi-user extension of the basic model since the encoder and decoder pairs are not assumed to trust other pairs, unlike assumed in the literature. Strong secrecy constraints imposed jointly on all secret keys, which is more stringent than separate secrecy leakage constraints for each secret key considered in the literature, are satisfied. Inner bounds for maximum key rate, and minimum privacy-leakage and storage rates are proposed for any finite number of entities. Inner and outer bounds for degraded and less-noisy BCs are given to illustrate cases with strong privacy. A multi-enrollment model that is used for common physical unclonable functions (PUFs) is also considered to establish inner and outer bounds for key-leakage-storage regions that differ only in the Markov chains imposed. For this special case, the encoder and decoder measurement channels have the same channel transition matrix and secrecy leakage is measured for each secret key separately. We illustrate cases for which it is useful to have multiple enrollments as compared to a single enrollment and vice versa.

## I. INTRODUCTION

Physical identifiers such as fine variations of ring oscillator (RO) outputs or random start-up values of static random access memories (SRAMs) that depend on uncontrollable manufacturing variations, are safer and cheaper alternatives to key storage in a non-volatile memory [2], [3]. Such physical identifiers for digital devices such as Internet-of-Things (IoT) devices are called physical unclonable functions (PUFs) [2]. We use the basic source model for key agreement from [4], [5] to find achievable rate regions for key agreement with PUFs and biometric identifiers. In this classic model, an encoder observes a source output to generate a secret key and sends public side information, i.e., *helper data*, to a decoder, so the decoder can reliably reconstruct the same secret key by observing another source output and the helper data. The main constraints are that the information leaked about the secret key, i.e., *secrecy leakage*, is negligible and the information

leaked about the identifier output, i.e., *privacy leakage*, is small [6]. Furthermore, the amount of public storage should also be minimized to limit the hardware cost [7].

Suppose the encoder generates a key from a noisy measurement of a hidden (or remote) source output, and a decoder has access to another noisy measurement of the same source and the helper data to reconstruct the same key. We call this model the *generated-secret* (GS) model with a hidden source. This model is introduced in [8] as an extension of the visible (noiseless) source outputs observed by the encoder, considered in [6]. Similarly, for the *chosen-secret* (CS) model, an embedded (or chosen) key and noisy identifier measurements are combined by the encoder to generate the public helper data. We consider both models to address different applications.

Multiple enrollments of a hidden source using noisy measurements are considered in [9], where weakly secure secret keys are generated without privacy leakage and storage constraints. Furthermore, there is a causality assumption in [9] on the availability of the helper data, i.e., any decoder has access to all previously-generated helper data. This assumption is not necessarily realistic as a decoder of, e.g., an IoT device that embodies a PUF should be low complexity and the amount of data to process increases linearly with the number of enrollments. In addition, any manipulation in any of the helper data can cause the complete multi-enrollment system to fail. A classic method used for key agreement, i.e., the fuzzy commitment scheme (FCS) [10], is used in [11] in combination with an SRAM PUF to enroll the noisy outputs of the same SRAM multiple times. The symmetry condition in [11, Eq. (16)] conditioned on a fixed SRAM cell state is entirely similar to the symmetry satisfied by binary-input symmetric output (BISO) channels; see e.g., [12, p. 613], [8, Eq. (14)]. For SRAM outputs that satisfy this symmetry, the secrecy leakage about each separate secret key is shown to be zero. In [13, Theorem 1] the secret-key capacity of the two-enrollment key agreement problem is established for measurement channels with the same channel transition matrix. However, these multi-enrollment models do not consider the privacy leakage and storage constraints, there is no constraint on the independence of the secret keys of different enrollments, and the secrecy leakage constraint is weak and

An extension of this paper appeared in IEEE Transactions on Information Forensics and Security in [1].

is not applied jointly on all secret keys. Furthermore, optimal random linear code constructions that achieve the boundaries of the key-leakage-storage regions are given in [14], where the classic code constructions such as the FCS are shown to be strictly suboptimal. Therefore, the multi-enrollment models and constructions in the literature are strictly suboptimal and not necessarily realistic. We therefore list stronger secrecy constraints jointly on all entities, which approximates the reality better in combination with storage rate and joint privacy-leakage rate constraints. These constraints define the *multi-entity key agreement* problem, where the entities that use the same identifier do not have to trust other entities after key agreement. Thus, the multi-entity key agreement problem is a proper multi-user extension of single-enrollment models. We first consider the multi-entity key agreement problem and then analyze a special case of the multi-enrollment key agreement problem to illustrate scenarios for which a single enrollment can be more useful than multiple enrollments and vice versa.

Every measurement of an identifier is considered to be noisy due to, e.g., local temperature and voltage changes in the hardware of the PUF circuit or a cut on the finger. Noise components at the encoder and decoder measurements of a hidden source can be also correlated due to, e.g., the surrounding logic in the hardware [15] or constant fingertip moisture. This correlation between the noise sequences is modeled in [16] as a broadcast channel (BC) [17] with an input that is the hidden source output and with outputs that are the noisy encoder and decoder measurements. We use this model for multi-entity key agreement, where each entity (i.e., each encoder and decoder pair) observes noisy identifier outputs of the same hidden source through different BCs. We allow the BCs to be different as honest entities use different hardware implementations of the encoder and decoder pairs, which results in different correlations between noise components.

We also consider physically-degraded (PD) and less-noisy (LN) BCs to give finer inner and outer bounds to the key-leakage-storage regions for the GS and CS models of the multi-entity key agreement problem. For the considered PD and LN BCs, we prove that strong privacy can be achieved. We next list our main contributions below, and a longer version with further model figures and proofs is available in [1].

## II. MULTI-ENTITY KEY AGREEMENT MODELS

Consider hidden identifier outputs  $X^n$  that are i.i.d. according to a probability distribution  $P_X$ . The hidden (or remote) source with outputs  $X^n$  is common to all honest entities that enroll the same identifier, but they observe different noisy measurements of the same hidden source. If there are a finite number  $J$  of honest entities that use the same identifier, the  $j$ -th encoder and decoder pair observes noisy source measurements that are outputs of a BC  $P_{\tilde{X}_j Y_j | X}$ , with an abuse of notation, for all  $j \in [1 : J]$ , where  $\tilde{\mathcal{X}}_j$ ,  $\mathcal{Y}_j$ , and  $\mathcal{X}$  are finite sets.

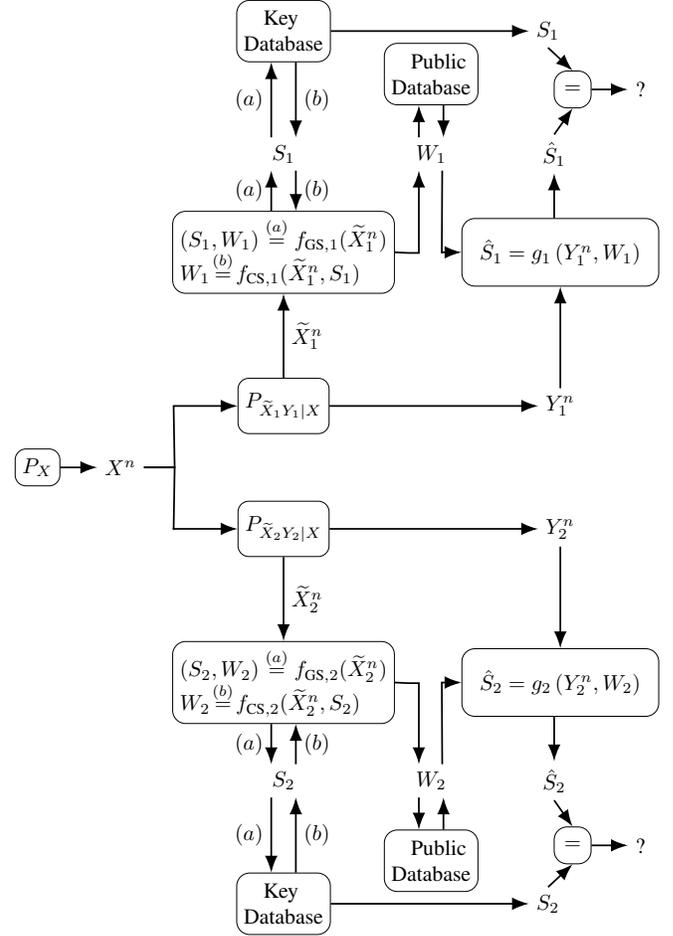


Fig. 1. Illustration of the multi-entity key agreement problem for  $J = 2$  entities with encoder and decoder measurements through BCs for (a) the GS model and (b) the CS model.

For the GS model, the  $j$ -th encoder generates helper data  $W_j$  and a secret key  $S_j$  from its observed sequence  $\tilde{X}_j^n$ . All secret keys are stored in a secure database, whereas helper data are stored in a public database so that an eavesdropper has access only to the helper data. Using the helper data  $W_j$  and its observed sequence  $Y_j^n$ , the  $j$ -th decoder generates the key estimate  $\hat{S}_j$ . Similar steps are applied for the CS model, except that each  $S_j$  should be embedded into the  $j$ -th encoder. These multi-entity models are shown in Fig. 1 for  $J=2$  entities.

Denote a set of secret keys as  $\mathcal{S}_{\mathcal{K}} = \{S_j : j \in \mathcal{K}\}$  and a set of helper data as  $\mathcal{W}_{\mathcal{K}} = \{W_j : j \in \mathcal{K}\}$  for any  $\mathcal{K} \subseteq [1 : J]$ . A (secret-key, privacy-leakage, storage), or key-leakage-storage, rate tuple is denoted as  $(R_s, R_\ell, R_w)$ . Similarly, we denote a set of secret-key rates, for any  $\mathcal{K} \subseteq [1 : J]$ , as  $\mathcal{R}_{s,\mathcal{K}} = \{R_{s,j} : j \in \mathcal{K}\}$  and a set of storage rates as  $\mathcal{R}_{w,\mathcal{K}} = \{R_{w,j} : j \in \mathcal{K}\}$ .

**Definition 1.** A key-leakage-storage rate tuple  $(\mathcal{R}_{s,[1:J]}, R_\ell, \mathcal{R}_{w,[1:J]})$  is achievable for the multi-entity GS and CS models with  $j$ -th encoder and decoder measurements

through a BC  $P_{\tilde{X}_j Y_j | X}$  if, given any  $\delta > 0$ , there is some  $n \geq 1$ , and  $J$  encoder and decoder pairs for which  $R_{s,j} = \frac{\log |\mathcal{S}_j|}{n}$  for all  $j \in [1 : J]$  and

$$\Pr \left[ \bigcup_{j \in [1:J]} \{S_j \neq \hat{S}_j\} \right] \leq \delta \quad (\text{reliability}) \quad (1)$$

$$\frac{1}{n} H(S_j) \geq R_{s,j} - \delta, \quad \forall j \in [1:J] \quad (\text{key uniformity}) \quad (2)$$

$$I(\mathcal{S}_{\mathcal{K}}; \mathcal{S}_{\mathcal{K}^c}) \leq \delta, \quad \forall \mathcal{K} \subseteq [1:J] \quad (\text{strong key ind.}) \quad (3)$$

$$\frac{1}{n} I(X^n; \mathcal{W}_{[1:J]}) \leq R_\ell + \delta \quad (\text{privacy}) \quad (4)$$

$$I(\mathcal{S}_{[1:J]}; \mathcal{W}_{[1:J]}) \leq \delta \quad (\text{strong secrecy}) \quad (5)$$

$$\frac{1}{n} \log |\mathcal{W}_j| \leq R_{w,j} + \delta, \quad \forall j \in [1:J] \quad (\text{storage}). \quad (6)$$

The *multi-entity key-leakage-storage* regions  $\mathcal{C}_{\text{gs}}$  for the GS model and  $\mathcal{C}_{\text{cs}}$  for the CS model are the closures of the set of all achievable rate tuples  $(\mathcal{R}_{s,[1:J]}, R_\ell, \mathcal{R}_{w,[1:J]})$ .

Both secret-key uniformity (2) and storage rate (6) constraints are  $J$  separate constraints. However, reliability (1), strong and mutual key independence (3), privacy-leakage rate (4), and secrecy leakage (5) constraints are joint constraints for all  $J$  honest entities. Suppose after a key generation, an honest entity has access only to its corresponding secret key and it does not have access to other entities' keys or sequences or even to the sequence it observed to generate its secret key.

The mutual key independence constraint in (3) is not imposed in the multi-enrollment key agreement problem considered in [11]. Furthermore, a normalized (weak) version of this constraint is imposed in the multi-enrollment key agreement problem considered in [9], where the  $j$ -th decoder is assumed to have access to the set of helper data  $\mathcal{W}_{[1:j]}$  for all  $j \in [1 : J]$ . The lack of the mutual key independence constraint and the assumption of availability of all previous helper data require that different encoder and decoder pairs should trust each other after key agreement. This can be the case, e.g., if all enrollments are made by the same entity. Therefore, the multi-entity key agreement problem imposes strictly more stringent constraints than the multi-enrollment key agreement problem.

The unnormalized secrecy leakage constraint (5) provides strong secrecy, a stronger notion than the weak secrecy considered in [6], [8], [9], [11]. Furthermore, (5) is more stringent than the set of individual secrecy leakage constraints  $I(S_j; \mathcal{W}_{[1:J]})$  imposed for all  $j \in [1 : J]$ , considered in [11] for symmetric PUFs in combination with the suboptimal FCS.

### III. INNER BOUNDS

We are interested in characterizing the optimal trade-off among the secret-key, privacy-leakage, and storage rates with strong secrecy for BC measurements at the encoders and decoders of any finite number  $J$  of entities that use the same hidden identifier outputs for the multi-entity key agreement

problem. We give achievable rate regions for the GS and CS models in Theorem 1. The proofs are given in [18, Section V].

Denote  $\mathcal{U}_{\mathcal{K}} = \{U_j : j \in \mathcal{K}\}$  and define a function  $\max\{\cdot, \cdot\}$  that gives the maximum of the input values as its output.

**Theorem 1.** (Inner Bounds for Multi-entity Models): *An achievable rate region  $\mathcal{R}_{\text{gs}}$  for the multi-entity GS model with  $J$  entities is the union over all  $P_{U_j | \tilde{X}_j}$  for all  $j \in [1 : J]$  of the rate tuples such that  $R_{s,j} \geq 0$  for all  $j \in [1 : J]$  and*

$$R_{s,j} \leq I(U_j; Y_j) - I(U_j; U_{[1:J] \setminus \{j\}}), \quad \forall j \in [1 : J] \quad (7)$$

$$R_\ell \geq \sum_{j=1}^J \max\{0, I(U_j; X) - I(U_j; Y_j)\}, \quad (8)$$

$$R_{w,j} \geq I(U_j; \tilde{X}_j) - I(U_j; Y_j), \quad \forall j \in [1 : J] \quad (9)$$

$$R_{s,j} + R_{w,j} \leq H(U_j | \mathcal{U}_{[1:J] \setminus \{j\}}), \quad \forall j \in [1 : J]. \quad (10)$$

An achievable rate region  $\mathcal{R}_{\text{cs}}$  for the multi-entity CS model with  $J$  entities is the union over all  $P_{U_j | \tilde{X}_j}$  for all  $j \in [1 : J]$  of the rate tuples such that  $R_{s,j} \geq 0$  for all  $j \in [1 : J]$ , (7), (8), and

$$R_{w,j} \geq I(U_j; \tilde{X}_j) - I(U_j; U_{[1:J] \setminus \{j\}}), \quad \forall j \in [1 : J] \quad (11)$$

$$R_{w,j} \leq H(U_j | \mathcal{U}_{[1:J] \setminus \{j\}}), \quad \forall j \in [1 : J]. \quad (12)$$

For the achievable rate regions  $\mathcal{R}_{\text{gs}}$  and  $\mathcal{R}_{\text{cs}}$ , we have

$$P_{\mathcal{U}_{[1:J]} | \tilde{\mathcal{X}}_{[1:J]} X \mathcal{Y}_{[1:J]}} = P_X \prod_{j=1}^J P_{U_j | \tilde{X}_j} P_{\tilde{X}_j Y_j | X}. \quad (13)$$

**Corollary 1.** *Suppose for all  $j \in [1 : J]$  that  $\tilde{X}_j - Y_j - X$  form a Markov chain, i.e.,  $X$  is a PD version of  $Y_j$  with respect to  $\tilde{X}_j$ , or  $P_{XY_j | \tilde{X}_j}$  is a LN BC with  $I(U_j; Y_j) \geq I(U_j; X)$  for all  $P_{U_j | \tilde{X}_j}$ . For the two cases, strong privacy, i.e.,*

$$R_\ell \geq 0 \quad (14)$$

can be achieved for the multi-entity GS and CS models in combination with the other bounds given in Theorem 1.

The proof of Corollary 1 follows from Theorem 1 because  $I(U_j; X) - I(U_j; Y_j) \leq 0$  for all  $j \in [1 : J]$  for BCs considered in Corollary 1. Corollary 1 illustrates that it is possible to obtain strong privacy, i.e., negligible unnormalized privacy leakage, without the requirement of a common randomness that is hidden from an eavesdropper which was assumed in [6], [19]. This is the case because the observation  $Y_j^n$  of each decoder is "better" than the observation  $\tilde{X}_j^n$  of the corresponding encoder with respect to the hidden source  $X^n$  for all entities.

**Remark 1.** The rate regions for our problem depend on the joint conditional probability distributions  $P_{XY_j | \tilde{X}_j}$  rather than only the marginal conditional distributions. Thus, the key-leakage-storage regions for the stochastically-degraded BCs are not necessarily equal to the regions for the corresponding PD BCs, unlike in the classic BC problem. Furthermore, since

$P_{\tilde{\mathcal{X}}_{[1:J]}|\mathcal{X}\mathcal{Y}_{[1:J]}}$  is fixed, the distinction between the LN BCs and essentially-less noisy BCs [20], is not necessary.

We next give simple outer bounds for the multi-entity key-leakage-storage regions  $\mathcal{C}_{\text{gs}}$  for the GS model and  $\mathcal{C}_{\text{cs}}$  for the CS model when the BCs  $P_{XY_j|\tilde{X}_j}$  for all  $j \in [1 : J]$  are PD BCs or LN BCs. These outer bounds give insights into the reason for different bounds on the secret-key rates. Based on these insights, we show a special multi-enrollment case in the next section with a less stringent secrecy constraint, for which the inner and outer bounds differ only in the Markov chains imposed and we illustrate that they match for simpler models.

**Lemma 1.** *Suppose one of the cases given in Corollary 1 is satisfied by the BCs  $P_{XY_j|\tilde{X}_j}$  for all  $j \in [1 : J]$ . An outer bound on the multi-entity key-leakage-storage region  $\mathcal{C}_{\text{gs}}$  is the union over all  $P_{U_j|\tilde{X}_j}$ , where  $U_j - \tilde{X}_j - (X, Y_j)$  form a Markov chain, for all  $j \in [1 : J]$  of the rate tuples such that  $R_{s,j} \geq 0$  for all  $j \in [1 : J]$ , (9), (14), and*

$$R_{s,j} \leq I(U_j; Y_j), \quad \forall j \in [1 : J]. \quad (15)$$

An outer bound to the multi-entity key-leakage-storage region  $\mathcal{C}_{\text{cs}}$  for the same BCs  $P_{XY_j|\tilde{X}_j}$  is the union over all  $P_{U_j|\tilde{X}_j}$ , where  $U_j - \tilde{X}_j - (X, Y_j)$  form a Markov chain, for all  $j \in [1 : J]$  of the rate tuples such that  $R_{s,j} \geq 0$  for all  $j \in [1 : J]$ , (14), (15), and

$$R_{w,j} \geq I(U_j; \tilde{X}_j), \quad \forall j \in [1 : J]. \quad (16)$$

The proof of Lemma 1 follows straightforwardly by following the steps in [8, Section VI], defining the auxiliary random variables  $U_{j,i} = (S_j, W_j, Y_j^{i-1})$  for all  $j \in [1 : J]$  and  $i \in [1 : n]$ , and by bounding  $I(X^n; \mathcal{W}_{[1:J]}) \geq 0$ ; therefore, we omit the proof.

The outer bounds do not include the inequalities in (10) and (12). Furthermore, the secret-key rate achieved by the inner bound in (7) is smaller than the outer bound given in (15), where the difference is the term  $-I(U_j; \mathcal{U}_{[1:J] \setminus \{j\}})$ . This term is a result of a constraint imposed to satisfy the strong and mutual key independence constraint given in (3). Therefore, we next consider a model without the constraint in (3) and use a secrecy-leakage constraint that is less stringent than the one in (5), i.e., replace (5) by  $I(S_j; \mathcal{W}_{[1:J]}) \leq \delta$  for all  $j \in [1 : J]$  which is also a strong secrecy metric. Due to the lack of a mutual key independence constraint, the model in the next section is not a multi-entity model but rather a multi-enrollment model. For a special case of this multi-enrollment key agreement problem, we establish inner and outer bounds for the key-leakage-storage regions that comprise the same bounds but for different Markov chains.

#### IV. BOUNDS FOR A MULTI-ENROLLMENT MODEL

Consider next the multi-enrollment model, where the strong and mutual key independence constraint (3) of the multi-entity model is not imposed. Assume further  $J = 2$  entities that

measure noisy outputs of the same hidden source  $X^n$  through separate channels that have the same channel transition matrices, i.e., for all  $j \in [1 : 2]$ ,  $\tilde{x}_j \in \tilde{\mathcal{X}}$ , and  $y_j \in \mathcal{Y}$  we have

$$P_{\tilde{X}_j Y_j | X}(\tilde{x}_j, y_j | x) = P_{\tilde{X}|X}(\tilde{x}_j | x) P_{\tilde{X}|X}(y_j | x). \quad (17)$$

This model is common for SRAM PUFs, for which each measurement channel is modeled as a BSC with the same crossover probability corresponding to a worst case scenario [21]. Using (17), we define a multi-enrollment model.

**Definition 2.** A key-leakage-storage rate tuple  $(\bar{R}_{s,1}, \bar{R}_{s,2}, \bar{R}_\ell, \bar{R}_{w,1}, \bar{R}_{w,2})$  is achievable for the multi-enrollment GS and CS models with measurements through a BC  $P_{\tilde{X}Y|X}(\tilde{x}, y | x)$  as in (17) if, given any  $\delta > 0$ , there is some  $n \geq 1$ , and two encoder and decoder pairs for which  $\bar{R}_{s,1} = \frac{\log |\mathcal{S}_1|}{n}$ ,  $\bar{R}_{s,2} = \frac{\log |\mathcal{S}_2|}{n}$ ,  $\bar{R}_{w,1} = \frac{H(W_1)}{n}$ ,  $\bar{R}_{w,2} = \frac{H(W_2)}{n}$ , and

$$\Pr \left[ \{S_1 \neq \hat{S}_1\} \cup \{S_2 \neq \hat{S}_2\} \right] \leq \delta \quad (\text{reliability}) \quad (18)$$

$$\frac{1}{n} H(S_j) = \bar{R}_{s,j} - \delta, \quad j = 1, 2 \quad (\text{key uniformity}) \quad (19)$$

$$\frac{1}{n} I(X^n; W_1, W_2) = \bar{R}_\ell + \delta \quad (\text{privacy}) \quad (20)$$

$$I(S_j; W_1, W_2) \leq \delta, \quad j = 1, 2 \quad (\text{strong secrecy}) \quad (21)$$

$$\frac{1}{n} \log |\mathcal{W}_j| = \bar{R}_{w,j} + \delta, \quad j = 1, 2 \quad (\text{storage}) \quad (22)$$

$$I(W_1; W_2) \leq \delta \quad (\text{storage ind.}) \quad (23)$$

The *multi-enrollment key-leakage-storage* regions  $\bar{\mathcal{C}}_{\text{gs}, J=2}$  for the GS model and  $\bar{\mathcal{C}}_{\text{cs}, J=2}$  for the CS model are the closures of the set of all achievable rate tuples.

We characterize in Theorem 2 inner and outer bounds for  $\bar{\mathcal{C}}_{\text{gs}, J=2}$  and  $\bar{\mathcal{C}}_{\text{cs}, J=2}$ . The proofs of Theorem 2 are given in [18, Section VI], where the reason for the necessity of the secrecy-leakage constraint in (21) that is less stringent than the joint secrecy-leakage constraint in (5) is given in Remark 2. Similarly, the reason for the necessity of the strong helper data (storage) independence constraint in (23) is discussed in Remark 4. The equalities in (19), (20), and (22) are required in the outer bounds in Theorem 2 to provide both upper and lower bounds on  $\bar{R}_\ell$  and  $\bar{R}_{w,j}$  in terms of entropy terms.

Define  $j' = 3 - j$  for  $j = 1, 2$ .

**Theorem 2.** (Inner Bounds for Multi-enrollment Models): *An achievable multi-enrollment key-leakage-storage region  $\bar{\mathcal{R}}_{\text{gs}, J=2}$  is the union over all  $P_{U_1|\tilde{X}_1}$  and  $P_{U_2|\tilde{X}_2}$  of the rate tuples such that  $\bar{R}_{s,j} \geq 0$  for  $j = 1, 2$  and*

$$\bar{R}_{s,j} \leq I(U_j; Y_j), \quad j = 1, 2 \quad (24)$$

$$\bar{R}_\ell \geq \sum_{j=1}^2 (I(U_j; X) - I(U_j; Y_j)), \quad (25)$$

$$\bar{R}_\ell \leq \sum_{j=1}^2 (I(U_j; X) - I(U_j; \tilde{X}_j) + \bar{R}_{w,j}), \quad (26)$$

$$\bar{R}_{w,j} \geq I(U_j; \tilde{X}_j) - I(U_j; Y_j), \quad j = 1, 2 \quad (27)$$

$$\bar{R}_{s,j} + \bar{R}_{w,j} \leq H(U_j), \quad j = 1, 2 \quad (28)$$

$$\bar{R}_{s,j} + \bar{R}_{w,j} + \bar{R}_{w,j'} \leq H(U_j, U_{j'}), \quad j = 1, 2. \quad (29)$$

An achievable multi-enrollment key-leakage-storage region  $\bar{\mathcal{R}}_{cs, J=2}$  is the union over all  $P_{U_1|\tilde{X}_1}$  and  $P_{U_2|\tilde{X}_2}$  of the rate tuples such that  $\bar{R}_{s,j} \geq 0$  for  $j = 1, 2$ , (24)-(26), and

$$\bar{R}_{w,j} \geq I(U_j; \tilde{X}_j), \quad j = 1, 2 \quad (30)$$

$$\bar{R}_{w,j} \leq H(U_j), \quad j = 1, 2 \quad (31)$$

$$\bar{R}_{w,j} + \bar{R}_{w,j'} \leq H(U_j, U_{j'}) + \bar{R}_{s,j'}, \quad j = 1, 2. \quad (32)$$

For both achievable regions  $\bar{\mathcal{R}}_{gs, J=2}$  and  $\bar{\mathcal{R}}_{cs, J=2}$ , we have

$$\begin{aligned} & P_{U_1 U_2 \tilde{X}_1 \tilde{X}_2 X Y_1 Y_2}(u_1, u_2, \tilde{x}_1, \tilde{x}_2, x, y_1, y_2) \\ &= P_{U_1|\tilde{X}_1}(u_1|\tilde{x}_1) P_{U_2|\tilde{X}_2}(u_2|\tilde{x}_2) P_{\tilde{X}_1|X}(\tilde{x}_1|x) P_{\tilde{X}_2|X}(\tilde{x}_2|x) \\ & \quad \times P_{\tilde{X}_1|X}(y_1|x) P_{\tilde{X}_2|X}(y_2|x) P_X(x). \end{aligned} \quad (33)$$

(Outer Bounds for Multi-enrollment GS and CS Models): An outer bound for  $\bar{\mathcal{C}}_{gs, J=2}$  is the union over all  $P_{U_1|\tilde{X}_1}$  and  $P_{U_2|\tilde{X}_2}$  of the rate tuples such that  $\bar{R}_{s,j} \geq 0$ , (24) - (29), and  $U_j - \tilde{X}_j - X - Y_j$  form a Markov chain for  $j = 1, 2$ . An outer bound for  $\bar{\mathcal{C}}_{cs, J=2}$  is the union over all  $P_{U_1|\tilde{X}_1}$  and  $P_{U_2|\tilde{X}_2}$  of the rate tuples such that  $\bar{R}_{s,j} \geq 0$ , (24) - (26), (30) - (32), and  $U_j - \tilde{X}_j - X - Y_j$  form a Markov chain for  $j = 1, 2$ .

The inner and outer bounds differ because the outer bounds define rate regions for the Markov chains  $U_1 - \tilde{X}_1 - X - Y_1$  and  $U_2 - \tilde{X}_2 - X - Y_2$ , which are larger than the rate regions defined by the inner bounds that satisfy (33). For instance, in the achievability proof of Theorem 2, we apply the properties of the Markov chain  $U_2 - \tilde{X}_2 - U_1$ , which does not form a Markov chain for the choice of  $U_1$  and  $U_2$  in the outer bounds. Therefore, inner and outer bounds do not match in general.

**Corollary 2.** Choosing  $U_1 = \tilde{X}_1$  and  $U_2 = \tilde{X}_2$ , it is straightforward to show that inner and outer bounds in Theorem 2 match if we do not impose any storage or privacy constraints, i.e., impose only (18), (19), and (21). This result improves on the secret-key capacity region given in [13, Theorem 1] for a weak secrecy constraint.

**Example 1.** Consider the RO PUF model from [22, Section 4.1] where a transform-coding method is applied to conservatively model the measurement channels  $P_{Y_1|X} = P_{\tilde{X}_1|X}$  as independent BSCs with the same crossover probability of  $p_A$  and where the hidden source output is  $\text{Bern}(\frac{1}{2})$ . We; therefore, can apply the achievability results from Theorem 2 to this RO PUF model. Using [8, Theorem 3] to evaluate the boundary tuples of  $\bar{\mathcal{R}}_{gs, J=2}$ , it suffices to consider probability distributions  $P_{U_j|\tilde{X}_j}$  for  $j = 1, 2$  such that  $P_{\tilde{X}_j|U_j}$  are BSCs. Consider the projection of the boundary tuples of  $\bar{\mathcal{R}}_{gs, J=2}$  onto key-leakage plane, i.e., (24) and (25). We

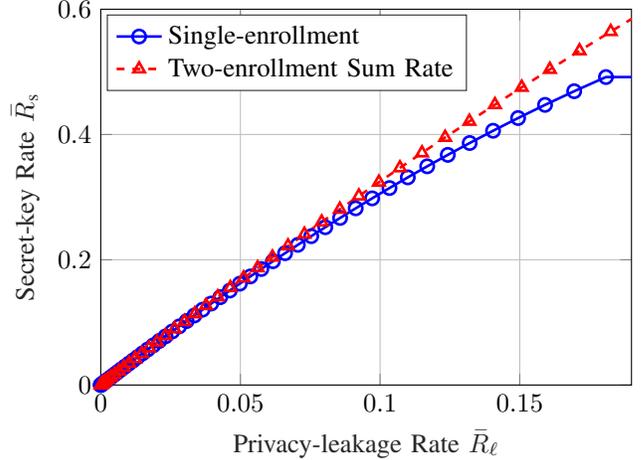


Fig. 2. Privacy-leakage vs. secret-key rate projection of the boundary tuples of the single- and two-enrollment RO PUF models with BSCs ( $p_A = 0.06$ ).

plot in Fig. 2 single-enrollment results where the privacy-leakage rate is measured with respect to single helper data, and two-enrollment results for the sum rate of the two keys, both for  $p_A = 0.06$  [22]. To achieve a total secret-key rate of  $I(\tilde{X}_1; Y_1) = I(\tilde{X}_2; Y_2)$ , the privacy-leakage rate for the two-enrollment model is approximately 13.5% less than the rate for the single-enrollment model. This gain follows from the information bottleneck problem that arises from (24) and (25) to find the boundary tuples.

**Example 2.** Consider uniform binary antipodal measurements over an additive white Gaussian noise (AWGN) channel. Define the signal power as  $P_s$  and the noise power as  $P_n$ , so we have a signal-to-noise ratio (SNR) of  $\text{SNR} = P_s/P_n$ . If a matched filter, which maximizes the SNR at the sampling instant for the AWGN channel, is applied at the encoder and decoder, the bit error probability  $P_b$  is given by

$$P_b = Q\left(\sqrt{\text{SNR}}\right). \quad (34)$$

The channel between input binary symbols and outputs of the matched filter is a BISO channel. Using [8, Theorem 3], we have that  $P_{\tilde{X}_j|U_j}$  for  $j = 1, 2$  that are BSCs suffice to obtain the boundary tuples of  $\bar{\mathcal{R}}_{gs, J=2}$ . The value  $p_A = 0.06$  used in Example 1 corresponds to an SNR of approximately 3.83dB.

In Fig. 3, the privacy-leakage rate vs. secret-key rate boundary tuples are depicted for two cases. First, a two-enrollment model at  $\text{SNR} = 3.83\text{dB}$  with a sum rate for two secret keys is depicted, where each enrollment has a signal power of  $P_s$ . We plot also a single-enrollment model with the signal power of  $2P_s$ , i.e., we have an SNR of approximately 6.84dB. Fig. 3 shows for the two cases with the same total signal power of  $2P_s$ , unlike in Example 1, that the single enrollment boundary tuple can result in a gain of approximately 228.55% at its top left corner point in terms of the secret-key rate achieved for a

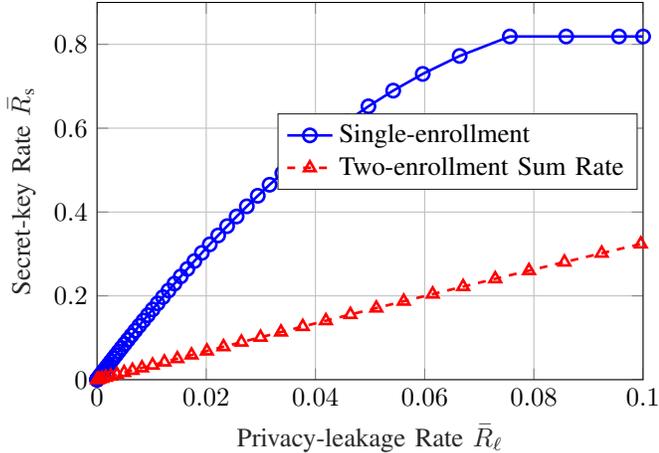


Fig. 3. Privacy-leakage vs. secret-key rate projection of the boundary tuples of the single- and two-enrollment RO PUF models with different SNRs.

given privacy-leakage rate. For such channels with a fixed total signal power; thus, the single-enrollment model can result in significant gains in terms of the achieved secret-key rate as compared to the two-enrollment model for small  $\bar{R}_\ell$  values.

## V. CONCLUSION

We derived inner bounds for the multi-entity key-leakage-storage regions for GS and CS models with strong secrecy, a hidden identifier source, and correlated noise components at the encoder and decoder measurements that are modeled as BCs. The inner bounds are valid for any finite number of entities that use the same hidden source to agree on a secret key. We argued that the mutual key independence constraint we impose makes the proposed multi-entity key agreement problem a proper multi-user extension of the classic single-enrollment key agreement problem, unlike the multi-enrollment key agreement problem considered in the literature. A set of degraded and less-noisy BCs was shown to provide strong privacy without a need for a common randomness. We also established inner and outer bounds for the key-leakage-storage regions for a two-enrollment model with measurement channels that are valid for SRAM and RO PUFs. Inner and outer bounds were shown to differ only in the Markov chains imposed and they match if the storage and privacy-leakage rate constraints are removed. Two examples illustrated that depending on the constraints of the practical scenario, a single or multiple enrollments might perform better in terms of the secret-key vs. privacy-leakage rate ratio. In future work, we will find a set of symmetric probability distributions for which the strong helper data independence constraint in the two-enrollment model can be eliminated.

## ACKNOWLEDGMENT

O. Günlü was supported by the German Federal Ministry of Education and Research (BMBF) within the national initiative

for “Post-Shannon Communication (NewCom)” under the Grant 16KIS1004.

## REFERENCES

- [1] O. Günlü, “Multi-entity and multi-enrollment key agreement with correlated noise,” *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1190–1202, 2021.
- [2] B. Gassend, “Physical random functions,” Master’s thesis, M.I.T., Cambridge, MA, Jan. 2003.
- [3] O. Günlü, “Key agreement with physical unclonable functions and biometric identifiers,” Ph.D. dissertation, TU Munich, Germany, Nov. 2018, published by Dr. Hut Verlag in Feb. 2019.
- [4] 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.
- [5] U. M. Maurer, “Secret key agreement by public discussion from common information,” *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.
- [6] 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.
- [7] I. Csiszár and P. Narayan, “Common randomness and secret key generation with a helper,” *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 344–366, Mar. 2000.
- [8] O. Günlü and G. Kramer, “Privacy, secrecy, and storage with multiple noisy measurements of identifiers,” *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 11, pp. 2872–2883, Nov. 2018.
- [9] L. Kusters and F. M. J. Willems, “Secret-key capacity regions for multiple enrollments with an SRAM-PUF,” *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 9, pp. 2276–2287, Sep. 2019.
- [10] A. Juels and M. Wattenberg, “A fuzzy commitment scheme,” in *ACM Conf. Comp. Commun. Security*, New York, NY, Nov. 1999, pp. 28–36.
- [11] L. Kusters, T. Ignatenko, F. M. J. Willems, R. Maes, E. van der Sluis, and G. Selimis, “Security of helper data schemes for SRAM-PUF in multiple enrollment scenarios,” in *IEEE Int. Symp. Inf. Theory*, Aachen, Germany, June 2017, pp. 1803–1807.
- [12] I. Land, S. Huettinger, P. A. Hoeher, and J. B. Huber, “Bounds on information combining,” *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 612–619, Feb. 2005.
- [13] L. Kusters, O. Günlü, and F. M. Willems, “Zero secrecy leakage for multiple enrollments of physical unclonable functions,” in *Symp. Inf. Theory Sign. Process. Benelux*, Twente, The Netherlands, May–June 2018, pp. 119–127.
- [14] O. Günlü, O. İscan, 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.
- [15] O. Günlü, “Design and analysis of discrete cosine transform based ring oscillator physical unclonable functions,” Master’s thesis, Techn. Univ. München, Munich, Germany, Oct. 2013.
- [16] O. Günlü, R. F. Schaefer, and G. Kramer, “Private authentication with physical identifiers through broadcast channel measurements,” in *IEEE Inf. Theory Workshop*, Visby, Sweden, Aug. 2019, pp. 1–5.
- [17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: John Wiley & Sons, 2012.
- [18] O. Günlü, “Multi-entity and multi-enrollment key agreement with correlated noise,” Oct. 2020, [Online]. Available: [arxiv.org/pdf/2005.08210.pdf](https://arxiv.org/pdf/2005.08210.pdf).
- [19] 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.
- [20] C. Nair, “Capacity regions of two new classes of two-receiver broadcast channels,” *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4207–4214, Sep. 2010.
- [21] R. Maes, P. Tuyls, and I. Verbauwhede, “A soft decision helper data algorithm for SRAM PUFs,” in *IEEE Int. Symp. Inf. Theory*, Seoul, South Korea, June 2009, pp. 2101–2105.
- [22] O. Günlü, T. Kernetzky, O. İscan, V. Sidorenko, G. Kramer, and R. F. Schaefer, “Secure and reliable key agreement with physical unclonable functions,” *Entropy*, vol. 20, no. 5, May 2018.