Common Randomness Amplification: A Constructive View

Grégory Demay Ueli Maurer
Department of Computer Science,
ETH Zürich, Switzerland
Email: {demayg,maurer}@inf.ethz.ch

***Abstract*—Common randomness is an important resource in many areas such as game theory and cryptography. We discuss the general problem of common randomness amplification between two distrustful parties connected by a communication channel and sharing some initial randomness. In this setting, both parties wish to agree on a common value distributed according to a target distribution by using their initial amount of common randomness and exchanging messages. Our results show that no protocol which is secure in a composable sense can significantly amplify the entropy initially shared by the parties.**

I. Introduction

A. Randomness as a Resource

Playing any probabilistic game over the Internet requires some randomness shared among the players. This common randomness is necessary to emulate what could have had happened if the players were physically present: shuffle cards, throw a dice, etc. The common randomness used does not need to be secret, but it is crucial for fairness that it cannot be influenced by any party.

We see common randomness as a *resource*, which we model as a system with an interface to every party in consideration. Ideal resources (such as common randomness) are generally not available as a system in the physical world, but they must be constructed by cryptographic means using available resources (such as a communication channel). Constructive cryptography introduced in [1] (see also [2]), is a new approach to cryptography in which cryptographic protocols are seen as constructions of resources from other resources. The definition of the term *construction* depends on the setting, i.e., on who can potentially be dishonest and whether the security should be information-theoretic or computational. A composition theorem of constructive cryptography guarantees that constructive steps compose.

In this paper, we focus on the general problem of *common randomness amplification* between two distrustful parties in the information-theoretic case. That is, two distrustful parties having access to an ideal communication channel and knowing an initial common random value, wish to agree on a new common random value which has higher entropy than what they initially shared.

B. Related Work

Blum in [3] gave the first coin tossing protocol under the assumption that one-way functions exist. Under the same assumption, Lindell presented in [4] a constant-round protocol for securely tossing polynomially many coins in parallel. The security of both protocols is stated in a stand-alone security model (the so-called "malicious model" described in [5]) where only sequential composition is guaranteed. Maintaining security under parallel composition requires a stricter notion of security, such as in the Universal Composability (UC) [6] or in the Abstract Cryptography [2] framework.

In [7], the authors studied the problem of extending given coin tosses. For a composable security notion, they showed that the possibility of such a task in the computational case depends on the amount of initial given coin tosses. In the information-theoretic case, they showed that there was no efficient protocol which could extend a coin toss and be secure in a composable sense.

C. Contributions and Outline

We introduce and study the general problem of common randomness amplification for two distrustful parties in the information-theoretic case. Our results use the concept of secure construction defined in [1], [2], which will be briefly restated in Section II-B. We define the resources used throughout the paper in Section II-C. The security definition for amplifying common randomness is stated in Section III. In Section III-A, we show that if one wishes to securely construct a source of common randomness, then communication has limited utility. This observation will lay the ground for our impossibility result in Section III-B.

II. Preliminaries

A. Notation

We denote sets by calligraphic letters or capital greek letters (e.g., $\mathcal{X}$, Σ). For any integer $n \in \mathbb{N}$, let $[n] := \{1, \ldots, n\}$. Throughout this paper, we consider only discrete random variables. A discrete random variable will be denoted by an upper-case letter X, its range by the corresponding calligraphic letter $\mathcal{X}$, and a realization of the random variable X will be denoted by the corresponding lower-case letter x. A tuple of n random variables $(X_1, \ldots, X_n)$ will be denoted by X^n. The probability distribution of a random variable X will be denoted as P_X. For two probability distributions P_X and Q_X, their *statistical distance* is denoted by $\mathsf{d}(\mathsf{P}_X, \mathsf{Q}_X)$ and is defined

978-1-4673-0223-4/12/$31.00 ©2012 IEEE

as follows

$$\mathsf{d}\left(\mathsf{P}_X, \mathsf{Q}_X\right) := \frac{1}{2} \sum_{x \in \mathcal{X}} \left| \mathsf{P}_X\left(x\right) - \mathsf{Q}_X\left(x\right) \right|.$$

Auxiliary lemmas about the statistical distance are postponed to the Appendix.

We will use standard information-theoretic notations such as in [8]. A superscript will be added when the probability distribution needs to be made explicit, for example $I^{\mathsf{R}}(Y; Z \mid X)$ will denote the mutual information between Y and Z given X, where the random variables considered are distributed according to the joint distribution R_{XYZ}. The binary entropy function will be denoted by $h(\cdot)$, where

$$h(x) := -x \log_2 x - (1-x) \log_2 (1-x), \ \forall x \in [0,1].$$

The data-processing inequality, as well as Fano's inequality will be used, and we refer to [8, Th 2.8.1, Th 2.10.1] for their statements. As a shorthand, we define the following function $F : [0,1] \times \mathbb{N} \setminus \{0\} \to [0, \infty]; \ F(\varepsilon, c) := h(\varepsilon) + \varepsilon \log_2 c$. Observe that $F(\cdot, c)$ is a non-decreasing function over $[0, \frac{1}{2}]$, for any $c \in \mathbb{N} \setminus \{0\}$.

B. Constructive Cryptography

We use the concept of abstract systems [2], [1] to formulate our results, and partly follow the concise exposition of [9]. At the highest level of abstraction, a system is an object with interfaces via which it interacts with other systems. Every two systems can be composed by connecting one interface of each system, and the resulting object is again a system. Also, we assume that every two different systems are mutually independent.

We consider three distinct types of systems: *resources*, *converters*, and *distinguishers*. Resources are denoted by upper-case boldface letters such as $\mathbf{R}$ and $\mathbf{S}$. In this paper, we always consider resources with two interfaces, the left interface will be referred to as Alice's, while the right interface will be referred to as Bob's. In our scenario, either Alice or Bob could be dishonest, and the case where both are dishonest does not need to be considered. Resources $\mathbf{R}$ and $\mathbf{S}$ could also be used in parallel, and the resulting resource, denoted $\mathbf{R} \| \mathbf{S}$, is again a 2-interface resource, where each of the interfaces of $\mathbf{R} \| \mathbf{S}$ allows access to the corresponding interface of both subsystems $\mathbf{R}$ and $\mathbf{S}$.

Converters are systems having one *inside* and one *outside* interface, and are denoted by lower-case greek letters, such as α, β, σ. The set of all converters will be denoted by Σ. A converter α can be attached to a resource $\mathbf{R}$ by connecting the inside interface of α to one of the two interfaces of $\mathbf{R}$. For example, if α is attached to the left interface of $\mathbf{R}$, then the resulting resource is denoted by $\alpha\mathbf{R}$, and is again a 2-interface system whose left interface is now the outside interface of α. Similarly, attaching the converter β to the right interface of $\mathbf{R}$ is denoted by $\mathbf{R}\beta$.

A distinguisher $\mathbf{D}$ is a system that connects to all interfaces of a resource $\mathbf{R}$ and outputs at a separate interface a single bit denoted B. The complete interaction between $\mathbf{D}$ and $\mathbf{R}$ defines a random experiment, and the probability that $\mathbf{D}$ outputs 1 in this random experiment is denoted by $\mathsf{P}^{\mathbf{DR}}(B = 1)$. The *distinguishing advantage* of $\mathbf{D}$ in distinguishing the system $\mathbf{R}$ from $\mathbf{S}$ is defined as $\Delta^{\mathbf{D}}(\mathbf{R}, \mathbf{S}) := \left| \mathsf{P}^{\mathbf{DR}}(B=1) - \mathsf{P}^{\mathbf{DS}}(B=1) \right|$. The set of all distinguishers is denoted by $\mathcal{D}$, and we define $\Delta^{\mathcal{D}}(\mathbf{R}, \mathbf{S}) := \sup_{\mathbf{D} \in \mathcal{D}} \Delta^{\mathbf{D}}(\mathbf{R}, \mathbf{S})$. Note that $\Delta^{\mathcal{D}}$ defines a pseudo-metric, and for convenience we will use the following notation, $\mathbf{R} \approx_\varepsilon \mathbf{S} :\Leftrightarrow \Delta^{\mathcal{D}}(\mathbf{R}, \mathbf{S}) \leq \varepsilon$, where $\varepsilon \in [0,1]$.

A protocol, in our case a pair of converters, is used to construct a specific ideal resource from available real resources, where the meaning of "construct" is now made precise.

Definition 1. A two-party protocol $\pi = (\alpha, \beta) \in \Sigma^2$, where only one party could be dishonest, securely constructs a resource $\mathbf{S}$ from a resource $\mathbf{R}$ within ε, denoted $\mathbf{R} \xrightarrow{(\pi, \varepsilon)} \mathbf{S}$, if and only if

$$\alpha \mathbf{R} \beta \approx_\varepsilon \mathbf{S}, \tag{1}$$

$$\exists \sigma \in \Sigma : \ \alpha \mathbf{R} \approx_\varepsilon \mathbf{S} \sigma, \tag{2}$$

$$\exists \tau \in \Sigma : \ \mathbf{R} \beta \approx_\varepsilon \tau \mathbf{S}. \tag{3}$$

Note that as a specific instantiation of abstract cryptography, [2, Th. 2] ensures us that our security definition is *generally composable*. That is, a protocol which is secure according to Definition 1 will remain secure under arbitrary sequential or parallel composition.

C. Resources Considered

We will consider two different resources of randomness: a symmetric source of randomness depicted in Figure 1, and a correlated source of randomness depicted in Figure 2.

Definition 2. A symmetric source of randomness for the distribution P_X, denoted by $[\mathsf{P}_X]_2$, is a 2-interface resource which outputs at both left and right interfaces the same random variable X distributed according to P_X.

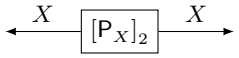


Figure 1: The symmetric source of randomness $[\mathsf{P}_X]_2$.

Definition 3. A correlated source of randomness for the joint distribution P_{YZ}, denoted by $[\mathsf{P}_{YZ}]$, is a 2-interface resource which outputs Y at the left interface and Z at the right interface, where the pair (Y, Z) is distributed according to P_{YZ}.

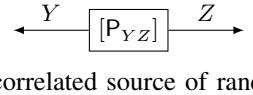


Figure 2: The correlated source of randomness $[\mathsf{P}_{YZ}]$.

Note that the resources defined above are *different* from usual common randomness functionalities found in the literature, such as the coin-tossing functionality defined in [7], or

the common random string functionality defined in [6]. Simply put, these functionalities add more constraints for the parties as to when or how they should receive their output from the resource.

The following lemma gives a trivial lower bound on the distinguishing advantage between a symmetric and a correlated source of randomness.

Lemma 1. *For any distributions* P_{YZ} *and* P_W*,*

$$\Delta^{\mathcal{D}}\left([\mathsf{P}_{YZ}], [\mathsf{P}_W]_2\right) \geq \mathsf{P}\left(Y \neq Z\right).$$

Proof: Consider the distinguisher $\mathbf{D}$ which outputs 1 if and only if the two random variables received are not equal. Then,

$$\Delta^{\mathcal{D}}\left([\mathsf{P}_{YZ}], [\mathsf{P}_W]_2\right) \geq \Delta^{\mathbf{D}}\left([\mathsf{P}_{YZ}], [\mathsf{P}_W]_2\right) = \mathsf{P}\left(Y \neq Z\right).$$

■

We will also make use of a perfect communication channel, which simply forwards messages.

Definition 4. A bi-directional communication channel, denoted $\leftrightarrow$, is a 2-interface resource which forwards every input at the left (respectively, right) interface to the right (respectively, left) interface.

III. Common Randomness Amplification

We are interested in the 2-party problem, called *common randomness amplification*, consisting of agreeing on a common random W distributed according to P_W, given solely an ideal communication channel and some common random X distributed according to P_X. The term amplification refers to the fact that we require the Shannon entropy $H(W)$ to be greater than $H(X)$, the entropy initially shared among both parties.

Definition 5. A two-party protocol $\pi = (\alpha, \beta) \in \Sigma^2$ is said to securely amplify common randomness within ε if

$$\left(\leftrightarrow \parallel [\mathsf{P}_X]_2\right) \xrightarrow{(\pi,\varepsilon)} [\mathsf{P}_W]_2 \text{ and } H(W) > H(X).$$

A. Constructing a Symmetric Source of Randomness

For the remainder of the paper, let $\pi = (\alpha, \beta) \in \Sigma^2$ be a protocol such that $\left(\leftrightarrow \parallel [\mathsf{P}_X]_2\right) \xrightarrow{(\pi,\varepsilon)} [\mathsf{P}_W]_2$ for some $\varepsilon \in [0, \frac{1}{2}]$. Definition 1 implies that

$$\alpha\left(\leftrightarrow \parallel [\mathsf{P}_X]_2\right)\beta \approx_\varepsilon [\mathsf{P}_W]_2, \tag{4}$$

$$\exists \sigma \in \Sigma: \ \alpha\left(\leftrightarrow \parallel [\mathsf{P}_X]_2\right) \approx_\varepsilon [\mathsf{P}_W]_2\,\sigma, \tag{5}$$

$$\exists \tau \in \Sigma: \ \left(\leftrightarrow \parallel [\mathsf{P}_X]_2\right)\beta \approx_\varepsilon \tau\,[\mathsf{P}_W]_2. \tag{6}$$

The system $\alpha\left(\leftrightarrow \parallel [\mathsf{P}_X]_2\right)\beta$ is depicted in Figure 3. Without loss of generality, we will assume the protocol (α, β) to send an even number of messages M^n, $M_i \in \mathcal{M}$, where messages with odd indices are sent by Alice, and messages with even indices are sent by Bob. The output of the protocol (α, β) will be denoted by (Y, Z), where both random variables are assumed to be over $\mathcal{W}$. The joint distribution defined by this random experiment is denoted by R_{XM^nYZ}.

We now state some implications from the security definition.

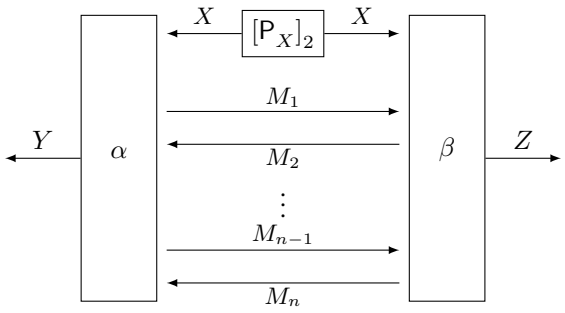


Figure 3: The protocol (α, β) receives the same random X and outputs (Y, Z) after having exchanged n messages M^n. The random variables involved in this random experiment are distributed according to R_{XM^nYZ}.

a) Implications of the correctness condition (4)*:* Note that (4) and Lemma 1 imply the probability that Y is different from Z in the random experiment defined by R_{XM^nYZ} is upper-bounded by ε , which with Fano's inequality gives $H^{\mathsf{R}}(Y \mid Z) \leq F(\varepsilon, |\mathcal{W}|)$. Moreover, the distribution R_{XM^nYZ} is such that Y and Z are conditionally independent given (X, M^n). Applying the data-processing inequality gives $I^{\mathsf{R}}(Y; XM^n) \geq I^{\mathsf{R}}(Y; Z)$, and thus

$$H^{\mathsf{R}}(Y \mid XM^n) \leq H^{\mathsf{R}}(Y \mid Z) \leq F(\varepsilon, |\mathcal{W}|). \tag{7}$$

b) Implications of the simulatability condition (5)*:* We now consider a specific distinguisher trying to distinguish $\alpha\left(\leftrightarrow \parallel [\mathsf{P}_X]_2\right)$ from $[\mathsf{P}_W]_2\,\sigma$ as follows. It emulates internally[1] β at the right interface of the connected system. Such a distinguisher is shown in Figure 4. When connected to $\alpha\left(\leftrightarrow \parallel [\mathsf{P}_X]_2\right)$ (see Figure 4a), such a distinguisher would see the random variables X, M^n, Y, Z distributed according to R_{XM^nYZ}; whereas when connected to $[\mathsf{P}_W]_2\,\sigma$ (see Figure 4b) such a distinguisher would see the random variables X, M^n, W, Z distributed according to some different joint distribution, say S_{XM^nWZ}. The maximum distinguishing advantage of such a distinguisher is $\mathsf{d}(\mathsf{R}_{XM^nYZ}, \mathsf{S}_{XM^nWZ})$, and (5) ensures that

$$\mathsf{d}(\mathsf{R}_{XM^nYZ}, \mathsf{S}_{XM^nWZ}) \leq \varepsilon. \tag{8}$$

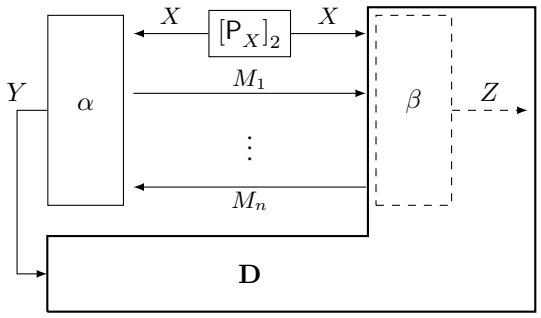
Note that in the random experiment defined by the joint distribution S_{XM^nWZ}, messages with even indices are actually sent by the (emulated) protocol β, and the random variable Z is also computed and output by β. Thus,

$$\begin{aligned} &\mathsf{S}_{M_i|XM^{i-1}W} = \mathsf{R}_{M_i|XM^{i-1}}, \text{ for even } i \in [n]; \\ &\mathsf{S}_{Z|XM^nW} = \mathsf{R}_{Z|XM^n}. \end{aligned} \tag{9}$$

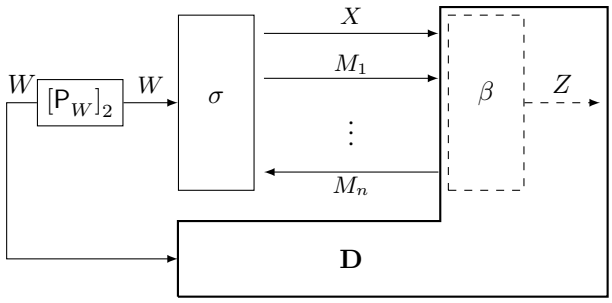
Consequently the distribution S_{XM^nWZ} is such that W and M_i are conditionally independent given (X, M^{i-1}), for an even i, or in other words,

$$I^{\mathsf{S}}(W; M_i \mid XM^{i-1}) = 0, \text{ for even } i \in [n]. \tag{10}$$

[1]Formally this assumes some closure properties on the distinguisher class $\mathcal{D}$ stated in [2, Def. 16].



(a) Real World. The distinguisher $\mathbf{D}$ emulates internally the protocol β at the right interface of $\alpha\left(\leftrightarrow \| [\mathsf{P}_X]_2\right)$. The distinguisher $\mathbf{D}$ sees the random variables X, M^n, Y, Z distributed according to R_{XM^nYZ}.



(b) Ideal World. The distinguisher $\mathbf{D}$ emulates internally the protocol β at the right interface of $[\mathsf{P}_W]_2\,\sigma$. The distinguisher $\mathbf{D}$ sees the random variables X, M^n, W, Z distributed according to S_{XM^nWZ}.

Figure 4: A distinguisher could emulate internally the protocol β in order to distinguish $\alpha\left(\leftrightarrow \| [\mathsf{P}_X]_2\right)$ from $[\mathsf{P}_W]_2\,\sigma$.

c) Implications of the simulatability condition (6)*:* Similarly, we consider a specific distinguisher which emulates internally the protocol α at the left interface of the connected system in order to distinguish $\left(\leftrightarrow \| [\mathsf{P}_X]_2\right)\beta$ from $\tau\,[\mathsf{P}_W]_2$. This setting is completely symmetric to the one in the previous paragraph, except when connected to $\tau\,[\mathsf{P}_W]_2$ such a distinguisher would see the random variables X, M^n, Y, W distributed according to some joint distribution denoted by T_{XM^nYW}. Thus,

$$\mathsf{d}(\mathsf{T}_{XM^nYW}, \mathsf{R}_{XM^nYZ}) \leq \varepsilon, \tag{11}$$

$$\begin{aligned} &\mathsf{T}_{M_i|XM^{i-1}W} = \mathsf{R}_{M_i|XM^{i-1}}, \text{ for odd } i \in [n]\,; \\ &\mathsf{T}_{Y|XM^nW} = \mathsf{R}_{Y|XM^n}, \end{aligned} \tag{12}$$

$$I^{\mathsf{T}}(W; M_i \mid XM^{i-1}) = 0, \text{ for odd } i \in [n]\,. \tag{13}$$

B. Impossibility of Common Randomness Amplification

We now prove that there is no information-theoretically secure protocol which amplifies common randomness. The main idea is that any secure construction of the type $\left(\leftrightarrow \| [\mathsf{P}_X]_2\right) \xrightarrow{(\pi,\varepsilon)} [\mathsf{P}_W]_2$ requires $H(W)$ to be bounded by $H(X)$ (plus some function of ε). For the sake of clarity, we first proceed to the perfect case ($\varepsilon = 0$) in Theorem 1 and then deal with the general case in Theorem 2.

Theorem 1. *There is no protocol which perfectly amplifies common randomness, i.e.,*

$$(\leftrightarrow \| [\mathsf{P}_X]_2) \xrightarrow{(\pi,0)} [\mathsf{P}_W]_2 \implies H(W) \leq H(X)\,.$$

Proof: Equations (8) and (11) imply for $\varepsilon = 0$ that

$$\mathsf{R}_{XM^nYZ} = \mathsf{S}_{XM^nWZ} = \mathsf{T}_{XM^nYW}.$$

Thus, any information-theoretic measure involving these distributions are the same, and (7) implies $H^{\mathsf{R}}(Y \mid XM^n) = 0$ for $\varepsilon = 0$. Since $H(W) = H^{\mathsf{R}}(Y)$, we have

$$\begin{aligned} H(W) &\leq H^{\mathsf{R}}(X) + H^{\mathsf{R}}(Y \mid X) \\ &= H(X) + I^{\mathsf{R}}(Y; M^n \mid X) \\ &= H(X) + \sum_{i\in[n]} I^{\mathsf{S}}(W; M_i \mid XM^{i-1}). \end{aligned}$$

The proof is finished upon noticing that (10) and (13), together with the fact that $\mathsf{S}_{XM^nWZ} = \mathsf{T}_{XM^nYW}$, imply

$$I^{\mathsf{S}}(W; M_i \mid XM^{i-1}) = 0, \text{ for all } i \in [n]\,.$$

■

Theorem 2. *There is no statistically secure protocol with n messages which significantly amplifies common randomness, i.e., for any $\varepsilon \in \left[0, \frac{1}{4(n+1)}\right]$,*

$$(\leftrightarrow \| [\mathsf{P}_X]_2) \xrightarrow{(\pi,\varepsilon)} [\mathsf{P}_W]_2 \implies H(W) \leq H(X) + f(\varepsilon),$$

where $f(\varepsilon) := 3(h(2\varepsilon) + 2\varepsilon \log_2 |\mathcal{W}|) + h(2(n+1)\varepsilon) + 2(n+1)\varepsilon \log_2 |\mathcal{W}|$.

Proof: Consider a mixture M of distributions between T and S defined as follows

$$\mathsf{M}_{XM^nYW} := \mathsf{T}_{XW} \cdot \prod_{\substack{i\in[n],\\ i \text{ odd}}} \mathsf{T}_{M_i|XM^{i-1}W} \cdot \prod_{\substack{i\in[n],\\ i \text{ even}}} \mathsf{S}_{M_i|XM^{i-1}W} \cdot \mathsf{T}_{Y|XM^nW}.$$

In the following, $\mathsf{M}(Y \neq W)$ denotes the probability that Y is different from W in the random experiment defined by the joint distribution M_{XM^nYW}.

d) Upper-bound of $\mathsf{M}(Y \neq W)$*:* Note that (4), (5), and (6) imply that

$$\alpha\tau\,[\mathsf{P}_W]_2 \approx_\varepsilon \alpha\left(\leftrightarrow \| [\mathsf{P}_X]_2\right)\beta \approx_\varepsilon [\mathsf{P}_W]_2\,,$$

and, consequently, $\mathsf{d}(\mathsf{T}_{YW}, \mathsf{P}_{WW}) \leq 2\varepsilon$.

The triangle inequality, (8), and (11) imply that $\mathsf{d}(\mathsf{T}_{XM^nYW}, \mathsf{S}_{XM^nWZ}) \leq 2\varepsilon$. We now show that the distribution M_{XM^nYW} is statistically close to T_{XM^nYW}, as follows. We can see the distribution M_{XM^nYW} as a mixture between the distributions T_{XM^nYW} and S_{XM^nWZ} involving $n+1$ random variables $(X, W, M_1), M_2, \ldots, M_n, Y$. Then applying Lemma A.3 and using the fact that n is assumed to be even, we have

$$\begin{aligned} \mathsf{d}(\mathsf{T}_{XM^nYW}, \mathsf{M}_{XM^nYW}) &\leq n \cdot \mathsf{d}(\mathsf{T}_{XM^nYW}, \mathsf{S}_{XM^nWZ}) \\ &\leq 2n\varepsilon. \end{aligned}$$

Thus, the last equation and $\mathsf{d}\left(\mathsf{T}_{YW}, \mathsf{P}_{WW}\right) \leq 2\varepsilon$ give us

$$\begin{aligned}\mathsf{M}\left(Y \neq W\right) &= \mathsf{d}(\mathsf{M}_{YW}, \mathsf{P}_{WW}) \\ &\leq \mathsf{d}(\mathsf{M}_{YW}, \mathsf{T}_{YW}) + \mathsf{d}(\mathsf{T}_{YW}, \mathsf{P}_{WW}) \\ &\overset{(a)}{\leq} \mathsf{d}(\mathsf{M}_{XM^nYW}, \mathsf{T}_{XM^nYW}) + \mathsf{d}(\mathsf{T}_{YW}, \mathsf{P}_{WW}) \\ &\leq 2(n+1)\varepsilon, \end{aligned} \tag{14}$$

where (a) follows from Lemma A.1.

e) Upper-bound of $H^{\mathsf{M}}\left(W \mid X\right)$: By construction of the distribution M and (10), (13) , we have

$$I^{\mathsf{M}}(W; M_i \mid XM^{i-1}) = 0, \text{ for all } i \in [n],$$

which implies that $I^{\mathsf{M}}(W; M^n \mid X) = 0$. Thus,

$$\begin{aligned} H^{\mathsf{M}}\left(W \mid X\right) &= H^{\mathsf{M}}\left(W \mid XM^n\right) \\ &\leq H^{\mathsf{M}}\left(WY \mid XM^n\right) \\ &= H^{\mathsf{M}}\left(Y \mid XM^n\right) + H^{\mathsf{M}}\left(W \mid XM^nY\right) \\ &\leq H^{\mathsf{M}}\left(Y \mid XM^n\right) + H^{\mathsf{M}}\left(W \mid Y\right), \end{aligned} \tag{15}$$

where the last inequality follows from the fact that conditioning reduces entropy.

We now upper bound $H^{\mathsf{M}}\left(Y \mid XM^n\right)$. First, note that $\mathsf{M}_{Y|XM^nW} = \mathsf{T}_{Y|XM^nW} = \mathsf{R}_{Y|XM^n}$, which implies $H^{\mathsf{M}}\left(Y \mid X = x, M^n = m^n\right) = H^{\mathsf{R}}\left(Y \mid X = x, M^n = m^n\right)$. Second, the distribution M is such that $\mathsf{M}_{M^n|X} = \mathsf{R}_{M^n|X}$ and Lemma A.2 implies $\mathsf{d}\left(\mathsf{R}_{XM^n}, \mathsf{M}_{XM^n}\right) = \mathsf{d}\left(\mathsf{R}_X, \mathsf{M}_X\right)$. Since $\mathsf{M}_X = \mathsf{T}_X$, by Lemma A.2 and (11), we have

$$\mathsf{d}\left(\mathsf{M}_X, \mathsf{R}_X\right) \leq \mathsf{d}\left(\mathsf{T}_{XM^nYW}, \mathsf{R}_{XM^nYZ}\right) \leq \varepsilon. \tag{16}$$

Thus, $H^{\mathsf{M}}\left(Y \mid XM^n\right)$ and $H^{\mathsf{R}}\left(Y \mid XM^n\right)$ are the average of the same function under two probability distributions which are ε-close. This implies using the triangle inequality,

$$\left|H^{\mathsf{M}}\left(Y \mid XM^n\right) - H^{\mathsf{R}}\left(Y \mid XM^n\right)\right| \leq 2\varepsilon \log_2 |\mathcal{W}|,$$

which combined with (7) give,

$$H^{\mathsf{M}}\left(Y \mid XM^n\right) \leq F(\varepsilon, |\mathcal{W}|) + 2\varepsilon \log_2 |\mathcal{W}|.$$

Looking back at (15), we also need to upper bound $H^{\mathsf{M}}\left(W \mid Y\right)$, which can be done using Fano's inequality and (14) to obtain

$$\begin{aligned} H^{\mathsf{M}}\left(W \mid X\right) &\leq F(\varepsilon, |\mathcal{W}|) + 2\varepsilon \log_2 |\mathcal{W}| + F(2(n+1)\varepsilon, |\mathcal{W}|) \\ &\leq 2F(2\varepsilon, |\mathcal{W}|) + F(2(n+1)\varepsilon, |\mathcal{W}|), \end{aligned} \tag{17}$$

where we used the fact that $F(\cdot, |\mathcal{W}|)$ is a non-decreasing function on $[0, \frac{1}{2}]$, and by assumption $2(n+1)\varepsilon \leq \frac{1}{2}$.

f) Upper-bound of $H\left(W\right)$: Note that (16) and Theorem A.1 imply that

$$H^{\mathsf{M}}\left(X\right) \leq H^{\mathsf{R}}\left(X\right) + F(2\varepsilon, |\mathcal{W}|).$$

The last equation and (17) finish the proof as follows,

$$\begin{aligned} H^{\mathsf{M}}\left(W\right) &\leq H^{\mathsf{M}}\left(X\right) + H^{\mathsf{M}}\left(W \mid X\right) \\ &\leq H^{\mathsf{R}}\left(X\right) + 3F(2\varepsilon, |\mathcal{W}|) + F(2(n+1)\varepsilon, |\mathcal{W}|). \end{aligned}$$

■

APPENDIX

Theorem A.1. *[8, Th. 17.3.3] Let X and Y be two discrete random variables over $\mathcal{X}$ distributed according to P_X and P_Y, respectively, and such that $\mathsf{d}(\mathsf{P}_X, \mathsf{P}_Y) \leq \frac{1}{4}$. Then,*

$$|H\left(X\right) - H\left(Y\right)| \leq -2\mathsf{d}(\mathsf{P}_X, \mathsf{P}_Y) \log_2 \frac{2\mathsf{d}(\mathsf{P}_X, \mathsf{P}_Y)}{|\mathcal{X}|}.$$

Lemma A.1. *For any joint probability distributions P_{XY} and Q_{XY},*

$$\mathsf{d}\left(\mathsf{P}_{XY}, \mathsf{Q}_{XY}\right) \geq \max\left\{\mathsf{d}\left(\mathsf{P}_X, \mathsf{Q}_X\right); \mathsf{d}\left(\mathsf{P}_Y, \mathsf{Q}_Y\right)\right\}.$$

Lemma A.2. *For any distribution P_X and Q_X, and any conditional distribution $\mathsf{P}_{Y|X}$,*

$$\mathsf{d}(\mathsf{P}_X, \mathsf{Q}_X) = \mathsf{d}(\mathsf{P}_X \cdot \mathsf{P}_{Y|X}, \mathsf{Q}_X \cdot \mathsf{P}_{Y|X}).$$

Lemma A.3. *Let P_{X^n} and Q_{X^n} be two joint distributions. Consider a joint distribution M_{X^n}, which is a mixture between P_{X^n} and Q_{X^n},*

$$\begin{aligned} \mathsf{M}_{X_i|X^{i-1}} &:= \mathsf{P}_{X_i|X^{i-1}}, \text{ for odd } i \in [n], \\ \mathsf{M}_{X_i|X^{i-1}} &:= \mathsf{Q}_{X_i|X^{i-1}}, \text{ for even } i \in [n]. \end{aligned}$$

Then,

$$\mathsf{d}\left(\mathsf{P}_{X^n}, \mathsf{M}_{X^n}\right) \leq 2\left\lfloor \frac{n}{2} \right\rfloor \mathsf{d}\left(\mathsf{P}_{X^n}, \mathsf{Q}_{X^n}\right).$$

Proof: Due to space constraints, we only provide a proof sketch. When n is odd, Lemma A.2 implies $\mathsf{d}\left(\mathsf{P}_{X^n}, \mathsf{M}_{X^n}\right) = \mathsf{d}(\mathsf{P}_{X^{n-1}}, \mathsf{M}_{X^{n-1}})$. When n is even, one can show using the triangle inequality, Lemmas A.2 and A.1, that $\mathsf{d}\left(\mathsf{P}_{X^n}, \mathsf{M}_{X^n}\right) \leq 2\mathsf{d}(\mathsf{P}_{X^n}, \mathsf{Q}_{X^n}) + \mathsf{d}(\mathsf{P}_{X^{n-1}}, \mathsf{M}_{X^{n-1}})$.

Note that M_{X^n} can be rewritten as follows, $\mathsf{M}_{X_i|X^{i-1}} = f(i) \cdot \mathsf{P}_{X_i|X^{i-1}} + (1 - f(i)) \cdot \mathsf{Q}_{X_i|X^{i-1}}$, where $i \in [n]$, and $f(i) := i \pmod 2$. Thus, using the convexity of the statistical distance $\mathsf{d}\left(\mathsf{P}_{X^n}, \mathsf{M}_{X^n}\right) \leq \mathsf{d}(\mathsf{P}_{X^{n-1}}, \mathsf{M}_{X^{n-1}}) + 2(1 - f(n))\mathsf{d}(\mathsf{P}_{X^n}, \mathsf{Q}_{X^n})$. Since $\mathsf{d}\left(\mathsf{P}_{X_1}, \mathsf{M}_{X_1}\right) = 0$, the final result is obtained by induction and Lemma A.1. ■

REFERENCES

[1] U. Maurer, "Constructive cryptography – a new paradigm for security definitions and proofs," in *Theory of Security and Applications (TOSCA 2011)*, ser. Lecture Notes in Computer Science, S. Moedersheim and C. Palamidessi, Eds., vol. 6993. Springer-Verlag, Apr. 2011, pp. 33–56.

[2] U. Maurer and R. Renner, "Abstract cryptography," in *The Second Symposium in Innovations in Computer Science, ICS 2011*, B. Chazelle, Ed. Tsinghua University Press, Jan. 2011, pp. 1–21.

[3] M. Blum, "Coin flipping by telephone a protocol for solving impossible problems," *SIGACT News*, vol. 15, pp. 23–27, Jan. 1983.

[4] Y. Lindell, "Parallel coin-tossing and constant-round secure two-party computation," *Journal of Cryptology*, vol. 16, pp. 143–184, Mar. 2008.

[5] O. Goldreich, *Foundations of Cryptography: Volume 2, Basic Applications*. New York, NY, USA: Cambridge University Press, 2004.

[6] R. Canetti, "Universally composable security: A new paradigm for cryptographic protocols," Cryptology ePrint Archive, Report 2000/067, 2000.

[7] D. Hofheinz, J. Müller-Quade, and D. Unruh, "On the (im-)possibility of extending coin toss," in *Advances in Cryptology - EUROCRYPT 2006*, ser. Lecture Notes in Computer Science, S. Vaudenay, Ed. Springer Berlin / Heidelberg, 2006, vol. 4004, pp. 504–521.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, Jul. 2006.

[9] U. Maurer, A. Rüedlinger, and B. Tackmann, "Confidentiality and integrity: A constructive perspective," in *Theory of Cryptography — TCC 2012*, ser. Lecture Notes in Computer Science, R. Cramer, Ed., vol. 7194, IACR. Springer, 2012, pp. 209–229.