Common Randomness Generation from Sources with Countable Alphabet

Wafa Labidi †‡, Rami Ezzine †‡, Christian Deppe †, Moritz Wiese †‡ and Holger Boche †‡§¶
†Technical University of Munich
‡CASA – Cyber Security in the Age of Large-Scale Adversaries– Exzellenzcluster, Ruhr-Universität Bochum, Germany
§BMBF Research Hub 6G-life, Munich, Germany
¶Munich Center for Quantum Science and Technology (MCQST)
¶Munich Quantum Valley (MQV)
Email: {wafa.labidi, rami.ezzine, christian.deppe, wiese, boche}@tum.de

Abstract—We study a two-source model for common randomness (CR) generation in which the sender Alice and the receiver Bob generate a common random variable with a high probability of agreement by observing independent and identically distributed (i.i.d.) samples of correlated sources on countably infinite alphabets. The two parties are additionally allowed to communicate over a noisy memoryless channel. In our work, we establish a single-letter lower and upper-bound on the CR capacity for the proposed model. This is a challenging scenario because some of the finite alphabet properties, namely of the entropy cannot be extended to the countably infinite case. We use a generalized typicality criterion, called unified typicality, which can be applied to random variables on countably infinite alphabets.

A detailed version with all proofs, explanations, and more discussions can be found in [1].

I. INTRODUCTION

In the two-source model for common randomness (CR) generation, the sender Alice and the receiver Bob, aim to generate a common random variable (RV) with a high probability of agreement. One can take advantage of this resource in the identification (ID) scheme in order to achieve a huge performance gain. The ID scheme was proposed by Ahlswede and Dueck [2] in 1989. In contrast to the classical Shannon transmission [3], the resource CR allows a considerable increase in the ID capacity of channels [4], [5]. The ID approach is much more efficient than the classical transmission scheme for many new applications with high reliability and latency requirements including machine-to-machine and human-to-machine systems [6], industry 4.0 [7] and 6G communication systems [8], [9]. For this purpose, several 6G research projects [10][11] are studying CR generation for future communication networks and it is expected that on the basis of CR, the resilience requirements [8], [9] and security requirements [12] can be met. The aforementioned requirements are crucial for achieving trustworthiness [13]. The resource CR is not uniquely relevant in the ID scheme. In general, the availability of this resource plays an important role in distributed systems [14].

Ahlswede and Csizár initially introduced the problem of CR generation from correlated discrete sources in [5]. They considered in particular a two-source model for CR generation, in which the sender and the receiver communicate over a rate-limited discrete noiseless channel and derived a single-letter formula for the CR capacity for that model.

Later, the results on CR capacity have been extended in [15] to point-to-point Gaussian channels. It has been shown in [15] that the correlation-assisted secure ID capacity of Gaussian channels in the log-log scale is lower-bounded by its corresponding CR capacity. This lower bound can be greater than the secure ID capacity over Gaussian channels with randomized encoding derived in [16]. Later, the problem of CR generation over fading channels has been investigated in [17] and in [18].

Recently, the authors in [19] studied CR generation from Gaussian sources and showed that the CR capacity is infinite when the Gaussian sources are perfectly correlated. In such a scenario, no communication over the channel is necessary. The major motivation of the work in [19] was the drastic effects that the CR generated from the noiseless feedback in the model treated in [20][21] produce on the ID capacity.

However, to the best of our knowledge, very few studies [22] have addressed the problem of CR generation from sources with a countably infinite alphabet, and as far as we know, no research has focused on deriving the CR capacity for such models. An example of such a source model is the Poisson source model, which is highly relevant in molecular communication and optical communication systems. The transition to an infinite alphabet can have drastic consequences in terms of Shannon entropy convergence, variational distance convergence, etc. Some of the finite alphabet properties, namely of the entropy can not be extended to the countably infinite case. Notably, it has been shown that the Shannon entropy is in fact discontinuous at all probability distributions with countably infinite support [23], [24].

In our work, we establish a single-letter lower and upper-bound on the CR capacity of a model involving a memoryless source on a countably infinite alphabet with one-way communication over noisy memoryless channels. We use a generalized typicality criterion, called unified typicality [25], which can be applied to any sources on countable alphabets and make use of the conditional typicality lemma and conditional divergence lemma [25], [26] established for...
the proposed typicality criterion.

Outline: In Section II, we recall some auxiliary results related to unified typicality involved in our work. In Section III, we present the system model for CR generation, define an achievable CR rate and the CR capacity for the proposed model and give a single-letter lower and upper-bound on the CR capacity. In Section IV, we prove the lower bound on the CR capacity. The upper-bound on the CR capacity is established in Section V. Section VI concludes the paper.

II. PRELIMINARIES

$D(\cdot\mid\cdot)$ denotes the Kullback-Leibler divergence; $H(\cdot)$, $\mathbb{E}(\cdot)$ and $I(\cdot;\cdot)$ are the entropy, the expected value, and the mutual information, respectively; $P_X$ denotes the probability mass function of a RV $X$ on a finite or countably infinite alphabet; $\mid\cdot\mid$ denotes the cardinality of a finite set; we use the notation $X \circ Y \circ Z$ to indicate a Markov chain.

Unified typicality for finite and countably infinite alphabets has been established in [25]. Unified typicality is based on a new information divergence measure introduced in [25], unifying both weak typicality [3] and strong typicality [27].

Definition 1. Suppose $\nu > 0$ and $X^n = (X_1, X_2, \ldots, X_n)$ was emitted by the memoryless source $P_X \in \mathcal{P}(\mathcal{X})$ with $\mathcal{X}$ being a countably infinite alphabet and where $H(X) < \infty$.

The unified typical set $\mathcal{U}_\nu^n(P_X)$ w.r.t. $P_X$ is the set of sequences $x^n = (x_1, x_2, \ldots, x_n) \in \mathcal{X}^n$ such that

$$D(P_X \mid P_X) + |H(\tilde{X}) - H(X)| \leq \nu,$$

where the RV $\tilde{X}$ has the distribution $P_X$ on $\mathcal{X}$ with $P_X$ being the empirical distribution of the sequence $x^n$.

The authors in [25] demonstrated the "Unified Asymptotic Equipartition Property (AEP)" for unified typicality, which is similar to the AEP for weak and strong typicality.

Theorem 2 ([25]). Let $H(X)$ be finite. For any $\nu > 0$:

1) If $x^n \in \mathcal{U}_\nu^n(P_X)$, then

$$2^{-n(H(X)+\nu)} \leq P_X^n(x^n) \leq 2^{-n(H(X)) - \nu}.$$

2) For sufficiently large $n$,

$$\Pr\{X^n \in \mathcal{U}_\nu^n(P_X)\} > 1 - \nu.$$

3) For sufficiently large $n$,

$$(1 - \nu)2^{-n(H(X)) - \nu} \leq |\mathcal{U}_\nu^n(P_X)| \leq 2^{-n(H(X)+\nu)}.$$

Unified typicality w.r.t. a bivariate distribution has also been defined in [25].

Definition 3. Suppose $\nu' > 0$ and the sequence $(X^n, Y^n)$ was emitted by the memoryless bivariate source $P_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ with $\mathcal{X}$ and $\mathcal{Y}$ being countably infinite alphabets and where $H(XY) < \infty$. The unified jointly typical set $\mathcal{U}_\nu^n(P_{XY})$ w.r.t. $P_{XY}$ is the set of sequences $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$ such that

$$D(P_{XY} \mid P_{XY}) + |H(\tilde{X}, \tilde{Y}) - H(XY)| + |H(\tilde{X}) - H(X)| + |H(\tilde{Y}) - H(Y)| \leq \nu',$$

where the RV $(\tilde{X}, \tilde{Y})$ has the distribution $P_{\tilde{X}\tilde{Y}}$ on $\mathcal{X} \times \mathcal{Y}$ with $P_{\tilde{X}\tilde{Y}}$ being the empirical distribution of $(x^n, y^n)$.

Theorem 4 ([25]). Let $H(XY)$ be finite. For any $\nu' > 0$:

1) If $(x^n, y^n) \in \mathcal{U}_\nu^n(P_{XY})$, then

$$2^{-n(H(XY)+\nu')} \leq P_{XY}^n(x^n, y^n) \leq 2^{-n(H(XY)-\nu')}.$$

2) For sufficiently large $n$,

$$\Pr\{(X^n, Y^n) \in \mathcal{U}_\nu^n(P_{XY})\} > 1 - \nu'.$$

3) For sufficiently large $n$,

$$(1 - \nu')2^{-n(H(XY)-\nu')} \leq |\mathcal{U}_\nu^n(P_{XY})| \leq 2^{-n(H(XY)+\nu')}.$$

Unified joint typicality can be viewed as a special case of the usual unified typicality, where the sequence $(X, Y)$ is considered as a single RV $Z$. This can be easily shown by choosing suitable $\nu$ and $\nu'$. An interesting case is when the sequences $X^n$ and $Y^n$ are output by the statistically independent sources $P_X$ and $P_Y$, respectively. We prove the following Lemma based on Theorem 2 and Theorem 4. For the proof, we refer the reader to [1].

Lemma 5. Let $0 < \nu' < \nu$. Suppose that the sequences $X^n$ and $Y^n$ are output by the statistically independent sources $P_X$ and $P_Y$, respectively. For any $\nu > \nu' > 0$, the probability that $(\tilde{X}, \tilde{Y}) \in \mathcal{U}_\nu^n(P_{XY})$ for some joint distribution $P_{XY}$ with marginals $P_X$ and $P_Y$ is bounded by

$$\Pr\{(\tilde{X}, \tilde{Y}) \in \mathcal{U}_\nu^n(P_{XY})\} \geq \Pr\{(\tilde{X}, \tilde{Y}) \in \mathcal{U}_\nu^n(P_{XY})\} \leq (1 - \nu')2^{-n(I(X;Y)+2\nu'+\nu)}$$

A generalization to a multivariate distribution can be easily shown [25]. The authors in [26] introduced the following Markov lemma for countable alphabets.

Theorem 6 ([26]). Let $P_{U\times X\times Y} \in \mathcal{P}(\mathcal{U} \times \mathcal{X} \times \mathcal{Y})$ be a memoryless multivariate source with $\mathcal{U}$, $\mathcal{X}$ and $\mathcal{Y}$ countable alphabets and $H(UXY) < \infty$. We assume that $U \circ X \circ Y$ is a Markov chain and

$$\sum_u P_{U|x}(u|x)\log P_{U|x}(u|x)^2 < \kappa,$$

where the constant $\kappa$ is positive and finite. If for any $\tilde{\nu} > 0$ and any given $(x^n, y^n) \in \mathcal{U}_\nu^n(P_{XY})$, $U^n$ is drawn from $\Pi_{i=1}^n P_{U|x_i=x_i}$, then

$$\Pr\{U^n, x^n, y^n \in \mathcal{U}_\nu^n(P_{U\times XY})\} \geq 1 - \tilde{\nu},$$

for sufficiently large $n$ and sufficiently small $\nu$.

III. SYSTEM MODEL, DEFINITIONS, AND MAIN RESULT

In this section, we introduce our system model and extend the definition of an achievable CR rate to this system model. We then present the main result of this paper.
A. System Model and Definitions

Let a bivariate memoryless source \( P_{XY} \) with two components, with generic variables \( X \) and \( Y \) on the countable alphabets (finite and countably infinite) \( \mathcal{X} \) and \( \mathcal{Y} \), respectively, be given. The marginal distributions \( P_X \) and \( P_Y \) satisfy:

\[
\mathbb{E} \left[ \log^2(P_X(X)) \right], \mathbb{E} \left[ \log^2(P_Y(Y)) \right] < \infty. \tag{2}
\]

The assumption in (2) is a technical condition. The restrictiveness of assumption (2) is still an open problem that we are currently investigating. The outputs of \( X \) are observed only by Terminal \( A \) and those of \( Y \) only by Terminal \( B \). Both outputs have length \( n \). We further assume that the joint distribution of \( (X, Y) \) is known to both terminals. Terminal \( A \) can send information to Terminal \( B \) over a memoryless channel \( W \). The Shannon capacity of the channel \( W \) is denoted by \( C(W) \). There are no other resources available to any of the terminals. This is an extension of the standard two-source model introduced by Ahlswede and Csiszar in [5], where they considered the communication over a discrete memoryless noiseless channel with limited capacity. A CR-generation protocol [5] of block length \( n \) consists of:

1) a function \( \Phi \) that maps \( X^n \) into a random variable \( K \) generated by Terminal \( A \) defined on \( \mathcal{K} \) with \( |\mathcal{K}| \geq 3 \),
2) a function \( \Lambda \) that maps \( X^n \) into the input sequence \( T^n \),
3) a function \( \Psi \) that maps \( Y^n \) and the output sequence \( Z^n \) into a random variable \( L \) generated by Terminal \( B \) defined on \( \mathcal{L} \).

This protocol generates a pair of RVs \( (K, L) \) that is called permissible [5]. The system model is depicted in Fig. 1.

![Fig. 1: Memoryless countable-alphabet source model with one-way communication over a noisy memoryless channel](image)

**Definition 7.** A number \( H \) is called an achievable CR rate for the system model in Fig. 1 if there exists a non-negative constant \( c \) such that for every \( \epsilon > 0 \) and \( \gamma > 0 \) and for sufficiently large \( n \) there exists a permissible pair of random variables \( (K, L) \) such that

\[
\Pr\{K \neq L\} \leq \epsilon, \tag{3}
\]

\[
|\mathcal{K}| \leq 2^c n, \tag{4}
\]

\[
\frac{1}{n} H(K) > H - \gamma. \tag{5}
\]

**Definition 8.** The CR capacity \( C_{CR}(P_{XY}, W) \) for the system model in Fig. 1 is the maximum achievable CR rate.

Now, we present the main result of our work. We characterize the CR capacity of the system model in Fig. 1.

**Theorem 9.** Let

\[
U = \left\{ U \in \mathcal{X}_c : \mathbb{E} \left[ \log^2(P_{U|X}(U|X = x)|X = x) \right] < \infty, \forall x \in \mathcal{X} \right\}, \tag{6}
\]

where \( \mathcal{X}_c \) denotes the set of discrete RVs defined on countably infinite alphabets. Define the function

\[
L^{(X,Y)} : t \mapsto \sup_{U \in \mathcal{U}} \inf_{I(U;X)} \frac{I(U;X)}{I(U;X) - I(U;Y) \leq t}
\]

For the system model depicted in Fig. 1, the CR capacity satisfies

\[
C_{CR}(P_{XY}, W) \geq \sup_{\alpha > 0} L^{(X,Y)}(C(W) - \alpha) \tag{7}
\]

and

\[
C_{CR}(P_{XY}, W) \leq \inf_{\alpha > 0} L^{(X,Y)}(C(W) + \alpha) \tag{8}
\]

where \( C(W) \) is the Shannon capacity of \( W \).

Since we are dealing with RVs on countably infinite alphabets, discontinuity issues might occur. It is to note that some properties of finite alphabets, namely of Shannon entropy cannot be extended to the countably infinite case. The equality of the bounds in (8) and (7) depends on whether or not the function \( L^{(X,Y)} \) is continuous at the channel capacity \( C(W) \). When the sets \( \mathcal{X} \) and \( \mathcal{Y} \) are finite, by continuity, the two bounds in (8) and (7) coincide and Theorem 9 is reduced to the known result [5, Theorem 4.1].

IV. PROOF OF THE LOWER-BOUND IN (7)

Let \( \alpha > 0 \) be fixed arbitrarily. Let \( U \) be an arbitrary random variable on \( \mathcal{U} \) satisfying \( U \circ X \circ Y \) and \( I(U;X) - I(U;Y) \leq C(W) \). In the following, we show that \( H = I(U;X) \) is an achievable CR rate for our system model.

Let \( \epsilon, \gamma > 0 \). Let \( 0 < \nu < \nu_1 < \nu_2 < \nu_3 \leq \frac{\gamma}{2} \). We generate \( N_1 N_2 \) codewords \( U^n(i,j), \ i = 1, \ldots, N_1, \ j = 1, \ldots, N_2 \) by choosing the \( n \) \( (N_1 N_2) \) symbols \( u_l(i,j), \ l = 1, \ldots, n \), independently at random using \( P_U \). Each realization \( u^n(i,j) \) of \( U^n(i,j) \) is known to both terminals. For some \( \delta > \frac{\gamma}{2 \nu_1} \), let

\[
N_1 = 2^{|\{I(U;X) - I(U;Y) \leq C(W)\}|}, \quad N_2 = 2^{|\{I(U;X) > C(W)\}|}.
\]

The task of the encoder consists in finding a pair \((i,j)\) such that \((x^n, u^n(i,j))\) are jointly typical. If such a pair \((i^*, j)\) exists, we set \( f(x^n) = i^* \) and \( \Phi(x^n) = u^n(i^*, j) \) (either one if there are several). If not successful, then \( f(x^n) \) is set to \( N_1 + 1 \) and \( K = \Phi(x^n) \) is set to a constant sequence \( u^n_0 \).
different from all the $u^n(i,j)$s and known to both terminals. We choose $\delta$ to be sufficiently small such that

$$\frac{\log \|f\|}{n} = \frac{\log(N_1 + 1)}{n} \leq C(W) - \delta', \quad (10)$$

for some $\delta'$ where $\|f\|$ refers to the cardinality of the codomain of the function $f$. The message $i^* = f(x^n)$, with $i^* \in \{1, \ldots, N_1 + 1\}$, is mapped into a sequence $t^n$ using a suitable code sequence with rate $\frac{\log \|f\|}{n}$ satisfying (10) and with error probability lower than $\frac{\delta}{n}$ for sufficiently large $n$. The sequence $t^n$ is transmitted over the channel $W$. Let $z^n$ denote the channel output sequence. Terminal $B$ converts the channel output $z^n$ into $i^*$. From the knowledge of $i^*$ and $y^n$, the task of the decoder is to find $j$ such that $(u^n(i^*, j), y^n)$ are jointly typical. If such an index $j$ exists, let $\Psi(y^n, z^n) = u^n(i^*, j)$. If there is no such $u^n(i^*, j)$ or there are several, $L = \Psi(y^n, z^n)$ is set to $u^n_0$.

Let $I^* = f(X^n)$ be the RV modeling the message encoded by Terminal $A$ and let $I^*$ be the RV modeling the message decoded by $B$. We first prove that (3) is satisfied. We have:

$$\Pr[K \neq L] \leq \Pr[K \neq L | I^* = \hat{I}^*] + \Pr[I^* \neq \hat{I}^*].$$

We establish an upper-bound on $\Pr[K \neq L | I^* = \hat{I}^*]$. For this purpose, we consider the following events. The source sequences $(x^n, y^n)$ are not jointly typical:

$$E_1 = \{(X^n, Y^n) \notin U^n_0(P_{XY})\}.$$

The encoder cannot find a pair $(i, j)$ such that $(u^n(i, j), x^n) \in U^n_0(P_{UX})$:

$$E_2 = \cap_{j=1}^{N_1} \{ (U^n(i, j), X^n) \notin U^n_0(P_{UX}) \}.$$

The decoder outputs $j \neq j'$ such that $(u^n(i, j'), y^n) \in U^n_0(P_{UY})$:

$$E_3 = \cup_{j \neq j'} \{ (U^n(i, j'), Y^n) \in U^n_0(P_{UY}) \}.$$

The decoder fails to find $j$ such that $(u^n(i, j), x^n, y^n) \in U^n_0(P_{UXY})$:

$$E_4 = \cap_{j=1}^{N_2} \{ (U^n(i, j), X^n, Y^n) \notin U^n_0(P_{UXY}) \}.$$

By the union bound we have

$$\Pr[K \neq L | I^* \neq \hat{I}^*] \leq \Pr[E_1] + \Pr[E_2 \cap E_3] + \Pr[E_4] + \Pr[E_1 \cap E_2 \cap E_3 \cap E_4].$$

In the following, we compute an upper bound on the probability of the events $E_i$, $i = 1, \ldots, 4$.

$$\Pr[E_1] = P^n_{XY} \left( U^n_0(P_{XY}) \right)^c \leq \nu_1,$$

where (a) follows from Theorem 4 since the sequence $(x^n, y^n)$ is drawn from the distribution $P^n_{XY}$.

$$\Pr[E_2] \leq \Pr[X^n \notin U^n_0(P_{UX})] + \sum_{x^n \in U^n_0(P_{UX})} P^n_{X}(x^n) \Pr[E_2 | X^n = x^n] \leq \nu + \sum_{x^n \in U^n_0(P_{UX})} P^n_{X}(x^n) \Pr[E_2 | X^n = x^n].$$

(a)

(b)

(c)

where (a) follows from Theorem 2, (b) follows because, by conditioning on $X^n = x^n$, the $N_1 N_2$ events of the intersection are independent and from Lemma 5, (c) follows because $(1 - x^n) \leq \exp(-m \delta)$ and (e) follows from (9).

$$\Pr[E_3] \leq \sum_{j \neq j'} \Pr \left[ \left( U^n(i, j') \right), (Y^n) \in U^n_0(P_{UY}) \right] \leq N_2 \cdot 2^{-2n(1/2 - \delta/2)} = 2^{-2n(1/2 - \delta/2)},$$

where (a) follows from the union bound and (b) follows by conditioning.

$$\Pr[E_4] = \sum_{(x^n, y^n) \in U^n_0(P_{UXY})} \Pr \left[ U^n(i, j), X^n, Y^n \right] = \left( \sum_{(x^n, y^n) \in U^n_0(P_{UXY})} \Pr \left[ U^n(i, j), X^n, Y^n \right] \right) \leq \nu_2 N_2,$$

where (a) follows because for $(x^n, y^n)$ being not jointly typical, we have $\Pr[E^*](X^n, Y^n) \notin U^n_0(P_{UXY}) = 0$ and the event $E^*$ implies the event $\cap_{j=1}^{N_2} \{ (U^n(i, j), X^n, Y^n) \notin U^n_0(P_{UXY}) \}$. (b) follows because the $N_2$ events of the intersection are independent and (c) follows from Theorem 6.

Now, we know that \( \lim_{n \to \infty} \exp \left( -(1 - \nu_2) 2^{n(2\delta - 3\delta/2)} \right) = 2^{-n(2\delta - 3\delta/2)} + \nu_3 N_2 = 0 \). Therefore, for sufficiently large $n$

$$\exp \left( -(1 - \nu_2) 2^{n(2\delta - 3\delta/2)} \right) + 2^{-n(2\delta - 3\delta/2)} + \nu_3 N_2 \leq \frac{\epsilon}{4}. $$

2428
It follows that for sufficiently large $n$

$$\Pr[K \neq L|I^* = I^*] \leq \frac{c}{2}. \quad (11)$$

Thus, we have $\Pr[K \neq L] \leq \frac{c}{2} + \Pr[I^* \neq \hat{I}^*] \leq \epsilon$, where the first inequality follows from (11) and the second inequality follows from the definition of the code sequence. Thus, $(K, L)$ satisfy (3). Let $u^n(i, j)$ be any realization of the RV $U^n(i, j)$.

Now, we show that $(K, L)$ satisfy (4) and (5). Clearly, (4) is satisfied for $c = 2[I(U; X) + 2\delta]$, $n$ sufficiently large:

$$|K| = N_1 N_2 + 1 = 2^{\Omega(n[I(U; X) + 2\delta])} \leq 2^{2\Omega(n[I(U; X) + 2\delta])}.$$

For a fixed $u^n(i, j) \in \mathcal{U}^n$, it holds that

$$\Pr[K = u^n(i, j)] \leq \Pr([u^n(i, j), X^n] \in \mathcal{U}_a^n(P_{X^n})|X^n \in \mathcal{U}_a^n(P_X)) + \Pr(u^n(i, j), X^n) \in \mathcal{U}_a^n(P_{X^n})|X^n = x^n) = 0$$

and (b) follows from (3). Then, (5) is satisfied. This completes the proof of (7).

V. PROOF SKETCH OF THE UPPER-BOUND IN (8)

Let $H$ be any achievable CR rate. So, there exists a non-negative constant $c$ such that for every $\epsilon > 0$ and $\gamma > 0$ and for sufficiently large $n$ there exists a permissible pair of RVs $(K, L)$ w.r.t. a fixed CR-generation protocol of block-length $n$, as presented in Section III-A such that the requirements in (3), (4) and (5) are satisfied, with $\epsilon > 0$ being the constant in (3) and $\gamma > 0$ being the constant in (5). Let $J$ be a RV uniformly distributed on $\{1, \ldots, n, \delta\}$ and independent of $K, X^n$ and $Y^n$. We further define $U = (K, X_1, \ldots, X_{j-1}, Y_{j+1}, \ldots, Y_n)$. It holds that $U \not\sim X_j \not\sim Y_j$. Next, we show that $U \in \mathcal{U}$, where $\mathcal{U}$ is defined in (6).

Proposition 1. For a fixed block-length $n$ and all $x \in \mathcal{X}$:

$$\mathbb{E}[\log^2 P_{U|X_j}(u|X_j = x) | X_j = x] < \infty.$$  

Proof. Let $V = X_1, \ldots, X_{j-1}, Y_{j+1}, \ldots, Y_n, J$. Then, it holds that $U = (K, V)$. For any realization $u = (k, v)$ of $U$, where $v = (x_1, \ldots, x_{j-1}, y_{j+1}, \ldots, y_n, j)$, we have

$$P_{U|X_j}(u|x) = \frac{1}{n} P_{K|X_j,X_j}(k|v, x) \prod_{i=1}^{j-1} P_X(x_i) \prod_{i=j+1}^{n} P_Y(y_i),$$

where (a) follows because $X_i, i = 1 \ldots n$ are mutually independent and because $J$ is independent of $X^n$, and because $J$ is uniformly distributed on $\{1, \ldots, n\}$. Therefore, we have

$$\log P_{U|X_j}(u|x) = \log P_{K|X_j,X_j}(k|v, x) + \sum_{i=1}^{j-1} \log(P_X(x_i)) + \sum_{i=j+1}^{n} \log(P_Y(y_i)) + \log\left(\frac{1}{n}\right).$$

It follows that

$$\mathbb{E}\left[\left(\log P_{U|X_j}(u|x)\right)^2 \right] \leq 4 \left(\log^2 P_{K|X_j}(k|v, x) + \frac{1}{n}\right) \sum_{i=1}^{j-1} \log(P_X(x_i)) + \sum_{i=j+1}^{n} \log(P_Y(y_i)) + \frac{1}{n}.$$

where (a) follows because $x^2 = |x|^2$ for $x \in \mathbb{R}$ and because $|x+y|^2 \leq 2(|x|^2 + |y|^2)$, (b) follows from triangle inequality, from $\left(\sum_{i=1}^{n} x_i^2\right) \leq n \sum_{i=1}^{n} x_i^2$ and from $1 \leq j \leq n$.

For $K|S$ of length $n$,

$$\mathbb{E}\left[\left(\log P_{K|S}(k|s)\right)^2 \mid S = s\right] < \infty.$$  

Proof. Let $V = X_1, \ldots, X_{j-1}, Y_{j+1}, \ldots, Y_n, J$. Then, it holds that $U = (K, V)$. For any realization $u = (k, v)$ of $U$, where $v = (x_1, \ldots, x_{j-1}, y_{j+1}, \ldots, y_n, j)$, we have

$$P_{U|X_j}(u|x) = \frac{1}{n} P_{K|X_j,X_j}(k|v, x) \prod_{i=1}^{j-1} P_X(x_i) \prod_{i=j+1}^{n} P_Y(y_i),$$

where (a) follows because $X_i, i = 1 \ldots n$ are mutually independent and because $J$ is independent of $X^n$, and because $J$ is uniformly distributed on $\{1, \ldots, n\}$. Therefore, we have

$$\log P_{U|X_j}(u|x) = \log P_{K|X_j,X_j}(k|v, x) + \sum_{i=1}^{j-1} \log(P_X(x_i)) + \sum_{i=j+1}^{n} \log(P_Y(y_i)) + \log\left(\frac{1}{n}\right).$$

It follows that

$$\mathbb{E}\left[\left(\log P_{U|X_j}(u|x)\right)^2 \right] \leq 4 \left(\log^2 P_{K|X_j,X_j}(k|v, x) + \frac{1}{n}\right) \sum_{i=1}^{j-1} \log(P_X(x_i)) + \sum_{i=j+1}^{n} \log(P_Y(y_i)) + \frac{1}{n}.$$

where (a) follows because $x^2 = |x|^2$ for $x \in \mathbb{R}$ and because $|x+y|^2 \leq 2(|x|^2 + |y|^2)$, (b) follows from triangle inequality, from $\left(\sum_{i=1}^{n} x_i^2\right) \leq n \sum_{i=1}^{n} x_i^2$ and from $1 \leq j \leq n$.  

It follows using [28, Theorem 5.5] that

$$\mathbb{E}\left[\left(\log P_{U|X_j}(u|x)\right)^2 \right] \leq 4 \left(\log^2 P_{K|X_j,X_j}(K|V, x) | X_j = x) + \frac{1}{n}\right) \sum_{i=1}^{j-1} \log(P_X(x_i)) + \sum_{i=j+1}^{n} \log(P_Y(y_i)) + \frac{1}{n}.$$

Consider $S = (X_1, \ldots, X_{j-1}, Y_{j+1}, \ldots, Y_n, J)$ and let $s = (x_1, \ldots, x_{j-1}, y_{j+1}, \ldots, y_n, j, x_j)$ be any realization of $S$. We introduce now the following lemma proved in [1].

**Lemma 10.** For $|K| \geq 3$, it holds for sufficiently large $n$ that

$$\mathbb{E}\left[\log^2 P_{K|S}(k|s) | S = s\right] \leq \frac{1}{\ln(2)} n \epsilon^2 + \frac{1}{\ln(2)} \frac{4}{e^2} + 2 \frac{1}{\ln(2)} e^{nc}.$$  

Now, it follows from (13) using Lemma 10 that

$$\mathbb{E}\left[\left(\log P_{U|X_j}(u|x)\right)^2 \right] \leq \frac{1}{\ln(2)} n \epsilon^2 + \frac{1}{\ln(2)} \frac{4}{e^2} + 2 \frac{1}{\ln(2)} e^{nc}.$$  

which implies that $\mathbb{E}\left[\left(\log P_{U|X_j}(u|x)\right)^2 \right] \leq \frac{1}{\ln(2)} n \epsilon^2 + \frac{1}{\ln(2)} \frac{4}{e^2} + 2 \frac{1}{\ln(2)} e^{nc}$. This completes the proof of Proposition 1.

It is shown in [1] that $H(K) \leq n I(U; X_j)$. The proof is analogous to the one provided in [5] for discrete sources with a countable alphabet. It is also shown in [1] that $I(U; X_j) - I(U; Y_j) \leq C(W) + \kappa(n, \epsilon)$, where $\kappa(n, \epsilon) = \frac{1}{n} + \epsilon \epsilon$. Since the joint distribution of $X_j$ and $Y_j$ is equal to $P_{XY}$, $\frac{H(K)}{n}$ is upper-bounded by $I(U; X)$ subject to

$$I(U; X) - I(U; Y) \leq C(W) + \kappa(n, \epsilon),$$
where $U \in \mathcal{U}$ and where $U \circ X \circ Y$ with $U \in \mathcal{U}$, where $\mathcal{U}$ is defined in (6).

As a result, it holds using (5) that for sufficiently large $n$, and for every $\epsilon, \gamma > 0$, any achievable CR rate $H$ satisfies

$$H < \sup_{U \circ X \circ Y} I(U; X) + \gamma,$$

It follows that

$$H \leq \inf_{\epsilon, \gamma > 0} \left( \lim_{n \to \infty} L(X, Y) (C(W) + \kappa(n, \epsilon)) + \gamma \right)$$

$$= \inf_{\alpha > 0} L(X, Y) (C(W) + \alpha).$$

This completes the proof of (8).

VI. CONCLUSION

CR generation has striking applications in the ID scheme, a new approach in communications that is highly relevant in 6G Communication. CR allows a significant increase in the ID capacity. For this reason, CR generation for future communication networks is a central research question in large 6G research projects. CR is also highly relevant in the modular coding scheme for secure communication and a useful resource in coding over AVCs. In this paper, we investigated the problem of CR generation from correlated sources with countable alphabets aided by one-way communication over noisy memoryless channels. We established a single-letter lower and upper-bound for the CR capacity. As a future work, it would be interesting to investigate the problem of CR generation from sources with an arbitrary joint distribution.

ACKNOWLEDGMENTS

H. Boche and C. Deppe were supported by the German Federal Ministry of Education and Research (BMBF) within the national initiative on 6G Communication Systems through the research hub 6G-life under Grant 16KISK002, within the national initiative on Post Shannon Communication (NewCom) under Grant 16KISK003K. H. Boche has further received funding by the German Research Foundation (DFG) within Germany’s Excellence Strategy EXC-2092 – 390781972. M. Wiese was supported by DFG within the Gottfried Wilhelm Leibniz Prize under Grant BO1734/20-1, and within Germany’s Excellence Strategy EXC-2111-390814868. W. Labidi and R. Ezzine were supported by BMBF within NewCom under Grant 16KISK003K and within the national initiative on 6G Communication Systems through the research hub 6G-life under Grant 16KISK002.

REFERENCES

[1] W. Labidi, R. Ezzine, C. Deppe, M. Wiese, and H. Boche, “Common Randomness Generation from Sources with Countable Alphabet,” arXiv e-prints, p. arXiv:2210.04556, Oct. 2022.
[2] R. Ahlswede and G. Dueck, “Identification via channels,” IEEE Transactions on Information Theory, vol. 35, no. 1, pp. 15–29, 1989.
[3] C. E. Shannon, “A mathematical theory of communication,” Bell System Technical Journal, vol. 27, pp. 379–423, 623–656, July, October 1948.
[4] R. Ahlswede, “General theory of information transfer: Updated,” Discrete Applied Mathematics, vol. 156, pp. 1348–1388, 05 2008.
[5] R. Ahlswede and I. Csiszar, “Common randomness in information theory and cryptography. II. CR capacity,” IEEE Transactions on Information Theory, vol. 44, no. 1, pp. 225–240, 1998.
[6] H. Boche and C. Deppe, “Secure identification for wiretap channels; robustness, super-additivity and continuity,” IEEE Transactions on Information Forensics and Security, vol. 13, no. 7, pp. 1641–1655, 2018.
[7] Y. Lu, “Industry 4.0: A survey on technologies, applications and open research issues,” Journal of Industrial Information Integration, vol. 6, pp. 1 – 10, 2017.
[8] G. Fettweis and H. Boche, “6G: The personal tactile internet—and open questions for information theory,” IEEE BITS the Information Theory Magazine, vol. 1, no. 1, pp. 71–82, 2021.
[9] J. Cabrera, H. Boche, C. Deppe, R. F. Schaefer, C. Scheunert, and F. Fitzek, “6G and the post-shannon-theory,” in Shaping Future 6G Networks: Needs, Impacts and Technologies, T. M. Emmanuel Bertin, Noel Crespi, Ed. Wiley-Blackwell, 2021.
[10] F. Fitzek and H. Boche, “Research landscape – 6G networks research in europe: 6G-life: Digital transformation and sovereignty of future communication networks,” IEEE Network, vol. 35, no. 6, pp. 4–5, Nov 2021.
[11] F. Fitzek et. al., “6G activities in germany,” IEEE Future Networks, to be published 2022.
[12] M. Wiese and H. Boche, “Semantic security via seeded modular coding schemes and ramanujan graphs,” IEEE Transactions on Information Theory, vol. 67, no. 1, pp. 32–80, 2021.
[13] G. Fettweis and H. Boche, “On 6G and trustworthiness,” Communications of the ACM, vol. 65, no. 4, pp. 48–49, Apr 2022.
[14] M. Sudan, H. Tyagi, and S. Watanabe, “Communication for generating correlation: A unifying survey,” IEEE Transactions on Information Theory, vol. 66, no. 1, pp. 3–37, 2020.
[15] R. Ezzine, W. Labidi, H. Boche, and C. Deppe, “Common randomness generation and identification over gaussian channels,” in GLOBECOM 2020 - 2020 IEEE Global Communications Conference (GLOBECOM), 2020, pp. 1–6.
[16] W. Labidi, C. Deppe, and H. Boche, “Secure identification for Gaussian channels,” in ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 2872–2876.
[17] R. Ezzine, M. Wiese, C. Deppe, and H. Boche, “Common randomness generation over slow fading channels,” in 2021 IEEE International Symposium on Information Theory (ISIT), 2021, pp. 1925–1930.
[18] ——, “Outage common randomness capacity characterization of multiple-antenna slow fading channels,” in 2021 IEEE Information Theory Workshop (ITW), 2021, pp. 1–6.
[19] W. Labidi, R. Ezzine, C. Deppe, and H. Boche, “Common randomness generation from gaussian sources,” in 2022 IEEE International Symposium on Information Theory (ISIT), 2022, submitted.
[20] W. Labidi, H. Boche, C. Deppe, and M. Wiese, “Identification over the gaussian channel in the presence of feedback,” in 2021 IEEE International Symposium on Information Theory (ISIT), 2021, pp. 278–283.
[21] M. Wiese, W. Labidi, C. Deppe, and H. Boche, “Identification over additive noise channels in the presence of feedback,” IEEE Transactions on Information Theory, pp. 1–1, 2022.
[22] B. Ghazi and T. Jayram, Resource-Efficient Common Randomness and Secret-Key Schemes, pp. 1834–1853.
[23] S.-W. Ho and R. W. Yeung, “On the interplay between entropy and variational distance,” IEEE Transactions on Information Theory, vol. 56, no. 12, pp. 5906–5929, 2010.
[24] S.-W. Ho, “The interplay between entropy and the capacity of additive noise channels in the presence of feedback,” in Proceedings. International Symposium on Information Theory, 2005. ISIT 2005., 2005, pp. 159–163.
[25] S.-W. Ho and R. W. Yeung, “The interplay between entropy and variational distance,” IEEE Transactions on Information Theory, vol. 56, pp. 1 – 10, 2017.
[26] S.-W. Ho, “On information divergence measures and a unified typicality,” IEEE Transactions on Information Theory, vol. 56, no. 12, pp. 5893–5905, 2010.
[27] S.-W. Ho, “Markov lemma for countable alphabets,” in 2010 IEEE International Symposium on Information Theory, 2010, pp. 1448–1452.
[28] J. Wolfowitz, Coding Theorems of Information Theory., New York, NY, USA: Springer Berlin, Heidelberg, 1961.
[29] A. Klenke, Probability Theory: A Comprehensive Course, 3rd ed. Springer-Verlag, 2020.