Secret Key Generation with One Communicator and a One-Shot Converse via Hypercontractivity

Jingbo Liu   Paul Cuff   Sergio Verdú

Dept. of Electrical Eng., Princeton University, NJ 08544

{jingbo,cuff,verdu}@princeton.edu

Abstract—A new model of multi-party secret key agreement is proposed, in which one terminal called the communicator can transmit public messages to other terminals before all terminals agree on a secret key. A single-letter characterization of the achievable region is derived in the stationary memoryless case. The new model generalizes some other (old and new) models of key agreement. In particular, key generation with an omniscient helper is the special case where the communicator knows all sources, for which we derive a zero-rate one-shot converse for the secret key per bit of communication.

I. INTRODUCTION

A random number known only to several geographically distributed terminals is a resource that can be used for cryptographic purposes such as secure communications. Remarkably, the terminals can usually distill such a shared random number, or secret key, by communicating information about certain correlated random processes they observe individually, even though the communication is wiretapped by some eavesdropper. The fundamental limits on the maximal secret key rate can be studied using information theoretic tools [1][2][3].

In this paper we propose a new protocol of multi-party secret key agreement, called *secret key generation with one communicator*, as shown in Figure 1. Terminals[1] $\mathcal{Z}, \mathcal{X}_1, \ldots, \mathcal{X}_m$ observe general sources $Z, X_1, \ldots, X_m$, respectively. The communicator $\mathcal{Z}$ is allowed to send public messages $W_1, \ldots, W_m$ to $\mathcal{X}_1, \ldots, \mathcal{X}_m$, before all the $m + 1$ terminals agree on an integer K (the key). We assume that for each $l \in \{1, \ldots, m\}$ there is an eavesdropper wiretapping the communication link from the communicator to $\mathcal{X}_l$. Independence of K and W_l for each $l \in \{1, \ldots, m\}$ ensures security.

We derive a single-letter characterization of the achievable public communication rates and the key rate in the stationary memoryless case, which is a special case of the above formulation where we identify $Z, X_1, \ldots, X_m$ with the corresponding block symbols.

Of course, other related protocols of key generation have been studied in the literature. The canonical one-way protocol (Model S with forward communication in [2]) is a special case of the secret key generation with one communicator protocol where $m = 1$, in which case a key rate
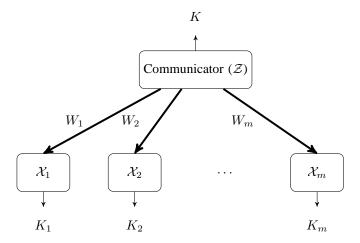
$$R < I(\mathsf{U}; \mathsf{X}_1) \qquad (1)$$

Figure 1: Key generation with one communicator

is achievable with a communication rate

$$R_1 > I(\mathsf{U}; \mathsf{Z}) - I(\mathsf{U}; \mathsf{X}_1) \qquad (2)$$

where Q_{ZX_1} is the per-letter distribution of the stationary memoryless source and $\mathsf{U} - \mathsf{Z} - \mathsf{X}_1$. However the $m = 1$ case does not assume the full complexity and difficulty of the general case, as we shall see later in terms of the single letter region and the coding scheme.

If no communication constraints are imposed, then the maximal key rate is (c.f. [4]):

$$\min_{1 \le l \le m} I(\mathsf{Z}; \mathsf{X}_l). \qquad (3)$$

While random binning Z^n shows the achievability of (3), it cannot be used for the rate constraint case, where the receivers do not need to be able the construct Z^n; indeed a main difficulty with the rate constraint is to decide what common message should the terminals be able to agree on.

Finally if the terminals only need to construct a common random number without any secrecy guarantee (the CR generation problem), then the rate region is also known [5, Theorem 4.2] (see also [6]): a CR rate of

$$R < I(\mathsf{U}; \mathsf{Z}) \qquad (4)$$

is achievable if

$$R_l > I(\mathsf{U}; \mathsf{Z}) - I(\mathsf{U}; \mathsf{X}_l), \quad \forall l = 1, \ldots, m \qquad (5)$$

where $\mathsf{U} - \mathsf{Z} - \mathsf{X}^m$, which can be shown using an extension of source coding with side information [7]. In contrast, the key generation problem in this paper requires much more involved achievability construction and analysis. Specifically, we use superposition coding in a novel way in order to convey the information of the key securely to the receivers. Whereas in the common usage of superposition coding the lower layer codeword is decoded before the upper layer [8], in our construction the index of the upper layer codeword is transmitted to the receiver to facilitate the decoding of the lower layer codeword. Moreover, we use the recent achievability technique of likelihood encoding [9] in order to simplify the security analysis considerably.

Particularly interesting is the special case of $Z = X^m$, which we call the *omniscient helper* problem. In this case, a zero-rate one-shot converse on the secret key per bit of communication can be derived using hypercontractivity[2], strengthening the best converse bound that can be obtained from Fano's inequality. This new converse, derived from first principles, also underlines the intimate interplay between key agreement and hypercontractivity.

II. PROBLEM SETUP AND MAIN RESULTS

Let $Q_{ZX_1^m}$ be the joint distribution of sources $Z, X_1, \ldots, X_m$. As in Figure 1, the Terminals $\mathcal{Z}, \mathcal{X}_1, \ldots, \mathcal{X}_m$ observe $Z, X_1, \ldots, X_m$, respectively, and the communicator $\mathcal{Z}$ computes the integers $W_1(Z), \ldots, W_m(Z)$ possibly stochastically and sends them to $\mathcal{X}_1, \ldots, \mathcal{X}_m$, respectively. Then, the $m + 1$ parties calculate integers $K(Z), K_1(X_1, W_1), \ldots, K_m(X_m, W_m)$ possibly stochastically.

In the case of stationary memoryless sources and block coding, we substitute $Z \leftarrow Z^n$ and $X_l \leftarrow X_l^n$ for each l, where n is the blocklength. The measures of reliable communication and secrecy are defined as follows:

$$\epsilon_n = \max_{1 \leq l \leq m} \mathbb{P}[K \neq K_l], \tag{6}$$

$$\nu_n = \max_{1 \leq l \leq m} \{\log |\mathcal{K}| - H(K|W_l)\}. \tag{7}$$

Definition 1. The $(m + 1)$-tuple $(R, R_1, \ldots, R_m)$ is said to be *achievable* if a sequence of key generation schemes can be designed to fulfill the following conditions:

$$\liminf_{n \to \infty} \frac{1}{n} \log |\mathcal{K}| \geq R; \tag{8}$$

$$\limsup_{n \to \infty} \frac{1}{n} \log |\mathcal{W}_l| \leq R_l, \quad l = 1, \ldots, m; \tag{9}$$

$$\lim_{n \to \infty} \epsilon_n = 0; \tag{10}$$

$$\lim_{n \to \infty} \nu_n = 0. \tag{11}$$

From the standard diagonalization argument [10], the achievable region is closed. Our main result is the following:

Theorem 2. *The set of achievable rates is the closure of*

$$\bigcup_{Q_{\mathsf{U}\mathsf{S}^m|\mathsf{Z}}} \left\{ \begin{array}{c} (R, R_1, \ldots, R_m): \quad R \leq \\ \min\{I(\mathsf{U}; \mathsf{Z}), I(\mathsf{U}\mathsf{S}_1; \mathsf{X}_1), \ldots, I(\mathsf{U}\mathsf{S}_m; \mathsf{X}_m)\} \\ R_l \geq I(\mathsf{U}\mathsf{S}_l; \mathsf{Z}|\mathsf{X}_l), \quad 1 \leq l \leq m \end{array} \right\}. \tag{12}$$

Remark 3. The region in Theorem 2 is not decreased if we restrict to the union over $Q_{\mathsf{U}|\mathsf{Z}} \prod_{l=1}^{m} Q_{\mathsf{S}_l|\mathsf{U}\mathsf{Z}}$.

Remark 4. Previous results of Ahlswede-Csiszár [2] and Csiszár-Narayan [4] shown in (1)-(3) are clearly special cases of Theorem 2.

A. Special Case: Omniscient Helper

As alluded to in the Introduction, the problem reduces to an interesting special case when the communicator knows all other sources. In this situation the communicator can be viewed as a helper, since the requirement that it can recover the key is vacuous because of its omniscience; the rate region in Theorem 2 can also be simplified as follows, since setting $\mathsf{S}_l = \mathsf{X}_l$ in (12) is optimal.

Theorem 5. *In the special case of $\mathsf{Z} = \mathsf{X}^m$, the set of achievable rates is the closure of*

$$\bigcup_{Q_{\mathsf{U}|\mathsf{X}^m}} \left\{ \begin{array}{c} (R, R_1, \ldots, R_m): \\ R \leq \min\{I(\mathsf{U}; \mathsf{X}^m), H(\mathsf{X}_1), \ldots, H(\mathsf{X}_m)\}; \\ R_l \geq I(\mathsf{U}; \mathsf{X}^m) - I(\mathsf{U}; \mathsf{X}_l), \quad 1 \leq l \leq m \end{array} \right\}. \tag{13}$$

The region in (13) has some special features:

Remark 6. The region for a product source can be strictly larger than the Minkowski sum of its factors. Indeed even with unconstrained communication rates, the supremum key rate is as in (3), where the minimum implies that a joint encoding can asymptotically strictly outperform separate encoding of the independent components.

Remark 7. The key rate can be positive even if X^m has independent coordinates. For example, when $m = 2$ and $\mathsf{X}_1 \perp \mathsf{X}_2$ are equiprobable binary random variables, a key rate of $R = 1$ is achievable if the helper sends $X_1{}^n \oplus X_2{}^n$ to $\mathcal{X}_2$ and, thus, the terminals agree on $X_1{}^n$.

Remark 8. Comparing the CR generation region (4),(5) and the key generation region (13), we see that in the omniscient helper problem the secrecy constraint does not increase the required communication rates as long as $R \leq \min_{1 \leq l \leq m} H(\mathsf{X}_l)$. In particular, this is unconditionally true for continuous sources with infinite entropy. But even when the rate regions coincide, the underlying achievability constructions are different; indeed the coding schemes for CR generation in ([5, Theorem 4.2] and [6]) do not provide security. The reason why secrecy can be gained with no extra cost for small R is that the helper shares sufficient secure randomness (the sources) with the other terminals to protect its messages.

III. One-Shot Achievability via Likelihood Encoder

We outline the derivation of a one-shot achievability bound using the recent proof technique of likelihood encoding [11][9]. This method adapts to general non-discrete, non-i.i.d. sources and simplifies the analysis of the secrecy constraint. Some standard notations in one-shot information theory, which may be found in reference [12], will be used in this section.

Theorem 9. *Suppose the sources have joint distribution Q_{ZX^m}. Fix an arbitrary $Q_{U|Z}$, $Q_{S_l U|Z}$, $1 \le l \le m$, integers $I_0, \ldots, I_m$ and $J_1 \ldots, J_m$. Then there exists a key generation scheme with $|\mathcal{K}| = I_0$, $|\mathcal{W}_l| = \prod_{i=1}^{l} I_i J_l$ which guarantees that*

$$\mathbb{P}[K \ne K_l] \le 2m(\epsilon + T + T_l), \tag{14}$$

$$\log |\mathcal{K}| - H(K|W_l)$$

$$\le \inf_{0 < \delta < (IJ_l)^{\frac{3}{2}} \exp(-1)} \left\{ 4m(2T + T_l + 2\delta) \log \frac{(IJ_l)^{\frac{3}{2}}}{\delta} \right\} \tag{15}$$

for each $1 \le l \le m$, where we have defined

$$I := \prod_{l=0}^{m} I_l, \tag{16}$$

$$T := \inf_{\gamma \in \mathbb{R}} \left\{ \mathbb{P}[\imath_{U;Z}(U;Z) > \gamma] + \frac{\exp(\frac{\gamma}{2})}{2\sqrt{I}} \right\}, \tag{17}$$

$$T_l := \inf_{\gamma \in \mathbb{R}} \left\{ \mathbb{P}[\imath_{S_l;Z|U}(S_l;Z|U) > \gamma] + \frac{\exp(\frac{\gamma}{2})}{2\sqrt{J_l}} \right\}, \tag{18}$$

$$\epsilon := \max_{1 \le l \le m} \inf_{\gamma \in \mathbb{R}} \{ \mathbb{P}[\imath_{US_l;X_l}(US_l;X_l) \le \log(I_0 - 1) + \gamma] + \exp(-\gamma) \}. \tag{19}$$

Proof Sketch:
- *Codebook construction*: for each $l = 1, \ldots, m$ define the set

$$\mathcal{I}_l := \{1, \ldots, I_l\}. \tag{20}$$

Construct a codebook $u(i_0, i_1, \ldots, i_m)$, $i_l \in \mathcal{I}_l$, $0 \le l \le m$, where each codeword is generated i.i.d. according to Q_U. Let $\mathcal{I} := \mathcal{I}_0 \times \mathcal{I}_1 \times \cdots \times \mathcal{I}_m$ and $I = |\mathcal{I}|$. For each $i \in \mathcal{I}$ and $1 \le l \le m$, independently generate a codebook

$$[s_l(i, j)]_{j \in \mathcal{J}_l} \tag{21}$$

where each codeword is generated i.i.d. according to $Q_{S_l|U=u(i)}$.
- *Encoding*: define μ_V as the equiprobable distribution on $\mathcal{I}$, and

$$\hat{P}_{Z|V=i} := Q_{Z|U=u(i)}, \quad \forall i; \tag{22}$$

$$\hat{P}_{ZV} := \hat{P}_{Z|V} \mu_V. \tag{23}$$

Moreover for each l, let $\mu_{\tilde{W}_l}$ be the equiprobable distribution on $\mathcal{J}_l$, and define

$$P^{(l)}_{Z|\tilde{W}_l=j V=i} := Q_{Z|S_l=s_l(i,j)U=u(i)}, \forall i, j; \tag{24}$$

$$P^{(l)}_{Z\tilde{W}_l V} := P^{(l)}_{Z|\tilde{W}_l V} \mu_{\tilde{W}_l} \mu_V. \tag{25}$$

Then the encoder is a stochastic map

$$\pi_{V\tilde{W}^m|Z} := \hat{P}_{V|Z} \prod_{l=1}^{m} P^{(l)}_{\tilde{W}_l|VZ} \tag{26}$$

that maps the observation $z \in \mathcal{Z}$ to $v \in \mathcal{I}$ and $\tilde{w}^m \in \mathcal{J}^m$. In other words, we first find v using a likelihood encoder with the likelihood function $\hat{P}_{Z=z|V}$ and then find $\tilde{w}_l$ using a likelihood encoder with the likelihood function $P^{(l)}_{Z=z|\tilde{W}_l V=v}$. Suppose $v = (v_0, v_1, \ldots, v_m)$ where $v_l \in \mathcal{I}_l$, $0 \le l \le m$. We identify $k = v_0$ and $w_l = (\tilde{w}_l, v_1^l)$ as the key for the communicator and the public messages. Note that the second components of w_l have a nested (aligned) structure, which is important for maximizing the key rate.
- *Error analysis*: The main idea is to use the soft covering lemma (c.f. [11, Theorem VII.1] or [13]) iteratively to show that the true distribution $\pi_{\tilde{W}_l VZ}$ is close to $\hat{P}^{(l)}_{\tilde{W}_l VZ}$ in total variation (expected over the codebook). By construction $\tilde{W}_l$ and V are independent under $\hat{P}^{(l)}$, implying that the individual message and the key are also nearly independent under π. Moreover, the decoding error probability of the receivers under $\hat{P}^{(l)}$ can be bounded directly by Shannon's achievability bound [14]. ■

Theorem 9 immediately implies the achievability part of the region (12) in the i.i.d. case: assume without loss of generality that the sources are ordered in such a way that $I(US_l; X_l)$ is non-increasing in l. We then identify $(S^m, U, X_1, X_2, \ldots, X_m, Z)$ in Theorem 9 as the block-coding counterpart $(S^{mn}, U^n, X_1^n, \ldots, X_m^n, Z^n)$ and let δ be exponentially converging to zero as $n \to \infty$, and

$$J_l := \exp(n(I(S_l; X_l|U) + \beta))), \quad l = 1, \ldots, m; \tag{27}$$

$$I_0 := \exp(n(\min\{I(U; Z), I(US_m; X_m)\} - \beta)); \tag{28}$$

$$I_l := \exp(n(\min\{I(U; Z), I(US_{l-1}; X_{l-1})\} - \min\{I(U; Z), I(US_l; X_l)\})), \quad l = 2, \ldots, m; \tag{29}$$

$$I_1 := \exp(n(I(U; Z) - \min\{I(U; Z), I(US_1; X_1)\})) \tag{30}$$

to show the achievability of rates

$$R := \min\{I(U; Z), I(US_m; X_m)\} - \beta; \tag{31}$$

$$R_l := \max\{I(S_l; Z|U), I(US_l; Z|X_l)\} + 3\beta, \quad 1 \le l \le m \tag{32}$$

for $\beta > 0$ arbitrary. This establishes the achievability of

$$\bigcup_{Q_{US^m|Z}} \left\{ \begin{array}{c} (R, R_1, \ldots, R_m): \quad R \le \\ \min\{I(U; Z), I(US_1; X_1), \ldots, I(US_m; X_m)\} \\ R_l \ge \max\{I(S_l; Z|U), I(US_l; Z|X_l)\}, \quad \forall l \end{array} \right\}. \tag{33}$$

Then the achievability of (12) follows by noting that the boundary of (33) can be achieved when S_l is chosen so that the two terms in the max are equal.

IV. CONVERSE

Due to space, this section only presents the main idea for the converse of Theorem 2.

A. Deterministic Encoder

We first consider the case where K and W^m are functions of Z^n (but $\mathcal{X}_l$ are allowed to calculate their keys randomly from $(W_l, X_l{}^n)$, for $1 \leq l \leq m$). Given a key generation scheme, denote by $K, K_1, \ldots, K_m$ the keys produced by $\mathcal{Z}, \mathcal{X}_1, \ldots \mathcal{X}_m$ and $W_1, W_2, \ldots, W_m$ the messages sent to $\mathcal{X}_1, \ldots, \mathcal{X}_m$. Define

$$U_i := (K, Z^{i-1}); \tag{34}$$
$$S_{li} := (W_l, X_l{}^{i-1}), \quad 1 \leq l \leq m, 1 \leq i \leq n \tag{35}$$

and let N be equiprobable on $\{1, \ldots, n\}$ independent of all previously defined random variables. We identify

$$\mathsf{U} = U_N, \quad \mathsf{S}_l = S_{lN}, \quad \forall l, \tag{36}$$

which fulfills that

$$(\mathsf{U}, \mathsf{S}_1, \ldots, \mathsf{S}_m) - \mathsf{Z} - (\mathsf{X}_1, \ldots, \mathsf{X}_m). \tag{37}$$

The bounds in Theorem 2 can be verified using entropic manipulations and Fano's inequality.

B. Stochastic Encoders

The converse for stochastic encoders cannot be obtained by simple modifications of the analysis in IV-A. Indeed, the bound in (12) no longer holds for stochastic encoders if we stick to the assignment of the auxiliary random variables in (34)-(36). An alternative approach is to view a stochastic encoder as a deterministic function of Z^n and V, where V is a random number satisfying $(X_1{}^n, \ldots, X_m{}^n) - Z^n - V$, and then employ the converse for deterministic encoders. We immediately see that any achievable rates $(R, R_1, \ldots, R_m)$ must satisfy

$$R \leq \min\{I(\mathsf{U}; \mathsf{ZV}), I(\mathsf{US}_1; \mathsf{X}_1), \ldots, I(\mathsf{US}_m; \mathsf{X}_m)\}; \tag{38}$$
$$R_l \geq \max\{I(\mathsf{S}_l; \mathsf{ZV}|\mathsf{U}), I(\mathsf{US}_l; \mathsf{ZV}|\mathsf{X}_l)\}, \quad 1 \leq l \leq m \tag{39}$$

for some $P_{\mathsf{USTV}|\mathsf{Z}}$. Then it is possible to show that the region specified by (38)-(39) is equivalent to the region specified in (12) upon optimization.

V. A ZERO-RATE ONE-SHOT CONVERSE

In this section we derive a novel one-shot bound, using hypercontractivity, on the maximum ratio of the log alphabet sizes of the key and the messages such that the key can be successfully generated in the omniscient helper problem. Since this ratio is supremized as the key rate and the communication rates tend to zero, such a converse bound may also be called a *zero-rate* converse. The bound is asymptotically tight in the case of abundant correlated sources but limited communication rates, and gives a *strong converse* as it shows that the total variation between the true and the correct distributions tends

to the maximal value under appropriate rate conditions. On the other hand, previous works have obtained one-shot converses using smooth Renyi entropy [15] or the meta-converse idea [16][17], for which the asymptotic tightness are achieved in the other extreme of limited correlated sources but unlimited communications.

An m-tuple of random variables $(X_1, \ldots, X_m)$ is said to be $(p_1, \ldots, p_m)$-hypercontractive for $p_l \in [1, \infty)$, $l = 1, \ldots, m$ if

$$\mathbb{E}\left[\prod_{l=1}^{m} f_l(X_l)\right] \leq \prod_{l=1}^{m} \|f_l(X_l)\|_{p_l} \tag{40}$$

for all bounded real-valued measurable functions f_l defined on $\mathcal{X}_l$, $l = 1, \ldots, m$. In [18], Nair showed that (40) is equivalent to the following inequality[3]

$$I(U; X^m) \geq \sum_{l=1}^{m} \frac{1}{p_l} I(U; X_l) \tag{41}$$

being valid for all $P_{U|X^m}$. Thus from Theorem 5 and (41), key generation cannot be accomplished asymptotically if

$$R < \sum_{l=1}^{m} \frac{1}{p_l}(R - R_l); \tag{42}$$

while if $r_1, \ldots, r_m$ satisfies the property that $1 \geq \sum_{l=1}^{m} \frac{1}{p_l}(1 - r_l)$ for all $(p_1, \ldots, p_m)$ such that $(X_1, \ldots, X_m)$ is $(p_1, \ldots, p_m)$-hypercontractive, then there exists $(R, R_1, R_2, \ldots, R_m)$ achievable such that $\frac{R_l}{R} = r_l$ for each $l = 1, \ldots, m$.

We prove a zero-rate one-shot converse for the omniscient helper problem. Consider the one-shot case. Suppose the (possibly stochastic) encoder for the public messages is specified by $P_{W^m|X^m}$ and the (possibly stochastic) decoder for the key is given by $\prod_{l=1}^{m} P_{K_l|X_lW_l}$. Let μ_{K^m} be the *correct* distribution under which $K_1 = K_2 = \cdots = K_m$ is equiprobably distributed on $\mathcal{K}$. Clearly, a small total variation $|P_{K^m} - \mu_{K^m}|$ implies both uniformity of the key distribution and a small probability of key disagreement.

Theorem 10. *In the omniscient helper problem, if the source X^m is $(p_1, \ldots, p_m)$-hypercontractive,[4] then*

$$\frac{1}{2}|P_{K^m} - \mu_{K^m}| \geq 1 - \frac{1}{|\mathcal{K}|} - \left[|\mathcal{K}|\prod_{l=1}^{m}\left(\frac{|\mathcal{W}_l|}{|\mathcal{K}|}\right)^{\frac{1}{p_l}}\right]^{\frac{1}{\sum p_l^{-1}}} \tag{43}$$

Remark 11. Theorem 10 only concerns the performance of CR generation, which will provide an obvious upper bound on the performance of key generation. For the omniscient helper problem, it turns out to be tight because the highest key-communication ratio is achieved with small rates (by convexity of the achievable region), in which regime the

---

[3]In [18] the equivalence is demonstrated for $m = 2$, but the method therein can be easily extended to the $m > 2$ case.

[4]In the i.i.d. case this is equivalent to the per-letter source X^m being $(p_1, \ldots, p_m)$-hypercontractive by the tensorization property [18].

secrecy constraint does not require higher communication rates (Remark 8).

Remark 12. Theorem 10 yields a stronger converse on the achievable ratio of the the log alphabet sizes of the key and the messages than Theorem 5, because:

- The converse from Theorem 5 is vacuous when the rates are zero. In contrast, Theorem 10 is still applicable when the log size of the key alphabet grows sub-linearly in the blocklength. In fact, as long as

$$\log|\mathcal{K}| - \sum_{l=1}^{m} \frac{1}{p_l}(\log|\mathcal{K}| - \log|\mathcal{W}_l|) \to -\infty \quad (44)$$

which is weaker than (42), Theorem 10 implies that $|P_{K^m} - \mu_{K^m}|$ converges to 2.

- Even if (42) holds, the converse of Theorem 5 relying on Fano's inequality does not guarantee that the error probability in (6) tends to 1. Moreover Theorem 5 uses relative entropy as the secrecy measure (7) (stronger than total variation), amounting to a weaker converse.

Proof: For any $k \in \mathcal{K}$,

$$\mathbb{P}\left[\bigcap_{l=1}^{m}\{K_l = k\}\right] \quad (45)$$

$$= \int_{\mathcal{X}^m} \sum_{w^m} \prod_{l=1}^{m} P_{K_l=k|X_lW_l=w_l} P_{W^m|X^m} \mathrm{d}P_{X^m} \quad (46)$$

$$\leq \int_{\mathcal{X}^m} \max_{w^m} \prod_{l=1}^{m} P_{K_l=k|X_lW_l=w_l} \mathrm{d}P_{X^m} \quad (47)$$

$$\leq \int_{\mathcal{X}^m} \prod_{l=1}^{m} \max_{w_l} P_{K_l=k|X_lW_l=w_l} \mathrm{d}P_{X^m} \quad (48)$$

$$\leq \prod_{l=1}^{m} \left[\int_{\mathcal{X}_l} (\max_{w_l} P_{K_l=k|X_lW_l=w_l})^{p_l} \mathrm{d}P_{X_l}\right]^{\frac{1}{p_l}} \quad (49)$$

$$\leq \prod_{l=1}^{m} \left[\int_{\mathcal{X}_l} \max_{w_l} P_{K_l=k|X_lW_l=w_l} \mathrm{d}P_{X_l}\right]^{\frac{1}{p_l}} \quad (50)$$

$$\leq \prod_{l=1}^{m} \left[\sum_{w_l} \int_{\mathcal{X}_l} P_{K_l=k|X_lW_l=w_l} \mathrm{d}P_{X_l}\right]^{\frac{1}{p_l}} \quad (51)$$

where

- (49) uses the definition of hypercontractivity;
- (50) uses $p_l > 1$ and $\max_{w_l} P_{K_l=k|X_lW_l=w_l} \leq 1$.

Raising both sides of (51) to the power of $\frac{1}{\sum_{i=1}^{m} p_i^{-1}}$, we obtain

$$\mathbb{P}\left[\bigcap_{l=1}^{m}\{K_l = k\}\right]^{\frac{1}{\sum_i p_i^{-1}}}$$

$$\leq \prod_{l=1}^{m} \left[\sum_{w_l} \int_{\mathcal{X}_l} P_{K_l=k|X_lW_l=w_l} \mathrm{d}P_{X_l}\right]^{\frac{p_l^{-1}}{\sum_i p_i^{-1}}} \quad (52)$$

But the function $t^m \mapsto \prod_{l=1}^{m} t_l^{\frac{p_l^{-1}}{\sum_i p_i^{-1}}}$ is a concave function on $[0, \infty)^m$, so by Jensen's inequality,

$$\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\prod_{l=1}^{m}\left[\sum_{w_l} \int_{\mathcal{X}_l} P_{K_l=k|X_lW_l=w_l} \mathrm{d}P_{X_l}\right]^{\frac{p_l^{-1}}{\sum_i p_i^{-1}}} \quad (53)$$

$$\leq \prod_{l=1}^{m}\left[\sum_{w_l} \int_{\mathcal{X}_l} \frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|} P_{K_l=k|X_lW_l=w_l} \mathrm{d}P_{X_l}\right]^{\frac{p_l^{-1}}{\sum_i p_i^{-1}}} \quad (54)$$

$$= \prod_{l=1}^{m}\left[\sum_{w_l} \int_{\mathcal{X}_l} \frac{1}{|\mathcal{K}|} \mathrm{d}P_{X_l}\right]^{\frac{p_l^{-1}}{\sum_i p_i^{-1}}} \quad (55)$$

$$= \prod_{l=1}^{m}\left(\frac{|\mathcal{W}_l|}{|\mathcal{K}|}\right)^{\frac{p_l^{-1}}{\sum_i p_i^{-1}}} \quad (56)$$

Combining (52) and (56) we obtain

$$\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\mathbb{P}\left[\bigcap_{l=1}^{m}\{K_l = k\}\right]^{\frac{1}{\sum_i p_i^{-1}}} \leq \prod_{l=1}^{m}\left(\frac{|\mathcal{W}_l|}{|\mathcal{K}|}\right)^{\frac{p_l^{-1}}{\sum_i p_i^{-1}}}. \quad (57)$$

Finally we invoke the following elementary bound:

$$\frac{1}{2}|P_{K^m} - \mu_{K^m}|$$

$$= \sum_{k=1}^{|\mathcal{K}|}\left|\mathbb{P}\left[\bigcap_{l=1}^{m}\{K_l = k\}\right] - \frac{1}{|\mathcal{K}|}\right| + 1 - \sum_{k=1}^{|\mathcal{K}|}\mathbb{P}\left[\bigcap_{l=1}^{m}\{K_l = k\}\right] \quad (58)$$

$$\geq 1 - \frac{1}{|\mathcal{K}|} - |\mathcal{K}|^{\frac{1}{\sum_l p_l^{-1}}-1}\sum_{k=1}^{|\mathcal{K}|}\mathbb{P}\left[\bigcap_{l=1}^{m}\{K_l = k\}\right]^{\frac{1}{\sum_l p_l^{-1}}} \quad (59)$$

and the proof is finished by combining (57) and (58). ∎

VI. DISCUSSION

It remains an enticing problem for future research to find out whether the achievable region is changed if we further require that the key has to be independent of all messages, instead of each message individually (see (7)). Such a stronger secrecy constraint is relevant when a powerful eavesdropper is able to intercept the messages to all the receivers. Our achievability proof does not guarantee this stronger level of secrecy, but for some specific sources it is possible to use structured codes to align different sub-codebooks so that the achievable rates do not change. Furthermore, in the unlimited communication case the key rate is not compromised by the stronger requirement either; see (3). Generally, an inner bound can be obtained by replacing S_l in (12) with S^l, the proof of which involves a vertical structure of superposition codebooks for $U, S_1, ..., S_m$, in contrast to the parallel structure of $S_1, ..., S_m$ in the achievability proof of Theorem 2.

Theorem 10 also gives an asymptotically tight strong converse bound for the canonical one-way protocol (Model S with forward communication in [2]). By setting $m = 2$ and

$R_1 = 0$, the resulting model immediately gives an upper-bound on the performance of a CR generation model where $\mathcal{X}_1$ communicates to $\mathcal{X}_2$. This in turn bounds performance of one-way key generation model because the public communication can be used as part of the CR. In the end we can show that the TV between P_{KW} and the correct distribution $\mu_{KW} = \mu_K \mu_W$ tends to 2 if $\log|\mathcal{K}| - \frac{s^*(X_1;X_2)}{1-s^*(X_1;X_2)}\log|\mathcal{W}| \to \infty$, where $s^*(X_1;X_2)$ is the strong data processing coefficient [18].

VII. Acknowledgments

References

[1] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Transactions on Information Theory*, vol. 39, pp. 733–742, Mar. 1993.

[2] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography. I. Secret sharing," *IEEE Transactions on Information Theory*, vol. 39, pp. 1121–1132, Apr. 1993.

[3] I. Csiszár and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Transactions on Information Theory*, vol. 46, pp. 344–366, Feb. 2000.

[4] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Transactions on Information Theory*, vol. 50, pp. 3047–3061, Dec. 2004.

[5] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography. II. CR capacity," *IEEE Transactions on Information Theory*, vol. 44, pp. 225–240, Jan. 1998.

[6] V. M. Prabhakaran and M. M. Prabhakaran, "Assisted common information with an application to secure two-party sampling," *IEEE Transactions on Information Theory*, vol. 60, no. 6, pp. 3413–3434, 2014.

[7] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, 1973.

[8] T. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 2–14, 1972.

[9] E. C. Song, P. Cuff, and H. V. Poor, "The likelihood encoder for lossy source compression," *arXiv:1408.4522*, Aug. 2014.

[10] *Te Sun Han, Information-Spectrum Method in Information Theory*. Springer, 2003.

[11] P. Cuff, "Distributed channel synthesis," *IEEE Transactions on Information Theory*, vol. 59, pp. 7071-7096, Nov. 2013.

[12] S. Verdú, "Non-asymptotic achievability bounds in multiuser information theory," in *50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1-8, 2012.

[13] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel," *IEEE Transactions on Information Theory*, vol. 52, pp. 1562-1575, Apr. 2006.

[14] C. E. Shannon, "Certain results in coding theory for noisy channels," *Information and Control*, vol. 1, pp. 6–25, Jan. 1957.

[15] R. Renner and S. Wolf, "Simple and tight bounds for information reconciliation and privacy amplification," in *Advances in Cryptology-ASIACRYPT 2005*, pp. 199–216, Springer, 2005.

[16] M. Hayashi, H. Tyagi, and S. Watanabe, "Secret key agreement: General capacity and second-order asymptotics," in *Proceedings of IEEE International Symposium on Information Theory (ISIT)*, pp. 1136–1140, 2014.

[17] H. Tyagi and S. Watanabe, "A bound for multiparty secret key agreement and implications for a problem of secure computing," in *Advances in Cryptology–EUROCRYPT 2014*, pp. 369–386, Springer, 2014.

[18] C. Nair, "Equivalent formulations of hypercontractivity using information measures," *IZS Workshop, 2014*, Feb. 2014.