Converses For Secret Key Agreement and Secure Computing

Himanshu Tyagi *Member, IEEE* and Shun Watanabe *Member, IEEE*

***Abstract*—We consider information theoretic secret key agreement and secure function computation by multiple parties observing correlated data, with access to an interactive public communication channel. Our main result is an upper bound on the secret key length, which is derived using a reduction of binary hypothesis testing to multiparty secret key agreement. Building on this basic result, we derive new converses for multiparty secret key agreement. Furthermore, we derive converse results for the oblivious transfer problem and the bit commitment problem by relating them to secret key agreement. Finally, we derive a necessary condition for the feasibility of secure computation by trusted parties that seek to compute a function of their collective data, using an interactive public communication that by itself does not give away the value of the function. In many cases, we strengthen and improve upon previously known converse bounds. Our results are single-shot and use only the given joint distribution of the correlated observations. For the case when the correlated observations consist of independent and identically distributed (in time) sequences, we derive *strong* versions of previously known converses.**

I. Introduction

Information theoretic cryptography relies on the availability of correlated random observations to the parties. Neither multiparty *secret key* (SK) agreement nor secure computation is feasible if the observation of the parties are mutually independent. In fact, SK agreement is not feasible even when the observations are independent across some partition of the set of parties[1]. As an extension of this principle, we can expect that the efficiency of a cryptographic primitive is related to "how far" the joint distribution of the observations is from a distribution that renders the observations independent (across some partition of the set of parties). We formalize this heuristic principle and leverage it to bound the efficiency of using correlated sources to implement SK agreement and secure computation. We present *single-shot* converse results; in particular, we do not assume that the observations of parties consist of long sequences generated by an *independent and identically distributed* (IID) random process[2].

In multiparty SK agreement, a set of parties observing correlated *random variables* (RVs) seek to agree on shared random bits that remain concealed from an eavesdropper with access to a correlated side information. The parties may communicate with each other over a noiseless public channel, but the transmitted communication will be available to the eavesdropper. The main tool for deriving our converse results is a reduction argument that relates multiparty SK agreement to binary hypothesis testing[3]. For an illustration of our main idea, consider the two party case when the eavesdropper observes only the communication between the legitimate parties and does not observe any additional side information. Clearly, if the observations of the legitimate parties are independent, a SK cannot be generated. We upper bound the length of SKs that can be generated in terms of "how far" is the joint distribution of the observations of the parties from a distribution that renders their observations independent. Specifically, for this special case, we show that the maximum length $S_\epsilon\left(X_1, X_2\right)$ of a SK (for a given secrecy index ϵ) is bounded above as

$$S_\epsilon\left(X_1, X_2\right) \leq -\log \beta_{\epsilon+\eta}\big(\mathrm{P}_{X_1X_2}, \mathrm{P}_{X_1} \times \mathrm{P}_{X_2}\big) + 2\log(1/\eta),$$

where $\beta_\epsilon\big(\mathrm{P}_{X_1X_2}, \mathrm{P}_{X_1} \times \mathrm{P}_{X_2}\big)$ is the optimal probability of error of type II for testing the null hypothesis $\mathrm{P}_{X_1X_2}$ with the alternative $\mathrm{P}_{X_1} \times \mathrm{P}_{X_2}$, given that the probability of error of type I is smaller than ϵ; this β_ϵ serves as a proxy for "distance" between $\mathrm{P}_{X_1X_2}$ and $\mathrm{P}_{X_1} \times \mathrm{P}_{X_2}$. Similarly, in the general case of an arbitrary number of parties with correlated side information at the eavesdropper, our main result in Theorem 3 bounds the secret key length in terms of the "distance" between the joint distribution of the observations of the parties and the eavesdropper and a distribution that renders the observations of the parties conditionally independent across some partition, when conditioned on the eavesdropper's side information. This bound is a manifestation of the aforementioned heuristic principle and is termed the *conditional independence testing* bound.

Our approach brings out a structural connection between SK agreement and binary hypothesis testing[4]. This is in the spirit of [52], where a connection between channel coding and binary hypothesis testing was used to establish an upper bound on the rate of good channel codes (see, also, [75], [28]). Also, our upper bound is reminiscent of the *measure*

Parts of this paper appeared in the proceedings of EUROCRYPT 2014 and of IEEE International Symposium on Information Theory 2015.

H. Tyagi is with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560012, India. Email: htyagi@ece.iisc.ernet.in

S. Watanabe is with the Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology, Tokyo 184-8588, Japan. Email: shunwata@cc.tuat.ac.jp

[1] With restricted interpretations of *feasibility*, these observations appear across the vast literature on SK agreement and secure computation; see, for instance, [44], [1], [16], [60], [40], [80], [48].

[2] Throughout this paper, IID observations refer to observations that are IID *in time*; at each instant t, the observations of the parties are correlated.

[3] This basic result was reported separately in [73].

[4] While the connection between SK agreement and hypothesis testing established in this paper is new, a similar connection between authentication and hypothesis testing is natural to expect and is well-known (see, for instance, [45]).

of entanglement for a quantum state proposed in [74], namely the minimum distance between the density matrix of the state and that of a disentangled state. This measure of entanglement was shown to be an upper bound on the entanglement of distillation in [74], where the latter is the largest proportion of maximally entangled states that can be distilled using a purification process [6].

Using our basic result, we obtain new converses for SK agreement, and also, for secure two-party computation by reducing SK agreement to oblivious transfer and bit commitment. In many cases, we strengthen and improve upon previously known results. Our main contributions are summarized below.

A. Secret key agreement

For two parties, the problem of SK agreement from correlated observations is well-studied. The problem was introduced by Maurer [44] and Ahlswede and Csiszár [1], who considered the case where the parties observe IID sequences. However, in certain applications it is of interest to consider observations arising from a single realization of correlated RVs. For instance, in applications such as biometric and hardware authentication (cf. [51], [20]), the correlated observations consist of different versions of the biometric and hardware signatures, respectively, recorded at the registration and the authentication stages. To this end, Renner and Wolf [60] derived bounds on the length of a SK that can be generated by two parties observing a single realization of correlated RVs, using one-side communication.

The problem of SK agreement with multiple parties, for the IID setup, was introduced in [16] (see also [9]). In this work, we consider the SK agreement problem for multiple parties observing a single realization of correlated RVs. Our conditional independence testing bound is a single-shot upper bound on the length of SKs that can be generated by multiple parties observing correlated data, using interactive public communication[5]. Unlike the single-shot upper bound in [60], which is restricted to two parties with one-way communication, we allow arbitrary interactive communication between multiple parties. Asymptotically our bound is tight – its application to the IID case recovers some previously known (tight) bounds on the asymptotic SK rates. In fact, we strengthen the previously known asymptotic results since we do not require the probability of error in SK agreement or the secrecy index to be asymptotically[6] 0. See Section IV for a detailed discussion.

B. Secure two-party computation

The problem of secure two-party computation was introduced by Yao in [83]. Two (mutually untrusting) parties seek to compute a function of their collective data, without sharing anything more about their data than what is given away by the function value. Several specific instances of this general problem have been studied. We consider the problems of *oblivious transfer* (OT) and *bit commitment* (BC), which constitute two basic primitives for secure two-party computation.

OT between two parties is a mode of message transmission "where the sender does not know whether the recipient actually received the information" [55]. In this paper, we consider the *one-out-of-two OT* problem [21] where the first party observes two strings K_0 and K_1 of length l each, and the second party seeks the value of the Bth string, $B \in \{0, 1\}$. The goal is to accomplish this task in such a manner that B and $K_{\overline{B}}$ remain concealed, respectively, from Party 1 and Party 2. This simply stated problem is at the heart of secure function computation as it is well-known [39] that any secure function computation task can be accomplished using the basic OT protocol repeatedly (for recent results on the complexity of secure function computation using OT, see [4]). Unfortunately, information theoretically secure OT is not feasible in the absence of additional resources. On the bright side, if the parties share a noisy communication channel or if they observe correlated randomness, OT can be accomplished (cf. [12], [13], [2], [48]). In this paper, we consider the latter case where, as an additional resource, the parties observe correlated RVs X_1 and X_2. Based on reduction arguments relating OT to SK agreement, we derive upper bounds on the length l of OT that can be accomplished for given RVs X_1, X_2. The resulting bound is, in general, tighter than that obtained in [79]. Furthermore, an application of our bound to the case of IID observations shows that the upper bound on the rate of OT length derived in [48] and [2][7] is strong, $i.e.$, the bound holds even without requiring asymptotically perfect recovery.

We now turn to the BC problem, the first instance of which was introduced by Blum in [7] as the problem of flipping a coin over a telephone, when the parties do not trust each other. A bit commitment protocol has two phases. In the first phase the committing party generates a random bit string K, its "coin flip". Subsequently, the two parties communicate with each other, which ends the first phase. In the second phase, the committing party reveals K. A bit commitment protocol must forbid the committing party from cheating and changing K in the second phase. As in the case of OT, information theoretically secure BC is not possible without additional resources. We consider a version where two parties observing correlated observations X_1 and X_2 want to implement information theoretically secure BC using interactive public communication. The goal is to maximize the length of the committed string K. By reducing SK agreement to BC, we derive an upper bound on BC length which improves upon the bound in [56]. Furthermore, for the case of IID observations, we derive a strong converse for BC capacity; the latter is the maximum rate of BC length and was characterized in [80].

C. Secure computation with trusted parties

In a different direction, we relate our result to the following problem of *secure function computation with trusted parties*

[5] A single-shot upper bound using Fano's inequality for the length of a multiparty SK, obtained as a straightforward extension of [16], [17], was reported in [72].

[6] Such bounds that do not require the probability of error to vanish to 0 are called *strong converse* bounds [15].

[7] The asymptotic bound in [2] was generalized in [57] usin the tension-bound technique introduced in [54]. It is not clear if our approach can derive a single-shot version of the general bound in [57].

introduced in [69] (for an early version of the problem, see [50]): Multiple parties observing correlated data seek to compute a function of their collective data. To this end, they communicate interactively over a public communication channel, which is assumed to be authenticated and error-free. It is required that the value of the function be concealed from an eavesdropper with access to the communication. When is such a secure computation of a given function feasible? In contrast to the traditional secure computation problem discussed above, this setup is appropriate for applications such as sensor networks where the legitimate parties are trusted and are free to extract any information about each other's data from the shared communication. Using the conditional independence testing bound, we derive a necessary condition for the existence of a communication protocol that allows the parties to reliably recover the value of a given function, while keeping this value concealed from an eavesdropper with access to (only) the communication. In [69], matching necessary and sufficient conditions for secure computability of a given function were derived for the case of IID observations. In contrast, our necessary condition for secure computability is single-shot and does not rely on the observations being IID.

D. Outline of paper

The next section reviews some basic concepts that will be used throughout this work. The conditional independence testing bound is derived in Section III. In the subsequent three sections, we present the implications of this bound: Section IV addresses strong converses for SK capacity; Section V addresses converse results for the OT and the BC problem; and Section VI contains converse results for the secure computation problem with trusted parties. The final section contains a brief discussion of possible extensions.

E. Notations

For brevity, we use abbreviations SK, RV, and IID for secret key, random variable, and independent and identically distributed, respectively; a plural form will be indicated by appending an 's' to the abbreviation. The RVs are denoted by capital letters and the corresponding range sets are denoted by calligraphic letters. The distribution of a RV U is given by P_U, when there is no confusion we drop the subscript U. The set of all parties $\{1, ..., m\}$ is denoted by $\mathcal{M}$. For a collection of RVs $\{U_1, .., U_m\}$ and a subset A of $\mathcal{M}$, U_A denotes the RVs $\{U_i,\, i \in A\}$. For a RV U, U^n denotes n IID repetitions of the RV U. Similarly, P^n denotes the distribution corresponding to the n IID repetitions generated from P. All logarithms in this paper are to the base 2.

II. Preliminaries

A. Secret keys

Consider SK agreement using interactive public communication by m (trusted) parties. The ith party observes a discrete RV X_i taking values in a finite set $\mathcal{X}_i$, $1 \le i \le m$.[8] Upon making these observations, the parties communicate interactively over a public communication channel that is accessible by an eavesdropper, who additionally observes a RV Z such that the RVs $(X_{\mathcal{M}}, Z)$ have a distribution $\mathrm{P}_{X_{\mathcal{M}}Z}$. We assume that the communication is error-free and each party receives the communication from every other party. Furthermore, we assume that the public communication is authenticated and the eavesdropper cannot tamper with it. Specifically, the communication is sent over r rounds of interaction. In the jth round of communication, $1 \le j \le r$, the ith party sends F_{ij}, which is a function of its observation X_i, a *locally generated* randomness[9] U_i and the previously observed communication

$$F_{11}, ..., F_{m1}, F_{12}, ..., F_{m2}, ..., F_{1j}, ..., F_{(i-1)j}.$$

The overall interactive communication $F_{11}, ..., F_{m1}, ..., F_{1r}, ..., F_{mr}$ is denoted by $\mathbf{F}$. Using their local observations and the interactive communication $\mathbf{F}$, the parties agree on a SK.

Formally, a SK is a collection of RVs $K_1, ..., K_m$, where the ith party gets K_i, that agree with probability close to 1 and are concealed, in effect, from an eavesdropper. Formally, the ith party computes a function K_i of $(U_i, X_i, \mathbf{F})$. Traditionally, the RVs $K_1, ..., K_m$ with a common range $\mathcal{K}$ constitute an (ϵ, δ)-SK if the following two conditions are satisfied (for alternative definitions of secrecy, see [44], [14], [16])

$$\mathrm{P}\left(K_1 = \cdots = K_m\right) \ge 1 - \epsilon, \tag{1}$$

$$d\left(\mathrm{P}_{K_1\mathbf{F}Z}, \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_{\mathbf{F}Z}\right) \le \delta, \tag{2}$$

where $\mathrm{P}_{\mathtt{unif}}$ is the uniform distribution on $\mathcal{K}$ and $d\left(\mathrm{P}, \mathrm{Q}\right)$ is the variational distance between P and Q given by

$$d\left(\mathrm{P}, \mathrm{Q}\right) = \frac{1}{2}\sum_x |\mathrm{P}\left(x\right) - \mathrm{Q}\left(x\right)|.$$

The first condition above represents the reliable *recovery* of the SK and the second condition guarantees *secrecy*. In this work, we use the following alternative definition of a SK, which conveniently combines the recoverability and the secrecy conditions (cf. [58]): The RVs $K_1, ..., K_m$ above constitute an ϵ-SK with common range $\mathcal{K}$ if

$$d\left(\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}, \mathrm{P}^{(\mathcal{M})}_{\mathtt{unif}} \times \mathrm{P}_{\mathbf{F}Z}\right) \le \epsilon, \tag{3}$$

where

$$\mathrm{P}^{(\mathcal{M})}_{\mathtt{unif}}\left(k_{\mathcal{M}}\right) = \frac{\mathbb{1}(k_1 = \cdots = k_m)}{|\mathcal{K}|}.$$

In fact, the two definitions above are closely related[10].

Proposition 1. *Given $0 \le \epsilon, \delta < 1$, if $K_{\mathcal{M}}$ constitute an (ϵ, δ)-SK under (1) and (2), then they constitute an $(\epsilon + \delta)$-SK under (3).*

Conversely, if $K_{\mathcal{M}}$ constitute an ϵ-SK under (3), then they

[8] The conditional independence testing bound given in Theorem 3 remains valid even for continuous valued RVs. However, in general, the resulting bound may not be achievable.

[9] The RVs $U_1, ..., U_m$ are mutually independent and independent jointly of $(X_{\mathcal{M}}, Z)$.

[10] Note that a SK agreement protocol that satisfies (3) *universally composable-emulates* an ideal SK agreement protocol (see [8] for a definition). The emulation is with emulation slack ϵ, for an environment of unbounded computational complexity.

constitute an (ϵ,ϵ)-SK under (1) and (2).

Therefore, by the composition theorem in [8], the complex cryptographic protocols using such SKs instead of perfect SKs are secure.[11]

We are interested in characterizing the maximum length $\log|\mathcal{K}|$ of an ϵ-SK.

Definition 1. Given $0 \le \epsilon < 1$, denote by $S_\epsilon\left(X_\mathcal{M}|Z\right)$ the maximum length $\log|\mathcal{K}|$ of an ϵ-SK $K_\mathcal{M}$ with common range $\mathcal{K}$.

Our upper bound is based on relating the SK agreement problem to a binary hypothesis testing problem; below we review some basic concepts in hypothesis testing that will be used.

B. Hypothesis testing

Consider a binary hypothesis testing problem with null hypothesis P and alternative hypothesis Q, where P and Q are distributions on the same alphabet $\mathcal{X}$. Upon observing a value $x \in \mathcal{X}$, the observer needs to decide if the value was generated by the distribution P or the distribution Q. To this end, the observer applies a stochastic test T, which is a conditional distribution on $\{0,1\}$ given an observation $x \in \mathcal{X}$. When $x \in \mathcal{X}$ is observed, the test T chooses the null hypothesis with probability $\mathrm{T}(0|x)$ and the alternative hypothesis with probability $\mathrm{T}(1|x) = 1 - \mathrm{T}(0|x)$. For $0 \le \epsilon < 1$, denote by $\beta_\epsilon(\mathrm{P},\mathrm{Q})$ the infimum of the probability of error of type II given that the probability of error of type I is less than ϵ, *i.e.*,

$$\beta_\epsilon(\mathrm{P},\mathrm{Q}) := \inf_{\mathrm{T}\,:\,\mathrm{P}[\mathrm{T}]\ge 1-\epsilon} \mathrm{Q}[\mathrm{T}], \tag{4}$$

where

$$\begin{aligned} \mathrm{P}[\mathrm{T}] &= \sum_x \mathrm{P}(x)\mathrm{T}(0|x), \\ \mathrm{Q}[\mathrm{T}] &= \sum_x \mathrm{Q}(x)\mathrm{T}(0|x). \end{aligned}$$

We note two important properties of the quantity $\beta_\epsilon(\mathrm{P},\mathrm{Q})$.

1) **Data processing inequality.** Let W be a stochastic mapping from $\mathcal{X}$ to $\mathcal{Y}$, *i.e.*, for each $x \in \mathcal{X}$, $W(\cdot|x)$ is a distribution on $\mathcal{Y}$. Then,
$$\beta_\epsilon(\mathrm{P},\mathrm{Q}) \le \beta_\epsilon(\mathrm{P}\circ W, \mathrm{Q}\circ W), \tag{5}$$
where $(\mathrm{P}\circ W)(y) = \sum_x \mathrm{P}\left(x\right) W(y|x)$.
2) **Stein's Lemma.** (cf. [43, Theorem 3.3]) For every $0 < \epsilon < 1$, we have
$$\lim_{n\to\infty} -\frac{1}{n}\log\beta_\epsilon(\mathrm{P}^n,\mathrm{Q}^n) = D(\mathrm{P}\|\mathrm{Q}), \tag{6}$$
where $D(\mathrm{P}\|\mathrm{Q})$ is the Kullback-Leibler divergence given by
$$D(\mathrm{P}\|\mathrm{Q}) = \sum_{x\in\mathcal{X}} \mathrm{P}(x)\log\frac{\mathrm{P}(x)}{\mathrm{Q}(x)},$$
with the convention $0\log(0/0) = 0$.

C. Remarks on evaluation of $\beta_\epsilon(\mathrm{P},\mathrm{Q})$

We close with a discussion on evaluating $\beta_\epsilon(\mathrm{P},\mathrm{Q})$. Note that the expression for $\beta_\epsilon(\mathrm{P},\mathrm{Q})$ in (4) is a linear program, solving which has a polynomial complexity in the size of the observation space. A simple manipulation yields the following computationally more tractable bound:

$$-\log\beta_\epsilon(\mathrm{P},\mathrm{Q}) \le \inf_\gamma \gamma - \log\left(\mathrm{P}_\gamma - \epsilon\right), \tag{7}$$

where

$$\mathrm{P}_\gamma = \mathrm{P}\left(\left\{x : \log\frac{\mathrm{P}(x)}{\mathrm{Q}(x)} \le \gamma\right\}\right).$$

When P and Q correspond to IID RVs, the tail probability in (7) can be numerically evaluated directly or can be approximated by the Béry-Esséen theorem (cf. [22]). On the other hand, numerical evaluation of the tail probability is rather involved when P and Q correspond to Markov chains. For this case, a computationally tractable and asymptotically tight bound on $\beta_\epsilon(\mathrm{P},\mathrm{Q})$ was established recently in [77]. Also, by setting $\gamma = D_\alpha(\mathrm{P},\mathrm{Q}) + \frac{1}{1-\alpha}\log(1-\epsilon-\epsilon')$, where $D_\alpha(\mathrm{P},\mathrm{Q})$ is the Rényi's divergence of order $\alpha > 1$ and is given by [61]

$$D_\alpha(\mathrm{P},\mathrm{Q}) = \frac{1}{\alpha-1}\log\sum_{x\in\mathcal{X}}\mathrm{P}(x)^\alpha\mathrm{Q}(x)^{1-\alpha},$$

the following simple bound on $\beta_\epsilon(\mathrm{P},\mathrm{Q})$ is obtained[12]:

$$-\log\beta_\epsilon(\mathrm{P},\mathrm{Q}) \le D_\alpha(\mathrm{P},\mathrm{Q}) + \frac{1}{\alpha-1}\log\frac{1}{1-\epsilon-\epsilon'} + \log\frac{1}{\epsilon'}. \tag{8}$$

A variant of this bound for the case of quantum observations was reported in [49, Theorem 1] (see, also, [27, Eqn. (2.63)]). For the classical case, the bound follows from the simple proof below: Denote by A_γ the set $\{x : \log\mathrm{P}(x)/\mathrm{Q}(x) \le \gamma\}$. Thus, for $\alpha > 1$,

$$\begin{aligned} 1 - \mathrm{P}_\gamma &= \sum_{x\in A_\gamma^c}\mathrm{P}(x) \\ &= \sum_{x\in A_\gamma^c}\mathrm{P}(x)^\alpha\mathrm{P}(x)^{1-\alpha} \\ &< \sum_{x\in A_\gamma^c}\mathrm{P}(x)^\alpha\mathrm{Q}(x)^{1-\alpha}2^{(1-\alpha)\gamma} \\ &\le 2^{(\alpha-1)D_\alpha(\mathrm{P},\mathrm{Q})+(1-\alpha)\gamma} \\ &= (1-\epsilon-\epsilon'), \end{aligned}$$

which further implies that $\mathrm{P}_\gamma > \epsilon + \epsilon'$. The bound (8) follows from (7). Note that while the bound (8) is not tight in general, as its corollary we obtain Stein's lemma (see (6)).

Finally, we remark that when the condition

$$\log\frac{\mathrm{P}(X)}{\mathrm{Q}(X)} = D(\mathrm{P}\|\mathrm{Q}) \tag{9}$$

is satisfied with probability 1 under P, the bound in (7) implies

$$-\log\beta_\epsilon(\mathrm{P},\mathrm{Q}) \le D(\mathrm{P}\|\mathrm{Q}) + \log(1/(1-\epsilon)). \tag{10}$$

[11] A perfect SK refers to unbiased shared bits that are independent of eavesdropper's observations.

[12] For other connections between β_ϵ and Rényi's divergence, see [27, Eqns. (3.37) and (3.38)], [53, Eqn. (29)].

D. Smooth min-entropy and smooth max-divergence

Given two RVs X and Y, a central question of information theoretic secrecy is (cf. [34], [35], [5]): How many unbiased, independent bits can be extracted from X that are unavailable to an observer of Y? When the underlying distribution is IID, the optimum rate of extracted bits can be expressed in terms of Shannon entropies and is given by $H(X|Y)$. However, for our single-shot setup, *smooth min-entropy* introduced in [60], [58] is a more relevant measure of randomness. We use the definition of smooth min-entropy introduced[13] in [58]; for a review of other variations, see [63].

We also review the *leftover hash lemma* [34], [5], which brings out the central role of *smooth min-entropy* in the answer to the question above. Also, as a "change of measure companion" for smooth min-entropy, we define *smooth max-divergence* and note that it satisfies the data processing inequality.

Definition 2. (min-entropy) The min-entropy of P is defined as

$$H_{\min}(\mathrm{P}) := \min_x \log \frac{1}{\mathrm{P}\left(x\right)}.$$

For distributions P_{XY} and Q_Y, the conditional min-entropy of P_{XY} given Q_Y is defined as

$$H_{\min}(\mathrm{P}_{XY}|\mathrm{Q}_Y) := \min_{x\,\in\,\mathcal{X},\,y\,\in\,\mathtt{supp}(\mathrm{Q}_Y)} \log \frac{\mathrm{Q}_Y\left(y\right)}{\mathrm{P}_{XY}\left(x,y\right)}.$$

Finally, the conditional min-entropy[14] of P_{XY} given Y is defined as

$$H_{\min}(\mathrm{P}_{XY}|Y) := \sup_{\mathrm{Q}_Y} H_{\min}(\mathrm{P}_{XY}|\mathrm{Q}_Y), \tag{11}$$

where the $\sup$ is over all Q_Y such that $\mathtt{supp}(\mathrm{P}_Y) \subseteq \mathtt{supp}(\mathrm{Q}_Y)$.

Note that

$$\begin{aligned}
&H_{\min}(\mathrm{P}_{XY}|Y)\\
&= -\inf_{\mathrm{Q}_Y}\max_{x,y}\log\frac{\mathrm{P}_{XY}\left(x,y\right)}{\mathrm{Q}_Y\left(y\right)}\\
&= -\inf_{\mathrm{Q}_Y}\max_{y}\log\frac{\mathrm{P}_Y\left(y\right)\max_x \mathrm{P}_{X|Y}\left(x|y\right)}{\mathrm{Q}_Y\left(y\right)}\\
&= -\log\sum_y \mathrm{P}_Y\left(y\right)\max_x \mathrm{P}_{X|Y}\left(x|y\right) - \inf_{\mathrm{Q}_Y}\max_y \log\frac{\tilde{\mathrm{P}}_Y\left(y\right)}{\mathrm{Q}_Y\left(y\right)}\\
&= -\log\sum_y \mathrm{P}_Y\left(y\right)\max_x \mathrm{P}_{X|Y}\left(x|y\right) - \inf_{\mathrm{Q}_Y} D_{\max}(\tilde{\mathrm{P}}_Y\|\mathrm{Q}_Y)\\
&= -\log\sum_y \mathrm{P}_Y\left(y\right)\max_x \mathrm{P}_{X|Y}\left(x|y\right), \qquad (12)
\end{aligned}$$

where

$$\tilde{\mathrm{P}}_Y\left(y\right) := \left(\sum_{y'} \mathrm{P}_Y\left(y'\right)\max_x \mathrm{P}_{X|Y}\left(x|y'\right)\right)^{-1} \mathrm{P}_Y\left(y\right) \times \max_x \mathrm{P}_{X|Y}\left(x\right),$$

and the final equality in (12) holds since the max-divergence $D_{\max}(\mathrm{P}\|\mathrm{Q})$ (see Definition 4 below) is nonnegative and equals 0 if and only if $\mathrm{P} = \mathrm{Q}$. This alternative form of conditional min-entropy was first derived in [41] for a more general, quantum setup (see, also, [36, Theorem 2(ii)]) and shows that $H_{\min}(\mathrm{P}_{XY}|Y)$ corresponds to the $-\log$ of the *average conditional guessing probability* for X given Y. However, the original form in (11) is more suited for our purpose.

The definition of min-entropy and conditional min-entropy remain valid for all subnormalized, nonnegative functions P_{XY}, *i.e.*, P_{XY} such that

$$\sum_{x,y} \mathrm{P}_{XY}\left(x,y\right) \le 1.$$

We need this extension and the concept of smoothing, defined next, to derive tight bounds.

Definition 3. (Smooth min-entropy) Given $\epsilon \ge 0$, the ϵ-smooth conditional minimum entropy of P_{XY} given Y is defined as

$$H^{\epsilon}_{\min}(\mathrm{P}_{XY}|Y) := \sup_{\tilde{\mathrm{P}}_{XY}:\, d\left(\mathrm{P}_{XY},\tilde{\mathrm{P}}_{XY}\right) \le \epsilon} H_{\min}(\tilde{\mathrm{P}}_{XY}|Y),$$

where the $\sup$ is over all subnormalized, nonnegative functions $\tilde{\mathrm{P}}_{XY}$. When Y is a constant, the ϵ-smooth min-entropy is denoted by $H^{\epsilon}_{\min}(\mathrm{P}_X)$.

We now state the leftover hash lemma, which says that we can extract $H^{\epsilon}_{\min}\left(\mathrm{P}_{XY}|Y\right)$ unbiased, independent bits from X that are effectively concealed from an observer of Y.

Lemma 2. (Leftover hash) [58] *Given a joint distribution* P_{XY}, *for every* $0 \le 2\epsilon < 1$ *and* $0 < \eta$ *there exists a mapping*[15] $K : \mathcal{X} \to \mathcal{K}$ *with* $\log|\mathcal{K}| = \lfloor H^{\epsilon}_{\min}\left(\mathrm{P}_{XY}|Y\right) - 2\log(1/2\eta)\rfloor$ *such that*

$$d\left(\mathrm{P}_{K(X)Y}, \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_Y\right) \le 2\epsilon + \eta.$$

Finally, we review smooth max-divergence, which was introduced first in [18] for a quantum setting. The method of smoothing in the following definition is slightly different from the one in [18] and is tailored to our purpose.

Definition 4. (Smooth max-divergence) The max-divergence between two distributions P and Q is defined as

$$D_{\max}(\mathrm{P}\|\mathrm{Q}) := \max_x \log\frac{\mathrm{P}\left(x\right)}{\mathrm{Q}\left(x\right)},$$

with the convention $\log(0/0) = 0$, and for $0 \le \epsilon < 1$, the ϵ-smooth max-divergence between P and Q is defined as

$$D^{\epsilon}_{\max}(\mathrm{P}\|\mathrm{Q}) := \inf_{\substack{\tilde{\mathrm{P}}\,\le\,\mathrm{P}:\\ \tilde{\mathrm{P}}(\mathcal{X})\,\ge\,1-\epsilon}} D_{\max}(\tilde{\mathrm{P}}\|\mathrm{Q}),$$

where the $\inf$ is over all subnormalized, nonnegative functions $\tilde{\mathrm{P}}$ such that $\tilde{\mathrm{P}}\left(x\right) \le \mathrm{P}\left(x\right)$ for all $x \in \mathcal{X}$ and $\sum_x \tilde{\mathrm{P}}\left(x\right) \ge 1-\epsilon$.

[13] A review of the notion of smooth min-entropy without the notations from quantum information theory can be also found in [76].

[14] There are other definitions of conditional min-entropy available in the literature. The form here is perhaps the most widely used and is appropriate for our purpose.

[15] A randomly chosen function from a 2-universal hash family suffices.

The following two properties of smooth max-divergence will be used:

1) **Data processing inequality.** For every stochastic mapping $W : \mathcal{X} \to \mathcal{Y}$,

$$D^{\epsilon}_{\max}(\mathrm{P} \circ W \| \mathrm{Q} \circ W) \leq D^{\epsilon}_{\max}(\mathrm{P}\|\mathrm{Q}). \qquad (13)$$

Indeed, for every $\tilde{\mathrm{P}}$ such that $\tilde{\mathrm{P}}\,(x) \leq \mathrm{P}\,(x)$ for all $x \in \mathcal{X}$ and $\sum_x \tilde{\mathrm{P}}\,(x) \geq 1 - \epsilon$, the following hold

$$\begin{aligned}&(\tilde{\mathrm{P}} \circ W)(\mathcal{Y}) \geq 1 - \epsilon,\\ &(\tilde{\mathrm{P}} \circ W)(y) \leq (\mathrm{P} \circ W)(y), \quad \forall\, y \in \mathcal{Y}.\end{aligned}$$

The property follows upon noting that for every $y \in \mathcal{Y}$

$$\begin{aligned}D_{\max}(\tilde{\mathrm{P}}\|\mathrm{Q}) &= \max_x \log \frac{\tilde{\mathrm{P}}\,(x)}{\mathrm{Q}\,(x)}\\ &\geq \log \frac{(\tilde{\mathrm{P}} \circ W)(y)}{(\mathrm{Q} \circ W)(y)},\end{aligned}$$

since $\max_i (a_i/b_i) \geq (\sum_i a_i / \sum_i b_i)$.

2) **Convergence to Kullback-Leibler divergence.** For IID distributions P^n and Q^n,

$$\lim_{n\to\infty} \frac{1}{n}\, D^{\epsilon}_{\max}(\mathrm{P}^n\|\mathrm{Q}^n) = D(\mathrm{P}\|\mathrm{Q}), \quad \forall\, 0 < \epsilon < 1.$$

The inequality '$\leq$' holds trivially if $D(\mathrm{P}\|\mathrm{Q}) = \infty$. Thus, it suffices to prove it under the assumption that $D(\mathrm{P}\|\mathrm{Q})$ is finite. To that end, consider $\tilde{\mathrm{P}}_n(\mathbf{x}) = \mathrm{P}^n(\mathbf{x})\mathbb{1}(\mathbf{x} \in \mathcal{T}_n)$, where $\mathcal{T}_n$ is the (strongly) typical set for P^n (cf. [15]). For a sequence $\mathbf{x} \in \mathcal{X}^n$ and an element $x \in \mathcal{X}$, denote by $N(x|\mathbf{x})$ the number of occurrences of x in $\mathbf{x}$. Then, every sequence $\mathbf{x} \in \mathcal{T}_n$ satisfies (cf. [15])

$$\left|\frac{N(x|\mathbf{x})}{n} - \mathrm{P}\,(x)\right| < \delta_n, \quad x \in \mathcal{X}, \qquad (14)$$

where $\delta_n \to 0$ as $n \to 0$ (for precise conditions, see the *δ-convention* in [15]). Note that $\tilde{\mathrm{P}}_n \leq \mathrm{P}^n$ and $\tilde{\mathrm{P}}_n(\mathcal{X}^n) = \mathrm{P}^n(\mathcal{T}_n) \geq 1 - \epsilon$ for all n sufficiently large. Thus,

$$\begin{aligned}\frac{1}{n}\, D^{\epsilon}_{\max}(\mathrm{P}^n\|\mathrm{Q}^n) &\leq \frac{1}{n}\, D_{\max}(\tilde{\mathrm{P}}_n\|\mathrm{Q}^n)\\ &= \max_{\mathbf{x}\in\mathcal{T}_n} \frac{1}{n} \log \frac{\mathrm{P}^n(\mathbf{x})}{\mathrm{Q}^n(\mathbf{x})}\\ &= \max_{\mathbf{x}\in\mathcal{T}_n} \frac{1}{n} \sum_{i=1}^{n} \log \frac{\mathrm{P}(x_i)}{\mathrm{Q}(x_i)}\\ &= \max_{\mathbf{x}\in\mathcal{T}_n} \sum_{x\in\mathcal{X}} \frac{N(x|\mathbf{x})}{n} \log \frac{\mathrm{P}\,(x)}{\mathrm{Q}\,(x)}\\ &\leq \sum_{x\in\mathcal{X}} \mathrm{P}\,(x) \log \frac{\mathrm{P}\,(x)}{\mathrm{Q}\,(x)} + o(1),\end{aligned}$$

where the last inequality follows from (14) under the assumption that $D(\mathrm{P}\|\mathrm{Q}) < \infty$.

For the inequality in the other direction, suppose we are given a $\tilde{\mathrm{P}}_n \leq \mathrm{P}^n$ with $\tilde{\mathrm{P}}_n(\mathcal{X}^n) \geq 1 - \epsilon$. Then,

$$\begin{aligned}\tilde{\mathrm{P}}_n(\mathcal{T}_n) &= \tilde{\mathrm{P}}_n(\mathcal{X}^n) - \tilde{\mathrm{P}}_n(\mathcal{T}_n^c)\\ &\geq 1 - \epsilon - \tilde{\mathrm{P}}_n(\mathcal{T}_n^c)\\ &\geq 1 - \epsilon - \mathrm{P}^n(\mathcal{T}_n^c)\\ &\geq (1-\epsilon)/2, \qquad (15)\end{aligned}$$

for all n sufficiently large. This further implies that there exists an $\mathbf{x}_0 \in \mathcal{T}_n$ such that

$$\tilde{\mathrm{P}}_n(\mathbf{x}_0) \geq \mathrm{P}^n(\mathbf{x}_0)(1-\epsilon)/2.$$

Indeed, if not, then $\tilde{\mathrm{P}}_n(\mathbf{x}) < \mathrm{P}^n(\mathbf{x})(1-\epsilon)/2$ for all $\mathbf{x} \in \mathcal{T}_n$, which further implies $\tilde{\mathrm{P}}_n(\mathcal{T}_n) < (1-\epsilon)/2$ contradicting (15). Thus,

$$\begin{aligned}\frac{1}{n} \max_{\mathbf{x}} \log \frac{\tilde{\mathrm{P}}_n(\mathbf{x})}{\mathrm{Q}^n(\mathbf{x})} &\geq \frac{1}{n} \log \frac{\tilde{\mathrm{P}}_n(\mathbf{x}_0)}{\mathrm{Q}^n(\mathbf{x}_0)}\\ &\geq \frac{1}{n} \log \frac{\mathrm{P}^n(\mathbf{x}_0)}{\mathrm{Q}^n(\mathbf{x}_0)} + \frac{1}{n} \log \frac{1-\epsilon}{2}.\end{aligned}$$

For the case $D(\mathrm{P}\|\mathrm{Q}) = \infty$, there exists $x_+ \in \mathcal{X}$ such that $\mathrm{P}(x_+) > 0$ and $\mathrm{Q}(x_+) = 0$. Since $\mathbf{x}_0 \in \mathcal{T}_n$, $N(x_+|\mathbf{x}_0) > 0$ and the right-side of the inequality above, too, is infinity. On the other hand, if $D(\mathrm{P}\|\mathrm{Q})$ is finite, using (14) for the sequence $\mathbf{x}_0 \in \mathcal{T}_n$, the right-side of the inequality above is further bounded below by $D(\mathrm{P}\|\mathrm{Q}) - o(1)$, which completes the proof.

III. The conditional independence testing bound

Converse results of this paper are based on an upper bound on the maximum length $S_\epsilon\,(X_{\mathcal{M}}|Z)$ of an ϵ-SK. We present this basic result here[16].

Consider a (nontrivial) partition $\pi = \{\pi_1, ..., \pi_l\}$ of the set $\mathcal{M}$. Heuristically, if the underlying distribution of the observations $\mathrm{P}_{X_{\mathcal{M}}Z}$ is such that $X_{\mathcal{M}}$ are conditionally independent across the partition π given Z, the length of a SK that can be generated is 0. Our approach is to bound the length of a generated SK in terms of "how far" is the distribution $\mathrm{P}_{X_{\mathcal{M}}Z}$ from another distribution $\mathrm{Q}^{\pi}_{X_{\mathcal{M}}Z}$ that renders $X_{\mathcal{M}}$ conditionally independent across the partition π given Z – the closeness of the two distributions is measured by $\beta_\epsilon\big(\mathrm{P}_{X_{\mathcal{M}}Z}, \mathrm{Q}^{\pi}_{X_{\mathcal{M}}Z}\big)$.

Specifically, for a partition π with $|\pi| \geq 2$ parts, let $\mathcal{Q}(\pi)$ be the set of all distributions $\mathrm{Q}^{\pi}_{X_{\mathcal{M}}Z}$ that factorize as follows:

$$\mathrm{Q}^{\pi}_{X_{\mathcal{M}}|Z}(x_1, \ldots, x_m|z) = \prod_{i=1}^{|\pi|} \mathrm{Q}^{\pi}_{X_{\pi_i}|Z}(x_{\pi_i}|z). \qquad (16)$$

Theorem 3 (**Conditional independence testing bound**). *Given* $0 \leq \epsilon < 1$, $0 < \eta < 1 - \epsilon$, *and a partition* π *of* $\mathcal{M}$. *It holds that*

$$\begin{aligned}&S_\epsilon\,(X_{\mathcal{M}}|Z)\\ &\leq \frac{1}{|\pi| - 1}\Big[-\log \beta_{\epsilon+\eta}\big(\mathrm{P}_{X_{\mathcal{M}}Z}, \mathrm{Q}^{\pi}_{X_{\mathcal{M}}Z}\big) + |\pi| \log(1/\eta)\Big],\end{aligned} \qquad (17)$$

for all $\mathrm{Q}^{\pi}_{X_{\mathcal{M}}Z} \in \mathcal{Q}(\pi)$.

Remark 1. Renner and Wolf [60] derived a bound on the length of a SK that can be generated by two parties using one-way communication. A comparison of this bound with the general

[16] The results of this section were presented in [73].

bound in Theorem 3 is unavailable, since the former involves auxiliary RVs and is difficult to evaluate.

Remark 2. For $m = 2$ and $Z =$ constant, the upper bound on the length of a SK in Theorem 3 is related closely to the *meta-converse* of Polyanskiy, Poor, and Verdú [52]. Indeed, a code for reliable transmission of a message M over a point-to-point channel yields a SK for the sender and the receiver; the length of this SK can be bounded by Theorem 3. However, the resulting bound is slightly weaker than the meta-converse and does not yield the correct third order asymptotic term (the coefficient of $\log n$) in the optimal size of transmission codes [64].

Remark 3. The proof of Theorem 3 below remains valid even when the secrecy condition (3) is replaced by the following more general condition:

$$d\left(\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}, \mathrm{P}^{(\mathcal{M})}_{\mathtt{unif}} \times \mathrm{Q}_{\mathbf{F}Z}\right) \leq \epsilon,$$

for *some* distribution $\mathrm{Q}_{\mathbf{F}Z}$. In particular, upper bound (17) holds even under the relaxed secrecy criterion above.

The key idea underlying our proof of Theorem 3 is a lower bound for $-\log \beta_\epsilon(\mathrm{P}, \mathrm{Q})$ for a binary hypothesis testing problem with observation space $\mathcal{K}^m$, null hypothesis P given by

$$\mathrm{P}_{K_1K_2\ldots K_m} = \mathrm{P}^{(\mathcal{M})}_{\mathtt{unif}}, \tag{18}$$

and the alternative hypothesis Q given by

$$\mathrm{Q}_{K_1K_2\ldots K_m} = \prod_{i=1}^{m} \mathrm{Q}_{K_i}, \tag{19}$$

namely the problem of testing if $K_1....K_m$ constitute a perfectly correlated uniform randomness or are they mutually independent.

Lemma 4. *For* $\mathrm{P}_{K_{\mathcal{M}}} = \mathrm{P}_{K_1\ldots K_m}$ *and* $\mathrm{Q}_{K_{\mathcal{M}}} = \mathrm{Q}_{K_1\ldots K_m}$ *given in* (18) *and* (19), *it holds for every* $0 < \eta < 1$ *that*

$$\log|\mathcal{K}| \leq \frac{1}{m-1}\Big[-\log\beta_\eta\big(\mathrm{P}_{K_{\mathcal{M}}}, \mathrm{Q}_{K_{\mathcal{M}}}\big) + m\log(1/\eta)\Big].$$

Proof: Consider the log-likelihood ratio test with threshold λ given by

$$\lambda = (m-1)\log|\mathcal{K}| - m\log(1/\eta),$$

$i.e.$, the deterministic test with the following acceptance region (for the null hypothesis)

$$\mathcal{A} := \left\{k_{\mathcal{M}} : \log\frac{\mathrm{P}_{K_{\mathcal{M}}}(k_{\mathcal{M}})}{\mathrm{Q}_{K_{\mathcal{M}}}(k_{\mathcal{M}})} \geq \lambda\right\}.$$

For this test, the probability of error of type II is bounded above as

$$\begin{aligned}\mathrm{Q}_{K_{\mathcal{M}}}(\mathcal{A}) &= \sum_{k_{\mathcal{M}}\in\mathcal{A}} \mathrm{Q}_{K_{\mathcal{M}}}(k_{\mathcal{M}}) \\ &\leq 2^{-\lambda} \sum_{k_{\mathcal{M}}\in\mathcal{A}} \mathrm{P}_{K_{\mathcal{M}}}(k_{\mathcal{M}}) \\ &\leq 2^{-\lambda} \\ &= |\mathcal{K}|^{1-m}\eta^{-m}.\end{aligned} \tag{20}$$

On the other hand, the probability of error of type I is given by

$$\begin{aligned}\mathrm{P}_{K_{\mathcal{M}}}(\mathcal{A}^c) &= \frac{1}{|\mathcal{K}|}|\{k : \mathbf{k} = (k, ..., k) \in \mathcal{A}^c\}| \\ &= \frac{1}{|\mathcal{K}|}\sum_k \mathbb{1}(\mathbf{k} \in \mathcal{A}^c) \\ &= \frac{1}{|\mathcal{K}|}\sum_k \mathbb{1}\left(\mathrm{Q}_{K_{\mathcal{M}}}(\mathbf{k})|\mathcal{K}|^m\eta^m > 1\right),\end{aligned} \tag{21}$$

where $\mathbf{k} := (k, \ldots, k)$ and the second equality holds since $\mathcal{A}^c$ consists of elements $k_{\mathcal{M}}$ satisfying

$$\frac{\mathrm{P}_{K_{\mathcal{M}}}(k_{\mathcal{M}})}{\mathrm{Q}_{K_{\mathcal{M}}}(k_{\mathcal{M}})} = \frac{\mathbb{1}(k_1 = \cdots = k_m)}{|\mathcal{K}|\mathrm{Q}_{K_{\mathcal{M}}}(\mathcal{M})} < 2^\lambda = |\mathcal{K}|^{m-1}\eta^m.$$

The inner sum can be further upper bounded as

$$\begin{aligned}\sum_k \mathbb{1}\left(\mathrm{Q}_{K_{\mathcal{M}}}(\mathbf{k})|\mathcal{K}|^m\eta^m > 1\right) &\leq \sum_k \left(\mathrm{Q}_{K_{\mathcal{M}}}(\mathbf{k})|\mathcal{K}|^m\eta^m\right)^{\frac{1}{m}} \\ &= |\mathcal{K}|\eta\sum_k \mathrm{Q}_{K_{\mathcal{M}}}(\mathbf{k})^{\frac{1}{m}} \\ &= |\mathcal{K}|\eta\sum_k\prod_{i=1}^m \mathrm{Q}_{K_i}(k)^{\frac{1}{m}} \\ &\leq |\mathcal{K}|\eta\prod_{i=1}^m\left(\sum_k \mathrm{Q}_{K_i}(k)\right)^{\frac{1}{m}} \\ &= |\mathcal{K}|\eta,\end{aligned} \tag{22}$$

where the first inequality above holds by $\mathbb{1}(\cdot) \leq 1$, and the second inequality above holds by Hölder's inequality. Upon combining (21) and (22) we obtain

$$\mathrm{P}_{K_{\mathcal{M}}}(\mathcal{A}^c) \leq \eta.$$

Thus, we have a test with probability of error of type I less than η and the probability of error of type II bounded as in (20). Therefore,

$$\beta_\eta\big(\mathrm{P}_{K_{\mathcal{M}}}, \mathrm{Q}_{K_{\mathcal{M}}}\big) \leq |\mathcal{K}|^{1-m}\eta^{-m},$$

which completes the proof. ■

The distribution $\mathrm{P}_{K_{\mathcal{M}}}$ in (18) corresponds to a perfect secret key shared by m parties. The next result extends Lemma 4 to the case where not only the key values $K_{\mathcal{M}}$ but also the communication $\mathbf{F}$ and the eavesdropper's side information Z are observed, and the null hypothesis $\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}$ corresponds to an ϵ-SK $K_{\mathcal{M}}$.

Lemma 5. *For an* ϵ*-SK* $K_{\mathcal{M}}$ *with a common range* $\mathcal{K}$ *generated using an interactive communication* $\mathbf{F}$, *let* $W_{K_{\mathcal{M}}\mathbf{F}|X_{\mathcal{M}}Z}$ *be the resulting conditional distribution*[17] *on* $(K_{\mathcal{M}}, \mathbf{F})$ *given* $(X_{\mathcal{M}}, Z)$. *Then, for every* $0 < \eta < 1 - \epsilon$ *and every* $\mathrm{Q}_{X_{\mathcal{M}}Z} = \prod_{i=1}^m \mathrm{Q}_{X_i|Z}\mathrm{Q}_Z$, *we have*

$$\begin{aligned}&\log|\mathcal{K}| \\ &\leq \frac{1}{m-1}\Big[-\log\beta_{\epsilon+\eta}\big(\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}, \mathrm{Q}_{K_{\mathcal{M}}\mathbf{F}Z}\big) + m\log(1/\eta)\Big],\end{aligned} \tag{23}$$

[17] The conditional distribution $W_{K_{\mathcal{M}}\mathbf{F}|X_{\mathcal{M}}Z}$ is defined only for $(x_{\mathcal{M}}, z)$ with $\mathrm{P}_{X_{\mathcal{M}}Z}(x_{\mathcal{M}}, z) > 0$.

where $\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}$ *is the marginal of* $(K_{\mathcal{M}}, \mathbf{F}, Z)$ *for the joint distribution*

$$\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}X_{\mathcal{M}}Z} = \mathrm{P}_{X_{\mathcal{M}}Z}W_{K_{\mathcal{M}}\mathbf{F}|X_{\mathcal{M}}Z},$$

and $\mathrm{Q}^{\pi}_{K_{\mathcal{M}}\mathbf{F}Z}$ *is the corresponding marginal for the joint distribution*

$$\mathrm{Q}_{K_{\mathcal{M}}\mathbf{F}X_{\mathcal{M}}Z} = \mathrm{Q}_{X_{\mathcal{M}}Z}W_{K_{\mathcal{M}}\mathbf{F}|X_{\mathcal{M}}Z}.$$

Also, we need the following basic property of interactive communication from [68], which will be used throughout this paper (see, also, [1, Lemma 2.2], [15, Lemma 17.18]).

Lemma 6 (**Interactive communication property**). *Given* $\mathrm{Q}_{X_{\mathcal{M}}Z} = \prod_{i=1}^{m} \mathrm{Q}_{X_i|Z}\mathrm{Q}_Z$ *and an interactive communication* $\mathbf{F}$*, the following holds:*

$$\mathrm{Q}_{X_{\mathcal{M}}|\mathbf{F}Z}\left(x_{\mathcal{M}}|f,z\right) = \prod_{i=1}^{m} \mathrm{Q}_{X_i|\mathbf{F}Z}\left(x_i|f,z\right),$$

$i.e.$*, conditionally independent observations remain so when conditioned additionally on an interactive communication. In particular, if* $\mathrm{Q}_{X_1X_2|Z} = \mathrm{Q}_{X_1|Z}\mathrm{Q}_{X_2|Z}$*, then*

$$\mathrm{Q}_{X_1X_2|\mathbf{F}Z} = \mathrm{Q}_{X_1|\mathbf{F}Z} \times \mathrm{Q}_{X_2|\mathbf{F}Z}.$$

Proof of Lemma 5. The proof is a simple modification of the proof of Lemma 4. First note that by Lemma 6 and the fact K_i is a function of (X_i, U_i) given $\mathbf{F}$, we have

$$\mathrm{Q}_{K_{\mathcal{M}}|\mathbf{F}Z} = \prod_{i=1}^{m} \mathrm{Q}_{K_i|\mathbf{F}Z}.$$

Thus, Lemma 4 applies with distribution $\mathrm{Q}_{K_{\mathcal{M}}|\mathbf{F}Z}$ in the role of Q for every $\mathbf{F} = f, Z = z$, and consequently, for every (f, z) there exists a set $\mathcal{A}_{f,z}$ such that

$$\mathrm{Q}_{K_{\mathcal{M}}|\mathbf{F}Z}\left(\mathcal{A}_{f,z}|f,z\right) \leq |\mathcal{K}|^{1-m}\eta^{-m}, \tag{24}$$

and

$$\mathrm{P}^{(\mathcal{M})}_{\mathtt{unif}}(\mathcal{A}^c_{f,z}) \leq \eta. \tag{25}$$

We consider the following test for a binary hypothesis testing problem with null hypothesis $\mathrm{P}^{(\mathcal{M})}_{\mathtt{unif}} \times \mathrm{P}_{\mathbf{F}Z}$ and alternative hypothesis $\mathrm{Q}_{K_{\mathcal{M}}\mathbf{F}Z}$: For an observed $(k_{\mathcal{M}}, f, z)$, we accept the null hypothesis if $k_{\mathcal{M}} \in \mathcal{A}_{f,z}$ and alternative otherwise. Using (24), the probability of error of type II is bounded above by

$$\sum_{f,z} \mathrm{Q}_{\mathbf{F}Z}\left(f,z\right)\mathrm{Q}_{K_{\mathcal{M}}|\mathbf{F}Z}\left(\mathcal{A}_{f,z}|f,z\right) \leq |\mathcal{K}|^{1-m}\eta^{-m},$$

and by (25), the probability of error of type I is bounded above by

$$\sum_{f,z} \mathrm{P}_{\mathbf{F}Z}\left(f,z\right)\mathrm{P}^{(\mathcal{M})}_{\mathtt{unif}}(\mathcal{A}_{f,z}) \leq \eta.$$

Finally, we consider the hypothesis testing problem with null hypothesis $\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}$ and alternative hypothesis $\mathrm{Q}_{K_{\mathcal{M}}\mathbf{F}Z}$ and apply the same test as above. Clearly, the probability of error of type II remains unchanged. Furthermore, in view of the secrecy condition (3), the probability of error of type I will increase by at most ϵ, which completes the proof. □

Proof of Theorem 3. We first consider the partition π with one element in each part, $i.e.$, $\pi_i = \{i\}$ for $1 \leq i \leq m$. For this case, it follows from Lemma 5 and the data processing inequality (5) with $\mathrm{P} = \mathrm{P}_{X_{\mathcal{M}}Z}$, $\mathrm{Q} = \mathrm{Q}^{\pi}_{X_{\mathcal{M}}Z}$, and $W = W_{K_{\mathcal{M}}\mathbf{F}|X_{\mathcal{M}}Z}$, that for an ϵ-SK $K_{\mathcal{M}}$ taking values in the set $\mathcal{K}$,

$$\begin{aligned}
&\log|\mathcal{K}| \\
&\leq \frac{1}{m-1}\Big[-\log\beta_{\epsilon+\eta}\big(\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}, \mathrm{Q}^{\pi}_{K_{\mathcal{M}}\mathbf{F}Z}\big) + m\log(1/\eta)\Big] \\
&\leq \frac{1}{m-1}\Big[-\log\beta_{\epsilon+\eta}\big(\mathrm{P}_{X_{\mathcal{M}}Z}, \mathrm{Q}^{\pi}_{X_{\mathcal{M}}Z}\big) + m\log(1/\eta)\Big],
\end{aligned} \tag{26}$$

for every $1 < \eta < 1 - \epsilon$.

To extend (26) to an arbitrary partition π, we claim that an ϵ-SK for the original model with m parties yields an ϵ-SK of the same length for a model with $|\pi|$ parties with the ith party observing X_{π_i}, $1 \leq i \leq |\pi|$, and the eavesdropper observing the RV Z as before. The result follows by applying the bound (26) to the new model with $|\pi|$ parties.

It only remains to prove the claim above. To that end, given an ϵ-SK $K_{\mathcal{M}}$ for the original model, we define an ϵ-SK for the new model (with the ith party observing X_{π_i}) as follows: The parties run the protocol for generating $K_{\mathcal{M}}$ with communication corresponding to any party $j \in \pi_i$ in the original model transmitted by the ith party in the new model. For each party i in the new model, we select a representative party $i_0 \in \pi_i$; for concreteness, let i_0 be the smallest index in the set π_i. An ϵ-SK for the new model is given by $(K'_1, ..., K'_{|\pi|})$ where $K'_i = K_{i_0}$ since, denoting by $\mathrm{P}^{(\pi)}_{\mathtt{unif}}$ the distribution

$$\mathrm{P}^{(\pi)}_{\mathtt{unif}}(k_1, ..., k_{|\pi|}) = \frac{1}{|\mathcal{K}|}\mathbb{1}(k_1 = \cdots = k_{|\pi|}),$$

we have by the monotonicity that

$$\begin{aligned}
&d\left(\mathrm{P}_{K'_1...K'_{|\pi|}\mathbf{F}Z}, \mathrm{P}^{(\pi)}_{\mathtt{unif}} \times \mathrm{P}_{\mathbf{F}Z}\right) \\
&\leq d\left(\mathrm{P}_{K_1...K_m\mathbf{F}Z}, \mathrm{P}^{(\mathcal{M})}_{\mathtt{unif}} \times \mathrm{P}_{\mathbf{F}Z}\right) \\
&\leq \epsilon,
\end{aligned}$$

which completes the proof. □

IV. Implications for secret key capacity

For the SK agreement problem, a special case of interest is when the observations consist of n length IID sequences, $i.e.$, the ith party observes $(X_{i1}, ..., X_{in})$ and the eavesdropper observes $(Z_1, ..., Z_n)$ such that the RVs $\{X_{\mathcal{M}t}, Z_t\}_{t=1}^n$ are IID. For this case, it is well known that a SK of length proportional to n can be generated; the maximum rate $(\log|\mathcal{K}_n|/n)$ of a SK is called the SK capacity [44], [1], [16].

To present the results of this section at full strength, we need to take recourse to the original definition of (ϵ, δ)-SK given in (1) and (2). In the manner of Definition 1, denote by $S_{\epsilon,\delta}(X_{\mathcal{M}}|Z)$ the maximum length of an (ϵ, δ)-SK. It follows from Proposition 1 that $S_{\epsilon,\delta}(X_{\mathcal{M}}|Z) \leq S_{\epsilon+\delta}(X_{\mathcal{M}}|Z)$.

Definition 5. **(SK capacity)** Given $0 < \epsilon, \delta < 1$, the (ϵ, δ)-SK capacity $C_{\epsilon,\delta}\left(X_{\mathcal{M}}|Z\right)$ is defined by

$$C_{\epsilon,\delta}\left(X_{\mathcal{M}}|Z\right) := \liminf_{n\to\infty} \frac{1}{n} S_{\epsilon,\delta}(X_{\mathcal{M}}^n|Z^n),$$

where the RVs $\{X_{\mathcal{M}t}, Z_t\}$ are IID for $1 \leq t \leq n$, with a common distribution $\mathrm{P}_{X_{\mathcal{M}}Z}$. The SK capacity $C\left(X_{\mathcal{M}}|Z\right)$ is defined as the limit

$$C\left(X_{\mathcal{M}}|Z\right) := \lim_{\epsilon,\delta\to 0} C_{\epsilon,\delta}\left(X_{\mathcal{M}}|Z\right).$$

For the case when the eavesdropper does not observe any side information, $i.e.$, $Z = constant$, the SK capacity for two parties was characterized by Maurer [44] and Ahlswede and Csiszár [1]. Later, the SK capacity for a multiparty model, with Z =constant was characterized by Csiszár and Narayan [16]. The general problem of characterizing the SK capacity for arbitrary Z remains open. Several upper bounds for SK capacity are known [44], [1], [46], [59], [16], [17], [26], which are tight for special cases.

In this section, we derive a single-shot version of the Gohari-Anantharam bound [26] on the SK capacity for two parties, which is the best known bound for this case. Furthermore, for multiple parties, we establish a strong converse for SK capacity, which shows that, surprisingly, we cannot improve the rate of a SK by relaxing the recoverability requirement (1) or the secrecy requirement (2).

A. Converse results for two parties

It was shown in [26] that for two parties,

$$C(X_1, X_2|Z) \leq \min_U I\left(X_1 \wedge X_2|U\right) + I(X_1, X_2 \wedge U|Z). \tag{27}$$

The proof in [26] relied critically on the assumption that the RVs $\{(X_{\mathcal{M}t}, Z_t)\}_{t=1}^n$ are IID and does not apply to the single-shot setup. The result below is a single-shot version of (27) and is proved by relying only on the structure of the SKs, without recourse to the *potential function approach*[18] of [26].

Theorem 7. *For* $0 < \epsilon, \delta$ *with* $\epsilon + 2\delta < 1$,

$$\begin{aligned} &S_{\epsilon,\delta}(X_1, X_2|Z) \\ &\leq S_{\epsilon,2\delta+\eta}(X_1, X_2|Z, U) + D_{\max}^{\xi}\left(\mathrm{P}_{X_1X_2ZU}\|\mathrm{P}_{X_1X_2Z}\mathrm{P}_{U|Z}\right) \\ &\qquad + 2\log(1/2(\eta-\xi)) + 1, \end{aligned}$$

for every RV U *and every* $0 \leq \xi < \eta < 1 - \epsilon - 2\delta$.

As corollaries, we obtain a single-shot version and a strong version of the upper bound in (27), which does not require perfect asymptotic recovery or perfect asymptotic secrecy.

Corollary 8 **(Single-shot bound for SK length).** *For* $0 < \epsilon, \delta$ *with* $\epsilon + 2\delta < 1$,

$$\begin{aligned} S_{\epsilon,\delta}(X_1, X_2|Z) &\leq -\log\beta_{\epsilon+2\delta+\eta}(\mathrm{P}_{X_1X_2ZU}, \mathrm{P}_{X_1|ZU}\mathrm{P}_{X_2ZU}) \\ &\quad + D_{\max}^{\eta_1}\left(\mathrm{P}_{X_1X_2ZU}\|\mathrm{P}_{X_1X_2Z}\mathrm{P}_{U|Z}\right) \\ &\quad + 4\log(1/(\eta-\eta_1-\eta_2)) + 1, \end{aligned}$$

for every RV U *and every* $0 \leq \eta_1 + \eta_2 < \eta < 1 - \epsilon - 2\delta$.

Corollary 9 **(Strong bound for SK capacity).** *For* $0 \leq \epsilon, \delta$ *with* $\epsilon + 2\delta < 1$,

$$C_{\epsilon,\delta}(X_1, X_2|Z) \leq \min_U I\left(X_1 \wedge X_2|U\right) + I(X_1, X_2 \wedge U|Z).$$

We conclude this section with proofs. The core of Theorem 7 is contained in the following lemma.

Lemma 10. *Let* (K_1, K_2) *be an* (ϵ, δ)*-SK taking values in* $\mathcal{K}$*, recoverable from a communication* $\mathbf{F}$*. Then,*

$$\begin{aligned} &H_{\min}^{\delta+\xi/2}\left(\mathrm{P}_{K_1\mathbf{F}ZU}|\mathbf{F}ZU\right) \\ &\geq \log|\mathcal{K}| - D_{\max}^{\xi}\left(\mathrm{P}_{K_1\mathbf{F}ZU}\|\mathrm{P}_{K_1\mathbf{F}Z}\mathrm{P}_{U|Z}\right), \end{aligned}$$

for every RV U *and every* $0 \leq \xi < 1 - \epsilon - 2\delta$.

Proof of Theorem 7. Let (K_1, K_2) be an (ϵ, δ)-SK taking values in $\mathcal{K}$. Then, by Lemma 10 and the data processing property of smooth max-divergence (13), we get

$$\begin{aligned} &H_{\min}^{\delta+\xi/2}\left(\mathrm{P}_{K_1\mathbf{F}ZU}|\mathbf{F}ZU\right) \\ &\geq \log|\mathcal{K}| - D_{\max}^{\xi}\left(\mathrm{P}_{X_1X_2ZU}\|\mathrm{P}_{X_1X_2Z}\mathrm{P}_{U|Z}\right). \end{aligned}$$

By the leftover hash lemma (see Section II-D), there exists a mapping K' of $\mathcal{K}$ taking at least $\log|\mathcal{K}| - D_{\max}^{\xi}\left(\mathrm{P}_{X_1X_2ZU}\|\mathrm{P}_{X_1X_2Z}\mathrm{P}_{U|Z}\right) - 2\log(1/2(\eta-\xi)) - 1$ values and satisfying

$$d\left(\mathrm{P}_{K'(K_1)\mathbf{F}ZU}, \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_{\mathbf{F}ZU}\right) \leq 2\delta + \eta.$$

Therefore, $(K'(K_1), K'(K_2))$ constitutes an $(\epsilon, 2\delta + \eta)$-SK for X_1 and X_2, when the eavesdropper observes (Z, U) and so,

$$\begin{aligned} &S_{\epsilon,2\delta+\eta}(X_1, X_2|Z, U) \\ &\geq \log|\mathcal{K}| - D_{\max}^{\xi}\left(\mathrm{P}_{X_1X_2ZU}\|\mathrm{P}_{X_1X_2Z}\mathrm{P}_{U|Z}\right) \\ &\qquad\qquad - 2\log\frac{1}{2(\eta-\xi)} - 1. \end{aligned}$$

□

Corollary 8 follows by Theorem 3.

Proof of Corollary 9. The result follows by Corollary 8 upon using Stein's lemma (see Section II-B), along with the convergence property of smooth max-divergence (see Section II-D). □

Proof of Lemma 10. By definitions of $H_{\min}^{\delta+\xi/2}$ and $D_{\max}^{\xi}$, it suffices to show that for every mapping $T : (k_1, f, z, u) \mapsto [0, 1]$ such that

$$\sum_{k_1,f,z,u} \mathrm{P}\left(k_1, f, z, u\right) T(k_1, f, z, u) \geq 1 - \xi, \tag{28}$$

here exist a subnormalized nonnegative function $\mathrm{Q}_{K_1\mathbf{F}ZU}$ and a distribution $\tilde{\mathrm{Q}}_{\mathbf{F}ZU}$ satisfying the following:

$$d\left(\mathrm{P}_{K_1\mathbf{F}ZU}, \mathrm{Q}_{K_1\mathbf{F}ZU}\right) \leq \delta + \xi/2 \tag{29}$$

and

$$H_{\min}\left(\mathrm{Q}_{K_1\mathbf{F}ZU}|\tilde{\mathrm{Q}}_{\mathbf{F}ZU}\right)$$

[18] In fact, a simple proof of (27) follows upon noting that for an optimum rate SK (K_1, K_2) recoverable from a communication $\mathbf{F}$, the SK capacity $C(X_1, X_2|Z)$ approximately equals $(1/n)H(K_1|\mathbf{F}, Z^n) \leq (1/n)H(K_1|\mathbf{F}, U^n, Z^n) + (1/n)I(K_1, \mathbf{F} \wedge U^n|Z^n)$, which is further bounded above by $C(X_1, X_2|U) + I(X_1, X_2 \wedge U|Z)$.

$$= \log|\mathcal{K}| - D_{\max}\left(\mathrm{P}_{K_1\mathbf{F}ZU}T \| \mathrm{P}_{K_1\mathbf{F}Z}\mathrm{P}_{U|Z}\right). \tag{30}$$

To that end, consider $\mathrm{Q}_{K_1\mathbf{F}ZU}$ given by

$$\begin{aligned}&\mathrm{Q}\left(k_1, f, z, u\right) \\ &:= \mathrm{P}_{\texttt{unif}}\left(k_1\right)\mathrm{P}\left(f, z\right)\mathrm{P}\left(u|k_1, f, z\right)T(k_1, f, z, u),\end{aligned} \tag{31}$$

which is a valid subnormalized nonnegative function since $T(k_1, f, z, u) \leq 1$. Furthermore, since

$$\mathrm{P}\left(k_1, f, z, u\right) = \mathrm{P}\left(k_1, f, z\right)\mathrm{P}\left(u|k_1, f, z\right),$$

we get (29) as follows:

$$\begin{aligned}&d\left(\mathrm{P}_{K_1\mathbf{F}ZU}, \mathrm{Q}_{K_1\mathbf{F}ZU}\right) \\ &\leq d\left(\mathrm{P}_{K_1\mathbf{F}Z}, \mathrm{P}_{\texttt{unif}}\mathrm{P}_{\mathbf{F}Z}\right) \\ &\quad + \sum_{k_1, f, z, u}\mathrm{P}\left(k_1, f, z, u\right)\left(1 - T(k_1, f, z, u)\right) \\ &\leq \delta + \frac{\xi}{2},\end{aligned}$$

where the first inequality is by the triangle inequality and the fact that $T(k_1, f, z, u) \leq 1$, and the last inequality uses the secrecy condition (2) and the assumption (28).

Next, for $\tilde{\mathrm{Q}}_{\mathbf{F}ZU}$ defined by

$$\tilde{\mathrm{Q}}(f, z, u) := \mathrm{P}\left(f, z\right)\mathrm{P}\left(u|z\right) \tag{32}$$

and $Q_{K_1\mathbf{F}ZU}$ defined in (31), observe that

$$\begin{aligned}&\frac{\mathrm{Q}\left(k_1, f, z, u\right)}{\tilde{\mathrm{Q}}(f, z, u)} \\ &= \mathrm{P}_{\texttt{unif}}\left(k\right)\left[\frac{\mathrm{P}\left(u|k_1, f, z\right)}{\mathrm{P}\left(u|z\right)}\right]T(k_1, f, z, u) \\ &= \mathrm{P}_{\texttt{unif}}\left(k\right)\left[\frac{\mathrm{P}\left(k_1, f, z, u\right)}{\mathrm{P}\left(k_1, f, z\right)\mathrm{P}\left(u|z\right)}\right]T(k_1, f, z, u),\end{aligned}$$

and so,

$$\begin{aligned}&H_{\min}\left(\mathrm{Q}_{K_1\mathbf{F}ZU}|\tilde{\mathrm{Q}}_{\mathbf{F}ZU}\right) \\ &= \log|\mathcal{K}| - \max_{k_1, f, z, u}\log\frac{\mathrm{P}\left(k_1, f, z, u\right)T(k_1, f, z, u)}{\mathrm{P}\left(k_1, f, z\right)\mathrm{P}\left(u|z\right)},\end{aligned}$$

which is the same as (30). □

B. Strong converse for multiple parties

Now we move to the m terminal case where the eavesdropper gets no side information, *i.e.*, $Z =$ constant. With this simplification, the SK capacity $C\left(X_{\mathcal{M}}\right)$ for multiple parties was characterized by Csiszár and Narayan [16]. Furthermore, they introduced the remarkable expression on the right-side of (33) below as an upper bound for $C\left(X_{\mathcal{M}}\right)$, and showed its tightness for $m = 2, 3$. Later, the tightness of the upper bound for arbitrary m was shown in [10]; we summarize these developments in the result below.

Theorem 11. *[16], [10] The SK capacity for the case when eavesdropper's side information $Z = constant$ is given by*

$$C\left(X_{\mathcal{M}}\right) = \min_{\pi}\frac{1}{|\pi| - 1}D\left(\mathrm{P}_{X_{\mathcal{M}}}\middle\|\prod_{i=1}^{|\pi|}\mathrm{P}_{X_{\pi_i}}\right), \tag{33}$$

where the min *is over all partitions π of $\mathcal{M}$.*

This generalized the classic result of Maurer [44] and Ahlswede and Csiszár [1], which established that for two parties, $C\left(X_1, X_2\right) = D\left(\mathrm{P}_{X_1X_2}\|\mathrm{P}_{X_1} \times \mathrm{P}_{X_2}\right) = I\left(X_1 \wedge X_2\right)$.

The converse part of Theorem 11 relied critically on the fact that $\epsilon_n + \delta_n \to 0$ as $n \to \infty$. Below we strengthen the converse and show that the upper bound for SK rates implied by Theorem 11 holds even when (ϵ_n, δ_n) is fixed. Specifically, for $0 < \epsilon, \delta$ with $\epsilon + \delta < 1$ and $Z =$ constant, an application of Theorem 3 to the IID RVs $X^n_{\mathcal{M}}$, with $\mathrm{Q}^{\pi}_{X^n_{\mathcal{M}}} = \prod_{i=1}^{|\pi|}\mathrm{P}^n_{X_{\pi_i}}$, yields

$$\begin{aligned}&S_{\epsilon,\delta}\left(X^n_1, ..., X^n_m\right) \\ &\leq \frac{1}{|\pi| - 1}\Bigg[-\log\beta_{\epsilon+\delta+\eta}\left(\mathrm{P}^n_{X_{\mathcal{M}}}, \prod_{i=1}^{|\pi|}\mathrm{P}^n_{X_{\pi_i}}\right) \\ &\qquad\qquad\qquad + |\pi|\log(1/\eta)\Bigg],\end{aligned}$$

where $\eta < 1 - \epsilon - \delta$. Therefore, using Stein's Lemma (see (6)) we get

$$\begin{aligned}&C_{\epsilon,\delta}\left(X_{\mathcal{M}}\right) \\ &\leq \frac{1}{|\pi| - 1}\liminf_{n\to\infty} -\frac{1}{n}\log\beta_{\epsilon+\delta+\eta}\left(\mathrm{P}^n_{X_{\mathcal{M}}}, \prod_{i=1}^{|\pi|}\mathrm{P}^n_{X_{\pi_i}}\right) \\ &= \frac{1}{|\pi| - 1}D\left(\mathrm{P}_{X_{\mathcal{M}}}\middle\|\prod_{i=1}^{|\pi|}\mathrm{P}_{X_{\pi_i}}\right).\end{aligned}$$

Also, note that if $\epsilon + \delta \geq 1$, the SK rate can be infinity. Indeed, consider a $(0, 1)$-SK where party 1 generates a RV K_1 uniformly over a set $\mathcal{K}$ and sends it to the other parties over the public communication channel, and a $(1, 0)$-SK where the ith party generates K_i uniformly over $\mathcal{K}$ using its local randomness U_i (without any public communication). If $\epsilon + \delta \geq 1$, the SK which equals the $(0, 1)$-SK above with probability $(1 - \epsilon)$ and the $(1, 0)$-SK above with probability ϵ constitutes an $(\epsilon, 1 - \epsilon)$-SK of length $\log|\mathcal{K}|$, and therefore, also an (ϵ, δ)-SK of the same length. Since $\mathcal{K}$ was arbitrary, the length of the resulting (ϵ, δ)-SK can be arbitrarily large.

Thus, we have established the following *strong converse* for the SK capacity when $Z =$ constant.

Corollary 12 (Strong converse for SK capacity). *Given $0 < \epsilon, \delta < 1$, the (ϵ, δ)-SK capacity $C_{\epsilon,\delta}\left(X_{\mathcal{M}}\right)$ is given by*

$$C_{\epsilon,\delta}\left(X_{\mathcal{M}}\right) = \min_{\pi}\frac{1}{|\pi| - 1}D\left(\mathrm{P}_{X_{\mathcal{M}}}\middle\|\prod_{i=1}^{|\pi|}\mathrm{P}_{X_{\pi_i}}\right), \ \textit{if } \epsilon + \delta < 1,$$

and

$$C_{\epsilon,\delta}\left(X_{\mathcal{M}}\right) = \infty, \qquad \textit{if } \epsilon + \delta \geq 1.$$

V. Implications for secure two-party computation

In this section, we consider secure computation by two (mutually untrusting) parties. First introduced by Yao in [83], these problems have propelled the research in cryptography over the last three decades. In particular, we will consider the *oblivious transfer* and the *bit commitment* problem, the two basic primitives for secure two-party computation. We will

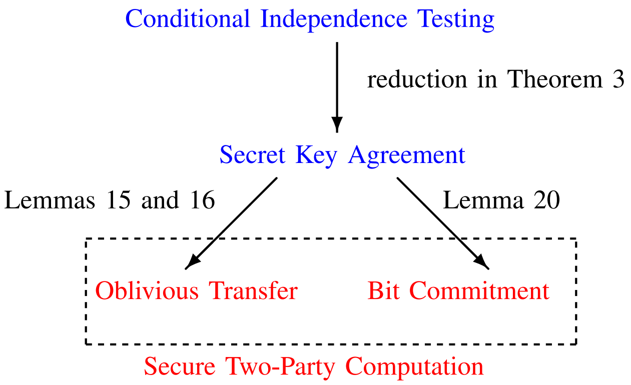

Fig. 1. Depiction of our reduction arguments.

look at the information theoretic versions of these problems where, as an additional resource, the parties observe correlated RVs X_1 and X_2. Our converse results are based on reduction arguments which relate these problems to the SK agreement problem, enabling the application of Theorem 3 (see Fig. 1).

A. Maximum common function and minimum sufficient statistic

To state our results, we need the notions of *maximum common function* and *minimum sufficient statistic*. The notion of maximum common function was introduced in [24] as a measure of "common information" of random variables X_1 and X_2. Its role in secrecy was first highlighted in [82] (see, also, [69], [47] for different roles of the maximum common function in secrecy and privacy.) Operationally, the maximum common function of X_1 and X_2 is defined as follows.

Definition 6 (Maximum Common Function). A *common function* of X_1 and X_2 is a random variable U for which there there exist functions $\phi_1(X_1)$ and $\phi_2(X_2)$ such that $\mathrm{P}\left(U = \phi_1(X_1) = \phi_2(X_2)\right) = 1$. The maximum common function[19] of X_1 and X_2, denoted by $\mathrm{mcf}(X_1, X_2)$, is a common function of X_1 and of X_2 such that every common function U of X_1 and X_2 is a function of $\mathrm{mcf}(X_1, X_2)$, *i.e.*, $H(U|\mathrm{mcf}(X_1, X_2)) = 0$.

In fact, [24] characterized $\mathrm{mcf}(X_1, X_2)$ and showed that it corresponds to the following equivalence relation on $\mathcal{X}_1$ (or a similarly defined equivalence relation on $\mathcal{X}_2$)

$$x_1 \sim x_1' \Leftrightarrow \exists\, x_{21}, ..., x_{2k} \in \mathcal{X}_2 \text{ and } x_{12}, ..., x_{1k} \in \mathcal{X}_1 \text{ s.t.}$$
$$\mathrm{P}_{X_1X_2}\left(x_{1i}, x_{2i}\right) \mathrm{P}_{X_1X_2}\left(x_{1(i+1)}, x_{2i}\right) > 0 \text{ for } 1 \le i \le k.$$

where $x_{11} = x_1$ and $x_{1(k+1)} = x_1'$. The role of minimum sufficient statistic in secrecy was highlighted in [82] as well. We give its operational definition below.

Definition 7 (Minimum Sufficient Satistics). A *sufficient statistic* for X_2 given X_1 is a random variable U such that there exists a function $U = g(X_1)$ such that the Markov chain X_1—U—X_2 holds. The minimum sufficient statistics for X_2 given X_1, denoted by $\mathrm{mss}(X_2|X_1)$, is a sufficient statistics for X_2 given X_1 such that it is a function of every sufficient statistic U for X_2 given X_1, i.e., $H(\mathrm{mss}(X_2|X_1)|U) = 0$.

An exact characterization of $\mathrm{mss}(X_2|X_1)$ is available, too, and it corresponds to the following equivalence relation on $\mathcal{X}_1$ (cf. [23], [37], [67]):

$$x_1 \sim x_1' \Leftrightarrow \mathrm{P}_{X_2|X_1}\left(x_2|x_1\right) = \mathrm{P}_{X_2|X_1}\left(x_2|x_1'\right), \quad \forall x_2 \in \mathcal{X}_2.$$

B. Oblivious transfer

We present bounds on the efficiency of implementing information theoretically secure one-out-of-two OT using correlated randomness. Suppose that party 1 generates K_0 and K_1, distributed uniformly over $\{0,1\}^l$, and party 2 generates B, distributed uniformly over $\{0,1\}$, as inputs to an OT protocol. The RVs K_0, K_1, and B are assumed to be mutually independent[20]. The goal of an OT protocol is for Party 2 to obtain K_B in such a manner that B is concealed from Party 1 and $K_{\overline{B}}$ is concealed from party 2, where $\overline{B} = 1 \oplus B$. Furthermore, Party i observes the RV X_i, $i = 1, 2$, as a resource to implement an OT protocol, where RVs (X_1, X_2) are independent jointly of (K_0, K_1, B). During the protocol, the parties are allowed to communicate interactively. In general, the parties are allowed to use local randomization; for simplicity of presentation, we restrict ourselves to protocols without local randomization. However, as pointed-out in Remark 6 below, our results remain valid even when local randomization is allowed.

Definition 8. (Oblivious transfer) An execution of a protocol realizing an $(\epsilon, \delta_1, \delta_2)$-OT (for a passive adversary[21]) of length l consists of an interactive communication $\mathbf{F}$ and an estimate $\hat{K} = \hat{K}(X_2, B, \mathbf{F})$ by Party 2 such that the following conditions are satisfied:

$$\mathrm{P}\left(K_B \neq \hat{K}\right) \le \epsilon, \tag{34}$$
$$d\left(\mathrm{P}_{K_{\overline{B}}X_2B\mathbf{F}}, \mathrm{P}_{K_{\overline{B}}} \times \mathrm{P}_{X_2B\mathbf{F}}\right) \le \delta_1, \tag{35}$$
$$d\left(\mathrm{P}_{BK_0K_1X_1\mathbf{F}}, \mathrm{P}_B \times \mathrm{P}_{K_0K_1X_1\mathbf{F}}\right) \le \delta_2, \tag{36}$$

where $\overline{B} = 1 \oplus B$. The first condition above denotes the reliability of OT, while the second and the third conditions ensure secrecy for party 1 and 2, respectively. Denote by $L_{\epsilon,\delta_1,\delta_2}(X_1, X_2)$ the largest l such that a protocol realizing an $(\epsilon, \delta_1, \delta_2)$-OT of length l exists.

When the underlying observations X_1, X_2 consist of n-length IID sequences X_1^n, X_2^n with common distribution $\mathrm{P}_{X_1X_2}$, it is known that $L_{\epsilon,\delta_1,\delta_2}(X_1^n, X_2^n)$ may grow linearly with n (cf. [48], [2]); the largest rate of growth is called the OT capacity.

Definition 9 (OT capacity). For $0 < \epsilon < 1$, the ϵ-OT capacity

[19] By definition, it is unique up to relabeling.

[20] Strictly speaking, OT refers to the problem where the strings K_0, K_1 and the bit B are fixed. The randomized version here is sometimes referred as *oblivious key transfer* (see [3], [81]) or *fully randomized oblivious transfer* (see [78], [42]), and they are equivalent to OT.

[21] Here, "passive adversary" refers to an "honest but curious" adversary that follows the protocol, but is curious to know the other party's input. Since we consider only converse results for OT, this assumption only strengthens our results and they remain valid for more powerful, active adversaries.

of (X_1, X_2) is defined[22] as

$$C_\epsilon(X_1, X_2) = \lim_{\delta_1, \delta_2 \to 0} \liminf_{n \to \infty} \frac{1}{n} L_{\epsilon, \delta_1, \delta_1}(X_1^n, X_2^n).$$

Then, the OT capacity is defined as

$$C(X_1, X_2) = \lim_{\epsilon \to 0} C_\epsilon(X_1, X_2).$$

The main result of this section is an upper bound on $L_{\epsilon,\delta_1,\delta_2}(X_1, X_2)$. Consequently, we recover the upper bound on $C(X_1, X_2)$ due to Ahlswede and Csiszár derived in [2]. In fact, we show that the upper bound is "strong" and applies to $C_\epsilon(X_1, X_2)$ for every $0 < \epsilon < 1$.

Heuristically, OT is feasible only when the observations of the two parties are correlated. However, no party should have an advantage over the other, and only that portion of correlated randomness observed by a party is useful which cannot be determined by the other party. Drawing on this heuristic, two different bounds for OT length are possible, each based on relating OT length to "how far" the joint distribution $\mathrm{P}_{X_1X_2}$ of the observed correlated randomness is from a useless distribution. The choice of the useless distribution is different in both bounds. In the first bound, we consider a distribution such that X_1 and X_2 are independent given $V_0 = \mathrm{mcf}(X_1, X_2)$. For such distributions, once the shared knowledge of each party is factored out, no correlation is available to facilitate OT. In the second bound, we consider distributions where $V_1 = \mathrm{mss}(X_2|X_1)$ can be determined by X_2. Note that for such distributions the factorization $\mathrm{P}_{V_1V_1X_2} = \mathrm{P}_{V_1|X_2}\mathrm{P}_{V_1|X_2}\mathrm{P}_{X_2}$ holds. Such distributions are useless for OT since the essential part of X_1 that is correlated with X_2, namely V_1, can be determined by X_2, thereby giving an advantage to Party 2. As in the case of SK agreement, we shall measure the distance between two distributions using β_ϵ. In fact, our proof entails reducing SK agreement to OT; the reductions used for the two bounds are different.

Theorem 13 (**Single-shot bound for OT length**). *For RVs X_1, X_2, $V_0 = \mathrm{mcf}(X_1, X_2)$ and $V_1 = \mathrm{mss}(X_2|X_1)$, the following inequalities hold:*

$$L_{\epsilon,\delta_1,\delta_2}(X_1, X_2) \leq -\log \beta_\eta \left(\mathrm{P}_{X_1X_2V_0}, \mathrm{P}_{X_1|V_0}\mathrm{P}_{X_2|V_0}\mathrm{P}_{V_0}\right) + 2\log(1/\xi), \quad (37)$$

$$L_{\epsilon,\delta_1,\delta_2}(X_1, X_2) \leq -\log \beta_\eta \left(\mathrm{P}_{V_1V_1X_2}, \mathrm{P}_{V_1|X_2}\mathrm{P}_{V_1|X_2}\mathrm{P}_{X_2}\right) + 2\log(1/\xi), \quad (38)$$

for all $\xi > 0$ with $\eta = \epsilon + \delta_1 + 2\delta_2 + \xi < 1$.

Corollary 14 (**Strong bound for OT capacity**). *For $0 < \epsilon < 1$, the ϵ-OT capacity of (X_1, X_2) satisfies*

$$C_\epsilon(X_1, X_2) \leq \min\{I(X_1 \wedge X_2|V_0),\, H(V_1|X_2)\}, \quad (39)$$

where $V_0 = \mathrm{mcf}(X_1, X_2)$ and $V_1 = \mathrm{mss}(X_2|X_1)$.

The proof of Theorem 13 entails reducing two SK agreement problems to OT[23]. The bound (37) is obtained by recovering K_B as a SK, while (38) is obtained by recovering $K_{\overline{B}}$ as a SK; we note these two reductions as separate lemmas below.

Lemma 15 (**Reduction 1 of SK agreement to OT**). *Consider SK agreement for two parties observing X_1 and X_2, respectively, with the eavesdropper observing $V_0 = \mathrm{mcf}(X_1, X_2)$. Given a protocol realizing an $(\epsilon, \delta_1, \delta_2)$-OT of length l, there exists a protocol for generating an $(\epsilon + \delta_1 + 2\delta_2)$-SK of length l. In particular,*

$$L_{\epsilon,\delta_1,\delta_2}(X_1, X_2) \leq S_{\epsilon+\delta_1+2\delta_2}(X_1, X_2|V_0).$$

Lemma 16 (**Reduction 2 of SK agreement to OT**). *Consider two party SK agreement where Party 1 observes X_1, Party 2 observes $(V_1, X_2) = (\mathrm{mss}(X_2|X_1), X_2)$ and the eavesdropper observes X_2. Given a protocol realizing an $(\epsilon, \delta_1, \delta_2)$-OT of length l, there exists a protocol for generating an $(\epsilon+\delta_1+2\delta_2)$-SK of length l. In particular,*

$$L_{\epsilon,\delta_1,\delta_2}(X_1, X_2) \leq S_{\epsilon+\delta_1+2\delta_2}(X_1, (V_1, X_2)|X_2).$$

Remark 4. Underlying the proof of $C(X_1, X_2) \leq I(X_1 \wedge X_2)$ in [2] was a reduction of SK agreement to OT, which is extended in our proof below to prove (37). In contrast, the proof of the bound $C(X_1, X_2) \leq H(X_1|X_2)$ in [2] relied on manipulations of entropy terms. Below we give an alternative reduction argument to prove (38).

Remark 5. In general, our bounds are stronger than those presented in [79]. For instance, the latter is loose when the observations consist of mixtures of IID RVs. Further, while both (38) and [79, Theorem 5] (specialized to OT) suffice to obtain the second bound in Corollary 14, in contrast to (37), [79, Theorem 2] does not yield the first bound in Corollary 14.

Remark 6. For simplicity of presentation, we did not allow local randomization in the formulation above. However, it can be easily included as a part of X_1 and X_2 by replacing X_i with (X_i, U_i), $i = 1, 2$, where $U_1, U_2, (X_1, X_2)$ are mutually independent. Since our proofs are based on reduction of SK agreement to OT, by noting that $\mathrm{mss}(X_2, U_2|X_1, U_1) = \mathrm{mss}(X_2|X_1)$ and that the availability of local randomness does not change our upper bound on SK length in Theorem 3, the results above remain valid even when local randomness is available.

Remark 7. The $(\epsilon, \delta_1, \delta_2)$-OT capacity $C_{\epsilon,\delta_1,\delta_2}(X_1, X_2)$ can be defined, without requiring δ_1, δ_2 to go to 0 as in the definition of $C_\epsilon(X_1, X_2)$. However, the right-side of (39) constitutes an upper bound for $C_{\epsilon,\delta_1,\delta_2}(X_1, X_2)$ only when $\epsilon + \delta_1 + 2\delta_2 < 1$, and establishing the validity of this bound for[24] $\epsilon + \delta_1 + 2\delta_2 \geq 1$ remains an open problem.

We prove Lemmas 15 and 16 next. The proof of Theorem 13 follows by Theorem 3, along with the Markov relation X_1—V_1—X_2 and the data processing inequality (5); the

[22] For brevity, we use the same notation for SK capacity and OT capacity; the meaning will be clear from the context. Similarly, the notation L, used here to denote the optimal OT length, is also used to denote the optimal BC length in the next section.

[23] A reduction of SK to OT in a computational secrecy setup appeared in [25].

[24] For $\epsilon + \delta_1 + \delta_2 \geq 1$, $C_{\epsilon,\delta_1,\delta_2}(X_1, X_2)$ can be shown to be unbounded in the manner of the discussion preceding Corollary 12.

corollary follows by Stein's Lemma (see Section II-B).

Proof. of Lemma 15. Let $\hat{K}$ be the estimate of K_B formed by Party 2. The following protocol generates an $(\epsilon+\delta_1+2\delta_2)$-SK of length l

(i) Party 1 generates two random strings K_0 and K_1 of length l, and Party 2 generates a random bit B. Two parties run the OT protocol, and Party 2 obtains an estimate $\hat{K}$ of K_B.
(ii) Party 2 sends B over the public channel.
(iii) Using B, Party 1 computes K_B.

We will show that the RVs $K_B, \hat{K}$ constitute an $(\epsilon+\delta_1+2\delta_2)$-SK. The reliability for this SK is guaranteed since both parties agree on K_B with probability greater than $1-\epsilon$. For establishing secrecy, note that if Party 2 sends $\overline{B}$ instead of B, the eavesdropper cannot determine K_B from $(\overline{B}, \mathbf{F})$ by the secrecy condition for Party 1. On the other hand, by the secrecy condition for Party 2, the overall observation $(K_0, K_1, X_1, \mathbf{F})$ of Party 1 has roughly the same distribution even when B is replaced by $\overline{B}$. Thus, the eavesdropper cannot determine K_B from $(B, \mathbf{F})$ as well.

Formally, by Proposition 1 and Remark 3, it suffices to show that for some distribution $Q_{V_0\mathbf{F}B}$ (see Remark 3),

$$d\left(\mathrm{P}_{K_B V_0 \mathbf{F} B}, \mathrm{P}_{\mathtt{unif}} \times \mathrm{Q}_{V_0 \mathbf{F} B}\right) \leq \delta_1 + 2\delta_2.$$

Observe that condition (36) is the same as

$$d\left(\mathrm{P}_{K_0 K_1 X_1 \mathbf{F}|B=0}, \mathrm{P}_{K_0 K_1 X_1 \mathbf{F}|B=1}\right) \leq 2\delta_2. \tag{40}$$

Let $\mathrm{Q}_{V_0\mathbf{F}B}\left(v, f, b\right) = \mathrm{P}_{V_0\mathbf{F}|B}\left(v, f|\overline{b}\right)\mathrm{P}_B\left(b\right)$. Then,

$$\begin{aligned}
&d\left(\mathrm{P}_{K_B V_0 \mathbf{F} B}, \mathrm{P}_{\mathtt{unif}} \times \mathrm{Q}_{V_0 \mathbf{F} B}\right)\\
&= \frac{1}{2}\sum_b d\left(\mathrm{P}_{K_b V_0 \mathbf{F}|B=b}, \mathrm{P}_{\mathtt{unif}} \times \mathrm{Q}_{V_0 \mathbf{F}|B=b}\right)\\
&= \frac{1}{2}\sum_b d\left(\mathrm{P}_{K_b V_0 \mathbf{F}|B=b}, \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_{V_0 \mathbf{F}|B=\overline{b}}\right)\\
&\leq \frac{1}{2}\sum_b \Big[d\left(\mathrm{P}_{K_b V_0 \mathbf{F}|B=\overline{b}}, \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_{V_0 \mathbf{F}|B=\overline{b}}\right)\\
&\quad + d\left(\mathrm{P}_{K_b V_0 \mathbf{F}|B=b}, \mathrm{P}_{K_b V_0 \mathbf{F}|B=\overline{b}}\right)\Big]\\
&= d\left(\mathrm{P}_{K_{\overline{B}} V_0 \mathbf{F} B}, \mathrm{P}_{\mathtt{unif}} \times \mathrm{P}_{V_0 \mathbf{F} B}\right)\\
&\quad + \frac{1}{2}\sum_b d\left(\mathrm{P}_{K_b V_0 \mathbf{F}|B=b}, \mathrm{P}_{K_b V_0 \mathbf{F}|B=\overline{b}}\right)\\
&\leq \delta_1 + 2\delta_2,
\end{aligned}$$

where the last inequality uses (35) and (40), together with the fact that V_0 is a function of X_2 as well as X_1. □

Proof. of Lemma 16. The following protocol generates an $(\epsilon+\delta_1+2\delta_2)$-SK of length l.

(i) Party 1 generates two random strings K_0 and K_1 of length l, and Party 2 generates a random bit B. Two parties run the OT protocol.
(ii) Upon observing $\mathbf{F}$, Party 2 samples $\tilde{X}_2$ according to the distribution
$\mathrm{P}_{X_2|V_1 B\mathbf{F}}\left(\cdot|V_1, \overline{B}, \mathbf{F}\right)$.
(iii) Party 2 sends B over the public channel.
(iv) Party 1 computes $K_{\overline{B}}$ and Party 2 computes $\tilde{K} = \hat{K}(\tilde{X}_2, \overline{B}, \mathbf{F})$.

We will show that the RVs $K_{\overline{B}}, \tilde{K}$ constitute an $(\epsilon+\delta_1+2\delta_2)$-SK. Heuristically, this protocol entails Party 2 emulating $\tilde{X}_2$, pretending that the protocol was executed for $\overline{B}$ instead of B. Since the communication of Party 1 is oblivious of the value of B, plugging $\tilde{X}_2$ into $\hat{K}$ will lead to an estimate of $K_{\overline{B}}$ provided that the emulated $\tilde{X}_2$ preserves the joint distribution.

By Proposition 1 and (35), it suffices to show that

$$\mathrm{P}\left(K_{\overline{B}} \neq \tilde{K}\right) \leq \epsilon + 2\delta_2. \tag{41}$$

To that end, note

$$\begin{aligned}
&\mathrm{P}\left(K_{\overline{B}} \neq \tilde{K}\right)\\
&= \frac{1}{2}\sum_{k,b,v,f} \mathrm{P}_{K_{\overline{b}} V_1 \mathbf{F}|B}\left(k, v, f|b\right) \times\\
&\qquad \mathrm{P}\left(\hat{K}(X_2, \overline{b}, f) \neq k \mid V_1 = v, B = \overline{b}, \mathbf{F} = f\right)\\
&\leq \frac{1}{2}\sum_{k,b,v,f} \mathrm{P}_{K_{\overline{b}} V_1 \mathbf{F}|B}\left(k, v, f|\overline{b}\right) \times\\
&\qquad \mathrm{P}\left(\hat{K}(X_2, \overline{b}, f) \neq k \mid V_1 = v, B = \overline{b}, \mathbf{F} = f\right) + 2\delta_2\\
&= \frac{1}{2}\sum_{k,b,v,f} \mathrm{P}_{K_b V_1 \mathbf{F}|B}\left(k, v, f|b\right) \times\\
&\qquad \mathrm{P}\left(\hat{K}(X_2, b, f) \neq k \mid V_1 = v, B = b, \mathbf{F} = f\right) + 2\delta_2\\
&= \mathrm{P}\left(K_B \neq \hat{K}\right) + 2\delta_2.
\end{aligned}$$

where the inequality uses (40) and the last equality uses the Markov relation $X_2 — V_1 B\mathbf{F} — K_0 K_1$, which holds in the view of the interactive communication property of Lemma 6; (41) follows by (34). □

C. Bit commitment

Two parties observing correlated observations X_1 and X_2 want to implement information theoretically secure BC using interactive public communication, *i.e.*, the first party seeks to report to the second the results of a series of coin tosses that it conducted at its end in such a manner that, at a later stage, Party 2 can detect if Party 1 was lying [7]. Formally, a BC protocol consists of two phases: the *commit phase* and the *reveal phase*. In the commit phase, Party 1 generates a random string K, distributed uniformly over $\{0,1\}^l$ and independent jointly of (X_1, X_2). Furthermore, the two parties communicate interactively with each other using an interactive communication $\mathbf{F}$. In the reveal phase, Party 1 "reveals" its data, *i.e.*, it sends X_1' and K', claiming these were its initial choices of X_1 and K, respectively. Subsequently, Party 2 applies a (randomized) test function $T = T(K', X_1', X_2, \mathbf{F})$, where $T = 0$ and $T = 1$, respectively, indicate $K' = K$ and $K' \neq K$.

Definition 10 (Bit commitment). An execution of a protocol realizing an $(\epsilon, \delta_1, \delta_2)$-BC of length l consists of an interactive communication $\mathbf{F}$ to be sent during the commit phase and a $\{0,1\}$-valued randomized test function T to be used in the reveal phase such that the following conditions are satisfied:

$$\mathrm{P}\left(T(K, X_1, X_2, \mathbf{F}) \neq 0\right) \leq \epsilon, \tag{42}$$

$$d\left(\mathrm{P}_{KX_2\mathbf{F}}, \mathrm{P}_K \times \mathrm{P}_{X_2\mathbf{F}}\right) \leq \delta_1, \tag{43}$$

$$\mathrm{P}\left(T(K', X_1', X_2, \mathbf{F}) = 0, K' \neq K\right) \leq \delta_2, \tag{44}$$

for any choice of RVs K' and X_1' that have the same range-sets as K and X_1, respectively[25], and satisfy

$$(K', X_1')\text{—}(K, X_1, \mathbf{F})\text{—}X_2.$$

The first condition above is the *soundness condition*, which captures the reliability of BC when Party 1 is honest. The next condition is the *hiding condition*, which ensures that Party 2 cannot ascertain the secret in the commit phase. Finally, the *binding condition* in (44) restricts the probability with which Party 1 can cheat in the reveal phase. Denote by $L_{\epsilon,\delta_1,\delta_2}(X_1, X_2)$ the largest l such that a protocol realizing an $(\epsilon, \delta_1, \delta_2)$-BC of length l exists.

For n-length IID sequences X_1^n, X_2^n generated from $\mathrm{P}_{X_1X_2}$, the largest rate of $L_{\epsilon,\delta_1,\delta_2}(X_1, X_2)$ is called the BC capacity.

Definition 11 (BC capacity). For $0 < \epsilon, \delta_1, \delta_2 < 1$, the $(\epsilon, \delta_1, \delta_2)$-BC capacity of (X_1, X_2) is defined as

$$C_{\epsilon,\delta_1,\delta_2}(X_1, X_2) = \liminf_{n\to\infty} \frac{1}{n} L_{\epsilon,\delta_1,\delta_2}(X_1^n, X_2^n).$$

The BC capacity is defined as

$$C(X_1, X_2) = \lim_{\epsilon,\delta_1,\delta_2\to 0} C_{\epsilon,\delta_1,\delta_2}(X_1, X_2).$$

The following result of Winter, Nascimento, and Imai [80] (see, also, [66, Chapter 8]) gives a simple formula for $C(X_1, X_2)$.

Theorem 17. *[80] For RVs X_1, X_2, let $V_1 = \mathrm{mss}(X_2|X_1)$. The BC capacity is given by*

$$C(X_1, X_2) = H(V_1|X_2).$$

In this section, we present an upper bound on $L_{\epsilon,\delta_1,\delta_2}(X_1, X_2)$, which in turn leads to a strong converse for BC capacity.

Theorem 18 (Single-shot bound for BC length). *Given $0 < \epsilon, \delta_1, \delta_2$, $\epsilon + \delta_1 + \delta_2 < 1$, for RVs X_1, X_2 and $V_1 = \mathrm{mss}(X_2|X_1)$, the following inequality holds:*

$$L_{\epsilon,\delta_1,\delta_2}(X_1, X_2) \leq -\log \beta_\eta \left(\mathrm{P}_{V_1V_1X_2}, \mathrm{P}_{V_1|X_2}\mathrm{P}_{V_1|X_2}\mathrm{P}_{X_2}\right) + 2\log(1/\xi),$$

for all ξ with $\eta = \epsilon + \delta_1 + \delta_2 + \xi$.

Corollary 19 (Strong converse for BC capacity). *For $0 < \epsilon, \delta_1, \delta_2$, $\epsilon + \delta_1 + \delta_2 < 1$, the $(\epsilon, \delta_1, \delta_2)$-BC capacity satisfies*

$$C_{\epsilon,\delta_1,\delta_2}(X_1, X_2) \leq H(V_1|X_2),$$

where $V_1 = \mathrm{mss}(X_2|X_1)$.

Theorem 18 is obtained by a reduction of SK agreement to BC, which is along the lines of [80], [33], [56]; the following lemma captures the resulting bound.

Lemma 20 (Reduction of SK to BC). *For $0 < \epsilon, \delta_1, \delta_2$, $\epsilon + \delta_1 + \delta_2 < 1$, it holds that*

$$L_{\epsilon,\delta_1,\delta_2}(X_1, X_2) \leq S_{\epsilon+\delta_1+\delta_2}(X_1, (V_1, X_2)|X_2),$$

where $V_1 = \mathrm{mss}(X_2|X_1)$.

Remark 8. While local randomization was not allowed in the foregoing discussion, as before (see Remark 6) our results do not change with the availability of local randomness.

Remark 9. For $\epsilon, \delta_1, \delta_2 > 0$, $\epsilon + \delta_1 + \delta_2 < 1$, the following bound on $L_{\epsilon,\delta_1,\delta_2}(X_1, X_2)$ was derived in [56, Lemma 4]:

$$L_{\epsilon,\delta_1,\delta_2}(X_1, X_2) \leq \frac{H(V_1|X_2) + h(\delta_1) + h(\epsilon + \delta_2)}{1 - \epsilon - \delta_1 - \delta_2},$$

where $h(\cdot)$ is the binary entropy function. However, this bound is weaker than Theorem 18, in general, and is not sufficient for deriving Corollary 19.

Theorem 18 follows by using Lemma 20 with Theorem 3, along with the Markov relation X_1—V_1—X_2 and the data processing inequality (5); the Corollary 19 follows by Stein's Lemma (see Section II-B). We prove Lemma 20 below.

Proof of Lemma 20. The reduction argument presented here is along the lines of [33, Proposition 9] (see, also, [56, Lemma 4]). Given an $(\epsilon, \delta_1, \delta_2)$-BC of length l, consider SK agreement by two parties observing X_1 and (V_1, X_2), respectively, with the eavesdropper observing X_2. To generate a SK, the parties run the commit phase of the BC protocol, *i.e.*, Party 1 generates $K \sim \mathtt{unif}\{0,1\}^l$ and the parties send the interactive communication $\mathbf{F}$. We show that the committed secret K constitutes a $(\epsilon + \delta_2, \delta_1)$-SK. Indeed, by the hiding condition (43), the SK K satisfies the secrecy condition (2) with $\delta = \delta_1$. To establish the reliability of this SK, we show that, roughly, K is the unique string which is compatible with $(V_1, X_2, \mathbf{F})$, namely that any other string will fail the test T, since otherwise a dishonest Party 1 can change the string in the reveal phase, contradicting the binding condition. Thus, Party 2 can obtain an estimate of K by finding the unique string that is compatible with $(V_1, X_2, \mathbf{F})$.

Formally, we complete the proof by showing that there exists $\hat{K} = \hat{K}(V_1, X_2, \mathbf{F})$ such that

$$\mathrm{P}\left(\hat{K} \neq K\right) \leq \epsilon + \delta_2. \tag{45}$$

To that end, for randomized test T, let $(\hat{k}, \hat{x}_1) = (\hat{k}(v, f), \hat{x}_1(v, f))$ be a function of (v, f) given by

$$\begin{aligned}(\hat{k}, \hat{x}_1) &= \operatorname*{argmax}_{k,x_1} \mathrm{P}\left(T(k, x_1, X_2, \mathbf{F}) = 0 \mid V_1 = v, \mathbf{F} = f\right) \\ &= \operatorname*{argmax}_{k,x_1} \sum_{x_2} \mathrm{P}_{X_2|V_1\mathbf{F}}\left(x_2|v, f\right) \mathrm{P}\left(T(k, x_1, x_2, f) = 0\right),\end{aligned}$$

and let $(\hat{K}, \hat{X}_1) = (\hat{k}(V_1, \mathbf{F}), \hat{x}_1(V_1, \mathbf{F}))$. Note that while the estimated secret $\hat{K}$ is a function of (v, f) and does not depend on X_2 directly, the latter is needed to facilitate the communication $\mathbf{F}$ in the emulation of the commit phase. For $(\hat{K}, \hat{X}_1)$ as above, we get

$$\mathrm{P}\left(T(\hat{K}, \hat{X}_1, X_2, \mathbf{F}) = 0\right)$$

[25] Note that this restriction is valid since a dishonest Party 1 seeks to replace K with K' in the reveal phase, without being caught by Party 2.

$$\begin{aligned}
&= \sum_{v,f} \mathrm{P}_{V_1\mathbf{F}}\left(v,f\right) \sum_{x_2} \mathrm{P}_{X_2|V_1\mathbf{F}}\left(x_2|v,f\right) \\
&\quad \mathrm{P}\left(T(\hat{k}(v,f),\hat{x}_1(v,f),x_2,f)=0\right) \\
&\geq \sum_{v,f} \mathrm{P}_{V_1\mathbf{F}}\left(v,f\right) \sum_{k,x_1} \mathrm{P}_{K,X_1|V_1\mathbf{F}}\left(k,x_1|v,f\right) \\
&\quad \sum_{x_2} \mathrm{P}_{X_2|V_1\mathbf{F}}\left(x_2|v,f\right) \mathrm{P}\left(T(k,x_1,x_2,f)=0\right) \\
&= \mathrm{P}\left(T(K,X_1,X_2,\mathbf{F})=0\right) \\
&\geq 1-\epsilon,
\end{aligned}$$

where the first inequality uses the definition of $(\hat{k}(v,f),\hat{x}_1(v,f))$ and the second equality uses the Markov relation KX_1—$V_1\mathbf{F}$—X_2, which holds in the view of the interactive communication property of Lemma 6. The inequality above, along with the binding condition (44), yields

$$\begin{aligned}
1-\epsilon &\leq \mathrm{P}\left(\hat{K}=K\right) + \mathrm{P}\left(T(\hat{K},\hat{X}_1,X_2,\mathbf{F})=0, \hat{K}\neq K\right) \\
&\leq \mathrm{P}\left(\hat{K}=K\right) + \delta_2,
\end{aligned}$$

which completes the proof of (45). □

We conclude this section by observing a simple application of Theorem 18 in bounding the efficiency of reduction of BC to OT. For a detailed discussion, see [56].

Example 1 (**Reduction of BC to OT**). Suppose two parties have at their disposal an OT of length n. Using this as a resource, what is the length l of $(\epsilon,\delta_1,\delta_2)$-BC that can be constructed?

Denoting by K_0, K_1 the OT strings, and by B the OT bit of Party 2, let $X_1=(K_0,K_1)$ and $X_2=(B,K_B)$. Note that (9) holds with $\mathrm{P}=\mathrm{P}_{X_1X_1X_2}$ and $\mathrm{Q}=\mathrm{P}_{X_1|X_2}\mathrm{P}_{X_1X_2}$, and

$$D(\mathrm{P}_{X_1X_1X_2}\|\mathrm{P}_{X_1|X_2}\mathrm{P}_{X_1X_2}) = n.$$

Therefore, by Theorem 18 and (10), we get

$$l \leq n + \log(1/(1-\epsilon-\delta_1-\delta_2-\eta)) + 2\log(1/\eta),$$

where $0<\eta<1-\epsilon-\delta_1-\delta_2$. This bound on efficiency of reduction is stronger than the one derived in [56, Corollary 2] (fixing $n=n'=1$ in that bound). In particular, it shows an additive loss of logarithmic order in $(1-\epsilon-\delta_1-\delta_2)$, while [56, Corollary 2] shows a multiplicative loss of linear order.

VI. Implications for secure computation with trusted parties

In this section, we present a connection of our result to a problem of secure function computation with trusted parties, where the parties seek to compute a function of their observations using a communication that does not reveal the value of the function by itself (without the observations at the terminals). This is in contrast to the secure computation treated in Section V where the communication is secure but the parties are required not to get any more information than the computed function value. This problem was introduced in [69] where a matching necessary and sufficient condition was given for the feasibility of secure computation in the asymptotic case with IID observations. Here, using Theorem 3, we derive a necessary condition for the feasibility of such secure computing for general observations (not necessarily IID).

Formally, consider $m \geq 2$ parties observing RVs $X_1,...,X_m$ taking values in finite sets $\mathcal{X}_1,...,\mathcal{X}_m$, respectively. Upon making these observations, the parties communicate interactively in order to *securely compute* a function $g : \mathcal{X}_1 \times ... \times \mathcal{X}_m \to \mathcal{G}$ in the following sense: The ith party forms an estimate $G_{(i)}$ of the function based on its observation X_i, local randomization U_i and interactive communication $\mathbf{F}$, *i.e.*, $G_{(i)} = G_{(i)}(U_i,X_i,\mathbf{F})$. For $0\leq\epsilon,\delta<1$, a function g is (ϵ,δ)-*securely computable* if there exists a protocol satisfying

$$\mathrm{P}\left(G=G_{(1)}=\cdots=G_{(m)}\right) \geq 1-\epsilon, \tag{46}$$

$$d\left(\mathrm{P}_{G\mathbf{F}},\mathrm{P}_G\times\mathrm{P}_{\mathbf{F}}\right) \leq \delta, \tag{47}$$

where $G=g\left(X_{\mathcal{M}}\right)$. The first condition captures the reliability of computation and the second condition ensures the secrecy of the protocol. Heuristically, for secrecy we require that an observer of (only) $\mathbf{F}$ must not get to know the computed value of the function. We seek to characterize the (ϵ,δ)-securely computable functions g.

In [69], an asymptotic version of this problem was addressed. The parties observe $X_1^n,...,X_m^n$ and seek to compute $G_t = g\left(X_{1t},...,X_{mt}\right)$ for each $t\in\{1,...,n\}$; consequently, the RVs $\{G_t, 1\leq t\leq n\}$ are IID. A function g is securely computable if the parties can form estimates $G_{(1)}^{(n)},...,G_{(m)}^{(n)}$ such that

$$\begin{gathered}
\mathrm{P}\left(G^n=G_{(1)}^{(n)}=\cdots=G_{(m)}^{(n)}\right) \geq 1-\epsilon_n, \\
d\left(\mathrm{P}_{G^n\mathbf{F}},\mathrm{P}_{G^n}\times\mathrm{P}_{\mathbf{F}}\right) \leq \epsilon_n,
\end{gathered}$$

where $\lim_{n\to\infty}\epsilon_n=0$. The following characterization of securely computable functions g is known.

Theorem 21. *[69] For the asymptotic case described above, a function g is securely computable if $H(G)<C$, where $H(G)$ is the entropy of the RV $G=g(X_{\mathcal{M}})$ and $C=C(X_{\mathcal{M}})$ is the SK capacity given in Theorem 11.*

Conversely, if a function g is securely computable, then $H(G)\leq C$.

Heuristically, the necessary condition above follows upon observing that if the parties can securely compute the function g, then they can extract a SK of rate $H(G)$ from RVs G^n. Therefore, $H(G)$ must be necessarily less than the maximum rate of a SK that can be generated, namely the SK capacity C. Using this heuristic, we present a necessary condition for a function g to be (ϵ,δ)-securely computable.

Corollary 22. *For $0\leq\epsilon,\delta<1$ with $\epsilon+\delta<1$, if a function g is (ϵ,δ)-securely computable, then*

$$\begin{aligned}
&H_{\min}^{\xi}(\mathrm{P}_G) \\
&\leq \frac{1}{|\pi|-1}\left[-\log\beta_\mu\left(\mathrm{P}_{X_{\mathcal{M}}},\mathrm{Q}_{X_{\mathcal{M}}}^{\pi}\right)+|\pi|\log(1/\eta)\right] \\
&\qquad + 2\log(1/2\zeta)+1,\ \forall\,\mathrm{Q}_{X_{\mathcal{M}}}^{\pi}\in\mathcal{Q}(\pi),
\end{aligned} \tag{48}$$

for every $\mu=\epsilon+\delta+2\xi+\zeta+\eta$ with $\xi,\zeta,\eta>0$ such that $\mu<1$, and for every partition π of $\mathcal{M}$.

Proof. The proof is based on extracting an ϵ-SK from the RV G that the parties share. Specifically, Lemma 2 with $X = G$, $Y =$ const, and condition (47) imply that there exists $K = K(G)$ with $\log|\mathcal{K}| = \lfloor H_{\min}^{\xi}(\mathrm{P}_G) - 2\log(1/2\zeta)\rfloor$ and satisfying

$$\begin{aligned} & d\left(\mathrm{P}_{K(G)\mathbf{F}}, \mathrm{P}_{\mathrm{unif}} \times \mathrm{P}_{\mathbf{F}}\right) \\ & \quad \leq \; d\left(\mathrm{P}_{K(G)\mathbf{F}}, \mathrm{P}_{K(G)} \times \mathrm{P}_{\mathbf{F}}\right) \\ & \qquad\qquad + d\left(\mathrm{P}_{K(G)} \times \mathrm{P}_{\mathbf{F}}, \mathrm{P}_{\mathrm{unif}} \times \mathrm{P}_{\mathbf{F}}\right) \\ & \quad \leq \; d\left(\mathrm{P}_{G\mathbf{F}}, \mathrm{P}_{G} \times \mathrm{P}_{\mathbf{F}}\right) + d\left(\mathrm{P}_{K(G)}, \mathrm{P}_{\mathrm{unif}}\right) \\ & \quad \leq \; \delta + 2\xi + \zeta. \end{aligned}$$

Thus, in the view of Proposition 1, the RV K constitutes[26] an $(\epsilon + \delta + 2\xi + \zeta)$-SK. An application of Theorem 3 gives (48). □

We conclude this section with two illustrative examples.

Example 2. **(Computing functions of independent observations using a perfect SK).** Suppose the ith party observes U_i, where the RVs $U_1, ..., U_m$ are mutually independent. Furthermore, all parties share a κ-bit perfect SK K which is independent of $U_{\mathcal{M}}$. How many bits κ are required to (ϵ, δ)-securely compute a function $g\left(U_1, ..., U_m\right)$?

Note that the data observed by the ith party is given by $X_i = (U_i, K)$. A simple calculation shows that for every partition π of $\mathcal{M}$,

$$\beta_\epsilon\left(\mathrm{P}_{X_{\mathcal{M}}}, \prod_{i=1}^{|\pi|} \mathrm{P}_{X_{\pi_i}}\right) \geq (1-\epsilon)\kappa^{1-|\pi|},$$

and therefore, by Corollary 22 a necessary condition for g to be (ϵ, δ)-securely computable is

$$\begin{aligned} H_{\min}^{\xi}(\mathrm{P}_G) \leq \kappa + \frac{1}{|\pi|-1}\left(|\pi|\log\frac{1}{\eta} + \log\frac{1}{1-\mu}\right) \\ + 2\log\frac{1}{2\zeta} + 1, \qquad (49) \end{aligned}$$

for every $\xi, \zeta, \eta > 0$ satisfying $\mu = \epsilon + \delta + 2\xi + \zeta + \eta < 1$. Note that the finest partition, i.e., $|\pi| = m$, gives the tightest bound in (49).

For the special case when $U_i = B_i^n$, a sequence of independent, unbiased bits, and

$$g\left(B_1^n, ..., B_m^n\right) = B_{11} \oplus ... \oplus B_{m1}, ..., B_{1n} \oplus ... \oplus B_{mn},$$

i.e., the parties seek to compute the (element-wise) parities of the bit sequences, it holds that $H_{\min}^{\xi}(\mathrm{P}_G) \geq n$. Therefore, (ϵ, δ)-secure computation is feasible only if $n \leq \kappa + O(1)$. We remark that this necessary condition is also (almost) sufficient. Indeed, if $n \leq \kappa$, all but the mth party can reveal all their bits $B_1^n, \dots, B_{m-1}^n$ and the mth party can send back $B_1^n \oplus \dots \oplus B_m^n \oplus K_n$, where K_n denotes any n out of κ bits of K. Clearly, this results in a secure computation of g.

Example 3. **(Secure transmission).** Two parties sharing a κ-bit perfect SK K seek to exchange a message M securely.[27] To this end, they communicate interactively using a communication $\mathbf{F}$, and based on this communication Party 2 forms an estimate $\hat{M}$ of the message M by Party 1. This protocol accomplishes (ϵ, δ)-secure transmission if

$$\begin{aligned} \mathrm{P}\left(M = \hat{M}\right) \geq 1 - \epsilon, \\ d\left(\mathrm{P}_{M\mathbf{F}}, \mathrm{P}_M \times \mathrm{P}_{\mathbf{F}}\right) \leq \delta. \end{aligned}$$

The classic result of Shannon [62] implies that $(0,0)$-secure transmission is feasible only if κ is at least $\log\|M\|$, where $\|M\|$ denotes the size of the message space.[28] But, can we relax this constraint for $\epsilon, \delta > 0$? In this example, we will give a necessary condition for the feasibility of (ϵ, δ)-secure transmission by relating it to the previous example.

Specifically, let the observations of the two parties consist of $X_1 = (M, K)$, $X_2 = K$. Then, (ϵ, δ)-secure transmission of M is tantamount to securely computing the function $g(X_1, X_2) = M$. Therefore, using (49), (ϵ, δ)-secure transmission of M is feasible only if

$$H_{\min}^{\xi}(\mathrm{P}_M) \leq \kappa + 2\log\frac{1}{\eta} + \log\frac{1}{1-\mu} + 2\log\frac{1}{2\zeta} + 1, \quad (50)$$

for every $\xi, \zeta, \eta > 0$ satisfying $\mu = \epsilon + \delta + 2\xi + \zeta + \eta < 1$.

Condition (50) brings out a trade-off between κ and $\epsilon + \delta$ (cf. [38, Problems 2.12 and 2.13]). For an illustration, consider a message M consisting of a RV Y taking values in a set $\mathcal{Y} = \{0,1\}^n \cup \{0,1\}^{2n}$ and with the following distribution:

$$\mathrm{P}_Y(y) = \left\{ \begin{array}{ll} \frac{1}{2} \cdot \frac{1}{2^n} & y \in \{0,1\}^n \\ \frac{1}{2} \cdot \frac{1}{2^{2n}} & y \in \{0,1\}^{2n} \end{array} \right. .$$

For $\epsilon + \delta = 0$, we know that secure transmission will require κ to be more than the *worst-case message length* $2n$. But perhaps by allowing $\epsilon + \delta$ to be greater than 0, we can make do with fewer SK bits; for instance, perhaps κ equal to $H(M) = (3/2)n + 1$ will suffice (note that the *average message length* equals $(3/2)n$). The necessary condition above says that this is not possible if $\epsilon + \delta < 1/2$. Indeed, since $H_{\min}^{\xi}(\mathrm{P}_Y) \geq 2n$ for $\xi = 1/4$, we get from (50) that the message $M = Y$ can be (ϵ, δ)-securely transmitted only if $2n \leq \kappa + O(1)$, where the constant depends on ϵ and δ.

VII. Discussion

In this work, we focused on converse results and presented single-shot upper bounds on the efficiency of using correlated randomness for SK agreement and secure computation protocols. When the underlying observations are IID, the resulting upper bounds were shown to be tight in several cases. It is natural to ask how tight are these bounds for IID observations of fixed, finite length. For the SK agreement problem, it is possible to mimic the approach in [44], [1], [16], [60] to obtain protocols that first use communication for *information reconciliation* and then extract SKs using *privacy amplification*. The challenge in the multiparty setup is to identify the appropriate *information to be reconciled*. For the case of two parties observing IID sequences, relying on Theorem 3, recently the second-order asymptotic term in the maximum length of a SK was established in [29], [30].

[26] Strictly speaking, the estimates $K_1, ..., K_m$ of K formed by different parties constitute the $(\epsilon + \delta + 2\xi + \zeta)$-SK in the sense of (3).

[27] A message M is a RV with known distribution P_M.

[28] This is a slight generalization of Shannon's original result; see [38, Theorem 2.7] for a proof.

Coming up with finite-length schemes that match the converse bounds for the various secure computation problems studied above is work in progress.

Our converse results in Sections V and VI entail reducing SK agreement to the secure computation task at hand, followed by an application of Theorem 3. The strength of Theorem 3 lies in its validity for interactive communication. The admissibility of interactive communication makes this bound useful in cryptography where interaction is natural to consider, and it is foreseeable, and indeed tempting, that this approach can lead to converse bounds for other problems in information theoretic secrecy and cryptography; an instance arises in [31].

In fact, our bound can find applications in problems involving interactive communication without any secrecy requirements. For instance, it is used in [71] to derive a lower bound for the length of the interactive communication needed for two parties to exchange their correlated data. Furthermore, it is used in [70] to derive a lower bound on the communication complexity for simulating protocols.

Note that similar to [52], [28] the choice of Q in Theorem 3 is arbitrary. In the applications to capacity results considered in this paper and in deriving the second-order asymptotics for the two party SK agreement problem in [29], Q equal to the product of marginals of P suffices. However, in a more involved application of Theorem 3, such as that in [71], [70], a judicious choice of Q is needed.

A quantum version of the two party secret key agreement problem of [44], [1] has been considered in [19], [11]. An extension of Theorem 3 to the case of quantum observations can be used to obtain converse results for such problems. In the classical case, for two parties with IID observations, Theorem 3 shows that the (ϵ, δ)-SK capacity is bounded above by

$$\begin{aligned} &\min_{\mathrm{Q}_{X_1|Z}\mathrm{Q}_{X_2|Z}\mathrm{Q}_Z} D(\mathrm{P}_{X_1X_2Z}\|\mathrm{Q}_{X_1|X}\mathrm{Q}_{X_2|Z}\mathrm{Q}_Z) \\ &= I(X_1 \wedge X_2|Z), \end{aligned} \tag{51}$$

where the equality follows from the Topsøe identity [65]. On the other hand, in the quantum case, the identity (51) does not hold [32]. Thus, a direct extension of Theorem 3 to quantum observations will not yield the quantum conditional mutual information bound for SK capacity derived in [11]. Finding an appropriate extension of Theorem 3 to the case of quantum observations is an interesting direction for future research.

Acknowledgment

Main part of this work was carried out when SW was visiting University of Maryland, College Park. SW would like to thank Prakash Narayan for his kind hospitality and valuable discussions. The authors would like to thank Prakash Narayan for helpful comments on an earlier draft of this paper. Also, the authors would also like to thank an anonymous reviewer whose detailed comments helped in improving the presentation of the paper. HT was supported in part by the U.S. National Science Foundation under the grant CCF1117546. SW was supported in part by Japan Society for the Promotion of Science Postdoctoral Fellowships for Research Abroad.

References

[1] R. Ahlswede and I. Csiszár, “Common randomness in information theory and cryptography–part i: Secret sharing,” *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, July 1993.

[2] ——, “On oblivious transfer capacity,” *Information Theory, Combinatorics, and Search Theory*, pp. 145–166, 2013.

[3] D. Beaver, “Precomputing oblivious transfer,” in *Advances in Cryptology - CRYPTO*, 1995, pp. 97–109.

[4] A. Beimel, Y. Ishai, R. Kumaresan, and E. Kushilevitz, “On the cryptographic complexity of the worst functions,” in *In TCC*, 2014, pp. 317–342.

[5] C. H. Bennett, G. Brassard, C. Crépeau, and U. M. Maurer, “Generalized privacy amplification,” *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1915–1923, November 1995.

[6] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, “Mixed-state entanglement and quantum error correction,” *Phys. Rev. A*, vol. 54, pp. 3824–3851, November 1996.

[7] M. Blum, “Coin flipping by telephone a protocol for solving impossible problems,” *SIGACT News*, vol. 15, no. 1, pp. 23–27, Jan. 1983.

[8] R. Canetti, “Universally composable security: a new paradigm for cryptographic protocols,” *Proc. Annual Symposium on Foundations of Computer Science (also, see Cryptology ePrint Archive, Report 2000/067)*, pp. 136–145, 2001.

[9] N. Cerf, S. Massar, and S. Schneider, “Multipartite classical and quantum secrecy monotones,” *Physical Review A*, vol. 66, no. 4, p. 042309, October 2002.

[10] C. Chan and L. Zheng, “Mutual dependence for secret key agreement,” *Proc. Annual Conference on Information Sciences and Systems (CISS)*, 2010.

[11] M. Christandl, A. Ekert, M. Horodecki, P. Horodecki, J. Oppenheim, and R. Renner, “Unifying classical and quantum key distillation,” in *TCC*, 2007, pp. 456–478.

[12] C. Crépeau and J. Kilian, “Weakening security assumptions and oblivious transfer,” in *Advances in Cryptology - CRYPTO*, 1990, pp. 2–7.

[13] C. Crépeau, “Efficient cryptographic protocols based on noisy channels,” in *Advances in Cryptology EUROCRYPT*, 1997, pp. 306–317.

[14] I. Csiszár, “Almost independence and secrecy capacity,” *Prob. Pered. Inform.*, vol. 32, no. 1, pp. 48–57, 1996.

[15] I. Csiszár and J. Körner, *Information theory: Coding theorems for discrete memoryless channels. 2nd edition.* Cambridge University Press, 2011.

[16] I. Csiszár and P. Narayan, “Secrecy capacities for multiple terminals,” *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3047–3061, December 2004.

[17] ——, “Secrecy capacities for multiterminal channel models,” *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2437–2452, June 2008.

[18] N. Datta, “Min- and max-relative entropies and a new entanglement monotone,” *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2816–2826, June 2009.

[19] I. Devetak and A. Winter, “Distillation of secret key and entanglement from quantum states,” *Proc. Roy. Soc. London A*, vol. 461, no. 2053, pp. 207–235, January 2005.

[20] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, “Fuzzy extractors: How to generate strong keys from biometrics and other noisy data,” *SIAM Journal on Computing*, vol. 38, no. 1, pp. 97–139, 2008.

[21] S. Even, O. Goldreich, and A. Lempel, “A randomized protocol for signing contracts,” *Communications of ACM*, vol. 28, no. 6, pp. 637–647, Jun. 1985.

[22] W. Feller, *An Introduction to Probability Theory and its Applications, Volume II. 2nd edition.* John Wiley & Sons Inc., UK, 1971.

[23] M. Fitzi, S. Wolf, and J. Wullschleger, “Pseudo-signatures, broadcast, and multi-party computation from correlated randomness,” in *Advances in Cryptology - CRYPTO*, 2004, pp. 562–578.

[24] P. Gács and J. Körner, “Common information is far less than mutual information,” *Problems of Control and Information Theory*, vol. 2, no. 2, pp. 149–162, 1973.

[25] Y. Gertner, S. Kannan, T. Malkin, O. Reingold, and M. Viswanathan, “The relationship between public key encryption and oblivious transfer,” in *Proc. Annual Symposium on Foundations of Computer Science (FOCS)*, 2000, pp. 325–335.

[26] A. A. Gohari and V. Anantharam, “Information-theoretic key agreement of multiple terminals: Part i,” *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3973 – 3996, August 2010.

[27] M. Hayashi, *Quantum Information: An Introduction.* Springer-Verlag Berlin Heidelberg, 2006.

[28] M. Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1753–1768, July 2003.

[29] M. Hayashi, H. Tyagi, and S. Watanabe, "Secret key agreement: General capacity and second-order asymptotics," *arXiv:1411.0735*, 2014.

[30] ——, "Secret key agreement: General capacity and second-order asymptotics," *Proc. IEEE International Symposium on Information Theory*, pp. 1136–1140, 2014.

[31] ——, "Strong converse for a degraded wiretap channel via active hypothesis testing," *Proc. Conference on Communication, Control, and Computing (Allerton)*, 2014.

[32] B. Ibinson, N. Linden, and A. Winter, "Robustness of quantum markov chains," *Commun. Math. Phys.*, pp. 289–304, 2008.

[33] H. Imai, K. Morozov, A. C. Nascimento, and A. Winter, "Efficient protocols achieving the commitment capacity of noisy correlations," in *Proc. IEEE International Symposium on Information Theory*, 2006, pp. 1432–1436.

[34] R. Impagliazzo, L. A. Levin, and M. Luby, "Pseudo-random generation from one-way functions," in *Proc. ACM Symposium on Theory of Computing (STOC)*, 1989, pp. 12–24.

[35] R. Impagliazzo and D. Zuckerman, "How to recycle random bits," in *Proc. Annual Symposium on Foundations of Computer Science (FOCS)*, 1989, pp. 248–253.

[36] M. Iwamoto and J. Shikata, "Information theoretic security for encryption based on conditional Rényi entropies," in *Information Theoretic Security*. Springer International Publishing, 2014, pp. 103–121.

[37] S. Kamath and V. Ananthram, "A new dual to the Gács-Körner common information defined via the Gray-Wyner system," *Proc. Conference on Communication, Control, and Computing (Allerton)*, pp. 1340–1346, 2010.

[38] J. Katz and Y. Lindell, *Introduction to Modern Cryptography*. Chapman & Hall/CRC, 2007.

[39] J. Kilian, "Founding crpytography on oblivious transfer," in *Proc. Symposium on Theory of Computing (STOC)*, 1988, pp. 20–31.

[40] ——, "More general completeness theorems for secure two-party computation," in *Proc. Symposium on Theory of Computing (STOC)*, 2000, pp. 316–324.

[41] R. König, R. Renner, and C. Schaffner, "The operational meaning of min- and max-entropy," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4337–4347, Sept 2009.

[42] R. König, S. Wehner, and J. Wullschleger, "Unconditional security from noisy quantum storage," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1962–1984, March 2012.

[43] S. Kullback, *Information Theory and Statistics*. Dover Publications, 1968.

[44] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[45] ——, "Authentication theory and hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1350–1356, July 2000.

[46] U. M. Maurer and S. Wolf, "Unconditionally secure key agreement and the intrinsic conditional information," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 499–514, March 1999.

[47] P. Narayan, H. Tyagi, and S. Watanabe, "Common randomness for secure computing," *To appear, Proc. IEEE International Symposium on Information Theory*, 2015.

[48] A. C. A. Nascimento and A. Winter, "On the oblivious-transfer capacity of noisy resources," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2572–2581, 2008.

[49] T. Ogawa and H. Nagaoka, "Strong converse and stein's lemma in quantum hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2428–2433, Nov 2000.

[50] A. Orlitsky and A. E. Gamal, "Communication with secrecy constraints," in *Proc. ACM Symposium on Theory of Computing (STOC)*, 1984, pp. 217–224.

[51] R. S. Pappu, "Physical one-way functions," *Ph. D. Dissertation, Massachussetts Institute of Technology*, 2001.

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

[53] Y. Polyanskiy and S. Verdú, "Arimoto channel coding converse and Rényi divergence," *Proc. Conference on Communication, Control, and Computing (Allerton)*, pp. 1327–1333, 2010.

[54] V. Prabhakaran and M. Prabhakaran, "Assisted common information with applications to secure two-party computation," *IEEE Trans. Inf. Theory*, vol. 60, no. 6, pp. 3413–3434, June 2014.

[55] M. O. Rabin, "How to exchange secrets with oblivious transfer," Cryptology ePrint Archive, Report 2005/187, 2005, http://eprint.iacr.org/.

[56] S. Ranellucci, A. Tapp, S. Winkler, and J. Wullschleger, "On the efficiency of bit commitment reductions," in *Proc. ASIACRYPT*, 2011, pp. 520–537.

[57] K. S. Rao and V. M. Prabhakaran, "A new upperbound for the oblivious transfer capacity of discrete memoryless channels," in *Proc. IEEE Information Theory Workshop*, 2014, pp. 35–39.

[58] R. Renner, "Security of quantum key distribution," *Ph. D. Dissertation, ETH Zurich*, 2005.

[59] R. Renner and S. Wolf, "New bounds in secret-key agreement: The gap between formation and secrecy extraction," in *Proc. EUROCRYPT*, 2003, pp. 562–577.

[60] ——, "Simple and tight bounds for information reconciliation and privacy amplification," in *Proc. ASIACRYPT*, 2005, pp. 199–216.

[61] A. Rényi, "On measures of entropy and information," *Proc. Fourth Berkeley Symposium on Mathematics Statistics and Probability, Vol. 1 (Univ. of Calif. Press)*, pp. 547–561, 1961.

[62] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.

[63] M. Tomamichel, "A framework for non-asymptotic quantum information theory," *Ph. D. Dissertation, ETH Zurich*, 2012, arXiv:1203.2142.

[64] M. Tomamichel and V. Y. F. Tan, "A tight upper bound for the third-order asymptotic for most discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7041–7051, Nov. 2013.

[65] F. Topsøe, "An information theoretical identity and a problem involving capacity," *Studia Sci. Math. Hungary*, pp. 291–292, 1967.

[66] P. Tuyls, B. Škorić, and T. Kevenaar (Eds), *Security with Noisy Data*. Springer, 2007.

[67] H. Tyagi, "Common information and secret key capacity," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5627–5640, 2013.

[68] H. Tyagi and P. Narayan, "How many queries will resolve common randomness?" *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5363–5378, September 2013.

[69] H. Tyagi, P. Narayan, and P. Gupta, "When is a function securely computable?" *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6337–6350, October 2011.

[70] H. Tyagi, S. Venkatakrishnan, P. Viswanath, and S. Watanabe, "Information complexity density and simulation of protocols," *arXiv:1504.05666*, 2015.

[71] H. Tyagi, P. Viswanath, and S. Watanabe, "Interactive communication for data exchange," *IEEE International Symposium on Information Theory*, pp. 1806–1810, 2015.

[72] H. Tyagi and S. Watanabe, "Secret key capacity for multipleaccess channel with public feedback," *Proc. Conference on Communication, Control, and Computing (Allerton)*, pp. 1–7, 2013.

[73] ——, "A bound for multiparty secret key agreement and implications for a problem of secure computing," in *Proc. EUROCRYPT*, 2014, pp. 369–386.

[74] V. Vedral and M. B. Plenio, "Entanglement measures and purification procedures," *Phys. Rev. A*, vol. 57, pp. 1619–1633, March 1998.

[75] L. Wang and R. Renner, "One-shot classical-quantum capacity and hypothesis testing," *Phys. Rev. Lett.*, vol. 108, no. 20, p. 200501, May 2012.

[76] S. Watanabe and M. Hayashi, "Non-asymptotic analysis of privacy amplification via Rényi entropy and inf-spectral entropy," in *Proc. IEEE International Symposium on Information Theory*, 2013, pp. 2715–2719.

[77] ——, "Finite-length analysis on tail probability for Markov chain and application to simple hypothesis testing," *arXiv:1401.3801*, 2014.

[78] S. Wehner, M. Curty, C. Schaffner, and H.-K. Lo, "Implementation of two-party protocols in the noisy-storage model," *Physical Review A*, vol. 81, no. 5, p. 052336, May 2010.

[79] S. Winkler and J. Wullschleger, "On the efficiency of classical and quantum secure function evaluation," *IEEE Trans. Inf. Theory*, vol. 60, no. 6, pp. 3123–3143, June 2014.

[80] A. Winter, A. C. A. Nascimento, and H. Imai, "Commitment capacity of discrete memoryless channels," in *Proc. Cryptography and Coding*, 2003, pp. 35–51.

[81] S. Wolf and J. Wullschleger, "Oblivious transfer is symmetric," in *Proc. EUROCRYPT*, 2006, pp. 222–232.

[82] ——, "New monotones and lower bounds in unconditional two-party computation," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2792–2797, June 2008.

[83] A. C. Yao, "Protocols for secure computations," in *Proc. Annual Symposium on Foundations of Computer Science (FOCS)*, 1982, pp. 160–164.

Himanshu Tyagi received the Bachelor of Technology degree in electrical engineering and the Master of Technology degree in communication and information technology, both from the Indian Institute of Technology, Delhi, India in 2007. He received the Ph.D. degree in Electrical and Computer Engineering from the University of Maryland, College Park. From 2013 to 2014, he was a postdoctoral researcher at the Information Theory and Applications (ITA) Center, University of California, San Diego. Since January 2015, he has been an Assistant Professor at the Indian Institute of Science in Bangalore.

Shun Watanabe (M'09) received the B.E., M.E., and Ph.D. degrees from the Tokyo Institute of Technology in 2005, 2007, and 2009, respectively. During April 2009 to February 2015, he was an Assistant Professor in the Department of Information Science and Intelligent Systems at the University of Tokushima. During April 2013 to March 2015, he was a visiting Assistant Professor in the Institute for Systems Research at the University of Maryland, College Park. Since February 2015, he has been an Associate Professor in the Department of Computer and Information Sciences at Tokyo University of Agriculture and Technology. His current research interests are in the areas of information theory, quantum information theory, cryptography, and computer science.