Converses For Secret Key Agreement and Secure Computing

Himanshu Tyagi and Shun Watanabe

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 computing by trusted parties that seek to compute a function of their collective data, using 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 computing 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 computing. We present *single-shot* converse

[1]With restricted interpretations of *feasibility*, these observations appear across the vast literature on SK agreement and secure computing; see, for instance, [34], [1], [14], [45], [32], [62], [36].

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 and 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 security 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. This is in the spirit of [39], 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, [58], [24]). Also, our upper bound is reminiscent of the *measure of entanglement* for a quantum state proposed in [57],

[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 [56].

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 [57], 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 computing 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. SK agreement

For two parties, the problem of SK agreement from correlated observations is well-studied. The problem was introduced by Maurer [34] 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. [38], [17]), 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 [45] 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 [14] (also, see [9] for an early formulation). 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[4]. Unlike the single-shot upper bound in [45], 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 security index to be asymptotically[5] 0. See Section IV for a detailed discussion.

[4]A single-shot upper bound using Fano's inequality for the length of a multiparty SK, obtained as a straightforward extension of [14], [15], was reported in [55].

[5]Such bounds that do not require the probability of error to vanish to 0 are called *strong converse* bounds [13].

B. Secure computing

Secure computing by two parties was introduced by Yao in [65]. 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 computing.

OT between two parties is a mode of message transmission "where the sender does not know whether the recipient actually received the information" [41]. In this paper, we consider the *one-of-two OT* problem [18] 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 the first and the second party. This simply stated problem is at the heart of secure function computation as it is well-known [31] 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. [11],[2] [36]). 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 [61]. 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 [36] and [2] 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 [42]. 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 [62].

C. Secure computing with trusted parties

In a different direction, we relate our result to the following problem of *secure function computation with trusted parties* introduced in [53] (for an early version of the problem, see [37]): 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 computing 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 [53], 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 bit commitment problem; and Section VI contains converse results for the secure computing 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 $\mathcal{X}_i$, $1 \leq i \leq m$.[6] 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, $1 \leq j \leq r$, the ith party sends F_{ij}, which is a function of its observation X_i, a *locally generated* randomness[7] 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 [34], [12], [14])

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

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

[6]The conditional independence testing bound given in Theorem 3 remains valid for RVs taking countably many values.

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

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

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

The first condition above represents the reliable *recovery* of the SK and the second condition guarantees *security*. In this work, we use the following alternative definition of a SK, which conveniently combines the recoverability and the security conditions (cf. [43]): The RVs $K_1,...,K_m$ above constitute an ϵ-SK with common range $\mathcal{K}$ if

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

where

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

In fact, the two definitions above are closely related[8].

Proposition 1. *Given $0\leq\epsilon,\delta\leq 1$, 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 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.[9]

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

Definition 1. Given $0\leq\epsilon<1$, 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.

[8]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.

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

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 $T(1|x) = 1 - T(0|x)$. For $0 \leq \epsilon < 1$, 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.,

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

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 $x \in \mathcal{X}$, $W(\cdot \mid x)$ is a distribution on $\mathcal{Y}$. Then,

$$\beta_\epsilon(\mathrm{P},\mathrm{Q}) \leq \beta_\epsilon(\mathrm{P}\circ W, \mathrm{Q}\circ W), \tag{5}$$

where $(\mathrm{P}\circ W)(y) = \sum_x \mathrm{P}\,(x)\,W(y \mid x)$.

2) **Stein's Lemma.** (cf. [33, 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 $0\log(0/0) = 0$.

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:

$$-\log \beta_\epsilon(\mathrm{P},\mathrm{Q}) \leq \inf_\gamma \gamma - \log\left(\mathrm{P}\left(\log\frac{\mathrm{P}(X)}{\mathrm{Q}(X)} \leq \gamma\right) - \epsilon\right). \tag{7}$$

When P and Q correspond to IID RVs, the tail probability in (7) can be numerically evaluated directly or can be approximated by the Bérry-Esséen theorem (cf. [19]). 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 [60]. 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 given by [46]

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

the following simple, closed-form bound on $\beta_\epsilon(\mathrm{P},\mathrm{Q})$ is obtained[10]:

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

While this bound is not tight in general, as its corollary we obtain Stein's lemma (see (6)).

Finally, we remark that when the condition

$$\log\frac{\mathrm{P}(X)}{\mathrm{Q}(X)} = D(\mathrm{P}\|\mathrm{Q}) \tag{8}$$

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

$$-\log\beta_\epsilon(\mathrm{P},\mathrm{Q}) \leq D(\mathrm{P}\|\mathrm{Q}) + \log(1/(1-\epsilon)). \tag{9}$$

C. Smooth minimum entropy and smooth maximum divergence

Given two RVs X and Y, a central question of information theoretic secrecy is (cf. [27], [28], [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 minimum entropy* introduced in [45], [43] is a more relevant measure of randomness. We use the definition of smooth min entropy introduced[11] in [43]; for a review of other variations, see [48].

[10]For other connections between β_ϵ and Rényi's divergence, see [40].

[11]A review of the notion of smooth minimum entropy without the notations from quantum information theory can be also found in [59].

We also review the *leftover hash lemma* [27], [5], which brings out the central role of *smooth minimum entropy* in the answer to the question above. Also, as a "change of measure companion" for smooth minimum entropy, we define *smooth maximum divergence* and note that it satisfies the data processing inequality.

Definition 2. (Minimum entropy) The minimum entropy of P is defined as

$$H_{\min}(\mathrm{P}) := \min_{x} \log \frac{1}{\mathrm{P}\,(x)}.$$

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

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

Finally, the conditional minimum entropy[12] of P_{XY} given Y is defined as

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

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

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

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

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

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

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

where the $\sup$ is over all subnormalized, nonnegative functions $\tilde{\mathrm{P}}_{XY}$. When Y is a constant, the ϵ-smooth minimum 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.

[12]There is no consensus on the definition of conditional minimum entropy. The form here is appropriate for our purpose.

Lemma 2. (Leftover hash) [43] *Given a joint distribution* P_{XY}*, for every* $0 \leq 2\epsilon < 1$ *and* $0 < \eta$ *there exists a mapping*[13] $K : \mathcal{X} \to \mathcal{K}$ *with* $\log|\mathcal{K}| = \lfloor H^{\epsilon}_{\min}(\mathrm{P}_{XY}|Y) - 2\log(1/2\eta)\rfloor$ *such that*

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

Finally, we review smooth maximum divergence, which was introduced first in [16] for a quantum setting. The method of smoothing in the following definition is slightly different from the one in [16] and is tailored to our purpose.

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

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

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

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

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

The following two properties of smooth maximum divergence will be used:

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

$$D^{\epsilon}_{\max}(\mathrm{P} \circ W\|\mathrm{Q} \circ W) \leq D^{\epsilon}_{\max}(\mathrm{P}\|\mathrm{Q}). \tag{10}$$

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

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

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

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 '$\leq$' follows upon choosing $\tilde{\mathrm{P}}_n(\mathbf{x}) = \mathrm{P}^n(\mathbf{x})\mathbb{1}(\mathbf{x} \in \mathcal{T}_n)$, where $\mathcal{T}_n$ is the typical set for P^n. For the other direction, given a $\tilde{\mathrm{P}}_n \leq \mathrm{P}^n$ with $\tilde{\mathrm{P}}_n(\mathcal{X}^n) \geq 1-\epsilon$, we have

$$\tilde{\mathrm{P}}_n(\mathcal{T}_n) \geq (1-\epsilon)/2, \tag{11}$$

for all n sufficiently large. Thus,

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

where the third inequality is by log-sum inequality [13, Lemma 3.1], and the last inequality uses (11). The proof is completing upon noting that

$$\log \mathrm{Q}^n(\mathcal{T}_n) = -nD(\mathrm{P}\|\mathrm{Q}) + o(n).$$

III. The conditional independence testing bound

Converse results of this paper are based on an upper bound on the maximum length $S_\epsilon\left(X_{\mathcal{M}}|Z\right)$ of an ϵ-SK. We present this basic result here[14].

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

[14]The results of this section were presented in [56].

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:

$$\mathrm{Q}^{\pi}_{X_{\mathcal{M}}|Z}(x_1, \ldots, x_m|z) = \prod_{i=1}^{|\pi|} \mathrm{Q}^{\pi}_{X_{\pi_i}|Z}(x_{\pi_i}|z). \tag{12}$$

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*

$$S_{\epsilon}\left(X_{\mathcal{M}}|Z\right) \leq \frac{1}{|\pi| - 1}\Bigg[- \log \beta_{\epsilon+\eta}\big(\mathrm{P}_{X_{\mathcal{M}}Z}, \mathrm{Q}^{\pi}_{X_{\mathcal{M}}Z}\big) + |\pi| \log(1/\eta)\Bigg] \tag{13}$$

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

Remarks. (i) Renner and Wolf [45] 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 bound in Theorem 3 is unavailable, since the former involves auxiliary RVs and is difficult to evaluate.

(ii) 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 [39]. 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 [49].

(iii) The proof of Theorem 3 below remains valid even when the security 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 (13) holds even under the relaxed security criterion above.

To prove Theorem 3, we first relate the SK length to the exponent of the probability of error of type II in a binary hypothesis testing problem where an observer of $(K_{\mathcal{M}}, \mathbf{F}, Z)$ seeks to find out if the

underlying distribution was $\mathrm{P}_{X_{\mathcal{M}}Z}$ or $\mathrm{Q}^{\pi}_{X_{\mathcal{M}}Z}$. This result is stated next.

Lemma 4. *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 on $(K_{\mathcal{M}}, \mathbf{F})$ given $(X_{\mathcal{M}}, Z)$. Then, for every $0 < \eta < 1 - \epsilon$ and every $\mathrm{Q}^{\pi}_{X_{\mathcal{M}}Z} \in \mathcal{Q}(\pi)$, we have*

$$\log|\mathcal{K}| \leq \frac{1}{|\pi|-1}\bigg[-\log\beta_{\epsilon+\eta}\big(\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}, \mathrm{Q}^{\pi}_{K_{\mathcal{M}}\mathbf{F}Z}\big) + |\pi|\log(1/\eta)\bigg], \tag{14}$$

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

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

Also, we need the following basic property of interactive communication from [52], which will be used throughout this paper (see, also, [13, Lemma 17.18]).

Lemma 5 (Interactive communication property). *Given $\mathrm{Q}^{\pi}_{X_{\mathcal{M}}Z} \in \mathcal{Q}(\pi)$ and an interactive communication $\mathbf{F}$, the following holds:*

$$\mathrm{Q}^{\pi}_{X_{\mathcal{M}}|\mathbf{F}Z}(x_{\mathcal{M}}|f,z) = \prod_{i=1}^{|\pi|} \mathrm{Q}^{\pi}_{X_{\pi_i}|\mathbf{F}Z}(x_{\pi_i}|f,z),$$

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 4. We establish (14) by constructing a test for the hypothesis testing problem with null hypothesis $\mathrm{P} = \mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}$ and alternative hypothesis $\mathrm{Q} = \mathrm{Q}^{\pi}_{K_{\mathcal{M}}\mathbf{F}Z}$. Specifically, we use a deterministic test[15] with the following acceptance region (for the null hypothesis)[16]:

$$\mathcal{A} := \left\{(k_{\mathcal{M}}, f, z) : \log\frac{\mathrm{P}^{(\mathcal{M})}_{\mathrm{unif}}(k_{\mathcal{M}})}{\mathrm{Q}^{\pi}_{K_{\mathcal{M}}|\mathbf{F}Z}(k_{\mathcal{M}}|f,z)} \geq \lambda_{\pi}\right\},$$

[15]In fact, we use a simple threshold test on the log-likelihood ratio but with $\mathrm{P}^{(\mathcal{M})}_{\mathtt{unif}} \times \mathrm{P}_{\mathbf{F}Z}$ in place of $\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}$, since the two distributions are close to each other by the security condition (3).

[16]The values $(k_{\mathcal{M}}, f, z)$ with $\mathrm{Q}^{\pi}_{K_{\mathcal{M}}|\mathbf{F}Z}(k_{\mathcal{M}}|f,z) = 0$ are included in $\mathcal{A}$.

where

$$\lambda_\pi = (|\pi| - 1) \log |\mathcal{K}| - |\pi| \log(1/\eta).$$

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

$$\begin{aligned} \mathrm{Q}^{\pi}_{K_{\mathcal{M}}\mathbf{F}Z}(\mathcal{A}) &= \sum_{f,z} \mathrm{Q}^{\pi}_{\mathbf{F}Z}(f,z) \sum_{\substack{k_{\mathcal{M}}: \\ (k_{\mathcal{M}},f,z)\in\mathcal{A}}} \mathrm{Q}^{\pi}_{K_{\mathcal{M}}|\mathbf{F}Z}(k_{\mathcal{M}}|f,z) \\ &\leq 2^{-\lambda_\pi} \sum_{f,z} \mathrm{Q}^{\pi}_{\mathbf{F}Z}(f,z) \sum_{k_{\mathcal{M}}} \mathrm{P}^{(\mathcal{M})}_{\mathrm{unif}}(k_{\mathcal{M}}) \\ &= |\mathcal{K}|^{1-|\pi|} \eta^{-|\pi|}. \end{aligned} \tag{15}$$

On the other hand, the probability of error of type I is bounded above as

$$\begin{aligned} \mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}(\mathcal{A}^c) &\leq d\left(\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}, \mathrm{P}^{(\mathcal{M})}_{\mathrm{unif}} \times \mathrm{P}_{\mathbf{F}Z}\right) + \mathrm{P}^{(\mathcal{M})}_{\mathrm{unif}} \times \mathrm{P}_{Z\mathbf{F}}(\mathcal{A}^c) \\ &\leq \epsilon + \mathrm{P}^{(\mathcal{M})}_{\mathrm{unif}} \times \mathrm{P}_{\mathbf{F}Z}(\mathcal{A}^c), \end{aligned} \tag{16}$$

where the first inequality follows from the definition of variational distance, and the second is a consequence of the security condition (3) satisfied by the ϵ-SK $K_{\mathcal{M}}$. The second term above can be expressed as follows:

$$\begin{aligned} \mathrm{P}^{(\mathcal{M})}_{\mathrm{unif}} \times \mathrm{P}_{\mathbf{F}Z}(\mathcal{A}^c) &= \sum_{f,z} \mathrm{P}_{\mathbf{F}Z}(f,z) \frac{1}{|\mathcal{K}|} \sum_{k} \mathbb{1}\left((\mathbf{k},f,z) \in \mathcal{A}^c\right) \\ &= \sum_{f,z} \mathrm{P}_{\mathbf{F}Z}(f,z) \frac{1}{|\mathcal{K}|} \sum_{k} \mathbb{1}\left(\mathrm{Q}^{\pi}_{K_{\mathcal{M}}|\mathbf{F}Z}(\mathbf{k}|f,z)|\mathcal{K}|^{|\pi|}\eta^{|\pi|} > 1\right), \end{aligned} \tag{17}$$

where $\mathbf{k} = (k, \ldots, k)$. The inner sum can be further upper bounded as

$$\begin{aligned} \sum_{k} \mathbb{1}\left(\mathrm{Q}^{\pi}_{K_{\mathcal{M}}|\mathbf{F}Z}(\mathbf{k}|f,z)|\mathcal{K}|^{|\pi|}\eta^{|\pi|} > 1\right) &\leq \sum_{k} \left(\mathrm{Q}^{\pi}_{K_{\mathcal{M}}|\mathbf{F}Z}(\mathbf{k}|f,z)|\mathcal{K}|^{|\pi|}\eta^{|\pi|}\right)^{\frac{1}{|\pi|}} \\ &= |\mathcal{K}|\eta \sum_{k} \mathrm{Q}^{\pi}_{K_{\mathcal{M}}|\mathbf{F}Z}(\mathbf{k}|f,z)^{\frac{1}{|\pi|}} \\ &= |\mathcal{K}|\eta \sum_{k} \prod_{i=1}^{|\pi|} \mathrm{Q}^{\pi}_{K_{\pi_i}|\mathbf{F}Z}(\mathbf{k}|f,z)^{\frac{1}{|\pi|}}, \end{aligned} \tag{18}$$

where the previous equality uses Lemma 5 and the fact that given $\mathbf{F}$, K_{π_i} is a function of (X_{π_i}, U_{π_i}).

Next, an application of Hölder's inequality to the sum on the right-side of (18) yields

$$\begin{aligned}\sum_{k}\prod_{i=1}^{|\pi|}\mathrm{Q}^{\pi}_{K_{\pi_i}|\mathbf{F}Z}(\mathbf{k}|f,z)^{\frac{1}{|\pi|}} &\leq \prod_{i=1}^{|\pi|}\left(\sum_{k}\mathrm{Q}^{\pi}_{K_{\pi_i}|\mathbf{F}Z}(\mathbf{k}|f,z)\right)^{\frac{1}{|\pi|}}\\ &\leq \prod_{i=1}^{|\pi|}\left(\sum_{k_{\pi_i}}\mathrm{Q}^{\pi}_{K_{\pi_i}|\mathbf{F}Z}(k_{\pi_i}|f,z)\right)^{\frac{1}{|\pi|}}\\ &= 1. \end{aligned} \tag{19}$$

Upon combining (17)-(19) we obtain

$$\mathrm{P}^{(\mathcal{M})}_{\mathrm{unif}} \times \mathrm{P}_{\mathbf{F}Z}(\mathcal{A}^c) \leq \eta,$$

which along with (16) gives

$$\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}\left(\mathcal{A}^c\right) \leq \epsilon + \eta. \tag{20}$$

It follows from (20) and (15) that

$$\beta_{\epsilon+\eta}\big(\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}, \mathrm{Q}^{\pi}_{K_{\mathcal{M}}\mathbf{F}Z}\big) \leq |\mathcal{K}|^{1-|\pi|}\eta^{-|\pi|},$$

which completes the proof. □

Proof of Theorem 3. Using the data processing inequality (5) with $\mathrm{P} = \mathrm{P}_{X_{\mathcal{M}}Z}$, $\mathrm{Q} = \mathrm{Q}^{\pi}_{X_{\mathcal{M}}Z}$, and $W = W_{K_{\mathcal{M}}\mathbf{F}|X_{\mathcal{M}}Z}$, we get

$$\beta_{\epsilon+\eta}\big(\mathrm{P}_{X_{\mathcal{M}}Z}, \mathrm{Q}^{\pi}_{X_{\mathcal{M}}Z}\big) \leq \beta_{\epsilon+\eta}\big(\mathrm{P}_{K_{\mathcal{M}}\mathbf{F}Z}, \mathrm{Q}^{\pi}_{K_{\mathcal{M}}\mathbf{F}Z}\big),$$

which along with Lemma 4 gives Theorem 3. □

IV. Implications for SK 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 [34], [1], [14].

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 [34] and Ahlswede and Csiszár [1]. Later, the SK capacity for a multiparty model, with Z =constant was characterized by Csiszár and Narayan [14]. The general problem of characterizing the SK capacity for arbitrary Z remains open. Several upper bounds for SK capacity are known [34], [1], [35], [44], [14], [15], [23], which are tight for special cases.

In this section, we derive a single-shot version of the Gohari-Anantharam bound [23] 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 security requirement (2).

A. Converse results for two parties

It was shown in [23] that for two parties,

$$C(X_1, X_2|Z) \leq \min_{U} I\left(X_1 \wedge X_2|U\right) + I(X_1, X_2 \wedge U|Z). \tag{21}$$

The proof in [23] 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 (21) and is proved by relying only on the structure of the SKs, without recourse to the *potential function approach*[17] of [23].

Theorem 6. *For $0 < \epsilon, \delta$ with $\epsilon + 2\delta < 1$,*

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

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

[17]In fact, a simple proof of (21) 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 $C(X_1, X_2|U) + I(X_1, X_2 \wedge U|Z)$.

As corollaries, we obtain a single-shot version and a strong version of the upper bound in (21), which does not require perfect asymptotic recovery or perfect asymptotic security.

Corollary 7 (Single-shot bound for SK length). *For $0 < \epsilon, \delta$ with $\epsilon + 2\delta < 1$,*

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

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

Corollary 8 (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 6 is contained in the following lemma.

Lemma 9. *Let (K_1, K_2) be an (ϵ, δ)-SK taking values in $\mathcal{K}$, recoverable from a communication $\mathbf{F}$. Then,*

$$H^{\delta+\xi/2}_{\min}\left(\mathrm{P}_{K_1\mathbf{F}ZU}|\mathbf{F}ZU\right) \geq \log|\mathcal{K}| - D^{\xi}_{\max}\left(\mathrm{P}_{K_1\mathbf{F}ZU}\|\mathrm{P}_{K_1\mathbf{F}Z}\mathrm{P}_{U|Z}\right)$$

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

Proof of Theorem 6. Let (K_1, K_2) be an (ϵ, δ)-SK taking values in $\mathcal{K}$. Then, by Lemma 9 and the data processing property of smooth maximum divergence (10), we get

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

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

$$\log|\mathcal{K}| - D^{\xi}_{\max}\left(\mathrm{P}_{X_1X_2ZU}\|\mathrm{P}_{X_1X_2Z}\mathrm{P}_{U|Z}\right) - 2\log(1/2(\eta - \xi)) - 1 \leq S_{\epsilon,2\delta+\eta}(X_1, X_2|Z, U).$$

□

Corollary 7 follows by Theorem 3.

Proof of Corollary 8. The result follows by Corollary 7 upon using Stein's lemma (see Section II-B), along with the convergence property of smooth maximum divergence (see Section II-C). □

Proof of Lemma 9. 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

$$\sum_{k_1,f,z,u} \mathrm{P}\left(k_1, f, z, u\right) T(k_1, f, z, u) \geq 1 - \xi, \tag{22}$$

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

$$d\left(\mathrm{P}_{K_1\mathbf{F}ZU}, \mathrm{Q}_{K_1\mathbf{F}ZU}\right) \leq \delta + \xi/2, \tag{23}$$

$$H_{\min}\left(\mathrm{Q}_{K_1\mathbf{F}ZU}|\tilde{\mathrm{Q}}_{\mathbf{F}ZU}\right) = \log|\mathcal{K}| - D_{\max}\left(\mathrm{P}_{K_1\mathbf{F}ZU}T\|\mathrm{P}_{K_1\mathbf{F}Z}\mathrm{P}_{U|Z}\right). \tag{24}$$

To that end, note

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

and let

$$\mathrm{Q}\left(k_1, f, z, u\right) := \mathrm{P}_{\mathtt{unif}}\left(k_1\right)\left[\frac{\mathrm{P}\left(k_1, u|f, z\right)}{\mathrm{P}\left(k_1|f, z\right)\mathrm{P}\left(u|f, z\right)}\right]\mathrm{P}\left(f, z, u\right)T(k_1, f, z, u),$$

$$\tilde{\mathrm{Q}}_{\mathbf{F}ZU} := \mathrm{P}_{\mathbf{F}Z}\mathrm{P}_{U|Z}.$$

Since $T(k_1, f, z, u) \leq 1$, it follows that

$$\sum_{k_1,f,z,u} \mathrm{Q}\left(k_1, f, z, u\right) \leq 1,$$

and so Q is a valid subnormalized nonnegative function. Also,

$$\begin{aligned}
&2d\left(\mathrm{P},\mathrm{Q}\right)\\
&=\sum_{f,z,u}\mathrm{P}\left(f,z,u\right)\sum_{k_1}\left|\mathrm{P}\left(k_1|f,z,u\right)-\mathrm{Q}\left(k_1|f,z,u\right)\right|\\
&\leq\sum_{k_1,f,z,u}\mathrm{P}\left(k_1,f,z,u\right)\left|1-T(k_1,f,z,u)\right|\\
&\qquad+\sum_{f,z,u}\mathrm{P}\left(f,z,u\right)\sum_{k_1}\left|\mathrm{P}\left(k_1|f,z,u\right)T(k_1,f,z,u)-\mathrm{Q}\left(k_1|f,z,u\right)\right|\\
&\leq\xi+\sum_{f,z,u}\mathrm{P}\left(f,z,u\right)\sum_{k_1}\left|\mathrm{P}\left(k_1|f,z,u\right)T(k_1,f,z,u)-\mathrm{Q}\left(k_1|f,z,u\right)\right|, \qquad (25)
\end{aligned}$$

where the previous inequality uses (22). For the second term above, from the definition of Q and security condition (2), we have

$$\begin{aligned}
&\sum_{f,z,u}\mathrm{P}\left(f,z,u\right)\sum_{k_1}\left|\mathrm{P}\left(k_1|f,z,u\right)T(k_1,f,z,u)-\mathrm{Q}\left(k_1|f,z,u\right)\right|\\
&\leq\sum_{k_1,f,z,u}\mathrm{P}\left(f,z\right)\mathrm{P}\left(u|k_1,f,z\right)\left|\mathrm{P}\left(k_1|f,z\right)-\mathrm{P}_{\mathtt{unif}}\left(k_1\right)\right|\\
&\leq\sum_{k_1,f,z}\mathrm{P}\left(f,z\right)\left|\mathrm{P}\left(k_1|f,z\right)-\mathrm{P}_{\mathtt{unif}}\left(k_1\right)\right|\\
&\leq 2\delta,
\end{aligned}$$

which together with (25) yields (23).

Next, observe that

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

and so,

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

which is the same as (24). □

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 [14]. Furthermore, they introduced the remarkable expression on the right-side of (26) below as an upper bound for $C(X_{\mathcal{M}})$, 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 10. *[14], [10] The SK capacity C for the case when eavesdropper's side information $Z =$ constant is given by*

$$C(X_{\mathcal{M}}) = \min_{\pi} \frac{1}{|\pi| - 1} D\left(\mathrm{P}_{X_{\mathcal{M}}} \middle\| \prod_{i=1}^{|\pi|} \mathrm{P}_{X_{\pi_i}}\right), \tag{26}$$

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

This generalized the classic result of Maurer [34] and Ahlswede and Csiszár [1], which established that for two parties, $C(X_1, X_2) = D(\mathrm{P}_{X_1 X_2} \| \mathrm{P}_{X_1} \times \mathrm{P}_{X_2}) = I(X_1 \wedge X_2)$.

The converse part of Theorem 10 relied critically on the fact that $\epsilon_n + \delta_n \to 0$ as $n \to 0$. Below we strengthen the converse and show that the upper bound for SK rates implied by Theorem 10 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 $X_{\mathcal{M}}^n$, with $\mathrm{Q}_{X_{\mathcal{M}}^n}^{\pi} = \prod_{i=1}^{|\pi|} \mathrm{P}_{X_{\pi_i}}^n$, yields

$$S_{\epsilon,\delta}(X_1^n, ..., X_m^n) \leq \frac{1}{|\pi| - 1} \left[-\log \beta_{\epsilon+\delta+\eta} \left(\mathrm{P}_{X_{\mathcal{M}}}^n, \prod_{i=1}^{|\pi|} \mathrm{P}_{X_{\pi_i}}^n \right) + |\pi| \log(1/\eta) \right],$$

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

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

Also, note that if $\epsilon + \delta > 1$, the SK rate can be infinity. Indeed, even for $\epsilon + \delta = 1$ the SK rate is infinity, as is seen by time sharing between an infinite rate $(1, 0)$-SK of rate infinity and $(0, 1)$-SK of rate infinity.

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

Corollary 11 (Strong converse for SK capacity). *Given $0 < \epsilon, \delta < 1$, the (ϵ, δ)-SK capacity $C_{\epsilon,\delta}(X_{\mathcal{M}})$*

is given by

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

and

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

V. Implications for secure computing

In this section, we consider secure computing by two (mutually untrusting) parties. First introduced by Yao in [65], 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 computing. We will 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.

To state our results, we need the notions of *maximum common function* and *minimum sufficient statistic*; their role in bounding the performance of secure computing protocols was first highlighted in [64]. Specifically, for RVs X_1, X_2, denote by $\mathrm{mcf}(X_1,X_2)$ the maximum common function of X_1 and X_2 [21] (see, also, [54]). Also, denote by $\mathrm{mss}(X_2|X_1)$ the minimum sufficient statistic for X_2 given X_1, i.e., the minimal function $g(X_1)$ such that the Markov chain X_1—$g(X_1)$—X_2 holds. Specifically, $\mathrm{mss}(X_2|X_1)$ is given by the function resulting from the following equivalence relation on $\mathcal{X}_1$ (cf. [20], [29], [51]):

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

A. Oblivious transfer

We present bounds on the efficiency of implementing information theoretically secure one-of-two OT using correlated randomness. Formally, suppose the first party observes K_0 and K_1, distributed uniformly over $\{0,1\}^l$, and the second party observes a random bit B. The RVs K_0, K_1, and B are mutually independent. Furthermore, party i observes the RV X_i, $i=1,2$, where RVs (X_1,X_2) are independent jointly of (K_0,K_1,B). The second party seeks to compute K_B without giving away B to the first party. At the same time, the first party does not want to give away $K_{\overline{B}}$ to the second party.

Definition 6. (**Oblivious transfer**) An $(\epsilon, \delta_1, \delta_2)$-OT of length l consists of an interactive communication protocol $\mathbf{F}$ and $\hat{K} = \hat{K}(X_2, B, \mathbf{F})$ such that the following conditions hold[18]:

$$\mathrm{P}\left(K_B \neq \hat{K}\right) \leq \epsilon, \tag{27}$$

$$d\left(\mathrm{P}_{K_{\overline{B}} X_2 B \mathbf{F}}, \mathrm{P}_{K_{\overline{B}}} \times \mathrm{P}_{X_2 B \mathbf{F}}\right) \leq \delta_1, \tag{28}$$

$$d\left(\mathrm{P}_{B K_0 K_1 X_1 \mathbf{F}}, \mathrm{P}_B \times \mathrm{P}_{K_0 K_1 X_1 \mathbf{F}}\right) \leq \delta_2, \tag{29}$$

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

When the underlying observations X_1, X_2 consist of n-length IID sequences X_1^n, X_2^n with common distribution $\mathrm{P}_{X_1 X_2}$, it is known that $L_{\epsilon,\delta_1,\delta_2}(X_1^n, X_2^n)$ may grow linearly with n (cf. [36], [2]); the largest rate of growth is called the OT capacity.

Definition 7 (**OT capacity**)**.** For $0 < \epsilon < 1$, the ϵ-OT capacity of (X_1, X_2) is defined[19] as

$$C_\epsilon(X_1, X_2) = \liminf_n \sup_{\delta_{1n}, \delta_{2n}} \frac{1}{n} L_{\epsilon,\delta_{1n},\delta_{2n}}(X_1^n, X_2^n),$$

where the sup is over all $\delta_{1n}, \delta_{2n} \to 0$ as $n \to \infty$. The OT capacity is defined as

$$C(X_1, X_2) = \inf_{0<\epsilon<1} C_\epsilon(X_1, X_2).$$

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 $0 < \epsilon < 1$.

Theorem 12 (**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:*

$$L_{\epsilon,\delta_1,\delta_2}(X_1, X_2) \leq -\log \beta_\eta \left(\mathrm{P}_{X_1 X_2 V_0}, \mathrm{P}_{X_1|V_0} \mathrm{P}_{X_2|V_0} \mathrm{P}_{V_0}\right) + 2\log(1/\xi), \tag{30}$$

$$L_{\epsilon,\delta_1,\delta_2}(X_1, X_2) \leq -\log \beta_\eta \left(\mathrm{P}_{V_1 V_1 X_2}, \mathrm{P}_{V_1|X_2} \mathrm{P}_{V_1|X_2} \mathrm{P}_{X_2}\right) + 2\log(1/\xi), \tag{31}$$

[18]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], [63]) and is equivalent to OT.

[19]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.

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

Corollary 13 (Strong bound for OT capacity). *For* $0 < \epsilon < 1$*, the* ϵ*-OT capacity of* (X_1, X_2) *satisfies*

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

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

The proof of Theorem 12 entails reducing two SK agreement problems to OT[20]. The bound (30) is obtained by recovering K_B as a SK, while (31) is obtained by recovering $K_{\overline{B}}$ as a SK; we note these two reductions as separate lemmas below.

Lemma 14 (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 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 15 (Reduction 2 of SK agreement to OT). *Consider two party SK agreement where the first party observes* X_1*, the second party observes* $(V_1, X_2) = (\mathrm{mss}(X_2|X_1), X_2)$ *and the eavesdropper observes* X_2*. Given 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,*

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

Remarks. (i) 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 (30). 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 (31).

(ii) In general, our bounds are stronger than those presented in [61]. For instance, the latter is loose when the observations consist of mixtures of IID RVs. Further, while both (31) and [61, Theorem 5] (specialized to OT) suffice to obtain the second bound in Corollary 13, in contrast to (30), [61, Theorem 2] does not yield the first bound in Corollary 13.

(iii) For simplicity of presentation, we did not allow local randomization in the formulation above.

[20]A reduction of SK to OT in a computational security setup appeared in [22].

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.

(iv) An $(\epsilon, \delta_1, \delta_2)$-OT capacity can be defined, without requiring δ_{1n}, δ_{2n} to go to 0 as in the definition of $C_\epsilon(X_1, X_2)$. The problem of characterizing $(\epsilon, \delta_1, \delta_2)$-OT capacity for all $0 < \epsilon, \delta_1, \delta_2 < 1$ remains open.

We prove Lemmas 14 and 15 next. The proof of Theorem 12 follows by Theorem 3, along with the Markov relation X_1—V_1—X_2 and the data processing inequality (5); the corollary follows by Stein's Lemma (see Section II-B).

Proof. of Lemma 14. Let $\hat{K}$ be the estimate of K_B formed by the second party. The following protocol generates an $(\epsilon + \delta_1 + 2\delta_2)$-SK of length l

(i) The first party generates two random strings K_0 and K_1 of length l, and the second party generates a random bit B. Two parties run the OT protocol.

(ii) The second party sends B over the public channel.

(iii) Using B, the first party computes K_B. The RVs $K_B, \hat{K}$ constitute an $(\epsilon + \delta_1 + 2\delta_2)$-SK.

Since both parties agree on K_B with probability greater than $1 - \epsilon$, by Proposition 1 and remark (iii) following Theorem 3, it suffices to show that for some distribution $Q_{V_0 \mathbf{F} B}$ (see Remark (iii) following Theorem 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 (29) is the same as

$$d\left(\mathrm{P}_{K_0 K_1 X_1 \mathbf{F}|B=0}, \mathrm{P}_{K_0 K_1 X_1 \mathbf{F}|B=1}\right) \leq 2\delta_2. \tag{32}$$

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

$$\begin{aligned}
&d\left(\mathrm{P}_{K_BV_0\mathbf{F}B},\mathrm{P}_{\mathtt{unif}}\times\mathrm{Q}_{V_0\mathbf{F}B}\right)\\
&=\frac{1}{2}\sum_b d\left(\mathrm{P}_{K_bV_0\mathbf{F}|B=b},\mathrm{P}_{\mathtt{unif}}\times\mathrm{Q}_{V_0\mathbf{F}|B=b}\right)\\
&=\frac{1}{2}\sum_b d\left(\mathrm{P}_{K_bV_0\mathbf{F}|B=b},\mathrm{P}_{\mathtt{unif}}\times\mathrm{P}_{V_0\mathbf{F}|B=\bar{b}}\right)\\
&\leq\frac{1}{2}\sum_b\left[d\left(\mathrm{P}_{K_bV_0\mathbf{F}|B=\bar{b}},\mathrm{P}_{\mathtt{unif}}\times\mathrm{P}_{V_0\mathbf{F}|B=\bar{b}}\right)+d\left(\mathrm{P}_{K_bV_0\mathbf{F}|B=b},\mathrm{P}_{K_bV_0\mathbf{F}|B=\bar{b}}\right)\right]\\
&=d\left(\mathrm{P}_{K_{\overline{B}}V_0\mathbf{F}B},\mathrm{P}_{\mathtt{unif}}\times\mathrm{P}_{V_0\mathbf{F}B}\right)+\frac{1}{2}\sum_b d\left(\mathrm{P}_{K_bV_0\mathbf{F}|B=b},\mathrm{P}_{K_bV_0\mathbf{F}|B=\bar{b}}\right)\\
&\leq\delta_1+2\delta_2,
\end{aligned}$$

where the last inequality uses (28) and (32), together with the fact that V_0 is a function of X_2 as well as X_1. □

Proof. of Lemma 15. The following protocol generates an $(\epsilon+\delta_1+2\delta_2)$-SK of length l.

(i) The first party generates two random strings K_0 and K_1 of length l, and the second party generates a random bit B. Two parties run the OT protocol.

(ii) Upon observing $\mathbf{F}$, the second party samples $\tilde{X}_2$ according to the distribution $\mathrm{P}_{X_2|V_1B\mathbf{F}}\left(\cdot\mid V_1,\overline{B},\mathbf{F}\right)$.

(iii) The second party sends B over the public channel.

(iv) The first party computes $K_{\overline{B}}$ and the second party computes $\tilde{K}=\hat{K}(\tilde{X}_2,\overline{B},\mathbf{F})$. The RVs $K_{\overline{B}},\tilde{K}$ constitute an $(\epsilon+\delta_1+2\delta_2)$-SK.

Heuristically, this protocol entails the second party emulating $\tilde{X}_2$, pretending that the protocol was executed for $\overline{B}$ instead of B. Since the communication of the first party 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 (28), it suffices to show that

$$\mathrm{P}\left(K_{\overline{B}}\neq\tilde{K}\right)\leq\epsilon+2\delta_2. \tag{33}$$

To that end, note

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

where the inequality uses (32) 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 5; (33) follows by (27). □

B. 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, the second party can detect if the first party was lying [7]. Formally, a BC protocol consists of two phases: the *commit phase* and the *reveal phase*. In the commit phase, the first party 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. In the reveal phase, the first party "reveals" its data, i.e., it sends X_1' and K', claiming these were its initial choices of X_1 and K, respectively. Subsequently, the second party 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 8 (Bit commitment). An $(\epsilon, \delta_1, \delta_2)$-BC of length l consists of a secret $K \sim \mathtt{unif}\{0,1\}^l$, an interactive communication $\mathbf{F}$ (sent during the commit phase), and a $\{0,1\}$-valued randomized test function $T = T(K', X_1', X_2, \mathbf{F})$ such that the following hold:

$$\mathrm{P}\left(T(K, X_1, X_2, \mathbf{F}) \neq 0\right) \leq \epsilon, \tag{34}$$

$$d\left(\mathrm{P}_{K X_2 \mathbf{F}}, \mathrm{P}_K \times \mathrm{P}_{X_2 \mathbf{F}}\right) \leq \delta_1, \tag{35}$$

$$\mathrm{P}\left(T(K', X_1', X_2, \mathbf{F}) = 0, K' \neq K\right) \leq \delta_2, \tag{36}$$

where RVs X_1', K' are arbitrary. The first condition above is the *soundness condition*, which captures the reliability of BC. The next condition is the *hiding condition*, which ensures that the second party cannot ascertain the secret in the commit phase. Finally, the *binding condition* in (36) restricts the probability with which the first party can cheat in the reveal phase. Denote by $L_{\epsilon,\delta_1,\delta_2}(X_1, X_2)$ the largest length l of an $(\epsilon, \delta_1, \delta_2)$-BC.

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 9 (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

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

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 Winters, Nascimento, and Imai [62] (see, also, [50, Chapter 8]) gives a simple formula for $C(X_1, X_2)$.

Theorem 16. *[62] For RVs X_1, X_2, let $V_1 = \mathrm{mss}(X_2|X_1)$. The BC capacity is given by*

$$C(X_1, X_2) = H(V_1 \mid X_2).$$

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 17 (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 $V_1 = \mathrm{mss}(X_1|X_2)$, 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 18 (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*

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

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

Theorem 17 is obtained by a reduction of SK agreement to BC, which is along the lines of [62], [26], [42]; the following lemma captures the resulting bound.

Lemma 19 (Reduction of SK to BC). *For* $0 < \epsilon, \delta_1, \delta_2$, $\epsilon + \delta_1 + \delta_2 < 1$, *it holds that*

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

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

Remarks. (i) While local randomization was not allowed in the foregoing discussion, as before (see Remark (iii) following Lemma 15) our results do not change with the availability of local randomness.

(ii) 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 [42, 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 17, in general, and is not sufficient for deriving Corollary 18.

Theorem 17 follows by using Lemma 19 with Theorem 3, along with the Markov relation X_1—V_1—X_2 and the data processing inequality (5); the Corollary 18 follows by Stein's Lemma (see Section II-B). We prove Lemma 19 below.

Proof of Lemma 19. The reduction argument presented here is along the lines of [26, Proposition 9] (see, also, [42, 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., the first party generates $K \sim \mathtt{unif}\{0,1\}^l$ and the parties send the interactive communication $\mathbf{F}$. We show that the committed secret K constitues a $(\epsilon + \delta_2, \delta_1)$-SK. Indeed, by the hiding condition (35), the SK K satisfies the security condition (2) with $\delta = \delta_1$. We complete the proof by showing that there exists $\hat{K} = \hat{K}(V_1, X_2, \mathbf{F})$ such that

$$\mathrm{P}\left(\hat{K} \neq K\right) \leq \epsilon + \delta_2. \tag{37}$$

To that end, let $(\hat{K}, \hat{X}_1)$ be a function of (v, f) given by

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

Note that while the estimated secret $\hat{K}$ does not depend on X_2, 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

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

where the first inequality uses the definition of $(\hat{K}, \hat{X}_1)$ 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 5. The inequality above, along with the binding condition (36), 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 (37). □

We conclude this section by observing a simple application of Theorem 17 in bouding the efficiency of reduction of BC to OT. For a detailed discussion, see [42].

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 second party, let $X_1 = (K_0, K_1)$ and $X_2 = (B, K_B)$. Note that (8) 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 17 and (9), 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 [42, 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 [42, Corollary 2] shows a multiplicative loss of linear order.

VI. Implications for secure computing 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 computing 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 [53] where a matching necessary and sufficient condition was given for the feasibility of secure computing 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

$$\mathrm{P}\left(G = G_{(1)} = ... = G_{(m)}\right) \geq 1-\epsilon, \tag{38}$$

$$d\left(\mathrm{P}_{G\mathbf{F}}, \mathrm{P}_G \times \mathrm{P}_{\mathbf{F}}\right) \leq \delta, \tag{39}$$

where $G = g\left(X_{\mathcal{M}}\right)$. The first condition captures the reliability of computation and the second condition ensures the security of the protocol. Heuristically, for security 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 [53], 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

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

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

Theorem 20. *[53] 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 10.*

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

$$H_{\min}^{\xi}(\mathrm{P}_G) \leq \frac{1}{|\pi|-1}\Big[-\log\beta_{\mu}\big(\mathrm{P}_{X_{\mathcal{M}}}, \mathrm{Q}_{X_{\mathcal{M}}}^{\pi}\big) + \quad |\pi|\log(1/\eta)\Big] + 2\log(1/2\zeta) + 1,$$
$$\forall\, \mathrm{Q}_{X_{\mathcal{M}}}^{\pi} \in \mathcal{Q}(\pi), \qquad (40)$$

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 (39) 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

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

Thus, in the view of Proposition 1, the RV K constitutes[21] an $(\epsilon+\delta+2\xi+\zeta)$-SK. An application of

[21]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).

Theorem 3 gives (40). □

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 $g(U_1, ..., U_m)$?

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 21 a necessary condition for g to be (ϵ, δ)-securely computable is

$$H_{\min}^{\xi}(\mathrm{P}_G) \leq \kappa + \frac{1}{|\pi| - 1} \left(|\pi| \log(1/\eta) + \log(1/(1-\mu)) \right) + 2\log(1/2\zeta) + 1, \tag{41}$$

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

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

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

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 computing 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 $B_1^n, \ldots, B_{m-1}^n$ and the mth party can send back $B_1^n \oplus \ldots \oplus B_m^n \oplus K_n$, 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.[22] To this end, they communicate interactively using a communication $\mathbf{F}$, and based on this communication the second party forms an estimate $\hat{M}$ of the first party's message M. This protocol accomplishes (ϵ, δ)-secure transmission if

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

The classic result of Shannon [47] implies that $(0, 0)$-secure transmission is feasible only if κ is at

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

least $\log \|M\|$, where $\|M\|$ denotes the size of the message space.[23] 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 (41), (ϵ, δ)-secure transmission of M is feasible only if

$$H_{\min}^{\xi}(\mathrm{P}_M) \leq \kappa + 2\log(1/\eta) + \log(1/(1-\mu)) + 2\log(1/2\zeta) + 1, \tag{42}$$

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

Condition (42) brings out a trade-off between κ and $\epsilon + \delta$ (cf. [30, 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:

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

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 (42) 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 computing 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 [34], [1], [14], [45] 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 asymptotics of the maximum

[23] This is a slight generalization of Shannon's original result; see [30, Theorem 2.7] for a proof.

length of a SK was established in [25]. Coming up with finite-length schemes that match the converse bounds for the various secure computing problems studied above is work in progress.

Finally, note that our converse results in Sections V and VI entail reducing SK agreement to the secure computing task at hand, followed by an application of Theorem 3. It is foreseeable, and indeed tempting, that this approach can lead to converse bounds for other problems in cryptography.

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

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

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

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

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

[16] 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.

[17] 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.

[18] 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.

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

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] M. Hayashi, H. Tyagi, and S. Watanabe, "Secret key agreement: General capacity and second-order asymptotics," *To appear, Proc. IEEE International Symposium on Information Theory*, 2014.

[26] 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.

[27] 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.

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

[29] 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.

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

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

[32] ——, in *Proc. Symposium on Theory of Computing (STOC)*, 2000, pp. 316–324.

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

[34] 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.

[35] 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.

[36] 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.

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

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

[39] 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.

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

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

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

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

[44] 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.

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

[46] 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.

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

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

[49] 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.

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

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

[52] 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.

[53] 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.

[54] ——, "When is a function securely computable?" in *Proc. IEEE International Symposium on Information Theory*, 2011, pp. 2876–2880.

[55] 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.

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

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

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

[59] 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.

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

[61] S. Winkler and J. Wullschleger, "On the efficiency of classical and quantum secure function evaluation," 2012, arXiv:1205.5136.

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

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

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

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