The Conditional Common Information in Classical and Quantum Secret Key Distillation

Eric Chitambar, Ben Fortescue, and Min-Hsiu Hsieh, *Senior Member, IEEE*

***Abstract*— In this paper, we consider two extensions of the Gács–Körner common information to three variables, the *conditional common information* (cCI) and the *coarse-grained conditional common information* (ccCI). Both quantities are shown to be useful technical tools in the study of classical and quantum resource transformations. In particular, the ccCI is shown to have an operational interpretation as the optimal rate of secret key extraction from an eavesdropped classical source p_{XYZ} when Alice (X) and Bob (Y) are unable to communicate but share common randomness with the eavesdropper Eve (Z). Moving to the quantum setting, we consider two different ways of generating a tripartite quantum state from classical correlations p_{XYZ}: 1) coherent encodings $\sum_{xyz}\sqrt{p_{xyz}}|xyz\rangle$ and 2) incoherent encodings $\sum_{xyz}p_{xyz}|xyz\rangle\langle xyz|$. We study how well can Alice and Bob extract secret key from these quantum sources using quantum operations compared with the extraction of key from the underlying classical sources p_{XYZ} using classical operations. While the power of quantum mechanics increases Alice and Bob's ability to generate shared randomness, it also equips Eve with a greater arsenal of eavesdropping attacks. Therefore, it is not obvious who gains the greatest advantage for distilling secret key when replacing a classical source with a quantum one. We first demonstrate that the classical key rate of p_{XYZ} is equivalent to the quantum key rate for an incoherent quantum encoding of the distribution. For coherent encodings, we next show that the classical and quantum rates are generally incomparable, and in fact, their difference can be arbitrarily large in either direction. Finally, we introduce a "zoo" of entangled tripartite states all characterized by the conditional common information of their encoded probability distributions. Remarkably, for these states almost all entanglement measures, such as Alice and Bob's entanglement cost, squashed entanglement, and relative entropy of entanglement, can be sharply bounded or even exactly expressed in terms of the conditional common information. In the latter case, we thus present a rare instance in which the various entropic entanglement measures of a quantum state can be explicitly calculated.**

***Index Terms*— Quantum cryptography, quantum information, common information, secret key distillation.**

Manuscript received March 13, 2017; revised October 16, 2017, February 7, 2018, and June 17, 2018; accepted June 18, 2018. Date of publication June 29, 2018; date of current version October 18, 2018. E. Chitambar was supported by the National Science Foundation Early CAREER Award under Grant 1352326. M.-H. Hsieh was supported by an ARC Future Fellowship under Grant FT140100574. This paper was presented in part at the 2015 CRYPTO and in part at the 2016 Arctic Crypt.

E. Chitambar and B. Fortescue are with the Department of Physics and Astronomy, Southern Illinois University Carbondale, Carbondale, IL 62901 USA (e-mail: echitamb@siu.edu).

M.-H. Hsieh is with the Centre for Quantum Software and Information, Faculty of Engineering and Information Technology, University of Technology Sydney, Ultimo, NSW 2007, Australia (e-mail: min-hsiu.hsieh@uts.edu.au).

Communicated by M. M. Wilde, Associate Editor for Quantum Information Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2018.2851564

I. Introduction

RESOURCE transformations generally concern transforming noisy resources into useful ones. One particular resource transformation is the problem of *secret key distillation*, which studies the extraction of secret key $\Phi_{XY}\cdot q_Z$ from some initial tripartite correlation p_{XYZ}. Here, Φ_{XY} is a perfectly correlated and uniformly random bit shared between Alice and Bob and q_Z is an arbitrary distribution held by an eavesdropper. Often, the correlations p_{XYZ} are presented as a many-copy source p_{XYZ}^n, and Alice and Bob wish to know the optimal rate of secret bits per copy that they can distill from this source. It turns out that Alice and Bob can often enhance their distillation capabilities by openly disclosing some information about X and Y through public communication [1], [2]. In general, Alice and Bob's communication schemes can be interactive with one round of communication depending on what particular messages were broadcasted in previous rounds. Such interactive protocols are known to generate higher key rates than non-interactive protocols [2]. It is also possible to consider scenarios in which Alice and Bob are not allowed to communicate, yet they still have access to some publically shared randomness that is uncorrelated with their primary source p_{XYZ}. Clearly public communication is a stronger resource than public shared randomness since the former is able to generate the latter. In fact, whereas public communication can significantly improve Alice and Bob's ability to distill secret key, public shared randomness offers no advantage whatsoever, as we demonstrate in this paper.

We envoke a conditional form of the Gács and Körner [3] common information to fully analyze the problem of key distillation using shared randomness. Recall that the original task studied in [3] involves Alice and Bob constructing source codes for X and Y by computing some common variable J_{XY} from their individual variables. The value $H(J_{XY})$ quantifies the greatest asymptotic average sequence-length of matching code words per copy when Alice and Bob independently apply optimal source encodings to their respective variables. For the task of key distillation, Alice and Bob are likewise trying to convert their sources into matching sequences of optimal length. However, the key distillation problem is different in two ways. On the one hand there is the additional constraint that the common sequence should be nearly uncorrelated with Eve. On the other hand, unlike the Gács-Körner problem, it is not required that these sequences belong to faithful encodings of the sources X and Y. Nevertheless, we find that the *coarse-grained conditional common information* (ccCI), $H(J_{XY}|Z)$, quantifies the distillable key when Alice and Bob are unable to communicate with one another. As shown below, this is

0018-9448 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

also the rate even if Alice and Bob have access to auxiliary public randomness which is uncorrelated with their primary distribution.

We refer to $H(J_{XY}|Z)$ as the coarse-grained conditional common information in order to contrast it to an alternative tripartite extension of the Gács-Körner common information. While its definition will be made more precise in Sect. II-A, we define the *conditional common information* (cCI) to essentially be the average common information of all the conditional distributions $p_{XY|Z=z}$. This quantity, denoted as $H(J_{XY|Z}|Z)$, plays an important role in studying the task of secrecy reversibility [4] (see also Sect. II-F). The value of $H(J_{XY}|Z)$ can be obtained by coarse-graining of the common information in the conditional distributions $p_{XY|Z=z}$, and hence $H(J_{XY}|Z) \leqslant H(J_{XY|Z}|Z)$.

Next, we move onto the secret key distillation problem in the quantum regime. There are a variety of physical situations in which one might encounter a many-copy source of variables XYZ. Most notably is the task of quantum key distribution (QKD) in which the variables XYZ are generated through the inherently stochastic nature of quantum measurement [5]. Alice, Bob, and Eve share a tripartite quantum state of the form $|\Psi_{qqq}\rangle^{ABE} = \sum_{x,y,z} \sqrt{p(x,y,z)}|xyz\rangle^{ABE}$, where $p(x, y, z)$ describes a joint distribution for variables XYZ. We say that $|\Psi_{qqq}\rangle^{ABE}$ is a quantum encoding (or "quantum embedding") of distribution $p(x, y, z)$ since when the three parties measure their quantum system in the computational basis (i.e. in the $\{|x\rangle^A\}$, $\{|y\rangle^B\}$, and $\{|z\rangle^E\}$ basis respectively), their measurement outcomes are distributed according to $p(x, y, z)$. If this is done on multiple copies of $|\Psi_{qqq}\rangle$, the parties thus generate a many-copy source of XYZ from which Alice and Bob can distill secret key using public discussion and local processing.

Note that the above mentioned scenario only describes one particular way that Alice and Bob could use multiple copies of $|\Psi_{qqq}\rangle$ to obtain key. With quantum mechanics, more physical operations are allowed than just measuring in the computational basis. Alice and Bob could, for instance, put their local subsystems through some quantum channel (i.e. a trace-preserving, completely positive map), or they could engage in an interactive protocol of local quantum operations and classical communication (LOCC) [6]–[8]. Naively then, it appears that with such greater operational powers, Alice and Bob can always distill at least as much key from a quantum source of $|\Psi_{qqq}\rangle$ than from a classical source of the underlying distribution $p(x, y, z)$. However, in the quantum scenario, Eve also gains operational strength in her eavesdropping abilities. This begs the natural question: for the purpose of secret key distillation, who gains the greatest advantage when embedding a given distribution $p(x, y, z)$ into a multi-party quantum system, the honest parties or the adversary?

Answering this question is a central aim of this paper. Through the construction of specific examples, we show that the advantage can lie either with Alice and Bob or with Eve. Hence the adage "quantum is more powerful than classical" is really a matter of perspective when it comes to the task of secret key distillation. Furthermore, we prove that *quantum coherence* plays the essential role in affecting whether the quantum key rate differs from its classical counterpart. More precisely, in the state $|\Psi_{qqq}\rangle$ given above, the distribution $p(x, y, z)$ is encoded as a coherent superposition of the basis states $|xyz\rangle^{ABE}$. An alternative form of quantum encoding is an *incoherent* mixture of states $\rho_{ccc}^{ABE} = \sum_{x,y,z} p(x, y, z)|xyz\rangle\langle xyz|$. We prove that even when Alice and Bob are allowed to perform arbitrary LOCC on ρ_{ccc}, their optimal rate of key extraction is not improved over the corresponding classical key rate. This result identifies quantum superposition, i.e., quantum coherence [9], as a key ingredient that distinguishes classical from quantum secret key distillation, something that has not been fully understood before.

Finally, we show how the study of classical secret key distillation can have applications in the theory of quantum entanglement. Our result involves computing several entanglement measures of a quantum state based on the properties of its embedded classical distribution. Evaluating some of the most important entanglement measures for a general quantum state is a notoriously difficult problem due to the variational character of these measures. However, as we will demonstrate in this paper, when embedding quantum states with certain types of probability distributions, the entanglement can be bounded by the secret key rate of the underlying distribution; and in some cases the two are equivalent. In fact, the entanglement can be characterized entirely in terms of the conditional common information. This offers a remarkable demonstration of how cryptographic results in classical information can be used to uncover novel physical properties of quantum systems.

Before presenting these results in greater detail, we begin in Sect. II with a relatively self-contained overview of the necessary concepts. In particular, we present the Gács-Körner Common Information and its two generalizations: the conditional common information (cCI) and the coarse-grained conditional common information (ccCI). We then define classes of probability distributions that possess special properties such as the ability to dilute and compress secret correlations at equal rates. We describe a unified framework for local information processing and public communication in both classical and quantum key distillation protocols. Secret key is presented as a classical analog to quantum entanglement, and the tasks of secret key distillation/formation are described as the counterparts to entanglement distillation/formation. Sections III and IV contain our main results. In Sect. III, we provide an operational interpretation for the coarse-grained conditional common information in the framework of classical secret key distillation assisted only by common randomness. Sect. IV presents results comparing secret key distillation in the classical and quantum regimes. Finally, Sect. V offers some concluding remarks.

II. Preliminaries

A. The Gács-Körner Common Information and Its Tripartite Extensions

In this section, we introduce the Gács and Körner [3] bipartite common information and generalize it into two different conditional forms.

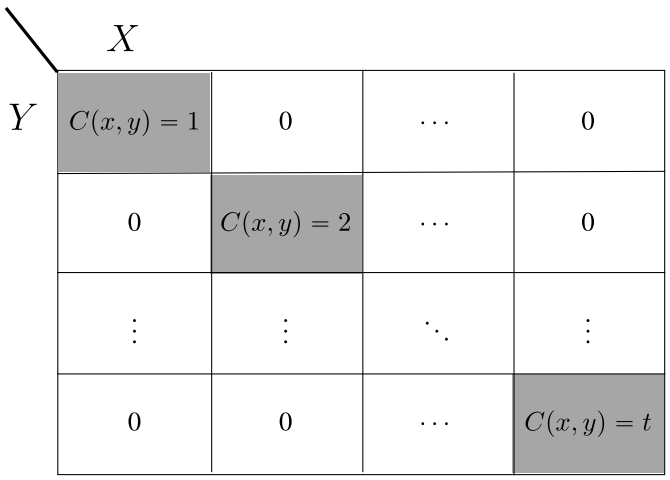

Fig. 1. With a suitable permutation, a distribution p_{XY} on $\mathcal{X} \times \mathcal{Y}$ can be arranged in a block diagonal form, where $\mathcal{X} = \bigcup_{i=1}^{t} \mathcal{X}_i$ and $\mathcal{Y} = \bigcup_{i=1}^{t} \mathcal{Y}_i$ for some t. A common partition of $p_{XY}(x, y)$ is said to be $C(x, y) = i$ if $(x, y) \in \mathcal{X}_i \times \mathcal{Y}_i$.

A *common partition of length* t for random variables XY with a distribution $p_{XY}(x, y)$ on $\mathcal{X} \times \mathcal{Y}$ are pairs of subsets $(\mathcal{X}_i, \mathcal{Y}_i)_{i=1}^{t}$ such that (see Figure 1):

(i) $\mathcal{X}_i \cap \mathcal{X}_j = \mathcal{Y}_i \cap \mathcal{Y}_j = \emptyset$ for $i \neq j$,
(ii) $p(\mathcal{X}_i|\mathcal{Y}_j) = p(\mathcal{Y}_i|\mathcal{X}_j) = \delta_{ij}$, and
(iii) if $(x, y) \in \mathcal{X}_i \times \mathcal{Y}_i$ for some i, then $p_X(x)p_Y(y) > 0$.

For a given common partition, we refer to the subsets $\mathcal{X}_i \times \mathcal{Y}_i$ as the "blocks" of the partition. The subscript i merely serves to label the different blocks, and for any fixed labeling, we associate a random variable $C(X, Y)$ called the *common partition variable* such that $C(x, y) = i$ if $(x, y) \in \mathcal{X}_i \times \mathcal{Y}_i$. Note that each party can determine the value of $C(X, Y)$ from his/her local information (i.e. $C(X, Y) = f(X) = g(Y)$ where $f(x) = i$ if $x \in \mathcal{X}_i$ and $g(y) = i$ if $y \in \mathcal{Y}_i$). A *maximal common partition* is a common partition of greatest length, and it is not difficult to see that every pair of random variables XY has a unique maximal common partition up to relabeling of blocks. We let J_{XY} denote the common partition variable of a maximal common partition, and we refer to J_{XY} as a **maximal common variable** of X and Y.

Proposition 1 [10]: The following are equivalent.

(a) J_{XY} is a maximal common variable of X and Y.
(b) J_{XY} belongs to the set

$$\operatorname*{argmax}_{K} \{H(K) : 0 = H(K|X) = H(K|Y)\},$$

and it is related to every other variable in this set by an invertible function. Note that the maximization here is (necessarily) restricted to all random variables with $|K| \leqslant \min\{|X|, |Y|\}$.
(c) If $f(X) = g(Y) = C$ for any variable C, then C is a function of J_{XY}.

Since every pair of variables is uniquely associated with a maximal common variable up to relabeling, entropic quantities like $H(J_{XY})$ are uniquely defined for any X and Y. In their original work, Gács and Körner [3] in fact identify $H(J_{XY})$ as the common information[1] of X and Y. Each maximal common variable J_{XY} ranges over the same finite set $\mathcal{J}$, and it provides a decomposition of p_{XY} as follows:

$$p(x, y) = \sum_{j \in \mathcal{J}} p(x, y|j)p(j), \tag{1}$$

where for any $x, x' \in \mathcal{X}$ and $y, y' \in \mathcal{Y}$, the conditional distributions satisfy $p(x, y|j)p(x, y'|j') = 0$ and $p(x, y|j)p(x', y|j') = 0$ if $j \neq j'$.

The following proposition provides a useful characterization of values x and x' that belong to the same block in a maximal common partition.

Proposition 2 [3]: If $J_{XY}(x) = J_{XY}(x')$, then there exists a sequence of values

$$x y_1 x_1 y_2 x_2 \cdots y_n x'$$

such that $p(x, y_1)p(y_1, x_1)p(x_1, y_2) \cdots p(y_n, x') > 0$.

We now arrive to the main quantities studied in this paper. For a tripartite distribution p_{XYZ}, the **coarse-grained conditional common information** (ccCI) is defined simply as the value $H(J_{XY}|Z)$. A different type of conditional common information can also be considered. One first defines a **maximal conditional common variable** $J_{XY|Z}$, which is a random variable depending on XYZ such that $J_{XY|Z=z}$ is a maximal common variable with respect to the conditional distribution $p_{XY|Z=z}$. Like the maximal common variable J_{XY}, the maximal conditional common variable is unique for every tripartite distribution p_{XYZ} up to a relabeling of the various $J_{XY|Z=z}$. Consequently, the quantity $H(J_{XY|Z}|Z)$ is unique, and we refer to it as the **conditional common information** (cCI).

Note that since $J_{XY|Z=z}$ is computed from both X and Y with the additional information that $Z = z$, maximality of $J_{XY|Z=z}$ ensures that J_{XY} is a function of $J_{XY|Z=z}$ for each $z \in \mathcal{Z}$. Therefore, $H(J_{XY}|Z) \leqslant H(J_{XY|Z}|Z)$ with equality iff $H(J_{XY|Z}|ZJ_{XY}) = 0$. When the equality condition holds, it means that for each $z \in \mathcal{Z}$, the value of $J_{XY|Z=z}$ can be determined from J_{XY} alone. In other words, there must exist maximal common variables $\{J_{XY|Z=z} : z \in \mathcal{Z}\}$ such that $J_{XY}(x) = i$ implies $J_{XY|Z=z}(x) = i$ whenever $p_{X|Z=z}(x) > 0$. For these maximal common variables, we therefore have $H(J_{XY|Z}|X) = H(J_{XY|Z}|Y) = 0$.

B. A Zoo of Tripartite Distributions

We next introduce families of different tripartite distributions that are important to our study. The maximal conditional common information $J_{XY|Z}$ is useful in characterizing these classes. A distribution p_{XYZ} is said to be ([4], [10]):

- **Block independent** (BI) if $I(X : Y|J_{XY|Z}Z) = 0$. Table I provides an example of a block independent distribution.
- **Uniform block independent** (UBI) if there exists a maximal conditional common variable $J_{XY|Z}$ such that

[1]We remark that a variant of common information was defined in [11], which also has important implications in tripartite secret correlations [12], [13].

TABLE I

AN EXAMPLE OF A BLOCK INDEPENDENT (BI) DISTRIBUTION. IN THIS EXAMPLE, $Z = \{0, 1, 2\}$ AND THE PROBABILITY DEPICTED IS THE CONDITIONAL PROBABILITY $p(x, y|z)$

$X \rightarrow$

$Z=0$	0	1	2
0	1/8	1/8	
1	1/8	1/8	
2			1/2

$Y \downarrow$

$Z=1$	0	1	2
0	1/6	1/6	
1	1/6	1/6	
2			1/3

$Z=2$	0	1	2
0	1/2		
1		1/8	1/8
2		1/8	1/8

TABLE II

AN EXAMPLE OF A UNIFORM BLOCK INDEPENDENT (UBI) DISTRIBUTION. IN THIS EXAMPLE, $Z = \{0, 1, 2\}$ AND THE PROBABILITY DEPICTED IS THE CONDITIONAL PROBABILITY $p(x, y|z)$

$X \rightarrow$

$Z=0$	0	1	2
0	1/8	1/8	
1	1/8	1/8	
2			1/2

$Y \downarrow$

$Z=1$	0	1	2
0	1/6	1/6	
1	1/6	1/6	
2			1/3

$Z=2$	0	1	2
0	1/4	1/4	
1			
2			1/2

TABLE III

AN EXAMPLE OF A UNIFORM BLOCK INDEPENDENT UNDER PUBLIC DISCUSSION (UBI-PD) DISTRIBUTION. IN THIS EXAMPLE, $Z = \{0, 1, 2\}$ AND THE PROBABILITY DEPICTED IS THE CONDITIONAL PROBABILITY $p(x, y|z)$. ALICE WHO HOLDS THE RANDOM VARIABLE X CAN ANNOUNCE WHETHER THE VALUE IS IN THE RANGE $\{0, 1, 2\}$ OR $\{3, 4, 5\}$. AFTER THIS PUBLIC ANNOUNCEMENT M, ALICE AND BOB WILL HOLD A UBI DISTRIBUTION $p_{(XM)(YM)(ZM)}$

$X \rightarrow$

$Z=0$	0	1	2
0	1/8	1/8	
1	1/8	1/8	
2			1/2

$Y \downarrow$

$Z=1$	0	1	2
0	1/6	1/6	
1	1/6	1/6	
2			1/3

$Z=2$	3	4	5
0	1/2		
1		1/8	1/8
2		1/8	1/8

both $I(X : Y|J_{XY|Z}Z) = 0$ and $H(J_{XY|Z}|X) = H(J_{XY|Z}|Y) = 0$. Note that for UBI distributions, the ccCI is equivalent to the cCI (i.e. $H(J_{XY}|Z) = H(J_{XY|Z}|Z)$). Table II provides an example of a uniform block independent distribution.

- **Uniform block independent under public discussion** (UBI-PD) if it is BI and there is a public communication protocol generating messages M such that $p_{(MX)(MY)(ZM)}$ is UBI and $I(M : J_{XY|Z}|Z) = 0$. Table III provides an example of a UBI-PD distribution.
- **Uniform block independent under public discussion and eavesdropper's local processing** (UBI-PD$\downarrow$) if there exists a channel $\overline{Z}|Z$ such that $p_{XY|\overline{Z}}$ is UBI with the required public communication M also satisfying $I(Z : J_{XY|\overline{Z}}|M\overline{Z}) = 0$. Table IV provides an example of a UBI-PD$\downarrow$ distribution.
- **Semi-unambiguous** [14] if $H(Z|XY) = 0$.
- **Unambiguous** [15] if $H(Z|XY) = 0$ and $H(XY|J_{XY|Z}Z) = 0$.

Being BI means that given Z, Alice and Bob share no more correlations besides their block number specified by some maximal conditional common variable $J_{XY|Z}$. For UBI distributions, the blocks of the conditional distributions $p_{XY|Z=z}$ can be ordered in such a way that is independent of Z, and their number can therefore be computed locally by Alice and Bob. Finally, for UBI-PD, the distribution becomes UBI once Alice and Bob exchange messages M which, from Eve's perspective, is independent of their block number. For semi-unambiguous distributions, the random variable Z can be uniquely determined by random variables X and Y; while for unambiguous distributions, each random variable can be uniquely determined by the other two random variables.

TABLE IV

AN EXAMPLE OF A UNIFORM BLOCK INDEPENDENT UNDER PUBLIC DISCUSSION AND EAVESDROPPER'S LOCAL PROCESSING (UBI-PD↓) DISTRIBUTION. IN THIS EXAMPLE, $Z = \{0, 1\}$ WITH $p_Z(0) = 1/5$ AND $p_Z(1) = 4/5$. THE RIGHTMOST TABLE IS A FULL COARSE-GRAINING OF Z, WHERE $\bar{Z}$ CONTAINS ONLY ONE VALUE 0. ALICE AND BOB CAN GENERATE $J_{XY|\bar{Z}}$ WITHOUT PUBLIC COMMUNICATION

$X \rightarrow$

$Y \downarrow$ / $Z=0$	0	1	2
0	1/2		
1		1/8	
2			3/8

$Z=1$	0	1	2
0	1/2		
1		1/8	5/32
2		5/32	1/16

$\bar{Z}|Z \Longrightarrow$

$\bar{Z}=0$	0	1	2
0	1/2		
1		1/8	1/8
2		1/8	1/8

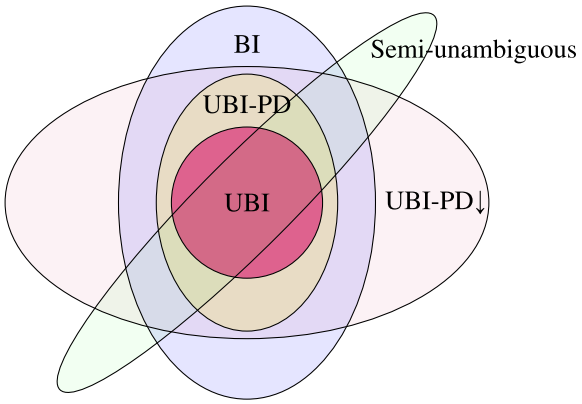

Fig. 2. The relations between known classical distributions.

The relations between these distributions are depicted in Figure 2.

C. A Unified Framework for Local Information Processing and Public Communication

In this section we review the definitions of classical and quantum local operations and public communication (LOPC).

1) Classical Operations: In the classical LOPC setting, each party is allowed to perform the following operations:

(i) Generate local random variables that are uncorrelated with the variables held by any other party.
(ii) Copy the value of any locally held variable.
(iii) Change the values of any locally held variables according to some function.
(iv) Broadcast the result of any computed function over an authenticated public channel.

Note that operations (i) - (iii) encompass any sort of noisy processing that a party may wish to perform. A general **classical LOPC protocol** $\mathcal{P}_c$ then consists of two phases: Phase I - a coordinated and multi-round exchange of public messages in which each message is a function of some party's local variables, and Phase II - each party processes his/her variables in a way that can depend on the particular messages exchanged in Phase I, thereby generating the output variables of the protocol. It is not difficult to see that the variables produced by any sequence of operations (i)–(iv) can always also be generated by a protocol following this two-phase format [1], [2], [16].

Consider now an arbitrary random variable G that is distributed according to p_G over alphabet $\mathcal{G}$. Formally, we will represent G as a quantum state:

$$\omega_G = \sum_{g \in \mathcal{G}} p(g)|g\rangle\langle g|, \tag{2}$$

where the $|g\rangle$ are orthonormal vectors for a vector space of dimension $|\mathcal{G}|$. Suppose that the party holding G announces some public message described by random variable M ranging over set $\mathcal{M}$. Then the resulting state has the form

$$\omega_{GM} = \sum_{g \in \mathcal{G}} \sum_{m \in \mathcal{M}} p(g, m)|g\rangle\langle g| \otimes |m\rangle\langle m|^C, \tag{3}$$

where we use the convention that system C holds the public communication accessible to all parties (including Eve).

2) Quantum Operations: In the quantum LOPC setting, each party is allowed to perform the following operations:

(i) Perform a local quantum instrument $(\mathcal{E}_m)_m$ [17], where each $\mathcal{E}_m$ is a completely positive (CP) map, and their sum $\sum_m \mathcal{E}_m$ is a trace-preserving map. Quantum instruments represent the most general type of quantum measurement. When performing the instrument on the state σ, the "measurement" outcome m is obtained with probability $p(m) = tr[\mathcal{E}_m(\sigma)]$, and the post-measurement state given this outcome is $\sigma_m = \mathcal{E}_m(\sigma)/p(m)$ for $p(m) > 0$.
(ii) Broadcast the result of any quantum measurement.

A general **quantum LOPC protocol** $\mathcal{P}_q$ is described by a multi-level "tree" of local instruments in which the choice of instrument performed at each node of the tree depends on the particular history of measurement outcomes leading up to that node (see [6], [8], and [18] for details). Classical operations (i)–(iv) above fall within the framework of local quantum instruments since evaluating a function is a special type of quantum measurement in which the measurement outcome is the function's value. Therefore, quantum LOPC generalizes the notion of classical LOPC.

In both classical and quantum LOPC protocols, we assume that only Alice and Bob engage in public discussion and Eve

is passively eavesdropping. With a slight abuse of notation, for a given classical/quantum protocol we use $\mathcal{P}_c/\mathcal{P}_q$ to denote both the particular protocol as well as the map associated with the protocol. For instance, when a quantum LOPC protocol is performed on state σ^{ABE}, the overall action can be expressed as

$$\sigma^{ABE} \to \widehat{\sigma}^{ABEC} = \mathcal{P}_q(\sigma^{ABE}) = \sum_m p(m)\sigma_m^{ABE} \otimes |m\rangle\langle m|^C,$$

where σ_m^{ABE} is the tripartite state generated when m is the total broadcasted message.

D. Classical and Quantum Secret Key Rates

Throughout this paper, we will assume that some basis for Alice, Bob, and Eve's system has been chosen and is fixed. Each of these is typically referred to as the *computational basis* for the given system and is denoted by $\{|x\rangle^A\}_{x=1}^{d_A}$, $\{|y\rangle^B\}_{y=1}^{d_B}$ and $\{|z\rangle^E\}_{z=1}^{d_E}$ respectively. Let p_{XYZ} be an arbitrary three-way joint probability distribution for random variables X, Y, and Z which takes on values $p(x, y, z)$. We introduce the following physical instantiations of p_{XYZ}:

- A coherent embedding (or qqq embedding):
$$|\Psi_{qqq}\rangle = \sum_{x,y,z} \sqrt{p(x, y, z)}|xyz\rangle^{ABE}. \quad (4)$$
- A one-sided incoherent embedding (or cqq embedding):
$$\rho_{cqq} = \sum_x p(x)|x\rangle\langle x|^A \otimes |\psi_x\rangle\langle\psi_x|^{BE}, \quad (5)$$
where $|\psi_x\rangle = \sum_{y,z} \sqrt{p(y, z|x)}|yz\rangle$.
- A two-sided incoherent embedding (or ccq embedding):
$$\rho_{ccq} = \sum_{x,y} p(x, y)|xy\rangle\langle xy|^{AB} \otimes |\psi_{xy}\rangle\langle\psi_{xy}|^E, \quad (6)$$
where $|\psi_{xy}\rangle = \sum_z \sqrt{p(z|xy)}|z\rangle$.
- An incoherent embedding (or ccc embedding):
$$\rho_{ccc} = \sum_{x,y,z} p(x, y, z)|xyz\rangle\langle xyz|^{ABE}. \quad (7)$$

Note that ρ_{ccc} corresponds to the state ω_{XYZ} introduced in Eq. (2). We can therefore think of ρ_{ccc} as either a classical or quantum object, the difference being dictated by whether it is processed using either classical or quantum LOPC.

The various embeddings can be related through a series of local physical transformations:

$$|\psi_{qqq}\rangle\langle\psi_{qqq}| \xrightarrow{(1)} \rho_{cqq} \xrightarrow{(2)} \rho_{ccq} \xrightarrow{(3)} \rho_{ccc}, \quad (8)$$

where (1) is attained by Alice performing a dephasing channel

$$\sigma \to \sum_x |x\rangle\langle x|\sigma|x\rangle\langle x|$$

and likewise for (2) and (3). One can also consider a dephasing exclusively on Eve's side. This corresponds to the state

$$\rho_{qqc} = \sum_z \sum_{x,x',y,y'} p(z)|\psi_z\rangle\langle\psi_z|^{AB} \otimes |z\rangle\langle z|^E, \quad (9)$$

$$\text{where } |\psi_z\rangle = \sum_{x,y} \sqrt{p(x, y|z)}|xy\rangle. \quad (10)$$

Secret Key Distillation: The scenario we consider is an identical, independent, and discrete (i.i.d.) source that is generating some particular embedding of p_{XYZ} for Alice, Bob and Eve. The goal of Alice and Bob is to distill secret key, which is shared randomness held independently of Eve's system. We denote the state corresponding to $\log s$ bits of perfectly shared randomness by

$$\Phi_s^{AB} = \frac{1}{s}\sum_{i=0}^{s-1} |ii\rangle\langle ii|^{AB}. \quad (11)$$

The notion of secret key rate is defined as follows.

Definition 3: For a distribution p_{XYZ}, we say that R is a (classical) **LOPC achievable key rate** if for every $\epsilon > 0$, there exists a classical LOPC protocol $\mathcal{P}_c$ acting on Alice and Bob's parts of $\sigma^{ABE} := \rho_{ccc}^{\otimes n}$ (for n sufficiently large) and generating the state $\widehat{\sigma}^{ABEC}$ (with C being the public communication system) such that

$$\tfrac{1}{2}\left\|\widehat{\sigma}^{ABEC} - \Phi_s^{AB} \otimes \widehat{\sigma}^{EC}\right\|_1 < \epsilon, \quad (12)$$

where $\widehat{\sigma}^{EC} = \mathrm{Tr}_{AB}\,\widehat{\sigma}^{ABEC}$ and $\frac{1}{n}\log s > R - \epsilon$. The supremum achievable key rate is denoted by $K_D(p_{XYZ})$. We say that R is a **ccc**, **ccq**, **cqq**, or **qqq LOPC achievable key rate** if there exists a quantum LOPC protocol $\mathcal{P}_q$ to replace $\mathcal{P}_c$ in Eq. (12), and we further take $\sigma^{ABE} := \rho_{ccc}^{\otimes n}$, $\sigma^{ABE} := \rho_{ccq}^{\otimes n}$, $\sigma^{ABE} := \rho_{cqq}^{\otimes n}$ or $\sigma^{ABE} := |\psi_{qqq}\rangle\langle\psi_{qqq}|^{\otimes n}$ respectively. The supremum achievable key rates in these scenarios are denoted by $K_D(\rho_{ccc})$, $K_D(\rho_{ccq})$, $K_D(\rho_{cqq})$ and $K_D(\psi_{qqq})$ respectively [19].

We are also interested in key distillation scenarios where the public communication is replaced with public shared randomness. This can be seen as a special subclass of LOPC protocols in which the only public communication consists of one message that distributes to all parties (including an eavesdropper) a random variable W which is uncorrelated with XYZ. We say that R is a **common randomness (c.r.) achievable key rate** if Eq. (12) is satisfied with the public communication just establishing common randomness in this way. We denote the supremum of all achievable c.r. key rates as $K_D^{c.r.}(p_{XYZ})$.

E. Quantum Entanglement

Quantum pure entanglement [20] is a resource shared between two or more quantum systems that is distinct from secret key. However, quantum entanglement and secret keys share many similarities [14], [15], [21]–[28]. Starting from a tripartite pure state $|\Psi_{qqq}\rangle$, Alice and Bob share one entangled bit (ebit) of quantum information in the state $|\Psi_{qqq}\rangle$ if it has the form

$$|\Psi_{qqq}\rangle^{ABE} = |\Phi_2\rangle^{AB} \otimes |\varphi\rangle^E, \quad (13)$$

where $|\Phi_2\rangle^{AB} := \sqrt{1/2}(|00\rangle + |11\rangle)^{AB}$ is a so-called ebit and $|\varphi\rangle^E = \sum_z \sqrt{p(z)}|z\rangle$ is any state held by Eve. On the surface, the tripartite state $|\Psi_{qqq}\rangle^{ABE} = |\Phi_2\rangle^{AB} \otimes |\varphi\rangle^E$ looks very similar to the state $\rho_{ccq}^{ABE} = \Phi_2^{AB} \otimes |\varphi\rangle\langle\varphi|^E$, which contains one bit of distillable secret key and is obtained from

$|\Psi_{qqq}\rangle^{ABE}$ through dephasing by Alice and Bob. However there is a critical difference between the two states. For $\Phi_2^{AB} \otimes |\varphi\rangle\langle\varphi|^E$, it is entirely consistent that there should exist some third party Sapna (S) who holds as side information the value of Alice and Bob's bit in Φ_2^{AB}.[2] In other words, we can envision a four-party state $\sigma^{ABES} = \Phi_3^{ABS} \otimes |\varphi\rangle\langle\varphi|^E$ with $\Phi_3^{ABS} = \frac{1}{2}(|000\rangle\langle 000| + |111\rangle\langle 111|)$. And while $K_D(\sigma^{ABE}) = 1$, there is no secrecy with respect to Sapna: $K_D(\sigma^{ABS}) = 0$. In contrast, Alice and Bob's entanglement in $|\Psi_{ABE}\rangle$ exists regardless of what side information is known. That is, if $\mathrm{Tr}_{ES}(\sigma^{ABES}) = |\Phi_2\rangle\langle\Phi_2|^{AB}$ for *any* state σ^{ABES}, then necessarily σ^{ABES} has the product-state form $\sigma^{ABES} = |\Phi_2\rangle\langle\Phi_2|^{AB} \otimes \sigma^{ES}$. This means that if Alice and Bob should dephase when holding the state σ^{ABES}, they will generate key that is secret from not only Eve but also Sapna: $|\Phi_2\rangle\langle\Phi_2|^{AB} \otimes \sigma^{ES} \rightarrow \Phi_2^{AB} \otimes \sigma^{ES}$. Therefore, pure-state entanglement is intrinsic to bipartite systems themselves and, unlike secret key, one does not need to introduce any third party to speak of its pure entanglement.[3]

Similar to the secret key rate K_D, one can also define for ρ^{AB} the entanglement distillation rate E_D [29]. This quantifies the asymptotic rate for which ebits can be obtained from ρ^{AB} using local operations and classical communication (LOCC). The operational class LOCC differs from quantum LOPC in that the former makes no explicit reference to a third party who records the "public" communication. It is a fundamental and challenging problem in quantum information to compute $E_D(\rho^{AB})$ for a given quantum state. Almost all meaningful measures of entanglement provide an upper bound for E_D [30], and three such measures are the relative entropy of entanglement [31], the squashed entanglement [32], and the entanglement of formation [33]:

- $E_r(\rho^{AB})$: the relative entropy of entanglement is

$$E_r(\rho^{AB}) = \min_{\sigma \in \mathcal{S}} S(\rho\|\sigma), \tag{14}$$

  where $\mathcal{S}$ is the set of separable density operators and $S(\rho\|\sigma) = -\mathrm{Tr}[\rho \log \sigma] - S(\rho)$ is the relative entropy;
- $E_{sq}(\rho^{AB})$: the squashed entanglement is

$$E_{sq}(\rho^{AB}) = \frac{1}{2} \inf_{\rho^{ABE}} I(A:B|E)_{\rho^{ABE}}, \tag{15}$$

  where the infimum is taken over all extensions ρ^{ABE} such that $\mathrm{Tr}_E\, \rho^{ABE} = \rho^{AB}$, and $I(A:B|E)_{\rho^{ABE}} = S(AE) + S(BE) - S(ABE) - S(E)$ is the conditional quantum mutual information of the state ρ^{ABE}.
- $E_F(\rho^{AB})$: the entanglement of formation is

$$E_F(\rho^{AB}) = \min \sum_i p(i) S(\mathrm{Tr}_A\, \varphi_i), \tag{16}$$

  with the minimization taken over all decompositions $\rho^{AB} = \sum_i p(i)|\varphi_i\rangle\langle\varphi_i|$.

[2]Note that throughout the paper, we will explicitly write out the projector $|\Phi_2\rangle\langle\Phi_2|$, and reserve the notation Φ_2 for a classical (diagonal) state.

[3]We remark that for quantum keys in a bipartite state with the strongest security condition, one also does not need to consider any additional adversarial party since he/she is assumed to hold the purification of the state.

The particular significance of these entanglement measures is that they provide upper bounds not only for the distillable entanglement but also for distillable key.

Theorem 4 [30], [34]–[36]: For an arbitrary tripartite state $|\Psi_{qqq}\rangle^{ABE}$ with $\rho^{AB} = \mathrm{Tr}_E\, |\Psi_{qqq}\rangle\langle\Psi_{qqq}|^{ABE}$, the rates $K_D(\Psi_{qqq})$ and $E_D(\rho^{AB})$ are both upper bounded by the relative entropy of entanglement $E_r(\rho^{AB})$ as well as the squashed entanglement $E_{sq}(\rho^{AB})$.

Unfortunately, each of the above entanglement measures involves a complicated minimization and in fact, their evaluation represents an NP-hard/NP-complete computational problem [37]. It is therefore not surprising that very few instances are known in which any of these measures can be explicitly computed. In this paper, we introduce a new class of quantum states for which all these measures can be evaluated. Our strategy will be based on the notion of *reversible secrecy*, which we describe next.

F. Reversible Entanglement and Secret Key

Dual to the task of entanglement distillation is the task of entanglement formation, which describes building a given state ρ^{AB} using LOCC and an initial supply of ebits. The **entanglement cost** E_C of a mixed state ρ^{AB} is the asymptotic optimal rate of ebit consumption for Alice and Bob to generate faithful copies of ρ^{AB} by LOCC [38]. The entanglement cost is obviously lower bounded by the distillable entanglement, and compared to the above entanglement measures, the following hierarchy holds [30]:

$$\begin{Bmatrix} E_D(\rho^{AB}) \\ K_D(\Psi_{qqq}) \end{Bmatrix} \leqslant \begin{Bmatrix} E_r(\rho^{AB}) \\ E_{sq}(\rho^{AB}) \end{Bmatrix} \leqslant E_C(\rho^{AB}) \leqslant E_F(\rho^{AB}), \tag{17}$$

where the first inequality references Theorem 4. A state ρ^{AB} is said to possess **reversible entanglement** if $E_D(\rho^{AB}) = E_C(\rho^{AB})$. Operationally this means that the entanglement in ρ^{AB} can be concentrated and diluted at equal rates.

Recently, the phenomenon of **reversible secrecy**, which is the classical analog to reversible entanglement, was studied in [4]. Here, one first identifies the key cost K_C of a distribution p_{XYZ} as the amount of secret correlations needed for Alice and Bob to asymptotically prepare their reduced distribution p_{XY} using a classical LOPC protocol that reveals to Eve in the public communication no greater amount of information about their variables than what she possesses in p_{XYZ} [39]. The distribution is said to possess reversible secrecy if $K_D(p_{XYZ}) = K_C(p_{XYZ})$. The following theorem provides a strong necessary condition on the structure of distributions having reversible secrecy, and well as a sufficient condition.

Theorem 5 [4]: (1) If $K_C(p_{XYZ}) = K_D(p_{XYZ})$ then there exists a channel for Eve $\overline{Z}|Z$ such that $p_{XY\overline{Z}}$ is BI. (2) If p_{XYZ} is UBI-PD$\downarrow$, then $K_C(p_{XYZ}) = K_D(p_{XYZ})$.

In [4], it is shown that the necessary condition (1) and sufficient condition (2) are equivalent whenever either Alice or Bob holds a binary random variable.

A key result proven in Theorem 12 below is that for a distribution p_{XYZ} having reversible secrecy and $|\Psi_{qqq}\rangle$ being

its qqq embedding, Eq. (17) can be further upper bounded as

$$\begin{aligned}\begin{cases} E_D(\rho^{AB}) \\ K_D(\Psi_{qqq}) \end{cases} &\leqslant \begin{cases} E_r(\rho^{AB}) \\ E_{sq}(\rho^{AB}) \end{cases} \leqslant E_C(\rho^{AB}) \\ &\leqslant E_F(\rho^{AB}) \leqslant K_D(p_{XYZ}). \end{aligned} \tag{18}$$

By identifying a class of distributions in Section II-A for which $K_D(\Psi_{qqq}) = K_D(p_{XYZ})$, this chain of inequalities becomes tight and we are thus able to compute the various entanglement measures of ρ^{AB}.

III. Key Distillation Using Auxiliary Public Randomness

The main result in this section links the secret key rate $K_D^{c.r.}(p_{XYZ})$ to the coarse-grained conditional common information (ccCI).

Theorem 6: $K_D^{c.r.}(p_{XYZ}) = H(J_{XY}|Z)$. Moreover, $H(J_{XY}|Z)$ is achievable with no additional common randomness.

Achievability: We will prove that $H(J_{XY}|Z)$ is an achievable rate without any auxiliary shared public randomness (i.e. W is constant). For n copies of p_{XYZ}, Alice and Bob extract their common information from each copy of p_{XYZ}. This will generate a sequence of J_{XY}^n, with Alice and Bob having identical copies of this sequence. It is now a matter of performing privacy amplification on this sequence to remove Eve's information [40]. The main construction is guaranteed to exist by Lemma 15 in Appendix. From this lemma, it follows that $H(J_{XY}|Z)$ is an achievable key rate.

Converse: The converse proof follows analogously to the converse proof of [41, Th. 2.6] (see also [16]). Starting from the definition of c.r. achievable key rate, Eq. (12) implies certain entropic constraints on the generated variables. Let us write

$$\widehat{\sigma}^{ABEC} = \sum_{k,k',z^n,w} p(k,k',z^n,w)|k,k'\rangle\langle k,k'|^{AB} \otimes |z^n,w\rangle\langle z^n,w|^{EC},$$

where W (taking on values w) is the public shared randomness initially uncorrelated with all other variables, K (taking on values k) is the final variable obtained by Alice locally processing (X^n, W), and K' (taking on values k') is the final variable obtained by Bob locally processing (Y^n, W). Let us consider first the scenario when K and K' are functions of (X^n, W) and (Y^n, W), respectively; we will consider the possibility of stochastic mappings later. From Eq. (12) in the definition of achievable key rate, we obtain

$$\epsilon > \Pr\{K \neq K'\}, \tag{19}$$

$$\epsilon > \tfrac{1}{2} \left\| p_{KZ^nW} - \mathbb{1}_{\mathcal{K}} \cdot p_{Z^nW} \right\|_1, \tag{20}$$

where $\mathbb{1}_{\mathcal{K}}$ is the uniform distribution over $\mathcal{K}$ with $\frac{1}{n}\log|\mathcal{K}| > R - \epsilon$. Applying Fano's Inequality and data processing to Eq. (19) yields

$$H(K|Y^nW) < h(\epsilon) + \epsilon \log|\mathcal{K}|, \tag{21}$$

where $h(x) = -x\log x - (1-x)\log(1-x)$ is the binary entropy function. Similarly, applying Fannes' Inequality to Eq. (20) gives

$$\log|\mathcal{K}| - H(K|Z^nW) < \epsilon \log|\mathcal{K}| + h(\epsilon). \tag{22}$$

For $\epsilon < 1/2$, these can be combined to obtain the bound

$$\log|\mathcal{K}| < H(K|Z^nW) - H(K|Y^nW) + 2h(\epsilon) + 2\epsilon\log|\mathcal{K}|, \tag{23}$$

and thus for any achievable $R < \frac{1}{n}\log|\mathcal{K}| + \epsilon$, we have

$$\begin{aligned} R &\leqslant \frac{1}{n}\log|\mathcal{K}| + \epsilon \\ &< \frac{1}{1-2\epsilon} \cdot \frac{1}{n}\left[H(K|Z^nW) - H(K|Y^nW)\right] + \frac{2h(\epsilon)}{1-2\epsilon} + \epsilon. \end{aligned} \tag{24}$$

To analyze the quantity $H(K|Z^nW) - H(K|Y^nW)$, we will use a standard trick, stated in Lemma 16, to obtain

$$\begin{aligned} &H(K|Z^nW) - H(K|Y^nW) \\ &\qquad = n[I(K : Y^{(J)}|UW) - I(K : Z^{(J)}|UW)], \end{aligned} \tag{25}$$

where $U := JY^{(<J)}Z^{(>J)}$. Notice that for any $i \in \{1, \cdots, n\}$ we have

$$X^{(<i)}X^{(>i)}Y^{(<i)}Z^{(>i)} - X^{(i)} - Y^{(i)}Z^{(i)}, \tag{26}$$

since the sampling is i.i.d. Here the notation $X - Y - Z$ means that X, Y, Z form a Markov chain so that $I(X : Z|Y) = 0$. Therefore, because K is a function of (X^n, W), we have

$$KU - X^{(J)}W - Y^{(J)}Z^{(J)}. \tag{27}$$

Removing the superscript "J" and taking $\epsilon \to 0$, we have the bound

$$R \leqslant I(K : Y|UW) - I(K : Z|UW) \tag{28}$$

such that $KU - XW - YZ$.

Next, Eq. (21) gives

$$\begin{aligned} h(2\epsilon) + 2\epsilon\log|\mathcal{K}| &> H(K|Y^nW) - H(K|X^nW) \\ &= n[I(K : X^{(J)}|JY^{(<J)}X^{(>J)}W) \\ &\quad - I(K : Y^{(J)}|JY^{(<J)}X^{(>J)}W)], \end{aligned} \tag{29}$$

where the first inequality follows because $H(K|X^nW)$ is nonnegative and the equality follows from Lemma 16. We want to put this in terms of U. To do this, note that

$$\begin{aligned} &I(K : X^{(J)}|JY^{(<J)}X^{(>J)}W) \\ &\quad = I(KY^{(<J)}X^{(>J)} : X^{(J)}|JW) \\ &\quad = I(KY^{(<J)}X^{(>J)}Z^{(>J)} : X^{(J)}|JW) \\ &\qquad - I(Z^{(>J)} : X^{(J)}|JKY^{(<J)}X^{(>J)}W) \\ &\quad = I(KUX^{(>J)} : X^{(J)}|JW) \\ &\quad = I(KU : X^{(J)}|JW) + I(X^{(>J)} : X^{(J)}|KUW), \end{aligned} \tag{30}$$

where the first equality follows from the chain rule and $I(Y^{(<J)}X^{(>J)} : X^{(J)}|JW) = 0$, and in the second equality

$$\begin{aligned} &I(Z^{(>J)} : X^{(J)}|JKY^{(<J)}X^{(>J)}W) \\ &\quad \leqslant I(Z^{(>J)} : KX^{(J)}|JY^{(<J)}X^{(>J)}W) \\ &\quad = I(Z^{(>J)} : X^{(J)}|JY^{(<J)}X^{(>J)}W) \\ &\quad = 0. \end{aligned} \tag{31}$$

The first equality (31) uses $I(Z^{(>J)} : K|JY^{(<J)}X^{(\geqslant J)}W) = 0$ since $K - JY^{(<J)}X^{(\geqslant J)}W - Z^{(>J)}$ is a Markov chain. Again this follows from the basic Markov condition $K - WX^n - Y^nZ^n$ and the sampling is i.i.d.. The second equality follows from i.i.d. sampling and W independence of X^n, Y^n, Z^n.

A similar analysis likewise gives

$$\begin{aligned} &I(K : Y^{(J)}|JY^{(<J)}X^{(>J)}W) \\ &\quad = I(KU : Y^{(J)}|JW) + I(X^{(>J)} : Y^{(J)}|KUW) \\ &\quad \leqslant I(KU : Y^{(J)}|JW) + I(X^{(>J)} : X^{(J)}|KUW), \end{aligned} \tag{32}$$

where the inequality follows from the Markov condition

$$X^{(>J)} - KUX^{(J)}W - Y^{(J)},$$

which can be derived from the more obvious Markov condition

$$KUX^n - JX^{(J)}W - Y^{(J)}.$$

Putting everything together yields

$$\begin{aligned} &h(\epsilon) + \epsilon(\log|\mathcal{K}| - 1) \\ &\quad > H(K|Y^nW) - H(K|X^nW) \\ &\quad > I(KU : X^{(J)}|JW) - I(KU : Y^{(J)}|JW) \\ &\quad = I(KU : X^{(J)}Y^{(J)}|JW) \\ &\qquad - I(KU : Y^{(J)}|JX^{(J)}W) - I(KU : Y^{(J)}|JW) \qquad (33) \\ &\quad = I(KU : X^{(J)}|JY^{(J)}W) + I(KU : Z^{(J)}|JY^{(J)}X^{(J)}W) \\ &\quad = I(KU : X^{(J)}Z^{(J)}|JY^{(J)}W), \qquad (34) \end{aligned}$$

where the second term in (33) is zero from the already proven Markov chain $KU - XW - YZ$, and in (34) we use the fact that $I(KU : Z^{(J)}|JY^{(J)}X^{(J)}W) = 0$. Removing the superscript "J" and taking $\epsilon \to 0$ necessitates the Markov chain $KU - YW - XZ$.

The double Markov chain $K - XW - Y$ and $K - YW - X$ implies that $I(K : XY|J_{XY}W) = 0$ (see Proposition 18 below). Since K is a function of (X, W), we have that $H(K|J_{XY}W) = 0$. Thus, K must also be a function of (Y, W). Continuing Eq. (28) gives the bound

$$\begin{aligned} R &\leqslant I(K : Y|UW) - I(K : Z|UW) \\ &= H(K|UW) - I(K : Z|UW) \\ &= H(K|ZUW) \leqslant H(K|ZW). \end{aligned} \tag{35}$$

We have therefore obtained the following:

$$R \leqslant \max H(K|ZW), \tag{36}$$

where the maximization is taken over all variables K such that $H(K|XW) = H(K|YW) = 0$.

We now apply Proposition 17 to Eq. (36). Suppose that K obtains the maximization in Eq. (36). Then, since K is a function of (J_{XY}, W), we have that

$$\begin{aligned} H(K|ZW) &\leqslant H(J_{XY}W|ZW) \\ &= H(J_{XY}|ZW) \leqslant H(J_{XY}|Z). \end{aligned} \tag{37}$$

This proves the desired upper bound under no local randomness.

Lastly, we consider the case when Alice and Bob obtain their final variables K and K' using stochastic processing. This can be modeled by introducing local random variables Q_A and Q_B for Alice and Bob, respectively. Then any stochastic processing depending on X and Y can be obtained by deterministic functions depending on $\hat{X} := (X, Q_A)$ and $\hat{Y} := (Y, Q_B)$. Repeating the above argument shows that $R \leqslant H(J_{\hat{X}\hat{Y}}|Z)$. It is straightforward to show that with Q_A and Q_B pairwise independent and independent of XY, we have $J_{\hat{X}\hat{Y}} = J_{XY}$.

Remark 7: The no-communication results discussed above are already implicit in the work of Csiszár and Narayan. In [41], they study various key distillation scenarios with Eve functioning as a helper and limited communication between Alice and Bob. Included in this is the no-communication scenario with and without helper. However, being very general in nature, Csiszár and Narayan's results involve optimizations over auxiliary random variables, and it is therefore still a nontrivial matter to discern Theorem 6 directly from their work. Additionally, they do not consider the scenario of just shared public randomness.

IV. Advantages in Quantum Versus Classical Key Distillation

In this section, we consider secret key distillation beyond classical sources. The main goal is to answer the following questions: who gains the greatest advantage when embedding a given distribution $p(x, y, z)$ into a multi-party quantum system, the honest parties or the adversary? We provide a partial answer to this question, and demonstrate interesting relationships between secret keys and quantum entanglement.

A. No advantages in Incoherent Embeddings

Theorem 8: $K_D(p_{XYZ}) = K_D(\rho_{ccc})$ for any distribution p_{XYZ}.

Proof: The inequality $K_D(\rho_{ccc}) \geqslant K_D(p_{XYZ})$ is immediate from the fact that every classical protocol $\mathcal{P}_c$ is a special type of quantum protocol $\mathcal{P}_q$. Now we turn to the converse $K_D(\rho_{ccc}) \leqslant K_D(p_{XYZ})$. The idea will be to show that every quantum LOPC protocol $\mathcal{P}_q$ distilling secret key can be transformed into a classical protocol $\mathcal{P}_c$ that distills the same amount of key. Suppose that $\frac{1}{2}\left\|\widehat{\sigma}^{ABEC} - \Phi_s^{AB} \otimes \widehat{\sigma}^{EC}\right\|_1 < \epsilon$ with $\widehat{\sigma}^{ABEC} = \mathcal{P}_q(\rho_{ccc}^{\otimes n})$. To perform the following analysis let's fix some notation. First, without loss of generality, let's assume that r is even with Alice (resp. Bob) measuring in all the odd-numbered (resp. even-numbered) rounds. We let $i_{\leqslant k}$ denote a particular sequence of the first k rounds with $i_{<k} := i_{\leqslant k-1}$. If we wish to refer to a specific outcome in the k^{th} round, we will denote this by i_k. Hence $i_{\leqslant k} = (i_1, i_2, \cdots, i_k)$ for some particular sequence. Finally, if, say, Alice is the measuring party in the k^{th} round, we denote her local instrument conditioned on outcome $i_{<k}$ by $(\mathcal{A}_{i_k}^{(i<k)})_{i_k}$. If we wish to speak of the full composition of Alice's CP maps corresponding to the outcome sequence $i_{\leqslant k}$, we will denote this simply by $\mathcal{A}^{(i_{\leqslant k})}$, with no subscript. That is (for odd-numbered k) we have

$$\mathcal{A}^{(i_{\leqslant k})} = \mathcal{A}_{i_k}^{(i_{<k})} \circ \mathcal{A}^{(i_{\leqslant k-2})} = \mathcal{A}_{i_k}^{(i_{<k})} \circ \mathcal{A}_{i_{k-2}}^{(i_{<k-2})} \circ \cdots \circ \mathcal{A}_{i_1},$$

and similarly Bob's action is described by

$$\mathcal{B}^{(i_{\leqslant k-1})}=\mathcal{B}_{i_{k-1}}^{(i_{<k-1})}\circ\mathcal{B}^{(i_{\leqslant k-3})}=\mathcal{B}_{i_{k-1}}^{(i_{<k-1})}\circ\mathcal{B}_{i_{k-3}}^{(i_{<k-3})}\circ\cdots\circ\mathcal{B}_{i_2}^{(i_{<2})}.$$

When Bob performs the measurement $\mathcal{B}_{i_r}^{(i<r)}$ in the final round, he announces his outcome and Alice is allowed to implement on her system one final trace-preserving map $\overline{\mathcal{A}}_{i_r}^{(i_{\leqslant r})}$. Thus, for one particular r-round outcome sequence i_r, Alice's total CP map is $\mathcal{A}^{(i_{\leqslant r})}=\overline{\mathcal{A}}_{i_r}^{(i_{\leqslant r})}\circ\mathcal{A}^{(i_{\leqslant r-1})}$ and Bob's total CP map is $\mathcal{B}^{(i_{\leqslant r})}=\mathcal{B}_{i_r}^{(i_{<r})}\circ\mathcal{B}^{(i_{\leqslant r-2})}$.

With the notation in hand, when performing protocol $\mathcal{P}_q$ on $\rho_{ccc}^{\otimes n}$, we can describe the overall state generated across all outcome branches by

$$\mathcal{P}_q(\rho_{ccc}^{\otimes n})=\sum_{\mathbf{x},\mathbf{y},\mathbf{z}}\sum_{i_{\leqslant r}}p^n(\mathbf{x},\mathbf{y},\mathbf{z})\\ \mathcal{A}^{(i_{\leqslant r-1})}\otimes\mathcal{B}^{(i_{\leqslant r})}(|\mathbf{xy}\rangle\langle\mathbf{xy}|)\otimes|\mathbf{z}\rangle\langle\mathbf{z}|^E\otimes|i_{\leqslant r}\rangle\langle i_{\leqslant r}|^C, \quad (38)$$

where the first sum is over $\mathcal{X}^n\times\mathcal{Y}^n\times\mathcal{Z}^n$ with $p^n(\mathbf{x},\mathbf{y},\mathbf{z})$ being the n-fold product distribution of p_{XYZ}, and the second sum is over all possible measurement sequences. If Alice and Bob dephase $\mathcal{P}_q(\rho_{ccc}^{\otimes n})$ in the computational basis, the resulting state will be at least ϵ-close to $\Phi_s^{AB}\otimes\widehat{\sigma}^{EC}$ by monotonicity of the trace norm. Hence it suffices to show that this dephased state $\Delta\left(\mathcal{P}_q(\rho_{ccc})\right)$ can be generated using classical LOPC. To see that this is possible, we repeatedly use the fact that the messages are generated locally from the expansion

$$\Delta\left(\mathcal{P}_q(\rho_{ccc}^{\otimes n})\right)=\sum_{\mathbf{x}',\mathbf{y}'}\sum_{\mathbf{x},\mathbf{y},\mathbf{z}}\sum_{i_{\leqslant r}}\Pr[\mathbf{x}',\mathbf{y}'|i_{\leqslant r},\mathbf{x},\mathbf{y},\mathbf{z}]p^n(\mathbf{x},\mathbf{y},\mathbf{z})\\ \cdot|\mathbf{x}'\mathbf{y}'\rangle\langle\mathbf{x}'\mathbf{y}'|^{AB}\otimes|\mathbf{z}\rangle\langle\mathbf{z}|^E\otimes|i_{\leqslant r}\rangle\langle i_{\leqslant r}|^C \quad (39)$$

where

$$\Pr[\mathbf{x}',\mathbf{y}'|i_{\leqslant r},\mathbf{x},\mathbf{y},\mathbf{z}]\\ =\frac{\langle\mathbf{x}'|\mathcal{A}^{(i_{\leqslant r-1})}(|\mathbf{x}\rangle\langle\mathbf{x}|)|\mathbf{x}'\rangle\cdot\langle\mathbf{y}'|\mathcal{B}^{(i_{\leqslant r})}(|\mathbf{y}\rangle\langle\mathbf{y}|)|\mathbf{y}'\rangle}{\Pr[i_{\leqslant r}|\mathbf{x},\mathbf{y},\mathbf{z}]} \quad (40)$$

and

$$\begin{aligned}&\Pr[i_{\leqslant r}|\mathbf{x},\mathbf{y},\mathbf{z}]\\ &=\mathrm{Tr}[\mathcal{A}^{(i_{\leqslant r})}\otimes\mathcal{B}^{(i_{\leqslant r})}(|\mathbf{xy}\rangle\langle\mathbf{xy}|)]\\ &=\mathrm{Tr}[\mathcal{A}^{(i_{\leqslant r-1})}\otimes\mathcal{B}^{(i_{\leqslant r})}(|\mathbf{xy}\rangle\langle\mathbf{xy}|)]\\ &=\prod_{\text{even }k}^{r}\frac{\mathrm{Tr}[\mathcal{A}^{(i_{\leqslant k-1})}\otimes\mathcal{B}^{(i_{\leqslant k})}(|\mathbf{xy}\rangle\langle\mathbf{xy}|)]}{\mathrm{Tr}[\mathcal{A}^{(i_{\leqslant k-1})}\otimes\mathcal{B}^{(i_{\leqslant k-2})}(|\mathbf{xy}\rangle\langle\mathbf{xy}|]}\\ &\quad\times\prod_{\text{odd }k}^{r}\frac{\mathrm{Tr}[\mathcal{A}^{(i_{\leqslant k})}\otimes\mathcal{B}^{(i_{\leqslant k-1})}(|\mathbf{xy}\rangle\langle\mathbf{xy}|)]}{\mathrm{Tr}[\mathcal{A}^{(i_{\leqslant k-2})}\otimes\mathcal{B}^{(i_{\leqslant k-1})}(|\mathbf{xy}\rangle\langle\mathbf{xy}|]}\\ &=\prod_{\text{even }k}^{r}\frac{\mathrm{Tr}[\mathcal{B}^{(i_{\leqslant k})}(|\mathbf{y}\rangle\langle\mathbf{y}|)]}{\mathrm{Tr}[\mathcal{B}^{(i_{\leqslant k-2})}(|\mathbf{y}\rangle\langle\mathbf{y}|]}\times\prod_{\text{odd }k}^{r}\frac{\mathrm{Tr}[\mathcal{A}^{(i_{\leqslant k})}(|\mathbf{x}\rangle\langle\mathbf{x}|)]}{\mathrm{Tr}[\mathcal{A}^{(i_{\leqslant k-2})}(|\mathbf{x}\rangle\langle\mathbf{x}|]}. \quad (41)\end{aligned}$$

Thus a classical protocol $\mathcal{P}_c$ generating $\Delta(\mathcal{P}_q(\rho_{ccc}^{\otimes n}))$ is the following:

1) In the first round, Alice measures her variable $\mathbf{x}$ and broadcasts message i_1 with probability $\Pr[i_1|\mathbf{x}]=\mathrm{Tr}[\mathcal{A}_{i_1}(|\mathbf{x}\rangle\langle\mathbf{x}|)]$.
2) In every subsequent even-numbered (resp. odd-numbered) round k, Bob (resp. Alice) consults the message history $i_{<k}$ and broadcasts i_k with probability

$$\Pr[i_k|i_{<k},\mathbf{y}]=\frac{\mathrm{Tr}[\mathcal{B}_{i_k}^{(i_{<k})}\circ\mathcal{B}^{(i_{\leqslant k-2})}(|\mathbf{y}\rangle\langle\mathbf{y}|)]}{\mathrm{Tr}[\mathcal{B}^{(i_{\leqslant k-2})}(|\mathbf{y}\rangle\langle\mathbf{y}|)]} \quad (42)$$

$$\left(\text{resp.}\quad\Pr[i_k|i_{<k},\mathbf{x}]=\frac{\mathrm{Tr}[\mathcal{A}_{i_k}^{(i_{<k})}\circ\mathcal{A}^{(i_{\leqslant k-2})}(|\mathbf{x}\rangle\langle\mathbf{x}|)]}{\mathrm{Tr}[\mathcal{A}^{(i_{\leqslant k-2})}(|\mathbf{x}\rangle\langle\mathbf{x}|)]}\right). \quad (43)$$

3) At the end of r rounds with the total message $i_{\leqslant r}$ having been generated, Alice and Bob process their variables using local channels $\mathbf{x}\to\mathbf{x}'$ and $\mathbf{y}\to\mathbf{y}'$ with transition probabilities given by

$$\Pr[\mathbf{x}'|i_{\leqslant r},\mathbf{x}]=\frac{\langle\mathbf{x}'|\mathcal{A}^{(i_{\leqslant r})}(|\mathbf{x}\rangle\langle\mathbf{x}|)|\mathbf{x}'\rangle}{\mathrm{Tr}[\mathcal{A}^{(i_{\leqslant r})}(|\mathbf{x}\rangle\langle\mathbf{x}|)]}, \quad (44)$$

$$\Pr[\mathbf{y}'|i_{\leqslant r},\mathbf{y}]=\frac{\langle\mathbf{y}'|\mathcal{B}^{(i_{\leqslant r})}(|\mathbf{y}\rangle\langle\mathbf{y}|)|\mathbf{y}'\rangle}{\mathrm{Tr}[\mathcal{B}^{(i_{\leqslant r})}(|\mathbf{y}\rangle\langle\mathbf{y}|)]}. \quad (45)$$

4) It can be seen that the state generated through this process is precisely $\Delta(\mathcal{P}_q(\rho_{ccc}^{\otimes n}))$.

■

B. Arbitrarily Large Advantages in Coherent Embeddings

Theorem 9: For any N, a distribution p_{XYZ} exists such that when embedding p_{XYZ} into a coherent quantum source, one of the following relationships holds:

(a) Eve gains an arbitrarily large advantage: $K_D(p_{XYZ})-K_D(\Psi_{qqq}^{ABE})>N$, or
(b) Alice and Bob gain an arbitrarily large advantage: $K_D(\Psi_{qqq}^{ABE})-K_D(p_{XYZ})>N$.

Remark 10: Since $K_D(\Psi_{qqq}^{ABE})\leqslant K_D(\rho_{qqc}^{ABE})$, item (b) implies that $K_D(\rho_{qqc}^{ABE})-K_D(p_{XYZ})>N$. In this case, the gain in Alice and Bob's key rate comes exclusively from the fact that the bipartite conditional distributions $p(x,y|z)$ are coherently embedded.

Remark 11: In the proof of (a) we actually demonstrate a much stronger result that $K_D(p_{XYZ})-E_F(\rho^{AB})>N$. This means that we can distill key from p_{XYZ} of a considerably higher rate than the rate of entanglement needed to generate the corresponding quantum state ρ^{AB}. To our knowledge, this is the first known result of its kind.

Proof: (a) We consider a very simple binary distribution $p(x,y,z)$ whose nonzero probabilities are given by $p(0,0,0)=p(1,1,0)=1/4$, $p(0,0,1)=\lambda/2$, and $p(1,1,1)=(1-\lambda)/2$. This is a UBI distribution since for each $z\in\{0,1\}$, the distribution $p(x,y|z)$ is diagonal and $J_{XY|z}$ is binary. Hence, one can verify the distribution is BI since $I(X:Y|J_{XY|Z=z}Z=z)=0$ for $z=\{0,1\}$. Furthermore, the condition $H(J_{XY|Z}|X)=H(J_{XY|Z}|Y)=0$ is satisfied whenever the conditional distributions $\{p(x,y|z)\}_{z\in\{0,1\}}$ have the same block diagonal structure for each z. Then by [4, Theorem 2], its key rate is precisely $K_D(p_{XYZ})=[1+h(\lambda)]/2$, where $h(x)=-x\log x-(1-x)\log(1-x)$. The corresponding qqq embedding has the form

$$|\Psi_{qqq}^{ABE}\rangle=\sqrt{1/2}[|\Phi\rangle^{AB}|0\rangle^E+(\sqrt{\lambda}|00\rangle+\sqrt{1-\lambda}|11\rangle)^{AB}|1\rangle^E].$$

Since ρ^{AB} is a two-qubit state, its entanglement of formation can be calculated using the celebrated concurrence formula [42] (see also [4]), and it is found to be

$$E_F(\rho^{AB}) = h\left([1+\sqrt{1-\left(1/2+\sqrt{\lambda(1-\lambda)}\right)^2})]/2\right). \quad (46)$$

This formula can be computed by noting that $\rho^{AB} = \mathrm{Tr}_E(|\psi_{qqq}\rangle\langle\psi_{qqq}|^{ABE})$ is essentially a single qubit density matrix ω as ρ^{AB} has support on the two-dimensional subspace $\{|00\rangle, |11\rangle\}$. The concurrence $C(\rho^{AB})$ is then given by $\sqrt{\lambda_{\max}}-\sqrt{\lambda_{\min}}$, were the λ_i are the eigenvalues of $\omega\sigma_y\omega^*\sigma_y = \omega\sigma_x\omega\sigma_x$ [42]. Direct calculation gives

$$\sqrt{\lambda_{\max}} - \sqrt{\lambda_{\min}} = 2\sum_z p(z)\sqrt{p(0,0|z)p(1,1|z)};$$

this establishes Eq. (46) after using the formula $E_F(\rho^{AB}) = h\left(1+\sqrt{1-C(\rho)^2}\right)$. A simple convexity argument then shows that $K_D(p_{XYZ}) > E_F(\rho^{AB})$ whenever $0 < \lambda < 1/2$. We now consider n copies of p_{XYZ}. Inspection reveals that $p_{XYZ}^{\otimes n}$ is also UBI for any n. Thus, $K_D(p_{XYZ}) = n[1 + h(\lambda)]/2$. On the other hand, the entanglement of formation is a sub-additive quantity such that $E_F\left((\rho^{AB})^{\otimes n}\right) \leqslant nE_F(\rho^{AB})$. Consequently, for any $0 < \lambda < 1/2$ we attain an arbitrarily large gap between $K_D(p_{XYZ}^n)$ and $E_F\left((\rho^{AB})^{\otimes n}\right)$. By Theorem 4 and the fact that $E_F(\rho^{AB}) \geqslant E_{sq}(\rho^{AB})$, this gap will be at least as large as the gap between $K_D(p_{XYZ}^n)$ and $K_D\left((\Psi_{qqq}^{ABE})^{\otimes n}\right)$.

(b) Consider the state $|\Psi\rangle^{ABE} = \sqrt{1/2}(|00\rangle + |1+\rangle)^{AB}|0\rangle^E$ where Eve is initially uncorrelated. This is a qqq encoding of a distribution $p_{XYZ} = p_{XY}\cdot q_Z$ whose mutual information is $I(X:Y) = 1 - h(1/3) \approx .311$. Since Eve has no side information, the classical secret key rate $K_D(p_{XYZ})$ is equal to the mutual information, a well-known result in secret key agreement [1], [2]. On the entanglement side, the reduced-state entropy characterizes the entanglement distillation rate for pure states [43]; hence $S(\rho^B) = E_D(\Psi^{ABE})$. One bit of entanglement can be converted into one bit of secret key, and thus $K_D(\Psi^{ABE}) \geqslant S(\rho_B)$. In fact, this inequality is tight since $S(\rho^B) = E_{sq}(\Psi^{AB}) \geqslant K_D(\Psi^{ABE})$. Because both the mutual information and von Neumann entropy are additive, a similar argument to part (a) shows that the gap between $S((\rho^B)^{\otimes n}) - K_D(p_{XYZ}^n)$ can be made arbitrarily large. ∎

C. Embedding Distributions With Reversible Secrecy

We now consider qqq embeddings of distributions with reversible secrecy, for which it turns out that the quantum embedding favors Eve over Alice and Bob. When adding UBI-PD and/or semi-unambiguous structure, relationships between key and quantum entanglement can be drawn.

Theorem 12:

(a) If p_{XYZ} has reversible secrecy (i.e. $K_C(p_{XYZ}) = K_D(p_{XYZ})$), then

$$K_D(p_{XYZ}) \geqslant E_{sq}(\rho^{AB}). \quad (47)$$

(b) If p_{XYZ} is UBI-PD (and hence reversible), then

$$K_D(p_{XYZ}) \geqslant E_F(\rho^{AB}). \quad (48)$$

(c) If p_{XYZ} has reversible secrecy and is semi-unambiguous, then

$$K_C(p_{XYZ}) = E_{sq}(\rho^{AB}) = K_D(\Psi^{ABE}) = K_D(p_{XYZ}) \quad (49)$$

(d) If p_{XYZ} is UBI-PD and semi-unambiguous, then

$$\begin{aligned} K_C(p_{XYZ}) &= K_D(p_{XYZ}) = K_D(\Psi^{ABE}) = E_F(\rho^{AB}) \\ &= E_C(\rho^{AB}) = E_D(\rho^{AB}) = E_r(\rho^{AB}) \\ &= E_{sq}(\rho^{AB}) = H(J_{XY|Z}|Z). \end{aligned}$$

Proof:

(a) By Theorem 5, if p_{XYZ} is reversible then there must be a channel $\overline{Z}|Z$ such that $p_{XY\overline{Z}}$ is block independent. In other words, there exists a decomposition

$$p_{XY|\overline{Z}}(x,y|\overline{z}) = \sum_{j_{\overline{z}}} p_X(x|j_{\overline{z}},\overline{z})p_Y(y|j_{\overline{z}},\overline{z})p(j_{\overline{z}}|\overline{z}) \quad (50)$$

where $p_X(\cdot|j_{\overline{z}},\overline{z})$ and $p_X(\cdot|j'_{\overline{z}},\overline{z})$ are disjoint distributions for $j_{\overline{z}} \neq j'_{\overline{z}}$, and likewise for Bob's conditional distributions. For each $\overline{z}$, define the local measurement channel acting on Alice's system

$$\omega \mapsto \Omega_A^{(\overline{z})}(\omega) := \sum_{j_{\overline{z}}} \sum_{\substack{x \text{ such that} \\ p(x|j_{\overline{z}},\overline{z})>0}} \langle x|\omega|x\rangle |j_{\overline{z}}\rangle\langle j_{\overline{z}}|. \quad (51)$$

Let $\Omega_B^{(\overline{z})}$ be defined similarly for Bob's system. These operations can be viewed as a measurement channel that creates a state that is a convex combination of common partitions of $p_{X,Y,\bar{Z}}$.

We next consider the decomposition $|\Psi_{ABE}\rangle = \sum_z \sqrt{p(z)}|\varphi_z\rangle|z\rangle$, in which $\langle xy|\varphi_z\rangle = \sqrt{p(x,y|z)}$. Note that $\rho^{AB} = \mathrm{Tr}_{\overline{Z}}\sigma^{AB\overline{Z}}$ where

$$\begin{aligned} \sigma^{AB\overline{Z}} &:= \sum_{\overline{z}}\sum_z p(z|\overline{z})p(\overline{z})|\varphi_z\rangle\langle\varphi_z| \otimes |\overline{z}\rangle\langle\overline{z}| \\ &= \sum_{\overline{z}} p(\overline{z})\sigma_{(\overline{z})}^{AB} \otimes |\overline{z}\rangle\langle\overline{z}|. \end{aligned} \quad (52)$$

On the state $\sigma^{AB\overline{Z}}$, we first dephase in the computational basis, and then apply $\Omega_A^{(\overline{z})} \otimes \Omega_B^{(\overline{z})}$ conditioned on $\overline{z}$. Doing so generates the state

$$\begin{aligned} &\widehat{\sigma}^{AB\overline{Z}} \\ &:= \sum_{\overline{z}} p(\overline{z}) \sum_{x,y} p(x,y|\overline{z})\Omega_A^{(\overline{z})}(|x\rangle\langle x|) \otimes \Omega_B^{(\overline{z})}(|y\rangle\langle y|) \otimes |\overline{z}\rangle\langle\overline{z}| \\ &= \sum_{\overline{z}} p(\overline{z}) \sum_{j_{\overline{z}}} p(j_{\overline{z}}|\overline{z})|j_{\overline{z}}j_{\overline{z}}\overline{z}\rangle\langle j_{\overline{z}}j_{\overline{z}}\overline{z}|. \end{aligned} \quad (53)$$

Hence,

$$\begin{aligned} E_{sq}(\rho_{AB}) &\leqslant \tfrac{1}{2}\sum_{\overline{z}} p(\overline{z})I(A:B)_{\sigma_{(\overline{z})}^{AB}} \\ &\leqslant \tfrac{1}{2}\sum_{\overline{z}} p(\overline{z})[S(\sigma_{(\overline{z})}^A) + S(\sigma_{(\overline{z})}^B)] \\ &\leqslant \tfrac{1}{2}\sum_{\overline{z}} p(\overline{z})[S(\widehat{\sigma}_{(\overline{z})}^A) + S(\widehat{\sigma}_{(\overline{z})}^B)] \\ &= \sum_{\overline{z}} p(\overline{z})H(J_{\overline{Z}}|\overline{Z}=\overline{z}) \\ &= K_D(p_{XYZ}). \end{aligned} \quad (54)$$

(b) If p_{XYZ} is UBI-PD then again by Theorem 5, we have $K_C(p_{XYZ}) = K_D(p_{XYZ}) = H(J_{XY|Z}|Z)$. Eq. (52) still holds in this case with $\overline{Z} = Z$, and σ_{ABZ} is obtained from $|\Psi_{ABE}\rangle$ by Eve dephasing in the computational basis. Block-independence of p_{XYZ} means that $S(\mathrm{Tr}_A\,|\varphi_z\rangle\langle\varphi_z|) = H(J_{XY|Z}|Z = z)$. Since $\{p(z), |\varphi_z\rangle\}$ provides a pure-state ensemble realizing ρ^{AB}, we see that $H(J_{XY|Z}|Z) \geqslant E_F(\rho^{AB})$.

(c) From Theorem 4 and [14], semi-unambiguous distributions are shown to satisfy the inequality $E_{sq}(\rho^{AB}) \geqslant K_D(\Psi^{ABE}_{qqq}) \geqslant K_D(p_{XYZ})$. Combining with part (a) gives the desired result.

(d) This follows from combining (b), (c) and Theorem 4 with the fact that $E_F(\rho^{AB}) \geqslant E_C(\rho^{AB}) \geqslant \max\{E_r(\rho^{AB}), E_{sq}(\rho^{AB})\}$. ∎

Remark 13: Theorem 12 (a) implies that Alice and Bob never gain an advantage over Eve when embedding a reversible distribution into a quantum source. In fact, it is not difficult to construct distributions in which Eve gains a non-zero advantage in the quantum setting. This can be seen by the chain of inequalities in Eq. (54). In particular, whenever $\sigma^{AB}_{\overline{z}}$ is not pure the inequality will be strict. This will hold, for instance, for any distribution p_{XYZ} with a non-trivial channel $\overline{Z}|Z$ such that $p_{XY\overline{Z}}$ is UBI-PD.

Remark 14: Theorem 12 (d) provides an alternative proof of the well-known fact that pure entangled state possess reversible entanglement [43]; i.e. that $E_C(|\psi\rangle\langle\psi|^{AB}) = E_D(|\psi\rangle\langle\psi|^{AB})$ for any $|\psi\rangle^{AB}$. This follows from the fact that any entangled state is a qqq embedding of a UBI distribution, the distribution $p(x, y)$ given by its Schmidt decomposition $|\psi\rangle^{AB} = \sum_{i=1}^{r} \sqrt{p(x,x)}|x, x\rangle^{AB}$ with an uncorrelated Eve. In fact, consider any mixed state of the form $\rho = \sum_{i=1}^{r} p_i|\psi_i\rangle\langle\psi_i|^{AB}$, in which $[\rho_i^A, \rho_j^A] = [\rho_i^B, \rho_j^B] = 0$ for all $i \neq j$, and either $\rho_i^A \perp \rho_j^A$ or $\rho_i^B \perp \rho_j^B$ for every $i \neq j$. Here we are denoting $\rho_i^A = \mathrm{Tr}_B\,|\psi_i\rangle\langle\psi_i|^{AB}$ and likewise for ρ_i^B. Such states possess reversible entanglement since they are the AB reduced state of a qqq embedding $|\psi\rangle^{ABE}$ that is generated by a UBI-PD and semi-ambiguous distribution. Indeed, the simultaneous commuting of the reduced states means that a common Schmidt basis can be found for all the $|\psi_i\rangle^{AB}$, which we take as the computational basis. Then by performing suitable local projections, the local orthogonality of the $|\psi_i\rangle$ enables Alice and Bob to non-destructively identify which $|\psi_i\rangle$ they have among the mixture. On the level of embedded distributions, this corresponds to a semi-unambiguous distribution since $H(Z|XY) = 0$. After Alice and Bob share the outcomes of their local projections, they hold one of the pure states $|\psi_i\rangle$ which is the embedding of a UBI distribution. Hence the original purification state $|\psi\rangle = \sum_{i=}^{r} \sqrt{p_i}|\psi_i\rangle^{AB}|i\rangle^E$ is the embedding of a semi-unambiguous and UBI-PD distribution. States having this form belong to a class of states known as LOCC-flagged, and it is conjectured that the class of LOCC-flagged distributions constitute the entire family of bipartite quantum states having reversible entanglement [4], [44], [45]. Theorem 12 (d) offers one direction for testing this conjecture. Namely one needs to consider whether the embedding of a UBI-PD distribution always leads to an LOCC-flagged state. As the structure of UBI-PD distributions can be quite complex, we leave this question for future research.

V. Conclusion

In this paper, we have studied the Gács-Körner common information and two ways that it can be extended into conditional form. Both types are useful tools in the resource theory of secret correlations. We have shown that the coarse-grained conditional common information $H(J_{XY}|Z)$ quantifies precisely the optimal rate of key extraction when Alice and Bob hold public shared randomness. In fact, achievability of this rate does not require use of shared randomness, a fact that cannot be taken for granted since in general, randomness can serve as a resource in distillation tasks [1], [15].

We have also considered the task of resource distillation in the quantum setting. Since secret key can be distilled from both quantum and classical states, a direct comparison can be made between the two scenarios. Quantum states that are diagonal in some fixed basis - such as ρ_{ccc} - lack coherence and are typically referred to as "classical" states since they possess the same entropic properties as classical probability distributions. However, as quantum objects, these states can undergo certain physical transformations that are impossible for classical states. We have shown that despite this enhanced dynamical ability, secret key distillation is equivalent for a classical distribution and its incoherent quantum embedding. The situation is much different when the embedding takes the form of a coherent superposition. We have presented examples demonstrating that quantum and classical key rates can be vastly different; sometimes it benefits Alice and Bob to have their correlations embedded in a quantum state, sometimes it benefits Eve. We have linked this investigation of quantum advantages to the problem of LOPC secrecy reversibility. By introducing different families of distributions that demonstrate secrecy reversibility, we are able to bound and in some cases explicitly compute the entanglement and distillable key of the embedded quantum states. In the latter case, the conditional common information $H(J_{XY|Z}|Z)$ emerges as the optimal rate of resource distillation. It is quite beautiful that notoriously difficult entanglement measures can be computed using exclusively a classical analysis of the states's underlying probability distribution. We hope this paper helps to advance our understanding of the relationship between classical secrecy and quantum entanglement.

Appendix
Technical Lemmas Used in Theorem 6

Lemma 15 ([16, Corollary 17.5]): Given is an i.i.d. source of two random variables J_{XY} and Z with J_{XY} ranging over set $\mathcal{J}$. Then, for any $\epsilon > 0$ and for all n sufficiently large with $\frac{1}{n}\log|\mathcal{K}| > H(J_{XY}|Z) - \epsilon$, there exists a constant $\gamma > 0$ and a mapping $\kappa : \mathcal{J}^n \to \mathcal{K} = \{1, 2, \cdots, |\mathcal{K}|\}$ such that

$$\log|\mathcal{K}| - H(\kappa(J^n_{XY})|Z^n) < 2^{-n\gamma}.$$

Proof: Let $K = \kappa(J_{XY}^n)$. Note that we can write

$$D(p_{KZ^n} \| \mathbb{1}_{\mathcal{K}} \cdot p_{Z^n}) = \log|\mathcal{K}| - H(K|Z^n),$$

where where $\mathbb{1}_{\mathcal{K}}$ is the uniform distribution over $\mathcal{K}$ and $D(p_X\|p_Y) = \sum_x p_X(x) \log \frac{p_X(x)}{p_Y(x)}$ is the relative entropy of distributions p_X and p_Y. Pinsker's Inequality then gives

$$\tfrac{1}{2}\|p_{KZ^n} - \mathbb{1}_{\mathcal{K}} \cdot p_{Z^n}\| \leqslant \sqrt{D(p_{KZ^n}\|\mathbb{1}_{\mathcal{K}} \cdot p_{Z^n})/2},$$

which implies that $H(J_{XY}|Z)$ is an achievable key rate. ∎

Lemma 16: Let J be uniformly distributed over the set $\{1, \cdots, n\}$ and let $A^{(i)}$ denote the i^{th} instance of A in A^n. Likewise, let $A^{(<i)} = A^{(1)} \cdots A^{(i-1)}$ and $A^{(>i)} = A^{(i+1)} \cdots A^{(n)}$ with $A^{(<1)} := \emptyset$ and $A^{(n+1)} := \emptyset$. Then for random variables P and Q and sequences of random variables A^n, B^n

$$H(P|A^nQ) - H(P|B^nQ) = n[I(P : B^{(J)}|TQ) - I(P : A^{(J)}|TQ)], \quad (55)$$

where $T = JA^{(>J)}B^{(<J)}$

Proof: See [16, Lemma 17.2] ∎

Proposition 17: If W is independent of XY and $H(K|XW) = H(K|YW) = 0$, then K is a function of (J_{XY}, W).

Proof: The fact that $H(K|XW) = H(K|YW) = 0$ implies the existence of two functions $f(X, W)$ and $g(Y, W)$ such that $\Pr[f(X, W) = g(Y, W)] = 1$. Consequently, if $p(x_1, y_1)p(x_1, y_2) > 0$, then $f(x_1, w) = g(y_1, w) = g(y_2, w)$ for all $w \in \mathcal{W}$ with $p(w) > 0$. Indeed, if, say, $f(x_1, w) \neq g(y_1, w)$, then $\Pr[f(X, W) \neq g(Y, W)] \geqslant p(x_1, y_1, w) = p(x_1, y_1)p(w) > 0$, where we have used the independence between XY and W. By the same reasoning, $p(x_1, y_1)p(y_1, x_2) > 0$ implies that $f(x_1, w) = f(x_2, w) = g(y_1, w)$ for all $w \in \mathcal{W}$. Turning to Proposition 2, if $J_{XY}(x) = J_{XY}(x')$, then there exists a sequence $xy_1x_1y_2x_2 \cdots y_nx'$ such that $p(xy_1)p(y_1x_1)p(x_1y_2) \cdots p(y_nx') > 0$. Therefore, as just argued, we must have that $f(x, w) = f(x', w)$ for all $w \in \mathcal{W}$. Hence K must be a function of (J_{XY}, W). ∎

Proposition 18 (Conditional Double Markov Chains ([16, Lemma 16.25])): Random variables $WXYZ$ satisfy the two Markov chains $X - YZ - W$ and $Y - XZ - W$ iff $I(XY : W|J_{XY|Z}Z) = 0$.

REFERENCES

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

[2] 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.

[3] P. Gács and J. Körner, "Common information is far less than mutual information," *Problems Control Inf. Theory*, vol. 2, no. 2, pp. 149–162, 1973.

[4] E. Chitambar, B. Fortescue, and M.-H. Hsieh, "Classical analog to entanglement reversibility," *Phys. Rev. Lett.*, vol. 115, p. 090501, Aug. 2015.

[5] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proc. IEEE Int. Conf. Comput., Syst., Signal Process.*, Dec. 1984, pp. 175–179.

[6] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter, "Everything you always wanted to know about LOCC (but were afraid to ask)," *Commun. Math. Phys.*, vol. 328, no. 1, pp. 303–326, 2014.

[7] E. Chitambar, M. H. Hsieh, and R. Duan, "When do local operations and classical communication suffice for two-qubit state discrimination?" *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1549–1561, Mar. 2014.

[8] E. Chitambar and M.-H. Hsieh, "Asymptotic state discrimination and a strict hierarchy in distinguishability norms," *J. Math. Phys.*, vol. 55, no. 11, p. 112204, 2014.

[9] E. Chitambar and M.-H. Hsieh, "Relating the resource theories of entanglement and quantum coherence," *Phys. Rev. Lett.*, vol. 117, p. 020402, Jul. 2016.

[10] E. Chitambar, B. Fortescue, and M.-H. Hsieh, "Distributions attaining secret key at a rate of the conditional mutual information," in *Proc. 35th Annu. Cryptol. Conf.*, Santa Barbara, CA, USA, Aug. 2015, pp. 443–462, doi: 10.1007/978-3-662-48000-7_22.

[11] A. D. Wyner, "The common information of two dependent random variables," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 2, pp. 163–179, Mar. 1975.

[12] A. Winter, "Secret, public and quantum correlation cost of triples of random variables," in *Proc. Int. Symp. Inf. Theory (ISIT)*, Sep. 2005, pp. 2270–2274.

[13] E. Chitambar, M.-H. Hsieh, and A. Winter, "The private and public correlation cost of three random variables with collaboration," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 2034–2043, Apr. 2016.

[14] M. Christandl, A. Ekert, M. Horodecki, P. Horodecki, J. Oppenheim, and R. Renner, "Unifying classical and quantum key distillation," in *Theory Cryptography* (Lecture Notes in Computer Science), vol. 4392, S. Vadhan, Ed. Berlin, Germany: Springer, 2007, pp. 456–478.

[15] M. Ozols, G. Smith, and J. A. Smolin, "Bound entangled states with a private key and their classical counterpart," *Phys. Rev. Lett.*, vol. 112, p. 110502, Mar. 2014.

[16] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[17] E. B. Davies and J. T. Lewis, "An operational approach to quantum probability," *Commun. Math. Phys.*, vol. 17, no. 3, pp. 239–260, 1970.

[18] M. Kleinmann, H. Kampermann, and D. Bruß, "Asymptotically perfect discrimination in the local-operation-and-classical-communication paradigm," *Phys. Rev. A, Gen. Phys.*, vol. 84, no. 4, p. 042326, Oct. 2011.

[19] I. Devetak and A. Winter, "Distillation of secret key and entanglement from quantum states," *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 461, pp. 207–235, Jan. 2005.

[20] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Rev. Mod. Phys.*, vol. 81, p. 865, Jun. 2009.

[21] D. Collins and S. Popescu, "Classical analog of entanglement," *Phys. Rev. A, Gen. Phys.*, vol. 65, no. 3, p. 032321, 2002.

[22] N. Gisin, R. Renner, and S. Wolf, "Linking classical and quantum key agreement: Is there a classical analog to bound entanglement?" *Algorithmica*, vol. 34, no. 4, pp. 389–412, 2002.

[23] A. Acín, L. Masanes, and N. Gisin, "Equivalence between two-qubit entanglement and secure key distribution," *Phys. Rev. Lett.*, vol. 91, p. 167901, Oct. 2003.

[24] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, "Information theories with adversaries, intrinsic information, and entanglement," *Found. Phys.*, vol. 35, no. 12, pp. 2027–2040, 2005.

[25] A. Acín and N. Gisin, "Quantum correlations and secret bits," *Phys. Rev. Lett.*, vol. 94, p. 020501, Jan. 2005.

[26] J. Oppenheim, R. W. Spekkens, and A. Winter. (2008). "A classical analogue of negative information." [Online]. Available: https://arxiv.org/abs/quant-ph/0511247

[27] J. Bae, T. Cubitt, and A. Acín, "Nonsecret correlations can be used to distribute secrecy," *Phys. Rev. A, Gen. Phys.*, vol. 79, no. 3, p. 032304, 2009.

[28] E. Chitambar and M.-H. Hsieh, "Round complexity in the local transformations of quantum and classical states," *Nature Commun.*, vol. 8, Dec. 2017, Art. no. 2086.

[29] E. M. Rains, "Rigorous treatment of distillable entanglement," *Phys. Rev. A, Gen. Phys.*, vol. 60, pp. 173–178, Jul. 1999.

[30] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, "General paradigm for distilling classical key from quantum states," *IEEE Trans. Inf. Theory*, vol. 55, no. 4, pp. 1898–1929, Apr. 2009.

[31] V. Vedral and M. B. Plenio, "Entanglement measures and purification procedures," *Phys. Rev. A, Gen. Phys.*, vol. 57, no. 3, pp. 1619–1633, Mar. 1998.

[32] M. Christandl and A. Winter, "'Squashed entanglement': An additive entanglement measure," *J. Math. Phys.*, vol. 45, no. 3, pp. 829–840, 2004.
[33] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, "Mixed-state entanglement and quantum error correction," *Phys. Rev. A, Gen. Phys.*, vol. 54, no. 5, pp. 3824–3851, 1996.
[34] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, "Secure key from bound entanglement," *Phys. Rev. Lett.*, vol. 94, no. 16, p. 160502, Apr. 2005.
[35] M. Christandl, "The structure of bipartite quantum states—Insights from group theory and cryptography," M.S. thesis, Selwyn College, Univ. Cambridge, Cambridge, U.K., 2006.
[36] M. M. Wilde, "Squashed entanglement and approximate private states," *Quantum Inf. Process.*, vol. 15, no. 11, pp. 4563–4580, Nov. 2016.
[37] Y. Huang, "Computing quantum discord is NP-complete," *New J. Phys.*, vol. 16, no. 3, p. 033027, 2014.
[38] P. M. Hayden, M. Horodecki, and B. M. Terhal, "The asymptotic entanglement cost of preparing a quantum state," *J. Phys. A, Math. Gen.*, vol. 34, no. 35, p. 6891, 2001.
[39] R. Renner and S. Wolf, "New bounds in secret-key agreement: The gap between formation and secrecy extraction," in *Advances in Cryptology* (Lecture Notes in Computer Science), vol. 2656, E. Biham, Ed. Berlin, Germany: Springer-Verlag, 2003, pp. 562–577.
[40] 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, Nov. 1995.
[41] I. Csiszár and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 344–366, Mar. 2000.
[42] W. K. Wooters, "Entanglement of formation of an arbitrary state of two qubits," *Phys. Rev. Lett.*, vol. 80, pp. 2245–2248, 1998.
[43] C. H. Bennett, H. Bernstein, S. Popescu, and B. Schumacher, "Concentrating partial entanglement by local operations," *Phys. Rev. A, Gen. Phys.*, vol. 53, no. 4, pp. 2046–2052, Apr. 1996.
[44] P. Horodecki, R. Horodecki, and M. Horodecki, "Entanglement and thermodynamical analogies," *Acta Phys. Slovaca*, vol. 48, pp. 141–156, May 1998.
[45] K. G. H. Vollbrecht, R. F. Werner, and M. M. Wolf, "Irreversibility of entanglement distillation for a class of symmetric states," *Phys. Rev. A, Gen. Phys.*, vol. 69, p. 062304, Jun. 2004.

Eric Chitambar biography not available at the time of publication.

Ben Fortescue biography not available at the time of publication.

Min-Hsiu Hsieh (M'09–SM'18) received his PhD degree in electrical engineering from the University of Southern California, Los Angeles, in 2008. From 2008-2010, he was a Researcher at the ERATO-SORST Quantum Computation and Information Project, Japan Science and Technology Agency, Tokyo, Japan. From 2010-2012, he was a Postdoctoral Researcher at the Statistical Laboratory, the Centre for Mathematical Sciences, the University of Cambridge, UK. He is now an Australian Research.