The Private and Public Correlation Cost of Three Random Variables with Collaboration

Eric Chitambar,[1] Min-Hsiu Hsieh,[2] and Andreas Winter[3]

[1] *Department of Physics and Astronomy, Southern Illinois University,*
Carbondale, Illinois 62901, USA

[2] *Centre for Quantum Computation & Intelligent Systems (QCIS),*
Faculty of Engineering and Information Technology (FEIT),
University of Technology Sydney (UTS), NSW 2007, Australia

[3] *ICREA & Física Teòrica: Informació i Fenòmens Quàntics*
Universitat Autònoma de Barcelona, ES-08193 Bellaterra (Barcelona), Spain

(12 September 2014)

Abstract

In this paper we consider the problem of generating arbitrary three-party correlations from a combination of public and secret correlations. Two parties – called Alice and Bob – share perfectly correlated bits that are secret from a collaborating third party, Charlie. At the same time, all three parties have access to a separate source of correlated bits, and their goal is to convert these two resources into multiple copies of some given tripartite distribution P_{XYZ}. We obtain a single-letter characterization of the trade-off between public and private bits that are needed to achieve this task. The rate of private bits is shown to generalize Wyner's classic notion of *common information* held between a pair of random variables. The problem we consider is also closely related to the task of *secrecy formation* in which P_{XYZ} is generated using public communication and local randomness but with Charlie functioning as an adversary instead of a collaborator. We describe in detail the differences between the collaborative and adversarial scenarios.

I. Introduction

Three-party correlations are central objects of interest in the discussion of public key agreement [Mau93], [AC93]. Two parties (Alice and Bob) have access to some source which generates multiple copies of three random variables XYZ. When n copies are generated, Alice sees X^n, Bob sees Y^n, and a third party (Charlie) sees Z^n. In the standard key agreement scenario, Charlie is viewed as untrustworthy eavesdropper and Alice and Bob wish to extract perfectly shared randomness from $X^nY^nZ^n$ using local randomness and public communication (LOPC). The security constraint is that at the end of this protocol, Charlie should be almost completely uncorrelated from Alice and Bob's shared randomness.

However, in many scenarios it may not be appropriate to assume that Charlie is a malicious eavesdropper. In fact, Charlie may actually be a helper or collaborator to Alice and Bob in their pursuit of obtaining private randomness from $X^nY^nZ^n$. For instance, one might imagine that Charlie represents some centralized hub that wishes to establish a secure link between two of its users. Distillation problems of this sort have been studied in Ref. [CN00]; see also Refs. [GW03], [SVW05], [Win07] for quantum analogues.

This paper considers the reverse of the scenario just described. Instead of asking how much secret key can be distilled from XYZ using LOPC, we ask how much secret key is needed to build XYZ using LOPC. This kind of problem has been studied much less, but goes back all the way to Wyner [Wyn75], and has received much more attention only in the last decade or so, from the *Reverse Shannon Theorem* [BSST02] (and its quantum generalization [B+14]), more generally to so-called coordination problems [Cuff08], [Cuff09].

Whether Charlie is an adversary or collaborator greatly changes the nature of the problem, as we shall see. First consider when Charlie is a collaborator. Alice and Bob initially share perfect randomness that is secret from Charlie, and using LOPC, they generate public communication U and variables $\hat{X}^n\hat{Y}^n$. However, in the spirit of collaboration, the public communication which they generate should also be usable by Charlie to generate $\hat{Z}^n$ so that ultimately $\hat{X}^n\hat{Y}^n\hat{Z}^n \approx X^nY^nZ^n$. We will refer to this as the *collaborative model* for generating XYZ. In a particular protocol, there will exist some trade-off between the amount of secret randomness Alice and Bob initially share versus the amount of public communication used to build $\hat{X}^n\hat{Y}^n\hat{Z}^n$. The main contribution of this paper is a single-letter characterization of this trade-off (Theorem 7).

On the other hand, when Charlie is an adversary, some care is needed to properly quantify the correlation costs of XYZ. This is because here Alice and Bob really only care about generating the marginal XY since, after all, Charlie is an adversary. Nevertheless, Z may contain some information about XY, and this should be somehow captured in the total cost for XYZ. In light of these considerations, Renner and Wolf have proposed the following notion of secrecy formation [RW03]. Starting from a source of pre-shared secret bits, Alice and Bob perform LOPC to generate three random variables $\hat{X}^n\hat{Y}^nU$, where again U describes the public communication conducted during the protocol. With $X^nY^nZ^n$ being the target distribution, the goal is for $\hat{X}^n\hat{Y}^nU$ to be approximately equivalent to some joint random variables of the form $X^nY^n\overline{Z}$, where $\overline{Z}$ is obtained by processing Z^n. This latter condition means that Charlie could simulate the entire communication Alice and Bob use to produce $\hat{X}^n\hat{Y}^n$ from his part Z^n. In the words of Renner and Wolf, this "formalizes the fact that the protocol communication U observed by [Charlie] does not give him more information than Z^n." We will refer to this as the *adversarial model* for generating XYZ.

In subsequent work, Horodecki *et al.* [HHHO05] discussed the hypothesis that the minimum rate of secret bits for generating XYZ in the adversarial sense is given by a quantity known as the intrinsic information [RW03]. If this were true, then optimal secrecy formation could alternatively be obtained by an asymptotic preparation of randomly chosen private correlations (i.e. distributions P_{XYZ} in which Charlie is completely uncorrelated from Alice and Bob). However, Horodecki *et al.* considered this hypothesis to be likely false, and indeed it was later shown to be so by one of us, with a single-letter secret key cost formula being derived [Win05]. In the present paper, we revisit the precise trade-off between public and secret correlation costs computed in Ref. [Win05] for the adversarial model, in particular the direct (achievability) part of the main result of that paper. Note that the direct part is also implicitly discussed in [HHHO05], since the formula of the secrecy cost is a convex hull over certain decompositions of the Alice-Bob distribution, decompositions characterized by Wyner's *common information* [Wyn75]. The converse (lower bound) presented in [HHHO05, Prop. 1] however was incomplete as it assumed a property known as "asymptotic continuity" (cf. [HHHH09]) of the common information, which was shown to be false in Ref. [Win05]; see also Witsenhausen [Wit76]. The correct optimality proof required a much more complex argument [Win05].

Both the collaborative and adversarial models generalize Wyner's notion of common information [Wyn75]. In Wyner's scenario, Alice and Bob simply want to produce X^nY^n using pre-shared randomness with no additional communication. The minimum amount of randomness per copy needed to approximately simulate X^nY^n is what Wyner identifies as the common information held between X and Y. This can be seen as a special case of the three-party problems when Z ranges over just a single value and the public communication rate is zero.

Interest in secrecy formation is largely inspired by the analogous notion of *entanglement formation* when dealing with quantum systems and quantum information. The entanglement cost of a quantum state is defined to be the asymptotic rate of pre-shared ebits that are needed to prepare many copies of the given state [HHT01]. In the quantum setting, a third party is not introduced into the definition of entanglement cost since by its very nature, quantum entanglement possesses an inherent shielding from external parties. This latter property is sometime referred to as the *monogamy of entanglement* [HHHH09].

The structure of this paper is as follows. We begin in Sect. II by describing how, for the task at hand, all public communication generated during an LOPC protocol can be equivalently replaced by pre-shared public correlations at the start of the protocol and no further communication. This provides a significant simplification to the problem since a general LOPC protocol can involve multiple rounds of communication. In Sect. III, we introduce in greater detail Wyner's model for generating bipartite random variables as well as the tripartite models when Charlie is acting either as a collaborator or as an adversary. The main result of this paper is presented in Sect. IV and its proof is given in Sects. V and VI. The Appendix contains a reformulation of the original protocol given in [Win05].

Throughout this paper, random variables will be denoted by capital italic letters $U, V, \cdots$, etc. The values of the these variables will be written in lower-case $u, v, \cdots$, etc., and a sequence of such values will be denoted as $\mathbf{u}, \mathbf{v}, \cdots$, etc. The distribution of a given random variable U will be interchangeably written by P_U and $P(U)$. When variables U and U' range over a common alphabet $\mathcal{U}$, their variational distance (up to a factor of 2) is given by

$$\|P_U - P_{U'}\|_1 := \sum_{u\in\mathcal{U}} |P_U(u) - P_{U'}(u)|.$$

Finally, when three random variables form a Markov chain, this will be denoted by $W - U - V$, and it indicates that $P(WV|U) = P(W|U)P(V|U)$. Equivalently, its conditional mutual information satisfies $I(W;V|U) = 0$.

II. Replacing Public/Secret Communication by Shared Correlations

Consider a general LOPC protocol in which Alice and Bob begin with R perfectly correlated bits. Specifically, Alice (resp. Bob) has variable V_A (resp. V_B) such that $V_A = V_B = V$ and $H(V) = R$. Alice, Bob and Charlie may also have sources of local randomness, but these can be built directly into the local processing of the variables by allowing for stochastic mappings. The protocol will then involve a sequence of publicly announced messages M_i where M_i is a function of $(V, M_{<i})$. Here, $M_{<i} = M_1 \cdots M_{i-1}$ denotes all previous messages. At the end of the protocol, the entire communication can be represented by the variable U. Alice and Bob then generate random variables $\hat{X}^n$ and $\hat{Y}^n$, both as the image of some stochastic map applied to UV. Thus, the entire protocol can be represented by random variables $\hat{X}^n\hat{Y}^nUV$ whose distribution satisfies

$$P(\hat{X}^n\hat{Y}^nUV) = P(\hat{X}^n\hat{Y}^n|UV)P(UV) = P(\hat{X}^n|UV)P(\hat{Y}^n|UV)P(UV). \tag{1}$$

The particular distribution $P(UV)$ depends on the nature of the LOPC protocol, and here we are using the fact that the computations of $\hat{X}^n$ and $\hat{Y}^n$ are done locally (i.e. independently of each other). In the collaborative model, Charlie also obtains $\hat{Z}^n$ as a function of U, and the distribution is given by

$$P(\hat{X}^n\hat{Y}^n\hat{Z}^nUV) = P(\hat{X}^n|UV)P(\hat{Y}^n|UV)P(\hat{Z}^n|U)P(UV). \tag{2}$$

Hence to simulate the random variables $\hat{X}^n\hat{Y}^n\hat{Z}^nUV$ with no communication, it suffices for all three parties to first share the random variable U (which represents the public correlations of the protocol); additionally, Alice and Bob share the variable V (representing the secret correlations of the protocol). Conversely, to each distribution $P(VU)$, an LOPC protocol exists with Alice and Bob first sharing secret bits $V_A = V_B = V$ and then broadcasting U according to $P(U|V)$.

Remark 1: In the next section, we introduce models in which U and V are uncorrelated. While this does not correspond directly to the most general LOPC process, when proving the converse in Section V, we will allow for correlated U and V. Hence, the upper bound we derive on secret and public correlation rates will also hold in the LOPC scenario. In Section VI, we show that these lower bounds can be obtained by public and private correlations U and V which are, in fact, independent.

Alternatively, one can also directly show, by an operational argument, that protocols with correlated U and V can always be asymptotically simulated by one where public and secret correlation are independent.

III. Three Models of Correlation Generation

We now describe three different models for generating dependent random variables. While both the models in Sects. III-A and III-B have been well-studied, our new contribution is the model described in Sect. III-C.

A. Wyner's Common Information

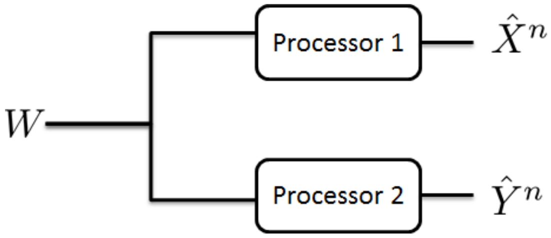

Fig. 1. Wyner's Common Information scenario.

In this subsection, we review Wyner's notion of common information as well as one of its operational interpretations [Wyn75].

Definition 2 (Wyner [Wyn75]): The common information $C(X:Y)$ between two random variables X and Y with joint distribution $Q(XY)$ is defined as

$$C(X:Y) = \min I(XY;W), \tag{3}$$

where the minimization is taken over all triples of random variables XYW so that

- the marginal distribution for X, Y is $Q(XY)$;
- $X - W - Y$ forms a Markov chain.

Furthermore, the minimum in Eq. (3) can be obtained with a random variable W ranging over sets of size no greater than $|\mathcal{X}||\mathcal{Y}|$.

To see why this quantity might capture the notion of "common information" between X and Y, consider the following task. Alice and Bob have access to a common source W, and acting independently of one another, they wish to process W in different ways so that their final joint distribution is a many-copy approximation of the target distribution $Q(XY)$ (see Fig. 1). The common information is the minimum rate of common randomness W needed to perform this task.

More precisely, we define an (n, R, ϵ) *source synthesis code* to consist of the following:

- a set $\mathcal{W}$ with cardinality $\lfloor 2^{nR} \rfloor$;
- conditional probability distributions $P_1^{(n)}(\mathbf{x}|w)$ and $P_2^{(n)}(\mathbf{y}|w)$, $w \in \mathcal{W}$, on $\mathcal{X}^n, \mathcal{Y}^n$, respectively;

such that

$$\left\| Q^{(n)} - \hat{P}^{(n)} \right\|_1 \leq \epsilon,$$

where

$$\hat{P}^{(n)}(\mathbf{x}, \mathbf{y}) := \frac{1}{|\mathcal{W}|} \sum_{w \in \mathcal{W}} P_1^{(n)}(\mathbf{x}|w) P_2^{(n)}(\mathbf{y}|w). \tag{4}$$

We say the rate R is *achievable* if for all $\epsilon > 0$ and n sufficiently large there exists a source synthesis code (n, R, ϵ). Define the correlation cost of XY as $C := \inf\{R : (n, R, \epsilon) \text{ is achievable}\}$.

Theorem 3 (Wyner [Wyn75]): For any pair XY of random variables, the minimum achievable rate of a source synthesis code is given by the common information:

$$C = C(X:Y).$$

The key ingredient in Wyner's achievability construction is a general result saying that for any two random variables U and W, the distribution of U^n can be reliably simulated by sampling from approximately $2^{nI(U:W)}$ sequences among the range of W^n and then applying the channel $P^n_{U|W}$ (see Lemma 12 below). Hence if $U = XY$ with W satisfying $X - W - Y$, then this simulation can be done locally, as depicted and in Fig. 1 and described in Eq. (4). This construction need not be limited to only two parties. For example, one can analogously define the common information of three variables XYZ with distribution $Q(XYZ)$ as

$$C(X:Y:Z) := \min I(XYZ;W), \tag{5}$$

where the minimization is taken over all variables $XYZW$ so that

- the marginal distribution for XYZ is $Q(XYZ)$;
- XYZ are conditionally independent variables given W.

Operationally, and in complete analogy to Wyner's Theorem 3, $C(X:Y:Z)$ is the smallest rate of shared random bits W that are needed to generate $Q(XYZ)$ when Alice, Bob and Charlie independently process W. In other words, there are now three channels $P_1^{(n)}(\mathbf{x}|w)$, $P_2^{(n)}(\mathbf{y}|w)$ and $P_3^{(n)}(\mathbf{z}|w)$ so that

$$Q^n(\mathbf{x},\mathbf{y},\mathbf{z}) \approx \frac{1}{|\mathcal{W}|}\sum_{w\in\mathcal{W}} P_1^{(n)}(\mathbf{x}|w)P_2^{(n)}(\mathbf{y}|w)P_3^{(n)}(\mathbf{z}|w)$$

with $\frac{1}{n}\log|\mathcal{W}| \leq C(X;Y;Z) + \delta$, for arbitrarily small δ. This is a special case of the more general three-party collaborative scenario that we will study below. Specifically, when the wires connected to V are removed in Fig. 3, we recover this scenario of Wyner's common information for three parties.

B. Key Cost in Three-Party Adversarial Scenario

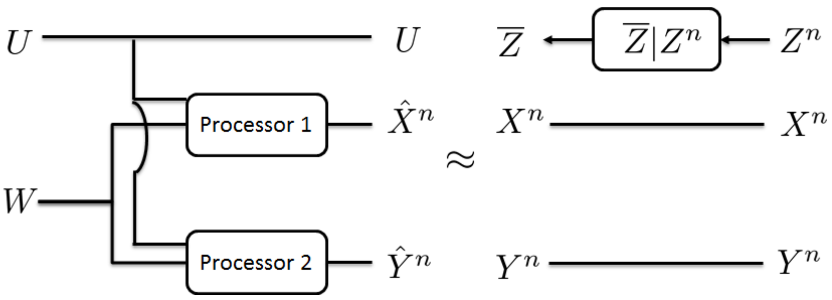

Fig. 2. Three-Party Adversarial Scenario.

We now consider three-party distributions generated by two sources of correlations. First is the adversarial model where Charlie is considered to be a malicious eavesdropper. Again, let $Q(XYZ)$ be the distribution whose correlation costs we wish to quantify. In the adversarial model, Alice and Bob start with some initially perfectly correlated bits that are secret from Charlie. Using LOPC, they wish to generate many copies of $Q(XY)$ so that the total public communication U produced in the protocol gives Charlie no more information about XY than what he has in the distribution $Q(XYZ)$. In other words, Charlie is able to apply some local processing $\overline{Z}|Z^n$ on her share part of $Q(X^nY^nZ^n)$ so that the resulting distribution is close to the distribution generated in the LOPC protocol. There will be some trade-off between the amount of initial secret correlations and the amount of public communication consumed in the protocol. Intuitively, the more perfectly correlated secret bits that Alice and Bob initially share, the less public communication they will need to generate $Q(XY)$.

By the discussion in Sect. II, we can simulate the entire protocol having public communication U by a protocol with no communication but initially shared public correlations. The resulting scenario is depicted in Fig. 2. The trade-off between public and private correlations in the task of secrecy formation is formally defined as follows.

Definition 4 (Renner & Wolf [RW03]): For distribution $Q(XYZ)$, an (n, R_P, R_K, ϵ) *secrecy formation code* is composed of the following:

- random variables (U,V) having joint distribution $P(UV)$ over the set $\mathcal{W}_P \times \mathcal{W}_K$ with cardinalities $|\mathcal{W}_P| = \lfloor 2^{nR_P} \rfloor$ and $|\mathcal{W}_K| = \lfloor 2^{nR_K} \rfloor$ respectively;

- conditional distributions on $\mathcal{X}^n$ and $\mathcal{Y}^n$,

$$P_1^{(n)}(\mathbf{x}|\mathbf{u},\mathbf{v}) \text{ and } P_2^{(n)}(\mathbf{y}|\mathbf{u},\mathbf{v}) \text{ for } \mathbf{u} \in \mathcal{W}_P,\ \mathbf{v} \in \mathcal{W}_K,$$

  which generate random variables $\widehat{X}^n\widehat{Y}^n$ with joint distribution

$$\widehat{P}(\mathbf{x},\mathbf{y}) := \sum_{\mathbf{u}\in\mathcal{W}_P}\sum_{\mathbf{v}\in\mathcal{W}_K} P_1^{(n)}(\mathbf{x}|\mathbf{u},\mathbf{v})P_2^{(n)}(\mathbf{y}|\mathbf{u},\mathbf{v})P(\mathbf{u},\mathbf{v});$$

- a channel $\bar{Z}|Z^n$ such that

$$\left\| Q(X^nY^n\bar{Z}) - \widehat{P}(\widehat{X}^n\widehat{Y}^nU)\right\|_1 \leq \epsilon. \tag{6}$$

The rate pair (R_P, R_K) is *achievable* if, for all $\epsilon > 0$, we can find an n sufficiently large such that there exists a secrecy formation code (n, R_P, R_K, ϵ).

The public-vs-secret tradeoff function is

$$R_K(R_P) = \inf\big\{R_K : (R_P, R_K) \text{ is achievable}\big\},$$

and the secret key cost of the triple XYZ is

$$K_c(X:Y|Z) := \lim_{R_P\to\infty} R_K(R_P).$$

Theorem 5 (Winter [Win05]): For the secrecy formation of the distribution $Q(XYZ)$, the rate pair (R_P, R_K) is achievable iff there exist random variables $XYZUV$ such that

$$R_K \geq I(XY;V|U) \quad \text{and} \quad R_P \geq I(Z;U), \tag{7}$$

where the random variables $XYZUV$ satisfy the properties

1) The XYZ marginal distribution is Q;
2) The following Markov chains hold:

$$XY - Z - U \quad \text{and} \quad X - UV - Y. \tag{8}$$

Furthermore the auxiliary random variables w.l.o.g. have bounded ranges: $|\mathcal{U}| \leq |\mathcal{Z}| + 1$ and $|\mathcal{V}| \leq |\mathcal{X}||\mathcal{Y}|$.

In particular,

$$K_c(X:Y|Z) = \min\big\{I(XY;V|U) : \text{Properties } \textit{1) and 2) hold}\big\}.$$

is the secret key cost of XYZ with unlimited public communication. ■

C. Key Cost in Three-Party Collaborative Scenario

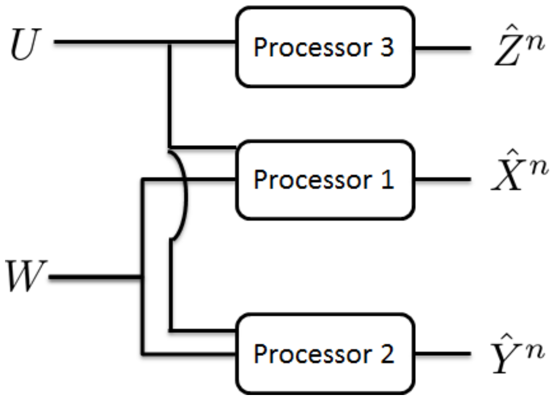

Fig. 3. Three-Party Collaborative Scenario.

We now shift perspectives and view Charlie as a collaborator instead of an adversary. As such, for a given distribution $Q(XYZ)$, Alice and Bob are no longer content with just generating variables $\hat{X}^n\hat{Y}^n$ using LOPC that are close to the target variables X^nY^n. They also want the public communication U to be sufficiently correlated with $\hat{X}^n\hat{Y}^n$ so that Charlie can locally process $\hat{Z}^n|U$ to jointly produce dependent variables $\hat{X}^n\hat{Y}^n\hat{Z}^n$ that are close to $X^nY^nZ^n$. Like in the adversarial setting, the public communication can be replaced with initially shared correlations between all the parties (see Fig. 3). There will also be a trade-off between public and private correlations in the following sense.

Definition 6: For a given tripartite probability distribution $Q(x,y,z)$ over $\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}$, we define an (n,R_P,R_K,ϵ) *split-source synthesis code* to be composed of the following:

- sets $\mathcal{W}_P$ and $\mathcal{W}_K$ with cardinalities $\lfloor 2^{nR_P}\rfloor$ and $\lfloor 2^{nR_K}\rfloor$, respectively;
- conditional probability distributions on $\mathcal{X}^n$, $\mathcal{Y}^n$ and $\mathcal{Z}^n$,

  $$P_1^{(n)}(\mathbf{x}|\mathbf{u},\mathbf{v}),\ P_2^{(n)}(\mathbf{y}|\mathbf{u},\mathbf{v}) \text{ and } P_3^{(n)}(\mathbf{z}|\mathbf{u}) \text{ for } \mathbf{u}\in\mathcal{W}_P,\ \mathbf{v}\in\mathcal{W}_K,$$

  which generate random variables $\widehat{X}^n\widehat{Y}^n\widehat{Z}^n$ with joint distribution

  $$\hat{P}^{(n)}(\mathbf{x},\mathbf{y},\mathbf{z}) = \frac{1}{|\mathcal{W}_P|}\frac{1}{|\mathcal{W}_K|}\sum_{\mathbf{u}\in\mathcal{W}_P}\sum_{\mathbf{v}\in\mathcal{W}_K} P_1^{(n)}(\mathbf{x}|\mathbf{u},\mathbf{v})P_2^{(n)}(\mathbf{y}|\mathbf{u},\mathbf{v})P_3^{(n)}(\mathbf{z}|\mathbf{u}), \tag{9}$$

such that $\|\hat{P}^{(n)} - Q^{(n)}\|_1 \le \epsilon$.

The rate pair (R_P,R_K) is *achievable* if, for all $\epsilon>0$, we can find an n sufficiently large such that there exists a split-source synthesis code (n,R_P,R_K,ϵ).

IV. Statement of Results

In this section we present our main result: a single-letter characterization of the trade-off between the public and private correlation rate pair in the collaborative scenario of Sect. III-C.

Theorem 7: For the split-source synthesis of the distribution $Q(XYZ)$, the rate pair (R_P,R_K) is achievable iff there exist random variables $XYZUV$ such that

$$R_K \ge I(XY;V|U) \quad \text{and} \quad R_P \ge I(XYZ;U), \tag{10}$$

where all random variables $XYZUV$ satisfy the following properties

1) The XYZ marginal distribution is Q;
2) $X-VU-Y$ and $XY-U-Z$ form Markov chains.

Furthermore, the random variables U and V in Eq. (10) can be restricted to sets of size no greater than $|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|$ and $|\mathcal{X}||\mathcal{Y}|$, respectively.

Fig. 4 illustrates the two-dimensional achievable rate region for the collaborative synthesis of a tripartite distribution. Theorem 7 determines the nontrivial corner point α in Fig. 4: indeed, the public correlation rate at α is given precisely by $C(XY:Z)$, the Wyner common information between XY and Z. Another corner point β is when $R_K=0$: here, the problem reduces to the three-party Wyner common information as described in Sect. III-A. Hence, the public correlation rate at β is given by $C(X:Y:Z)$. From this we see that Theorem 7 generalizes the notion of Wyner's common information (Theorem 3) when Z is trivial.

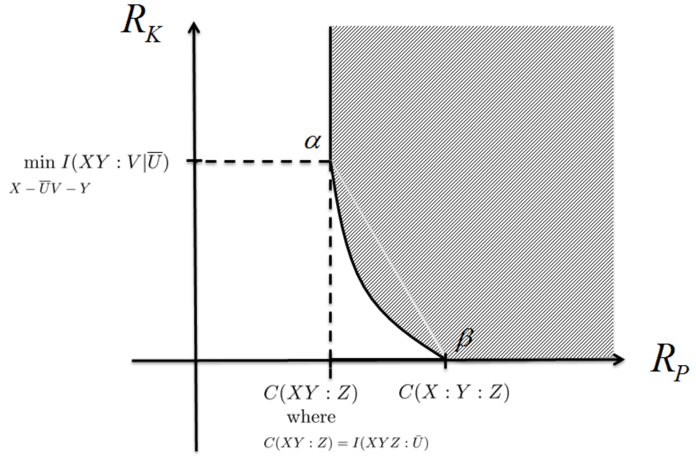

Fig. 4. Achievable rate region for the collaborative costs of a tripartite distribution using public and secret correlations. The rate pair of the point α (minimal public correlation) is given after Theorem 7, as is the point β (vanishing private correlation). Note that, as we can always substitute trivially one bit of private correlation by one bit of public correlation, the line connecting these points has slope at least as steep as -1.

Any rate pair on the line connecting α with β can be achieved by *time-sharing* the two protocols that achieve α and β, respectively. What can be said about this line connecting α and β? Denote the rate pairs at these points by $(R_P^{(\alpha)},R_K^{(\alpha)})$ and

$(R_P^{(\beta)}, R_K^{(\beta)})$ respectively. Since $C(XY:Z) \leq C(X:Y:Z)$, the slope of the connecting line will always be negative. On the other hand, since the secret correlations of any protocol can always be converted to public correlations, we have that the

$$R_P^{(\beta)} \leq R_P^{(\alpha)} + R_K^{(\alpha)}. \tag{11}$$

Hence the boundary line connecting α and β must have a slope not exceeding $-45°$. For some distributions, Eq. 11 is an equality while for others it is not. In the latter cases, the optimal exchange between private and public correlations is nontrivial and not a simple publication of the private correlations.

Example 1: Let P_{XYZ} be any distribution such that $H(Z|XY) = 0$, i.e. Z is a function of XY. Then it is easy to verify that the optimal R_K and R_P tradeoff is one-to-one, and thus the line connecting α and β has a slope of $-45°$. First consider the point α. Here the public correlation rate is given by $R_P^{(\alpha)} = \min_U I(XYZ;U)$ such that $XY - U - Z$. It is easy to show that $I(XYZ;U) \geq I(XY;Z)$, and this lower bound can be attained by $U = Z$ since $H(Z|XY) = 0$. Therefore, $R_P^{(\alpha)} = H(Z)$ and $R_K^{(\alpha)} = \min_V I(XY;V|Z)$ where $X - ZV - Y$. Let $\hat{V}$ denote the variable attaining this minimum. Now consider the point β where $R_K^{(\beta)} = 0$ and $R_P^{(\beta)} = \min_U I(XYZ;U)$ such that XYZ are conditionally independent given U. Then

$$\begin{aligned}
\min_U I(XYZ;U) &= \min_U \left(I(Z;U) + I(XY;U|Z)\right) \\
&\geq I(Z;XY) + \min_U I(XY;U|Z) \\
&\geq I(Z;XY) + I(XY;\hat{V}|Z) \\
&= H(Z) + I(XY;\hat{V}|Z),
\end{aligned} \tag{12}$$

where the first inequality is data processing since $Z - U - XY$; the second inequality is obtained since if XYZ are conditionally independent given U, then $X - ZU - Y$; and the last equality follows because $H(Z|XY) = 0$. Combining with Eq. (11), we see that

$$R_P^{(\beta)} = R_P^{(\alpha)} + R_K^{(\alpha)}.$$

Example 2: Next, we consider a very simple distribution P_{XYZ} over $\{0,1,2\}^{\times 3}$ with the only nonzero values being $P(x,y,z) = \frac{1}{5}$ for $P(2,2,z)$ with $z \in \{0,1,2\}$ and $P(x,y,2)$ with $x = y \in \{0,1\}$. This belongs to a more general class of "L-shaped" distributions studied by Witsenhausen [Wit76, Thm. 7]. From his result, the common information $C(XY:Z)$ is found to be

$$R_P^{(\alpha)} = C(XY:Z) = \tfrac{4}{5}\left(\log \tfrac{4}{5} - \log \tfrac{2}{5}\right) \approx .693. \tag{13}$$

This is computed from the optimal decomposition of P_{XYZ} into conditionally independent parts:

$$\tfrac{1}{5}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \tfrac{1}{10}\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}(0\ 0\ 1) + \tfrac{1}{10}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}(2\ 2\ 1).$$

Here we have grouped XY into one variable ranging over $\{0,1,2\}$ so that the (i,j) element of the matrix is $P_{XYZ}(i,i,j)$. From this decomposition, we see that

$$R_K^{(\alpha)} = \min_V I(XY;V|U) = -\tfrac{1}{2}\left(\tfrac{4}{5}\log\tfrac{2}{5} + \tfrac{1}{5}\log\tfrac{1}{5}\right). \tag{14}$$

For the corner point β, we observe that XYZ are conditionally independent given X. Hence,

$$R_P^{(\beta)} \leq H(X) = -\tfrac{3}{5}\log\tfrac{3}{5} - \tfrac{2}{5}\log\tfrac{1}{5}. \tag{15}$$

Hence,

$$R_P^{(\alpha)} + R_K^{(\alpha)} \approx 1.08 > R_P^{(\beta)} \approx .950, \tag{16}$$

and so the optimal private to public exchange for this distribution is not achieved by simply publicly revealing private correlations.

In the next two sections, we will prove Theorem 7, first the converse (Sect. V), then the direct part (Sect. VI).

V. Converse

Here we derive lower bounds that hold for more general models than a synthesis code. Specifically, we assume that $\hat{X}\hat{Y}\hat{Z}UV$ is given along with conditional probabilities $P_1^{(n)}$, $P_2^{(n)}$, and $P_3^{(n)}$ such that the generated distribution

$$\hat{P}^{(n)}(\mathbf{x},\mathbf{y},\mathbf{z}) = \sum_{u \in \mathcal{W}_P} \sum_{v \in \mathcal{W}_K} P(u,v) P_1^{(n)}(\mathbf{x}|u,v) P_2^{(n)}(\mathbf{y}|u,v) P_3^{(n)}(\mathbf{z}|u), \tag{17}$$

satisfies $\|\hat{P}^{(n)} - Q^{(n)}\|_1 \leq \epsilon$. Note that the local processing in the secrecy formation protocol imposes that $\hat{X}^n - VU - \hat{Y}^n$ and $\hat{X}^n\hat{Y}^n - U - \hat{Z}^n$ form Markov chains. However, unlike a synthesis code defined in Sect. III-C, we do not require that U and V are independent. This relaxation enables to simulate LOPC protocols as discussed in Remark 1.

Following the argument in [Win05], monotonicity and the chain rule allow us to write

$$\begin{aligned} R_K \geq \frac{1}{n}\log|\mathcal{W}_K| &\geq \frac{1}{n} I(\hat{X}^n\hat{Y}^n;V|U) \\ &= \sum_{j=1}^{n}\frac{1}{n} I(\hat{X}_j\hat{Y}_j;V|U\hat{X}_{<j}\hat{Y}_{<j}) \\ &= I(\hat{X}_J\hat{Y}_J;V|UJ\hat{X}_{<J}\hat{Y}_{<J}) \\ &= I(\hat{X}_J\hat{Y}_J;V|UJ\hat{X}_{<J}\hat{Y}_{<J}\hat{Z}_{<J}), \end{aligned} \tag{18}$$

where $J \in \{1,\cdots n\}$ is a uniformly distributed variable and the last equality follows from the conditional independence $\hat{X}^n\hat{Y}^n - U - \hat{Z}^n$ (see Proposition 8 below). We next introduce the following random variables $\hat{W} := J\hat{X}_{<J}\hat{Y}_{<J}\hat{Z}_{<J}$ and $W := JX_{<J}Y_{<J}Z_{<J}$ and the variables $\tilde{U} \in \mathcal{W}_P$, $\tilde{V} \in \mathcal{W}_K$, $\tilde{X} \in \mathcal{X}$, $\tilde{Y} \in \mathcal{Y}$, and $\tilde{Z} \in \mathcal{Z}$ defined through the joint distributions

$$P(\tilde{X}\tilde{Y}\tilde{Z}\tilde{V}\tilde{U}|W) = \hat{P}(\hat{X}_J\hat{Y}_J\hat{Z}_JVU|\hat{W}). \tag{19}$$

Then

$$\begin{aligned} \hat{P}(\hat{X}_J\hat{Y}_J\hat{Z}_J\hat{W}) = \hat{P}(\hat{X}_J\hat{Y}_J\hat{Z}_J|\hat{W})\hat{P}(\hat{W}) &= P(\tilde{X}\tilde{Y}\tilde{Z}|W)\hat{P}(\hat{W}) \\ &= P(\tilde{X}\tilde{Y}\tilde{Z}W) + P(\tilde{X}_J\tilde{Y}\tilde{Z}_J|W)[\hat{P}(\hat{W}) - P(W)]. \end{aligned} \tag{20}$$

At the same time, applying the triangle inequality to $\|\hat{P}^{(n)}(\hat{X}^n\hat{Y}^n\hat{Z}^n) - Q^{(n)}(X^nY^nZ^n)\|_1 \leq \epsilon$ allows us to conclude that

$$\|\hat{P}(\hat{X}_J\hat{Y}_J\hat{Z}_J\hat{W}) - P(X_JY_JZ_JW)\|_1 \leq \epsilon, \tag{21}$$

and therefore $\|P(\hat{W}) - P(W)\|_1 \leq \epsilon$. Combining the latter with Eqns. (20) and (21) yields

$$\|P(\tilde{X}\tilde{Y}\tilde{Z}W) - P(X_JY_JZ_JW)\|_1 \leq 2\epsilon. \tag{22}$$

Since $X^nY^nZ^n$ are i.i.d., the marginal distribution of $P(X_JY_JZ_JW)$ is $Q(XYZ)$. Hence, the previous inequality gives

$$\|P(\tilde{X}\tilde{Y}\tilde{Z}) - Q(XYZ)\|_1 \leq 2\epsilon. \tag{23}$$

Eq. (19) also gives that $P(\tilde{V}|\tilde{U}W) = \hat{P}(V|U\hat{W})$ and $P(\tilde{X}\tilde{Y}\tilde{Z}|\tilde{U}W) = \hat{P}(\hat{X}_J\hat{Y}_J\hat{Z}_J|U\hat{W})$. The Markov conditions $\hat{X}_J\hat{Y}_J - U\hat{W} - \hat{Z}_J$, $\hat{X}_J - VU\hat{W} - \hat{Y}_J$, and $\hat{Z}_J - U\hat{W} - V$ therefore imply

$$\tilde{X}\tilde{Y} - \tilde{\tilde{U}} - \tilde{Z} \quad \text{and} \quad \tilde{X} - \tilde{V}\tilde{\tilde{U}} - \tilde{Y}, \tag{24}$$

where $\tilde{\tilde{U}} := \tilde{U}W$.

From Eq. (19), we have $\hat{P}(\hat{X}_J\hat{Y}_JV|U\hat{W})P(U|\hat{W}) = P(\tilde{X}\tilde{Y}\tilde{V}|\tilde{U}W)P(\tilde{U}|W)$ which further implies that $\hat{P}(\hat{X}_J\hat{Y}_JV|U\hat{W}) = P(\tilde{X}\tilde{Y}\tilde{V}|\tilde{U}W)$ since $\hat{P}(U|\hat{W}) = P(\tilde{U}|W)$, again by Eq. (19). Thus, for each fixed value of w, we have that

$$I(\hat{X}_J\hat{Y}_J;V|U\hat{W}=w)\hat{P}(U|\hat{W}=w) = I(\tilde{X}\tilde{Y};\tilde{V}|\tilde{U}W=w)P(\tilde{U}|W=w).$$

Multiply both sides by $P(\hat{W}=w)$ and take the sum. Using the fact that $\|P(\hat{W}) - P(W)\|_1 \leq \epsilon$ and the triangle inequality lead to:

$$|I(\hat{X}_J\hat{Y}_J;V|U\hat{W}) - I(\tilde{X}\tilde{Y};\tilde{V}|\tilde{\tilde{U}})| \leq \epsilon\log|\mathcal{X}||\mathcal{Y}|,$$

hence,

$$R_K \geq I(\tilde{X}\tilde{Y};\tilde{V}|\tilde{\tilde{U}}) - \epsilon\log|\mathcal{X}||\mathcal{Y}|. \tag{25}$$

By the same arguments, we can bound the public communication as

$$\begin{aligned} R_P \geq \frac{1}{n}\log|\mathcal{W}_P| &\geq \frac{1}{n} I(\hat{X}^n\hat{Y}^n\hat{Z}^n;U) = I(\hat{X}_J\hat{Y}_J\hat{Z}_J : U|\hat{W}) \\ &\geq I(\tilde{X}\tilde{Y}\tilde{Z};\tilde{U}|W) - \epsilon\log|\mathcal{X}||\mathcal{Y}||\mathcal{Z}| \\ &= I(\tilde{X}\tilde{Y}\tilde{Z};\tilde{\tilde{U}}) - I(\tilde{X}\tilde{Y}\tilde{Z};W) - \epsilon\log|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|. \end{aligned}$$

To bound the term $I(\tilde{X}\tilde{Y}\tilde{Z};W) = H(\tilde{X}\tilde{Y}\tilde{Z}) - \sum_w H(\tilde{X}\tilde{Y}\tilde{Z}|W=w)P(W=w)$, we recall a well-known continuity relation: Any two random variables A and A' ranging over $\mathcal{A}$ with $\delta := \|P(A) - P(A')\|_1 \leq 1/2$ satisfy $|H(A) - H(A')| \leq -\delta\log\frac{\delta}{|\mathcal{A}|}$ [CK11]. Therefore, using Eq. (22) and the fact that $I(X_JY_JZ_J;W) = 0$, we readily obtain

$$R_P \geq I(\tilde{X}\tilde{Y}\tilde{Z};\tilde{\tilde{U}}) + 4\epsilon\log 2\epsilon - 5\epsilon\log|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|. \tag{26}$$

At this point we have constructed random variables $\tilde{X}\tilde{Y}\tilde{Z}\tilde{V}\tilde{\tilde{U}}$ that satisfy Eqns. (23)–(26). By Lemma 9, we can assume without loss of generality that $\tilde{V}$ and $\tilde{\tilde{U}}$ range over sets of size no greater than $|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|$. Hence, the set of random variables satisfying Eqns. (23)–(26) is compact, and therefore a limit point will exist which also satisfies these constraints when taking $\epsilon \to 0$. This proves the lower bound of Theorem 7. ■

Proposition 8: If n-part random variables A^n and B^n satisfy $A^n - C - B^n$, then the reduced variables A_j and B_k satisfy $A_j - CA_{<j}B_{<k} - B_k$ for any $1 \leq j, k \leq n$, where $A_{<j} = A_1 \dots A_{j-1}$ and likewise $B_{<k} = B_1 \dots B_{k-1}$.

Proof: Consider the marginal distribution $A_jA_{<j} - C - B_kB_{<k}$. Then

$$\begin{aligned} P(A_jB_k|CA_{<j}B_{<k}) &= \frac{P(A_jA_{<j}B_kB_{<k}|C)}{P(A_{<j}B_{<k}|C)}, \\ &= \frac{P(A_jA_{<j}|C)}{P(A_{<j}|C)}\frac{P(B_kB_{<k}|C)}{P(B_{<k}|C)}, \\ &= P(A_j|A_{<j}C)P(B_k|B_{<k}C). \end{aligned} \tag{27}$$

Therefore, $A_j - CA_{<j}B_{<k} - B_k$. ■

Lemma 9: Suppose that $XYZUV$ are random variables with UV ranging over $\mathcal{U}\times\mathcal{V}$ such that $XY - U - Z$ and $X - UV - Y$. Then there exists random variables $X'Y'Z'V'U'$ satisfying the same Markov chain and

$$I(X'Y'Z';U') = I(XYZ;U), \tag{28a}$$

$$P(X'Y'Z') = P(XYZ), \tag{28b}$$

$$I(X'Y';V'|U') \leq I(XY;V|U), \tag{28c}$$

with U' and V' ranging over sets $\mathcal{U}'$ and $\mathcal{V}'$ of sizes $|\mathcal{U}'| \leq |\mathcal{X}||\mathcal{Y}||\mathcal{Z}| + 1$ and $|\mathcal{V}'| \leq |\mathcal{X}||\mathcal{Y}||\mathcal{Z}|$. Furthermore, if $Z - U - V$ also holds, then the size of $\mathcal{V}'$ can be further reduced to $|\mathcal{V}'| \leq |\mathcal{X}||\mathcal{Y}|$.

Proof: For the given distribution $P(XYZUV)$, let $\{P(XYZ|vu)\}_{v\in\mathcal{V},u\in\mathcal{U}}$ and $\{P(XYZ|u)\}_{u\in\mathcal{U}}$ be the associated conditional distributions. For each fixed $u \in \mathcal{U}$, let Λ_u be the collection of conditional distributions over $\mathcal{V}$ such that $\lambda(v|u) \in \Lambda_u$ if $\sum_v P(XYZ|uv)\lambda(v|u) = P(XYZ|u)$. This represents a total of $N = |\mathcal{X}||\mathcal{Y}||\mathcal{Z}| - 1$ linear constraints on the $\lambda(v|u)$ (note if $Z - U - V$ also holds, then $\sum_v P(XYZ|uv)\lambda(v|u) = P(XYZ|u)$ reduces to $\sum_v P(XY|uv)\lambda(v|u) = P(XY|u)$ which represents a total of $N = |\mathcal{X}||\mathcal{Y}| - 1$ linear constraints on the $\lambda(v|u)$). Now Λ_u is convex and the set $\{\sum_v H(XY|u,v)\lambda(v|u) : \lambda \in \Lambda_u\}$ will obtain both its maximum and minimum at an extreme point of Λ_u. Then an application of Carathéodory's Theorem (Lemma 10) guarantees that such an extreme point is a distribution over $\mathcal{V}$ with no more than $N+1$ nonzero probability values [Kle63]. Hence by a conditional relabeling of the v, we have a subset $\mathcal{V}' \subset \mathcal{V}$ with $|\mathcal{V}'| \leq N + 1$ and a collection of conditional distributions $\lambda'(v|u)$ over $\mathcal{V}'$ such that

$$\sum_{v\in\mathcal{V}'} P(XYZ|uv)\lambda'(v|u) = \sum_{v\in\mathcal{V}} P(XYZ|u) \quad \forall u \in \mathcal{U}, \tag{29a}$$

$$\sum_{v\in\mathcal{V}'} H(XY|uv)\lambda'(v|u) \geq \sum_{v\in\mathcal{V}} H(XY|V,u) \quad \forall u \in \mathcal{U}. \tag{29b}$$

We now perform a similar argument by letting Γ be the set of all distributions over $\mathcal{U}$ such that $\gamma(u) \in \Gamma$ if $\sum_{u\in\mathcal{U}} P(XYZ|u)\gamma(u) = P(XYZ)$ and $\sum_{u\in\mathcal{U}} H(XYZ|U = u)\gamma(u) = H(XYZ|U)$. This represents $|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|$ linear constraints on γ, and we seek the minimum value of the set $\{\sum_u [H(XY|u) - H(XY|V,u)]\gamma(u) : \gamma \in \Gamma\}$. A second application of Carathéodory's Theorem ensures the existence of a distribution $\gamma'(u)$ ranging over $\mathcal{U}' \subset \mathcal{U}$ with $|\mathcal{U}'| \leq |\mathcal{X}||\mathcal{Y}||\mathcal{Z}| + 1$ for which

$$\sum_{u\in\mathcal{U}'} P(XYZ|u)\gamma'(u) = P(XYZ), \tag{30a}$$

$$\sum_{u\in\mathcal{U}'} H(XYZ|u)\gamma'(u) = H(XYZ|U), \tag{30b}$$

$$\sum_{u\in\mathcal{U}} [H(XY|u) - H(XY|V,u)]\gamma'(u) \leq H(XY|U) - H(XY|UV) = I(XY;V|U). \tag{30c}$$

This completes the construction of random variables $X'Y'Z'U'V'$ whose joint distribution is given by $P(X'Y'Z'U'V') := P(XYZ|uv)\lambda'(v|u)\gamma'(u)$. By its definition and by Eq. (29a), $X'Y'Z'U'V'$ inherits whatever Markov chain properties are present in $XYZUV$. Eq. (30a) gives $P(X'Y'Z') = P(XYZ)$, and combining this with Eq. (30b) yields $I(X'Y'Z';U') = I(XYZ;U)$. Finally, combining Eq. (30c) and Eq. (29b) gives

$$I(X'Y';V'|U') \leq I(XY;V|U), \tag{31}$$

concluding the proof. ■

Lemma 10 (Carathéodory's Theorem [Roc96]): Let S be a subset of $\mathbb{R}^n$ and $\mathrm{conv}(S)$ its convex hull. Then any $x \in \mathrm{conv}(S)$ can be expressed as a convex combination of at most $n+1$ elements of S. ■

Remark 11: An application of Carathéodory's Theorem shows that if the elements of $\mathrm{conv}(S)$ are further required to satisfy d linear constraints, then the resulting set is convex with extreme points being convex combinations of at most $d+1$ extreme points of $\mathrm{conv}(S)$ [Kle63].

VI. Achievability

Let $XYZUV$ be random variables with joint distribution $P(XYZUV)$ satisfying (1) $Q(XYZ) = P(XYZ)$ and (2) $X - VU - Y$ and $XY - U - Z$. In what follows, we let $T^n_{[U]_\delta}$ denote the set of all δ-typical sequences with respect to random variable U having distribution $P(U)$. Recall that a sequence $\mathbf{u} \in \mathcal{U}^n$ is δ-typical if $\left|\frac{N(u|\mathbf{u})}{n} - P(u)\right| \leq \delta$ for all $u \in \mathcal{U}$ [CK78].

Our code makes repeated use of Wyner's original code. The following is proven in [Wyn75], where here we have modified the statement using Pinsker's inequality, $D(P_1||P_2) \geq \frac{1}{2}\|P_1 - P_2\|_1^2$, to obtain a bound on the variational distance. See also later works by Han and Verdú [HV92], [HV93] and Ahlswede [A06], where more general versions were proved (Ref. [AW02] contains a quantum analogue).

Lemma 12 (Wyner [Wyn75, Thm. 6.3]): Let AB be random variables over $\mathcal{A} \times \mathcal{B}$ with joint distribution $P(AB)$, and let $R > I(A;B)$. For $\epsilon > 0$ and sufficiently large n, there exists a subset $\beta \subset T^n_{[B]_\delta} \subset \mathcal{B}^n$ of size $|\beta| = \lfloor 2^{nR} \rfloor$ such that for

$$\hat{P}^{(n)}(\mathbf{a}) = \frac{1}{|\beta|} \sum_{\mathbf{b} \in \beta} P^{(n)}(\mathbf{a}|\mathbf{b}) \quad \text{for} \quad \mathbf{a} \in \mathcal{A}^n, \tag{32}$$

it holds that $\|P^{(n)}(A^n) - \hat{P}^{(n)}(A^n)\|_1 \leq \epsilon$.

Identify $A := XYZ$ and $B := U$ in Lemma 12. Thus, for n sufficiently large, we can find a subset $\mathcal{W}_P \subset \mathcal{U}^n$ such that $|\mathcal{W}_P| = \lfloor 2^{nR_P} \rfloor$ with

$$R_P = I(XYZ : U) + \delta, \tag{33}$$

and

$$\left\| Q^{(n)}(\mathbf{x}, \mathbf{y}, \mathbf{z}) - \frac{1}{|\mathcal{W}_P|} \sum_{\mathbf{u} \in \mathcal{W}_P} P^{(n)}(\mathbf{x}, \mathbf{y}|\mathbf{u}) P^{(n)}(\mathbf{z}|\mathbf{u}) \right\|_1 \leq \epsilon, \tag{34}$$

where we have used the Markov chain $XY - U - Z$. Here, $P^{(n)}(\mathbf{z}|\mathbf{u})$ will be the encoder employed by the collaborative third party.

We next consider the term

$$P^{(n)}(\mathbf{x}, \mathbf{y}|\mathbf{u}) = \prod_{u \in \mathcal{U}} P^{(N(u|\mathbf{u}))}(\mathbf{x}_u, \mathbf{y}_u|u), \tag{35}$$

where $(\mathbf{x}_u, \mathbf{y}_u)$ is a sequence of length $N(u|\mathbf{u})$ that occurs with the event $U = u$. Knowing that $\mathbf{u} \in T^n_{[U]_\delta}$, we let $n_u := \lfloor n(P(u) + \delta) \rfloor$, and for each $u \in \mathcal{U}$, we apply Lemma 12 on the conditional distribution $P(XYV|U = u)$ with the choice $A := XY$ and $B := V$. This will generate a collection of codeword sets α_u, each with respective size $|\alpha_u| = \lfloor 2^{n_u(I(XY:V|U=u)+\delta)} \rfloor$. Furthermore,

$$\hat{P}^{(n_u)}(\mathbf{x}_u, \mathbf{y}_u|u) := \frac{1}{|\alpha_u|} \sum_{\mathbf{v} \in \alpha_u} P^{(n_u)}(\mathbf{x}_u, \mathbf{y}_u|u, \mathbf{v}_u) = \frac{1}{|\alpha_u|} \sum_{\mathbf{v} \in \alpha_u} P^{(n_u)}(\mathbf{x}_u|u, \mathbf{v}_u) P^{(n_u)}(\mathbf{y}_u|u, \mathbf{v}_u) \tag{36}$$

satisfies $||\hat{P}^{(n_u)}(\mathbf{x}_u, \mathbf{y}_u|u) - P^{(n_u)}(\mathbf{x}_u, \mathbf{y}_u|u)||_1 < \epsilon$. In the previous equation, the Markov chain $X - UV - Y$ has been employed. For each $u \in \mathcal{U}$ and $\mathbf{v}_u \in \alpha_u$, let $\hat{P}^{(N(u|\mathbf{u}))}(\mathbf{x}_u|u, \mathbf{v}_u)$ denote the marginal distribution obtained from $P^{(n_u)}(\mathbf{x}_u|u, \mathbf{v}_u)$ by summing over the last $n_u - N(u|\mathbf{v}_u)$ events. Let $\hat{P}^{(N(u|\mathbf{u}))}(\mathbf{y}_u|u, \mathbf{v}_u)$ be defined likewise. Thus,

$$||\hat{\hat{P}}^{N(u|\mathbf{u})}(\mathbf{x}_u, \mathbf{y}_u|u) - P^{N(u|\mathbf{u})}(\mathbf{x}_u, \mathbf{y}_u|u)||_1 < \epsilon,$$

where

$$\hat{\hat{P}}^{N(u|\mathbf{u})}(\mathbf{x}_u, \mathbf{y}_u|u) := \frac{1}{|\alpha_u|} \sum_{\mathbf{v}_u \in \alpha_u} \hat{P}^{(N(u|\mathbf{u}))}(\mathbf{x}_u|u, \mathbf{v}_u) \hat{P}^{(N(u|\mathbf{u}))}(\mathbf{y}_u|u, \mathbf{v}_u). \tag{37}$$

We now paste together the different codes to form the code set $\mathcal{W}_K = \prod_{u \in \mathcal{U}} \alpha_u$. For any typical $\mathbf{u}$ and $\mathbf{v} \in \mathcal{W}_K$, we define the local generators

$$\begin{aligned} \hat{\hat{P}}^{(n)}(\mathbf{x}|\mathbf{u}, \mathbf{v}) &:= \prod_{u \in \mathcal{U}} \hat{P}^{(N(u|\mathbf{u}))}(\mathbf{x}_u|u, \mathbf{v}_u) \\ \hat{\hat{P}}^{(n)}(\mathbf{y}|\mathbf{u}, \mathbf{v}) &:= \prod_{u \in \mathcal{U}} \hat{P}^{(N(u|\mathbf{u}))}(\mathbf{y}_u|u, \mathbf{v}_u), \end{aligned} \tag{38}$$

which satisfy

$$\left\| P^{(n)}(\mathbf{x},\mathbf{y}|\mathbf{u}) - \frac{1}{|\mathcal{W}_K|}\sum_{\mathbf{v}\in\mathcal{W}_K}\hat{\hat{P}}^{(n)}(\mathbf{x}|\mathbf{u},\mathbf{v})\hat{\hat{P}}^{(n)}(\mathbf{y}|\mathbf{u},\mathbf{v}) \right\|_1 < |\mathcal{U}|\epsilon. \tag{39}$$

The size of $\mathcal{W}_K$ is bounded by

$$\begin{aligned}\log|\mathcal{W}_K| = \sum_{u\in\mathcal{U}}\log|\alpha_u| &\le \sum_{u\in\mathcal{U}} n_u(I(XY:V|U=u)+\delta) \\ &\le n(I(XY:V|U)+O(\delta)).\end{aligned} \tag{40}$$

Combining this simulation of $P^{(n)}(\mathbf{x},\mathbf{y}|\mathbf{u})$ with Eq. (34) gives the final error bound

$$\left\| Q^{(n)}(\mathbf{x},\mathbf{y},\mathbf{z}) - \hat{P}^{(n)}(\mathbf{x},\mathbf{y},\mathbf{z}) \right\|_1 \le (|\mathcal{U}|+1)\epsilon. \tag{41}$$

where

$$\hat{P}^{(n)}(\mathbf{x},\mathbf{y},\mathbf{z}) = \frac{1}{|\mathcal{W}_P|}\frac{1}{|\mathcal{W}_K|}\sum_{\mathbf{u}\in\mathcal{W}_P}\sum_{\mathbf{v}\in\mathcal{W}_K}\hat{\hat{P}}^{(n)}(\mathbf{x}|\mathbf{u},\mathbf{v})\hat{\hat{P}}^{(n)}(\mathbf{y}|\mathbf{u},\mathbf{v})P^{(n)}(\mathbf{z}|\mathbf{u}).$$

Since $|\mathcal{U}| \le |\mathcal{X}||\mathcal{Y}||\mathcal{Z}|$ and $|\mathcal{V}| \le |\mathcal{X}||\mathcal{Y}|$, the bounds on R_P, R_K and $\|Q^{(n)} - \hat{P}^{(n)}\|_1$ can be made arbitrarily close to $I(XYZ:U)$, $I(XY:V|U)$ and zero, respectively.

■

VII. Conclusion

In this paper we have introduced the problem of tripartite correlation generation using public and private correlations. This can be seen as a collaborative alternative to the cryptographic problem of secrecy formation. We have found that despite the two different natures of the problem, the optimal secret correlation rates have a very similar structure. We have completely characterized the public-vs-private rate region for the collaborative scenario. One point of interest is when the public correlation rate is minimum (point α in Fig. 4), and another is when the secret correlation rate is zero (point β in Fig. 4). We have shown that the optimal exchange of private to public correlations does not always involve a trivial publicizing of private information. However, it is an interesting open problem to determine the slope of the line connecting α and β for a general distribution, in particular to understand what limits there are, if any, on the exchange rate of private to public correlation rate.

Further afield, following the example of Ref. [Win05], one could ask for the benefit of using entanglement instead of, or in addition to, the private and public shared randomness. We leave this and other questions for future investigations.

Acknowledgments

EC was supported by the National Science Foundation (NSF) Early CAREER Award No. 1352326. MH is supported by an ARC Future Fellowship under Grant FT140100574. AW was supported by the European Commission (STREP “RAQUEL”), the European Research Council (Advanced Grant “IRQUAT”), and the Spanish MINECO (project FIS2008-01236) with FEDER funds.

Appendix

Here we review the achievability component of Theorem 5. The coding for Alice and Bob is the same as described in Section VI. Let $XYZUV$ be random variables obtaining the minimum in Theorem 5, and let $P(XYZUV)$ denote their joint distribution so that the marginal on XYZ is $Q(XYZ)$. The public correlation (communication) is U^n (n i.i.d. realisations of U). Let us restrict attention to the typical subset for which the relative frequency of each letter u in U^n is close to $P(U=u)$; in particular, $\left|P(U=u) - \frac{N(u|\mathbf{u})}{n}\right| \le \delta$. For the set of positions where u occurs, we can employ Lemma 12.

Consider $A = XY|_{U=u}$ and $B = V|_{U=u}$ in Lemma 12. Thus for $n_u := \lfloor n(P(u)+\delta)\rfloor$ sufficiently large, we can find a subset $\alpha_u \subset \mathcal{V}^{n_u}$ such that $|\alpha_u| = \lfloor 2^{n_u(I(XY:V|U=u)+\delta)}\rfloor$ with

$$\left\| P^{(n_u)}(X^{n_u}Y^{n_u}|u) - \hat{P}^{(n_u)}(X^{n_u}Y^{n_u}|u) \right\|_1 \le \epsilon, \tag{42}$$

where

$$\hat{P}^{(n_u)}(X^{n_u}Y^{n_u}|u) := \frac{1}{|\alpha_u|}\sum_{\mathbf{v}\in\alpha_u} P^{(n_u)}(X^{n_u}Y^{n_u}|u,\mathbf{v}).$$

We paste these codes together and define local channels for Alice and Bob $\hat{\hat{P}}^{(n)}(X^n|\mathbf{u},\mathbf{v})\hat{\hat{P}}^{(n)}(Y^n|\mathbf{u},\mathbf{v})$ which, for each $\mathbf{u} \in T^n_{[U]_\delta}$, samples from the concatenated code and discards the extra letter occurrences not found in $\mathbf{u}$ (see Section VI). With $\mathcal{W}_K$ denoting the set of code words, this generates the simulation

$$\tilde{P}^{(n)}(X^nY^n|\mathbf{u}) := \frac{1}{|\mathcal{W}_K|}\sum_{\mathbf{v}\in\mathcal{W}_K}\hat{\hat{P}}^{(n)}(X^n|\mathbf{u},\mathbf{v})\hat{\hat{P}}^{(n)}(Y^n|\mathbf{u},\mathbf{v})$$

which satisfies

$$\left\| P^{(n)}(X^nY^n|\mathbf{u}) - \tilde{P}^{(n)}(X^nY^n|\mathbf{u}) \right\|_1 < |\mathcal{U}|\epsilon. \tag{43}$$

The size of $\mathcal{W}_K$ is bounded by

$$\begin{aligned} \log|\mathcal{W}_K| = \sum_{u\in\mathcal{U}} \log|\alpha_u| &\le \sum_{u\in\mathcal{U}} n_u(I(XY:V|U=u)+\delta) \\ &\le n(I(XY:V|U)+O(\delta)). \end{aligned} \tag{44}$$

Charlie's Simulation: Eq. (43) holds for every $\mathbf{u} \in T^{(n)}_{[U]_\delta}$. The next question is how we choose our code words $\mathbf{u}$, which represents the public communication. If we just took $T^n_{[U]_\delta}$ as the codebook, then the public correlation rate would be $H(U)$. But we can actually do better, and we will use Wyner's theorem again to construct a smaller codebook.

From Wyner, for n sufficiently large there exists a subset $\beta \subset T^n_{[U]_\delta}$ with $|\beta| \le 2^{n(I(Z:U)+\delta)}$ such that

$$\hat{P}^{(n)}(Z^n) := \frac{1}{|\beta|}\sum_{\mathbf{u}\in\beta} P^{(n)}(Z^n|\mathbf{u})$$

and

$$\left\| \hat{P}^{(n)}(Z^n) - Q^{(n)}(Z^n) \right\|_1 \le \epsilon. \tag{45}$$

Let $\tilde{U}$ be uniformly distributed over β and define the channel $\tilde{U}|Z^n$ by

$$\Phi^{(n)}(\mathbf{u}|\mathbf{z}) = \frac{1}{|\beta|}\frac{P^{(n)}(\mathbf{z}|\mathbf{u})}{\hat{P}^{(n)}(\mathbf{z})} \quad \text{for } \mathbf{u}\in\beta. \tag{46}$$

When Charlie applies $\Phi^{(n)}$ to his part of distribution $Q^{(n)}(X^nY^nZ^n)$, the new distribution is given by

$$\tilde{Q}^{(n)}(X^nY^nZ^n\tilde{U}) : = Q^{(n)}(X^nY^n|Z^n)\Phi^{(n)}(\tilde{U}|Z^n)Q^{(n)}(Z^n). \tag{47}$$

Note that the reduced distribution $\tilde{Q}^{(n)}(X^nY^n\tilde{U})$ is precisely what is obtained when Charlie attempts to simulate the public communication $\tilde{U}$ by acting on $X^nY^nZ^n$ with $\Phi^{(n)}$. Thus, we want to prove that $\tilde{Q}^{(n)}(X^nY^n\tilde{U})$ is close to the distribution generated by $\tilde{P}^{(n)}(\mathbf{x},\mathbf{y}|\mathbf{u})$ when $\mathbf{u}$ is chosen uniformly from β, which we denote by

$$\tilde{P}^{(n)}(X^nY^n\tilde{U}) := \frac{1}{|\beta|}\tilde{P}^{(n)}(X^nY^n|\tilde{U}).$$

To do this, we first bound the difference

$$\begin{aligned} &\left\| \frac{1}{|\beta|}P^{(n)}(X^nY^n|\tilde{U}) - \tilde{Q}^{(n)}(X^nY^n\tilde{U}) \right\|_1 \\ &\qquad = \left\| \frac{1}{|\beta|}P^{(n)}(X^nY^n|\tilde{U}) - \sum_{\mathbf{z}\in\mathcal{Z}^n} Q^{(n)}(X^nY^n|\mathbf{z})\Phi^{(n)}(\tilde{U}|\mathbf{z})Q^{(n)}(\mathbf{z}) \right\|_1 \\ &\qquad \le \left\| \frac{1}{|\beta|}P^{(n)}(X^nY^n|\tilde{U}) - \sum_{\mathbf{z}\in\mathcal{Z}^n} Q^{(n)}(X^nY^n|\mathbf{z})\Phi^{(n)}(\tilde{U}|\mathbf{z})\hat{P}^{(n)}(\mathbf{z}) \right\|_1 + \epsilon \\ &\qquad = \left\| \frac{1}{|\beta|}P^{(n)}(X^nY^n|\tilde{U}) - \frac{1}{|\beta|}\sum_{\mathbf{z}\in\mathcal{Z}^n} P^{(n)}(X^nY^n|\mathbf{z})P^{(n)}(\mathbf{z}|\tilde{U}) \right\|_1 + \epsilon \\ &\qquad = \left\| \frac{1}{|\beta|}P^{(n)}(X^nY^n|\tilde{U}) - \frac{1}{|\beta|}\sum_{\mathbf{z}\in\mathcal{Z}^n} P^{(n)}(X^nY^n|\tilde{U}\mathbf{z})P^{(n)}(\mathbf{z}|\tilde{U}) \right\|_1 + \epsilon = \epsilon. \end{aligned} \tag{48}$$

Here, we have used both Eqns. (46) and (45), and the last line follows from the Markov chain condition $XY - Z - U$. Therefore, combining with Eq. (43), we obtain the desired result that

$$\left\| \tilde{Q}^{(n)}(X^nY^n\tilde{U}) - \tilde{P}^{(n)}(X^nY^n\tilde{U}) \right\|_1 \le \epsilon(1+|\mathcal{U}|). \tag{49}$$

To summarize the protocol, consider any $\delta,\epsilon > 0$ and n sufficiently large. Either Alice or Bob locally generates the random variable $\tilde{U}$ which is uniformly distributed over a set of size $|\beta| \le 2^{n(I(Z;U)+\delta)}$. The value of $\tilde{U}$ is announced publicly. Sharing no more than $n(I(XY;V|U)+\delta)$ bits of secret correlation, Alice and Bob generate distribution $\tilde{P}^{(n)}(X^nY^n)$ which is jointly distributed with $\tilde{U}$ according to $\tilde{P}^{(n)}(X^nY^n\tilde{U})$. At the same time, we have shown the existence of a channel $\Phi^{(n)}$ such that when Charlie applies this to her part of $X^nY^nZ^n$, it generates the distribution $\tilde{Q}^{(n)}(X^nY^n\tilde{U})$ for which

$$\left\| \tilde{Q}^{(n)}(X^nY^n\tilde{U}) - \tilde{P}^{(n)}(X^nY^n\tilde{U}) \right\|_1 \le \epsilon(1+|\mathcal{U}|).$$

Therefore, we have satisfied the two components of the achievability criteria.

References

[A06] Rudolf Ahlswede. On Concepts of Performance Parameters for Channels. In *General Theory of Information Transfer and Combinatorics* (R. Ahlswede, L. Bäumer, N. Cai, H. Aydinian, V. Blinovsky, C. Deppe, and H. Mashurian, eds.), *Lecture Notes in Computer Science*, vol. 4123, pp. 639-17663, Springer Verlag, Berlin Heidelberg, 2006.

[AC93] Rudolf Ahlswede and Imre Csiszár. Common Randomness in Information Theory and Cryptography. I. Secret Sharing. *IEEE Transactions on Information Theory* 39(4):1121–1132, 1993. doi:10.1109/18.243431.

[AW02] Rudolf Ahlswede and Andreas Winter. Strong converse for identification via quantum channels. *IEEE Transactions on Information Theory*, 48(3):569-17579, 2002.

[BSST02] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Transactions on Information Theory* 48(10):2637–2655, 2002.

[B+14] Charles H. Bennett, Igor Devetak, Aram W. Harrow, Peter W. Shor, and Andreas Winter The Quantum Reverse Shannon Theorem and Resource Tradeoffs for Simulating Quantum Channels. *IEEE Transactions on Information Theory* 60(5):2926–2959, 2014.

[CK78] Imre Csiszár and Janos Körner. Broadcast channels with confidential messages. *IEEE Transactions on Information Theory* 24(3):339–348, 1978. doi:10.1109/TIT.1978.1055892.

[CK11] Imre Csiszár and Janos Körner. *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Second ed., Cambridge University Press, Cambridge, UK, 2011.

[CN00] Imre Csiszár and Prakash Narayan. Common randomness and secret key generation with a helper. *IEEE Transactions on Information Theory* 46(2):344–366, 2000. doi:10.1109/18.825796.

[Cuff08] Paul Cuff. Communication Requirements for Generating Correlated Random Variables. In *Proceedings of the International Symposium on Information Theory, ISIT 2008*, pp. 1393–1397, Toronto, Canada, 6-11 July 2008.

[Cuff09] Paul Cuff. *Communication in Networks for Coordinating Behaviour*. PhD thesis, Stanford University, 2009.

[GW03] Matteo Gregoratti and Reinhard F. Werner. Quantum lost and found. *Journal of Modern Optics* 50(6&7):913–933, 2003.

[HV92] Te-Sun Han and Sergio Verdú. New Results in the Theory of Identification via Channels. *IEEE Transactions on Information Theory* 38(1):1417-25, 1992

[HV93] Te-Sun Han and Sergio Verdú. Approximation Theory of Output Statistics. *IEEE Transactions on Information Theory* 39(3):75217-772, 1993.

[HHT01] Patrick M. Hayden, Michał Horodecki, and Barbara M. Terhal. The asymptotic entanglement cost of preparing a quantum state. *Journal of Physics A: Mathematical and General* 34(35):6891–6898, 2001.

[HHHH09] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Reviews in Modern Physics* 81(2):865, 2009. doi:10.1103/RevModPhys.81.865.

[HHHO05] Karol Horodecki, Michał Horodecki, Paweł Horodecki, and Jonathan Oppenheim. Information theories with adversaries, intrinsic information, and entanglement. *Foundations of Physics* 35(12):2027–2040, 2005. doi:10.1007/s10701-005-8660-5.

[Kle63] Victor Klee. On a theorem of Dubins. *Journal of Mathematical Analysis and Applications* 7(3):425–427, 1963. doi:10.1016/0022-247X(63)90063-5.

[Mau93] Ueli M. Maurer. Secret Key Agreement by Public Discussion From Common Information. *IEEE Transactions on Information Theory* 39(3):733–742, 1993. doi:10.1109/18.256484.

[Roc96] R. Tyrell Rockafellar. *Convex Analysis*. Princeton Mathematical Series. Princeton University Press, 1996.

[RW03] Renato Renner and Stefan Wolf. New bounds in secret-key agreement: The gap between formation and secrecy extraction. In *Proceedings of Advances in Cryptology, EUROCRYPT 2003* (Eli Biham, ed.), *Lecture Notes in Computer Science*, vol. 2656, pp. 562–577. Springer Verlag, Berlin Heidelberg, 2003. doi:10.1007/3-540-39200-9_35.

[SVW05] John A. Smolin, Frank Verstraete, and Andreas Winter. Entanglement of assistance and multipartite state distillation. *Physical Review A* 72:052317, 2007.

[Win05] Andreas Winter. Secret, public and quantum correlation cost of triples of random variables. In *Proceedings of the International Symposium on Information Theory, ISIT 2005*, pp. 2270–2274, Adelaide, South Australia, 5-9 September 2005 doi:10.1109/ISIT.2005.1523752.

[Win07] Andreas Winter. On Environment-Assisted Capacities of Quantum Channels. *Markov Processes and Related Fields* 13(1-2):297–314, 2007.

[Wit76] Hans S. Witsenhausen. Values and bounds for the common information of two discrete random variables. *SIAM Journal of Applied Mathematics* 31:313–333, 1976. doi:10.1137/0131026.

[Wyn75] Aaron D. Wyner. The Common Information of Two Dependent Random Variables. *IEEE Transactions on Information Theory* 21(2):163–179, 1975. doi:10.1109/TIT.1975.1055346.