Secure Distributed Lossless Compression with Side Information at the Eavesdropper

Joffrey Villard and Pablo Piantanida
Department of Telecommunications, SUPELEC
91192 Gif-sur-Yvette, France
Email: {joffrey.villard,pablo.piantanida}@supelec.fr

Abstract

This paper investigates the problem of secure distributed lossless compression in the presence of arbitrarily correlated side information at an eavesdropper. This scenario consists of two encoders (referred to as Alice and Charlie) that wish to reliably transmit their respective (correlated) sources to a legitimate receiver (referred to as Bob) while satisfying some requirement on the equivocation rate at the eavesdropper (referred to as Eve). Error-free rate-limited channels are assumed between the encoders and the legitimate receiver, one of which being perfectly observed by the eavesdropper, which also has access to a correlated source as side information. For instance, this problem can be seen as a generalization of the well-known Slepian-Wolf problem taking into account the security requirements. A complete characterization of the compression-equivocation rates region for the case of arbitrarily correlated sources is derived. It is shown that the statistical differences between the sources can be useful in terms of secrecy.

I. Introduction

Consider the problem of compressing correlated sources at sensor nodes in a distributed fashion where the sensors may wish to communicate with a fusion center on a wireless network. The correlation between the observations can be used to minimize the rates needed for the communication between the sensors and the fusion center (referred to as Bob). In addition to this, we assume that the encoders wish to leak the least possible amount of information about their sources to an eavesdropper (referred to as Eve), *e.g.* an untrusted sensor, who may observe one of the channels and capture information during the communication.

The above scenario involves many of the major information-theoretic issues on source and channel coding problems. In terms of source coding, Slepian and Wolf [1] introduced the problem of distributed lossless compression. This topic has been the focus of intense study and some remarkable progress has already been made in theoretical and practical aspects. On the other hand, extensive research has been done during the recent years on secure communications over noisy channels. Shannon in [2] introduced the information-theoretic notion of secrecy,

The work of J. Villard is supported by DGA (French Armement Procurement Agency). This research is partially supported by the FP7 Network of Excellence in Wireless COMmunications NEWCOM++.

where security is measured through the equivocation rate (*i.e.* the remaining uncertainty about the message) at the eavesdropper. The wiretap channel was introduced by Wyner [3], who showed that it is possible to send information at a positive rate with perfect secrecy as long as the channel of the eavesdropper is a degraded version of the legitimate user's one. Csiszàr and Körner [4] extend this result to the setting of general broadcast channels with any arbitrary equivocation rate. Several extensions of the wiretap and fading channels have been done (cf. [5], [6] and references therein). So far, very few work has been reported on source coding (or compression) problems with security constraints.

One can identify two approaches in the literature on secure source coding. In fact, it is assumed either that there already exists a secure rate-limited channel between Alice and Bob, which allows the system to use secret keys, or the decoders have access to side information about the source. In the scenario of secret key sharing, both lossless and lossy compression have been studied in various contexts [7]–[12]. For the second scenario where side information is available at both decoders, the case of lossless and lossy source coding has been recently studied in [13] and [14], respectively. The distributed compression setup *i.e.*, source coding with *coded* side information, has been studied in [15]–[17]. In their "one-sided helper" scenario, Tandon *et al.* [17] provide a complete characterization of the achievable region when only one source is to be perfectly estimated, and Eve does not have any side information. Gunduz *et al.* [16] study secure compression in a setup similar to the one considered here but they only provide inner and outer bounds for the achievable region.

In this paper, we investigate the problem of secure distributed lossless compression of memoryless sources in the presence of an eavesdropper with correlated side information who may observe one of the channels, as it is shown in Fig. 1. In this setting the channels between encoders and decoders are assumed to be noiseless so that they cannot provide any advantage to increase security. Our goal is to understand the minimum amount of information that needs to be revealed to Eve. We provide a complete characterization of the compression-equivocation rates region for the case of arbitrarily correlated sources.

The organization of this paper is as follows. Section II states definitions along with the main results, while Section III provides a specialization in case of uncoded side information at Bob. The sketch of the proofs are relegated to Section IV. Finally, Section V summarizes the paper.

Notation

For any sequence $(x_i)_{i\in\mathbb{N}^*}$, notation x_k^n stands for the collection $(x_k, x_{k+1}, \dots, x_n)$. x_1^n is simply denoted by x^n. The cardinality of an alphabet is denoted by $\|\cdot\|$. Logarithms are taken in base 2 and denoted by $\log(\cdot)$. Entropy is denoted by $H(\cdot)$, and mutual information by $I(\cdot;\cdot)$. We denote strongly typical and conditional typical sets by $T_\delta^n(X)$ and $T_\delta^n(Y|x^n)$ resp., and use the so-called *Delta-Convention* [18]. Let X, Y and Z be three random variables on some alphabets with probability distribution p. If $p(x|y,z) = p(x|y)$ for each x, y, z, then X, Y and Z form a Markov chain, which is denoted by $X -\!\!\!\!\circ\!\!\!\!- Y -\!\!\!\!\circ\!\!\!\!- Z$.

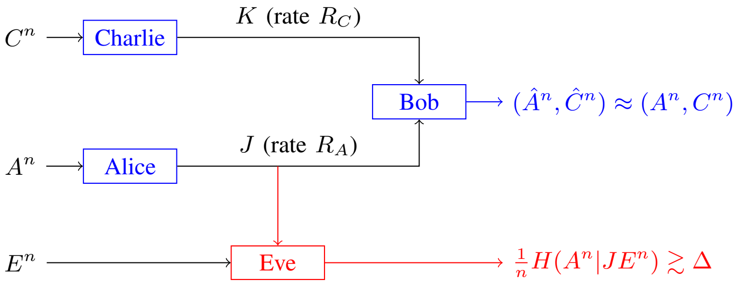
C^n
Charlie
K (rate R_C)
Bob
$(\hat{A}^n, \hat{C}^n) \approx (A^n, C^n)$
A^n
Alice
J (rate R_A)
E^n
Eve
$\frac{1}{n}H(A^n|JE^n) \gtrsim \Delta$

Figure 1. Distributed secure lossless compression with side information at the eavesdropper.

II. Definitions and Main Results

A. Problem Definition

In this section, we give a more rigorous formulation of the context depicted in Fig. 1. Let $\mathcal{A}$, $\mathcal{C}$ and $\mathcal{E}$ be three finite sets. Alice, Charlie and Eve observe sequences of random variables $(A_i)_{i\in\mathbb{N}^*}$, $(C_i)_{i\in\mathbb{N}^*}$ and $(E_i)_{i\in\mathbb{N}^*}$ respectively, which take values on $\mathcal{A}$, $\mathcal{C}$ and $\mathcal{E}$, resp. For each $i \in \mathbb{N}^*$, random variables A_i, C_i and E_i are distributed according to the joint distribution $p(a, c, e)$ on $\mathcal{A} \times \mathcal{C} \times \mathcal{E}$. Moreover, they are independent across time i.

Definition 1: An (n, R_A, R_C)-code for distributed compression in this setup is defined by

- An encoding function at Alice $f_A : \mathcal{A}^n \to \{1, \dots, 2^{nR_A}\}$,
- An encoding function at Charlie $f_C : \mathcal{C}^n \to \{1, \dots, 2^{nR_C}\}$,
- A decoding function at Bob $g : \{1, \dots, 2^{nR_A}\} \times \{1, \dots, 2^{nR_C}\} \to \mathcal{A}^n \times \mathcal{C}^n$.

Definition 2: A tuple $(R_A, R_C, \Delta) \in \mathbb{R}_+^3$ is said to be *achievable* if, for any $\varepsilon > 0$, there exists an $(n, R_A + \varepsilon, R_C + \varepsilon)$-code (f_A, f_C, g) such that:

$$\Pr\{g(f_A(A^n), f_C(C^n)) \neq (A^n, C^n)\} \leq \varepsilon \ ,$$
$$\frac{1}{n} H(A^n|f_A(A^n), E^n) \geq \Delta - \varepsilon \ .$$

The set of all such achievable tuples is denoted by $\mathcal{R}^*$ and is referred to as the *compression-equivocation rates region*.

Remark 1: Notice that the equivocation rate used in Definition 2 measures the uncertainty about the source A. In fact, our results can also apply when considering both sources A and C since the (joint) equivocation rate writes:

$$\frac{1}{n}H(A^n C^n|f_A(A^n), E^n) = \frac{1}{n}H(A^n|f_A(A^n), E^n) + \frac{1}{n}H(C^n|A^n E^n) \ .$$

Remark 2: Region $\mathcal{R}^*$ is closed and convex.

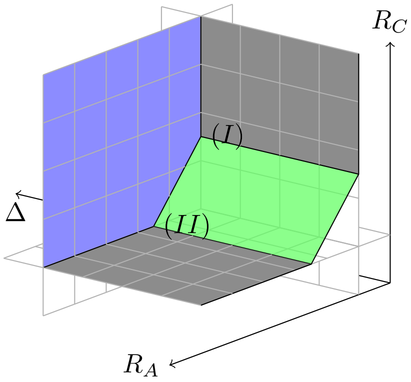

Figure 2. Achievable tuples (R_A, R_C, Δ).

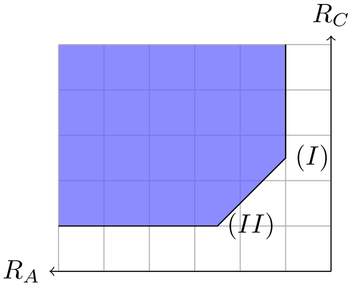

Figure 3. Projection on plane $\Delta = 0$.

B. Main Result

The following theorem provides a single-letter characterization of region $\mathcal{R}^*$.

Theorem 1: Region $\mathcal{R}^*$ writes as the closure of the set of all tuples $(R_A, R_C, \Delta) \in \mathbb{R}_+^3$ such that there exists a random variable U on some finite set $\mathcal{U}$ verifying the Markov chain $U -\!\!\circ\!\!- A -\!\!\circ\!\!- (C, E)$, and the following inequalities:

$$R_A \geq H(A|C) \ , \tag{1}$$

$$R_C \geq H(C|U) \ , \tag{2}$$

$$R_A + R_C \geq H(AC) \ , \tag{3}$$

$$\Delta \leq I(A;C|U) - I(A;E|U) \ , \tag{4}$$

The achievability of Theorem 1 is based on random binning at the encoders Alice and Charlie, and joint decoding at Bob. The detailed proof is relegated to Section IV-A. The above region can also be achieved using a time-sharing combination of two complementary families of codes. Since this approach may yield better intuition, its proof is sketched below. The proof of the converse part is given in Section IV-B.

Inequalities (1)–(3) resemble the ones of Slepian and Wolf [1, Section III]. They ensure perfect reconstruction of both variables A and C at Bob. The sum-rate constraint (3) captures the trade-off between rates R_A and R_C. The information must be transmitted by one or the other encoder. As a matter of fact, if there is no secrecy requirement, setting $U = A$ is optimal, and region $\mathcal{R}^*$ reduces to the Slepian-Wolf's one [1].

Let us now give some intuition on Equation (4). Depending on the distribution of (A, C, E), variable U can be tuned to make Bob *more capable* than Eve *i.e.*, maximize $I(A;C|U) - I(A;E|U)$. This quantity represents the gain (or the loss) at Eve in terms of equivocation rate. Note that Equation (4) also writes

$$\Delta \leq H(A|UE) - H(A|UC) \ .$$

The first term $H(A|UE)$ corresponds to the equivocation rate at Eve if she observes both variables U and E. Variable U is thus considered as a *common message i.e.*, as if Eve could decode it. As a matter of fact, Proposition 2 below shows that it is also optimal to encode U so that Eve can reliably estimate it.

At the same time, Equation (2) (which writes $R_C \geq H(C|A) + I(A;C|U)$) imposes a trade-off between the equivocation rate at Eve Δ and the rate of Charlie R_C. If the secrecy requirement is harsh, more information must be sent through the private channel (between Charlie and Bob).

The remaining rate of Alice (on the public channel) *i.e.*, $H(A|UC)$, is directly subtracted from the equivocation rate, meaning that it is treated as “raw” bits of A.

Remark 3: If the side information at Eve E is *less noisy* than C, then setting $U = A$ is optimal, and hence Slepian-Wolf coding achieves the whole region.

Sketch of proof of Theorem 1 (Time-sharing combination technique): We first construct two codes achieving corner points (I) and (II) illustrated in Fig. 2. Each corner point is achieved using a three-step communication scheme which aim is to reliably deliver variables (U, A) and C to Bob. At each step, the information previously received (and decoded) is used as side-information at Bob. *Random binning a la* Slepian-Wolf is performed to take advantage of this side information. These schemes correspond to the possible combinations of the set $\{U, A, C\}$, provided that U and C are decoded prior to A, as summarized in row #2 of Table I. For each scheme, the equivocation rate at Eve can be characterized following the argument of Appendix IV-A7. After Fourier-Motzkin elimination and classical manipulation, we can prove that the proposed schemes can achieve corner points (I) and (II), which coordinates are given in Table I.

As a matter of fact, points (I) and (II) correspond to identical equivocation rate level, say Δ. By a time-sharing combination of these schemes, each point on segment (I)–(II) is also achievable and presents equivocation rate Δ. Moreover, this segment can be easily described since the quantity $R_A + R_C$ is identical for both points (I) and (II) (see Fig. 3).

Segment (I)–(II) defines a region which is delimited by four hyperplanes given by the equations of Theorem 1.

■

The following proposition gives an upper bound on the cardinality of alphabet $\mathcal{U}$. The proof is given in Appendix A

Proposition 1: In the single-letter characterization of the compression-equivocation rates region $\mathcal{R}^*$ given by Theorem 1, it suffices to consider sets $\mathcal{U}$ such that $\|\mathcal{U}\| \leq \|\mathcal{A}\| + 1$.

Table I
THE TWO CORNER POINTS.

Corner point	(I)	(II)
Communication order	C, U, A	U, C, A
R_A	$H(A\|C)$	$I(U;A) + H(A\|UC)$
R_C	$H(C)$	$H(C\|U)$
Δ	$H(A\|UE) - H(A\|UC)$	$H(A\|UE) - H(A\|UC)$

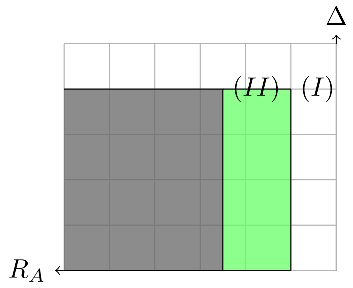

Figure 4. Projection on the plane $R_C = 0$.

Giving U to Eve is also optimal

With small changes in the proof of Theorem 1, we can prove the following proposition (see Appendix B).

Proposition 2: Region $\mathcal{R}^*$ writes as the closure of the set of all tuples $(R_A, R_C, \Delta) \in \mathbb{R}_+^3$ such that there exists a random variable U on some finite set $\mathcal{U}$ s.t. $U -\!\!\circ\!\!- A -\!\!\circ\!\!- (C, E)$ form a Markov chain and

$$\begin{aligned} R_A &\geq \big[I(U;C) - I(U;E)\big]_+ + H(A|C) \ , \\ R_C &\geq H(C|U) \ , \\ R_A + R_C &\geq H(AC) \ , \\ \Delta &\leq I(A;C|U) - I(A;E|U) \ . \end{aligned}$$

This new single-letter characterization means that it is also optimal to *help* Eve in decoding the auxiliary variable U. The corresponding additional rate $[I(U;C) - I(U;E)]_+$ does not lead to a lower equivocation rate at Eve. This should be considered with reference to known results on the wiretap channel [4], [6], where the so called *common message* can be chosen so that Eve also decodes it, without changing the achievable region.

III. Uncoded Side Information at Bob

In this paragraph, we consider the special case when Bob has access to *uncoded* side information *i.e.*, Bob and Charlie are collocated, or equivalently $R_C \to \infty$. The set of all achievable tuples in this setup is defined as:

$$\mathcal{R}^*_{\text{uncoded}} = \{(R_A, \Delta) : \exists\, R_C \text{ s.t. } (R_A, R_C, \Delta) \in \mathcal{R}^*\} \ .$$

The following corollary (which is a consequence of [14, Theorem 1]) directly follows from Theorem 1, removing constraints on R_C *i.e.*, from the optimality of point (I) (see Fig. 4).

Corollary 1: Region $\mathcal{R}^*_{\text{uncoded}}$ writes as the closure of the set of all tuples $(R_A, \Delta) \in \mathbb{R}_+^2$ such that there exists a random variable U on some finite set $\mathcal{U}$ s.t. $U -\!\!\circ\!\!- A -\!\!\circ\!\!- (C, E)$ form a Markov chain and

$$R_A \geq H(A|C) \ , \tag{5}$$

$$\Delta \leq I(A;C|U) - I(A;E|U) \ . \tag{6}$$

Remark 4: In case of a noiseless public channel of unlimited capacity *i.e.*, $R_A \to \infty$, Prabhakaran and Ramchandran studied in [13] the so-called *leakage rate*, defined as $\liminf \frac{1}{n} I(A^n; JE^n)$, which equals $H(A) - \Delta$. Their result "When Bob remains silent" [13, Theorem 1] thus follows as a special case of Corollary 1.

IV. Sketch of Proof of Theorem 1

A. Achievability

Let U be a random variable on some finite set $\mathcal{U}$, and $(R_A, R_C, \Delta) \in \mathbb{R}_+^3$. In this section, we describe a scheme which achieves tuple (R_A, R_C, Δ) under some sufficient conditions *i.e.*, for any $\varepsilon > 0$, we construct an $(n, R_A + \varepsilon, R_C + \varepsilon)$-code (f_A, f_C, g) such that:

$$\Pr\left\{g(f_A(A^n), f_C(C^n)) \neq (A^n, C^n)\right\} \leq \varepsilon \ ,$$

$$\frac{1}{n} H(A^n | f_A(A^n), E^n) \geq \Delta - \varepsilon \ .$$

In this scheme, Alice (resp. Charlie) perform random binning on U and A (resp. C). From their bin indices, Bob *jointly* decodes variables U and C, then A.

Let $\varepsilon > 0$, $R_1 > 0$, $R_2 > 0$ such that $R_1 + R_2 = R_A + \varepsilon$, and $S_1 \geq R_1$, $S_2 \geq R_2$, $S_C \geq R_C + \varepsilon$. Define $\gamma = \frac{\varepsilon}{8}$.

1) Codebook Generation at Alice: Randomly pick 2^{nS_1} sequences $u^n(s_1)$ from $T_\delta^n(U)$ and divide them into 2^{nR_1} equal size bins $\{B_1(r_1)\}_{r_1 \in \{1,\ldots,2^{nR_1}\}}$. Then, for each codeword $u^n(s_1)$, pick 2^{nS_2} sequences $a^n(s_1, s_2)$ from $T_\delta^n(A|u^n(s_1))$ and divide them into 2^{nR_2} equal size bins $\{B_2(s_1, r_2)\}_{r_2 \in \{1,\ldots,2^{nR_2}\}}$.

2) Codebook Generation at Charlie: Randomly pick 2^{nS_C} sequences $c^n(s)$ from $T_\delta^n(C)$ and divide them into $2^{n(R_C+\varepsilon)}$ equal size bins $\{B_C(r)\}_{r \in \{1,\ldots,2^{n(R_C+\varepsilon)}\}}$.

3) Encoding at Alice: Assume that sequence A^n is produced at Alice. Look for a codeword $u^n(s_1)$ such that $(u^n(s_1), A^n) \in T_\delta^n(U, A)$. Let $B_1(r_1)$ and $B_2(s_1, r_2)$ be the bins of $u^n(s_1)$ and $A^n = a^n(s_1, s_2)$, respectively. Alice sends the message $J = f_A(A^n) \triangleq (r_1, r_2)$ on her error-free channel.

4) Encoding at Charlie: Assume that sequence $C^n = c^n(s) \in B_C(r)$ is produced at Charlie. Charlie then sends the message $K = f_C(C^n) \triangleq r$ on his error-free channel.

5) Decoding at Bob: Assume that Bob receives $J = (r_1, r_2)$ from Alice and $K = r$ from Charlie. First look for the unique *jointly typical* sequences (u^n, c^n) with bin indices (r_1, r) *i.e.*, look for the unique indices (s_1, s) such that $(u^n(s_1), c^n(s)) \in (B_1(r_1) \times B_C(r)) \cap T_\delta^n(U, C)$. Then look for the unique index s_2 such that $a^n(s_1, s_2) \in B_2(s_1, r_2) \cap T_\delta^n(A|c^n(s))$. The estimate $g(J, K)$ is then defined as the decoded tuple $(a^n(s_1, s_2), c^n(s))$.

6) Errors and Constraints: Denoting by E the event "An error occurred during the encoding or decoding steps," we expand its probability (averaged over the set of all possible codebooks) as follows: $\Pr\{\mathsf{E}\} \leq P_t + P_e + P_d$, where each term corresponds to a particular error event, as detailed below. We derive sufficient conditions on the parameters that make each of these probabilities small.

1) From standard properties of typical sequences [18], there exists a sequence $\eta_n \xrightarrow[n\to\infty]{} 0$ such that $P_t \triangleq \Pr\{(A^n, C^n, E^n) \notin T_\delta^n(A, C, E)\} \leq \eta_n$. Consequently, $P_t \leq \gamma$ for some sufficiently large n.

2) In the first encoding step, Alice needs to find (at least) one codeword $u^n(s_1)$ such that $(u^n(s_1), A^n) \in T_\delta^n(U,A)$. Using standard properties of typical sequences [18], we can prove that, if $S_1 > I(U;A)$, then the corresponding error probability vanishes as n tends to infinity, and hence can be upper bounded by γ for some sufficiently large n.

   In the second encoding step, sequence A^n needs to appear in the 2^{nS_2} codewords $a^n(s_1,s_2) \in T_\delta^n(A|u^n(s_1))$. If $S_2 > H(A|U)$, this step will succeed with a probability larger than $1-\gamma$, for some sufficiently large n.

   Similarly, condition $S_C > H(C)$ is needed to ensure that the encoding step at Charlie succeeds with a probability larger than $1-\gamma$.

3) The decoding error probability P_d must be carefully handled. An error occurs when the decoded tuple differ from the original one (u,a,c). There are three meaningful possible events so that P_d writes:[1]

$$
\begin{aligned}
P_d &\triangleq \Pr\left\{\cancel{(c,u,a)}\right\} \\
&= \Pr\left\{\{\cancel{c},\check{u}\} \cup \{\check{c},\cancel{u}\} \cup \{\cancel{c},\cancel{u}\} \cup \{\check{c},\check{u},\cancel{a}\}\right\} \\
&\leq \Pr\{\cancel{c},\check{u}\} + \Pr\{\check{c},\cancel{u}\} + \Pr\{\cancel{c},\cancel{u}\} + \Pr\{\check{c},\check{u},\cancel{a}\} \ .
\end{aligned}
$$

   We now study each term of the r.h.s. of the above equation:

$$
\begin{aligned}
\Pr\{\cancel{c},\check{u}\} &= \Pr\left\{\exists\ s' \neq s \text{ s.t. } (u^n(s_1), c^n(s')) \in (B_1(r_1) \times B_C(r)) \cap T_\delta^n(U,C)\right\} \\
&\leq 2^{n(S_C - R_C - \varepsilon)} \Pr\left\{(U^n,C^n) \in T_\delta^n(U,C) \middle| \ U^n \in T_\delta^n(U), C^n \in T_\delta^n(C)\right\} \\
&\leq 2^{n(S_C - R_C - \varepsilon)}\ 2^{-n(I(U;C) - \eta_n)} \ ,
\end{aligned}
$$

   for some sequence $\eta_n \xrightarrow[n\to\infty]{} 0$. If $S_C - R_C - \varepsilon < I(U;C)$, then the above probability vanishes as n tends to infinity, and hence can be upper bounded by γ for some sufficiently large n.

   Similarly, if $S_1 - R_1 < I(U;C)$ (resp. $S_1 - R_1 + S_C - R_C - \varepsilon < I(U;C)$), then $\Pr\{\cancel{c},\check{u}\} \leq \gamma$ (resp. $\Pr\{\cancel{c},\cancel{u}\} \leq \gamma$) for some sufficiently large n.

$$
\begin{aligned}
\Pr\{\check{c},\check{u},\cancel{a}\} &= \Pr\left\{\exists\ s_2' \neq s_2 \text{ s.t. } a^n(s_1,s_2') \in B_2(s_1,r_2) \cap T_\delta^n(A|c^n(s))\right\} \\
&\leq 2^{n(S_2 - R_2)} \Pr\left\{(A^n,C^n) \in T_\delta^n(A,C) \middle| \ A^n \in T_\delta^n(A|u^n(s_1)), C^n \in T_\delta^n(C|u^n(s_1))\right\} \\
&\leq 2^{n(S_2 - R_2)}\ 2^{-n(I(A;C|U) - \eta_n)} \ ,
\end{aligned}
$$

   for some sequence $\eta_n \xrightarrow[n\to\infty]{} 0$. If $S_2 - R_2 < I(A;C|U)$, then the above probability vanishes as n tends to infinity, and hence $\Pr\{\check{c},\check{u},\cancel{a}\} \leq \gamma$, for some sufficiently large n.

In this paragraph, we found sufficient conditions that ensure $\Pr\{\mathsf{E}\} \leq 8\gamma$ *i.e.*, $\Pr\{g(f_A(A^n), f_C(C^n)) \neq (A^n,C^n)\} \leq \varepsilon$, for some sufficiently large n.

[1] We denote by $\check{x}$ the event "Sequence x^n has been correctly decoded", and $\cancel{x}$ its complement. Same notation holds for tuples.

7) Equivocation Rate at Eve: The equivocation rate at Eve can be lower bounded as follows:

$$\begin{aligned}\frac{1}{n}H(A^n|f_A(A^n),E^n) &\geq \frac{1}{n}H(A^n|r_1r_2E^n)\\ &= \frac{1}{n}\Big[H(A^n|r_1E^n) - I(A^n;r_2|r_1E^n)\Big]\\ &\overset{(a)}{\geq} \frac{1}{n}\Big[H(A^n|u^n(s_1)E^n) - H(r_2)\Big]\\ &\overset{(b)}{\geq} H(A|UE) - R_2\ ,\end{aligned}$$

where

- step (a) follows from the facts that bin index r_1 is a deterministic function of codeword $u^n(s_1)$, bin index r_2 is a deterministic function of A^n, and conditioning reduces the entropy,
- step (b), from the fact that the random variables are i.i.d., and $r_2 \in \{1,\ldots,2^{nR_2}\}$.

Condition $\Delta - \varepsilon \leq H(A|UE) - R_2$ is thus sufficient to achieve equivocation rate $\Delta - \varepsilon$ at Eve.

8) End of Proof: In this section, we proved that sufficient conditions for the achievability of a tuple (R_A, R_C, Δ) are given by the following system of inequalities, for each $\varepsilon > 0$:

$$\left\{\begin{array}{rcl} R_1 & > & 0\\ R_2 & > & 0\\ R_A+\varepsilon & = & R_1+R_2\\ R_C & \geq & 0\\ S_1 & \geq & R_1\\ S_2 & \geq & R_2\\ S_C & \geq & R_C+\varepsilon\\ S_1 & > & I(U;A)\\ S_2 & > & H(A|U)\\ S_C & > & H(C)\\ S_C-R_C-\varepsilon & < & I(U;C)\\ S_1-R_1 & < & I(U;C)\\ S_1-R_1+S_C-R_C-\varepsilon & < & I(U;C)\\ S_2-R_2 & < & I(A;C|U)\\ R_2+\Delta-\varepsilon & \leq & H(A|UE) \end{array}\right.$$

Fourier-Motzkin elimination then yields:

$$\left\{\begin{array}{rcl} R_A+\varepsilon & > & H(A|C)\\ R_C+\varepsilon & > & H(C|U)\\ R_A+R_C+2\varepsilon & > & H(AC)\\ \Delta-\varepsilon & < & I(A;C|U)-I(A;E|U) \end{array}\right.$$

This proves the achievability part of Theorem 1.

B. Converse

Let (R_A, R_C, Δ) be an achievable tuple and $\varepsilon > 0$. There exists an $(n, R_A + \varepsilon, R_C + \varepsilon)$-code (f_A, f_C, g) s.t.:

$$\begin{aligned} \Pr\left\{g(f_A(A^n), f_C(C^n)) \neq (A^n, C^n)\right\} &\leq \varepsilon \ , \\ \frac{1}{n} H(A^n | f_A(A^n), E^n) &\geq \Delta - \varepsilon \ . \end{aligned}$$

Denote by $J = f_A(A^n)$ and $K = f_C(C^n)$ the messages transmitted by Alice and Charlie, respectively. For each $i \in \{1, \ldots, n\}$, define random variable U_i as follows:

$$U_i = (J, C_{i+1}^n, E^{i-1}) \ . \tag{7}$$

Note that $U_i \mathrel{-\!\!\circ\!\!-} A_i \mathrel{-\!\!\circ\!\!-} (C_i, E_i)$ form a Markov chain.

Following the usual technique, we also define an independent random variable Q uniformly distributed over the set $\{1, \ldots, n\}$, and $A = A_Q$, $C = C_Q$, $E = E_Q$, $U = (Q, U_Q)$. Note that $U \mathrel{-\!\!\circ\!\!-} A \mathrel{-\!\!\circ\!\!-} (C, E)$ still form a Markov chain, and that (A, C, E) is distributed according to the joint distribution $p(a, c, e)$ *i.e.*, the original distribution of (A_i, C_i, E_i).

1) Rate at Alice: Following the argument of the converse for the Slepian-Wolf theorem [19, Section 15.4.2], we prove lower bounds on the rates:

$$\begin{aligned} n(R_A + \varepsilon) &\geq H(J) \\ &\overset{(a)}{\geq} H(J|C^n) \\ &\overset{(b)}{=} I(A^n; J|C^n) \\ &\overset{(c)}{=} H(A^n|C^n) - H(A^n|JKC^n) \\ &\overset{(d)}{\geq} nH(A|C) - nO(\varepsilon) \ , \end{aligned} \tag{8}$$

where

- step (a) follows from the fact that conditioning reduces the entropy,
- step (b) from $J = f_A(A^n)$,
- step (c) from $K = f_C(C^n)$,
- step (d) from the fact that random variables A_i and C_i are i.i.d., and Fano's inequality[2].

[2] Landau-like notation $O(\varepsilon)$ stands for a term X such that $0 \leq X \leq k\varepsilon$ for some constant $k > 0$.

2) Rate at Charlie: Using similar arguments with $K = f_C(C^n)$, we can obtain:

$$\begin{aligned} n(R_C + \varepsilon) &\geq H(K) \\ &\overset{(a)}{\geq} H(K|J) \\ &\overset{(b)}{=} I(K; C^n|J) \\ &= H(C^n|J) - H(C^n|JK) \\ &\overset{(c)}{\geq} \sum_{i=1}^{n} H(C_i|JC_{i+1}^n) - nO(\varepsilon) \\ &\overset{(d)}{\geq} \sum_{i=1}^{n} H(C_i|U_i) - nO(\varepsilon) \ , \end{aligned}$$

where

- step (a) follows from the fact that conditioning reduces the entropy,
- step (b) from $K = f_C(C^n)$,
- step (c) from the chain rule for conditional entropy, the fact that random variables C_i's are independent across time, and Fano's inequality,
- step (d) from the fact that conditioning reduces the entropy, and definition (7).

Now, using auxiliary random variable Q,

$$\begin{aligned} R_C + \varepsilon &\geq \frac{1}{n} \sum_{i=1}^{n} H(C_Q|U_Q, Q = i) - O(\varepsilon) \\ &= H(C|U) - O(\varepsilon) \ . \end{aligned}$$

3) Sum-Rate: A lower bound on the sum-rate can be derived as well:

$$\begin{aligned} n(R_A + R_C + 2\varepsilon) &\geq H(JK) \\ &\overset{(a)}{=} I(A^n C^n; JK) \\ &= H(A^n C^n) - H(A^n C^n|JK) \\ &\overset{(b)}{\geq} nH(AC) - nO(\varepsilon) \ , \end{aligned}$$

where

- step (a) follows from $J = f_A(A^n)$ and $K = f_C(C^n)$,
- step (b) from the fact that random variables A_i and C_i are i.i.d., and Fano's inequality.

4) Equivocation Rate at Eve:

$$
\begin{aligned}
&n(\Delta - \varepsilon) \\
&\qquad \le H(A^n|JE^n) \\
&\qquad = H(A^n|JK) + I(A^n;K|J) - I(A^n;E^n|J) \\
&\qquad \overset{(a)}{\le} nO(\varepsilon) + I(A^n;C^n|J) - I(A^n;E^n|J) \\
&\qquad \overset{(b)}{=} nO(\varepsilon) + I(A^n;C^n) - I(J;C^n) - I(A^n;E^n) + I(J;E^n) \\
&\qquad \overset{(c)}{=} nO(\varepsilon) + \sum_{i=1}^{n} \Big[I(A_i;C_i) - I(JC_{i+1}^n;C_i) - I(A_i;E_i) + I(JE^{i-1};E_i) \Big] \\
&\qquad \overset{(d)}{=} nO(\varepsilon) + \sum_{i=1}^{n} \Big[I(A_i;C_i) - I(JC_{i+1}^n;C_i) - I(A_i;E_i) + I(JE^{i-1};E_i) \\
&\qquad\qquad\qquad\qquad + I(E_i;C_{i+1}^n|JE^{i-1}) - I(C_i;E^{i-1}|JC_{i+1}^n) \Big] \\
&\qquad = nO(\varepsilon) + \sum_{i=1}^{n} \Big[I(A_i;C_i) - I(JC_{i+1}^n E^{i-1};C_i) - I(A_i;E_i) + I(JC_{i+1}^n E^{i-1};E_i) \Big] \\
&\qquad \overset{(e)}{=} nO(\varepsilon) + \sum_{i=1}^{n} \Big[I(A_i;C_i|U_i) - I(A_i;E_i|U_i) \Big] \ ,
\end{aligned}
$$

where

- step (a) follows from Fano's inequality, and $K = f_C(C^n)$,
- step (b) from the Markov chain $J \mathrel{-\!\!\circ\!\!-} A^n \mathrel{-\!\!\circ\!\!-} (C^n, E^n)$,
- step (c) from the chain rule for mutual information, and the fact that random variables A_i, C_i, and E_i are independent across time,
- step (d) from Csiszár and Körner equality [4],
- step (e) from definition (7), and the Markov chain $U_i \mathrel{-\!\!\circ\!\!-} A_i \mathrel{-\!\!\circ\!\!-} (C_i, E_i)$.

Using auxiliary random variable Q, we can prove that

$$\Delta - \varepsilon \le I(A;C|U) - I(A;E|U) + O(\varepsilon) \ .$$

C. End of Proof

We proved that, for each achievable tuple (R_A, R_C, Δ) and each $\varepsilon > 0$, there exists a random variable U such that $U \mathrel{-\!\!\circ\!\!-} A \mathrel{-\!\!\circ\!\!-} (C, E)$ form a Markov chain, and

$$
\begin{aligned}
R_A + O(\varepsilon) &\ge H(A|C) \ , \\
R_C + O(\varepsilon) &\ge H(C|U) \ , \\
R_A + R_C + O(\varepsilon) &\ge H(AC) \ , \\
\Delta - O(\varepsilon) &\le I(A;C|U) - I(A;E|U) \ .
\end{aligned}
$$

Letting ε tend to zero proves the converse part of Theorem 1.

V. Summary and Discussions

The problem of secure distributed compression of memoryless sources in the presence of an eavesdropper with correlated side information was investigated. A complete characterization of the compression-equivocation rates region was derived for the case of arbitrarily correlated sources. It was shown that the statistical properties of the sources can be exploited by the encoders to increase the equivocation rate at the eavesdropper.

As future and on-going work, it would be of interest to extend the results in the present work to the more general setting in which the legitimate decoder wishes to estimate the sources within certain distortion criteria *i.e.*, secure distributed *lossy* source coding in a setup similar to the Berger-Tung one [20].

Appendix A
Proof of Proposition 1

The proof of Proposition 1 follows standard cardinality bounding arguments [21, Appendix C]. We first rewrite the inequalities of Theorem 1 involving variable U:

$$\begin{aligned} R_C &\geq H(C|U) \ , \\ \Delta &\leq H(C|U) - H(C|A) - I(A;E|U) \ . \end{aligned}$$

Then consider the following $\|\mathcal{A}\| + 1$ continuous functions of $p(a|u)$:

$$\begin{aligned} &p(a|u) \ , \\ &H(C|U=u) \ , \\ &I(A;E|U=u) \ . \end{aligned}$$

From Fenchel-Eggleston-Carathéodory's theorem, there exists a random variable U' on $\mathcal{U}'$ with $\|\mathcal{U}'\| \leq \|\mathcal{A}\| + 1$ such that $p(a)$, $H(C|U)$, and $I(A;E|U)$ are preserved.

This proves Proposition 1.

Appendix B
Sketch of Proof of Proposition 2

A. Achievability

The proof of the achievability part follows the same argument that Appendix IV-A. A new constraint is added on the size of each bin $B_1(r_1)$ to the system of Section IV-A8:

$$S_1 - R_1 < I(U;E) \ .$$

This inequality ensures that Eve can reliably decode $u^n(s_1)$ from the bin index r_1 and her side information E^n. Fourier-Motzkin-Elimination then yields the expected inequalities.

B. Converse

The proof of the converse part follows the same argument that Appendix IV-B. In particular, definition (7) remains the same. The only difference lies in the lower bound for the rate:

$$\begin{aligned}
n(R_A+\varepsilon) &\geq H(J) \\
&\overset{(a)}{=} I(J;A^n|C^n)+I(J;C^n) \\
&\overset{(b)}{=} H(A^n|C^n)-H(A^n|JKC^n)+I(J;C^n) \\
&\overset{(c)}{\geq} -nO(\varepsilon)+\sum_{i=1}^{n}\Big[H(A_i|C_i)+I(JC_{i+1}^n;C_i)\Big] \\
&\overset{(d)}{=} -nO(\varepsilon)+\sum_{i=1}^{n}\Big[H(A_i|C_i)+I(JC_{i+1}^n;C_i)+I(E^{i-1};C_i|JC_{i+1}^n)-I(C_{i+1}^n;E_i|JE^{i-1})\Big] \\
&\overset{(e)}{\geq} -nO(\varepsilon)+\Big[\sum_{i=1}^{n}H(A_i|C_i)+I(JC_{i+1}^nE^{i-1};C_i)-I(JC_{i+1}^nE^{i-1};E_i)\Big] \\
&\overset{(f)}{=} -nO(\varepsilon)+\sum_{i=1}^{n}\Big[H(A_i|C_i)+I(U_i;C_i)-I(U_i;E_i)\Big]\ ,
\end{aligned}$$

where

- step (a) follows from $J=f_A(A^n)$,
- step (b) from $K=f_C(C^n)$,
- step (c) from Fano's inequality, the chain rule for conditional mutual information and the fact that random variables A_i, C_i are independent across time,
- step (d) from Csiszár and Körner equality [4],
- step (e) from the independence of the random variables $(A_k)_k$, $(C_k)_k$ and $(E_k)_k$ across time and the non-negativity of mutual information,
- step (f) from definition (7).

Using random variable Q and following the argument of Appendix IV-B, we proved the following lower bound:

$$R+\varepsilon \geq H(A|C)+I(U;C)-I(U;E)-O(\varepsilon)\ .$$

Since Equation (8) still holds, we proved the bound on R_A given by Proposition 2. Other steps of the proof remain unchanged.

REFERENCES

[1] D. Slepian and J. Wolf. Noiseless coding of correlated information sources. *IEEE Trans. Inf. Theory*, 19(4):471–480, 1973.

[2] C.E. Shannon. Communication theory of secrecy systems. *BSTJ*, 28:656–715, 1949.

[3] A.D. Wyner. The wire-tap channel. *BSTJ*, 54(8):1355–1387, 1975.

[4] I. Csiszar and J. Korner. Broadcast channels with confidential messages. *IEEE Trans. Inf. Theory*, 24(3):339–348, 1978.

[5] Special issue on information theoretic security. *IEEE Trans. Inf. Theory*, 54(6):2405–2818, 2008.

[6] Y. Liang, H.V. Poor, and S. Shamai. *Information theoretic security*. Now Publishers, 2009.

[7] H. Yamamoto. A source coding problem for sources with additional outputs to keep secret from the receiver or wiretappers. *IEEE Trans. Inf. Theory*, 29(6):918–923, 1983.

[8] H. Yamamoto. A rate-distortion problem for a communication system with a secondary decoder to be hindered. *IEEE Trans. Inf. Theory*, 34(4):835–842, 1988.

[9] H. Yamamoto. Coding theorems for Shannon's cipher system with correlated source outputs, and common information. *IEEE Trans. Inf. Theory*, 40(1):85–95, 1994.

[10] H. Yamamoto. Rate-distortion theory for the Shannon cipher system. *IEEE Trans. Inf. Theory*, 43(3):827–835, 1997.

[11] R. Liu and W. Trappe. *Securing wireless communications at the physical layer*. Springer, 2010.

[12] N. Merhav. On the Shannon cipher system with a capacity-limited key-distribution channel. *IEEE Trans. Inf. Theory*, 52(3):1269–1273, 2006.

[13] V. Prabhakaran and K. Ramchandran. On secure distributed source coding. In *Proc. ITW*, pages 442–447, 2007.

[14] J. Villard and P. Piantanida. Secure lossy source coding with side information at the decoders. In *Proc. Allerton*, 2010.

[15] D. Gunduz, E. Erkip, and H.V. Poor. Secure lossless compression with side information. In *Proc. ITW*, pages 169–173, 2008.

[16] D. Gunduz, E. Erkip, and H.V. Poor. Lossless compression with security constraints. In *Proc. ISIT*, pages 111–115, 2008.

[17] R. Tandon, S. Ulukus, and K. Ramchandran. Secure source coding with a helper. In *Proc. Allerton*, pages 1061–1068, 2009.

[18] I. Csiszar and J. Körner. *Information theory: coding theorems for discrete memoryless systems*. Akadémiai Kiado, Budapest, 1982.

[19] T.M. Cover and J.A. Thomas. *Elements of information theory (2nd Ed)*. Wiley-Interscience, 2006.

[20] T. Berger. Multiterminal source coding. In G. Longo, editor, *The information theory approach to communications*. Springer-Verlag, 1977.

[21] A. El Gamal and Y.-H. Kim. *Lecture Notes on Network Information Theory*. arXiv:1001.3404, 2010.