# A Sufficient Condition for Private Information Hiding of Two Correlated Sources Under Multiple Access Attacks*

Yadong Wang, Yangfan Zhong, Fady Alajaji and Tamás Linder
Department of Mathematics and Statistics
Queen's University, Kingston, ON, K7L 3N6, Canada
Email: {yadong, yangfan, fady, linder}@mast.queensu.ca

***Abstract*— Consider a multi-user private information-hiding scenario in which two information hiders separately embed correlated sources $(S_1, S_2)$ into a common host source $U$ (covertext). The $i$-th information hider embeds the secret source $S_i$ into the covertext $U$ subject to a distortion constraint $D_i$ $(i = 1, 2)$. The outputs (stegotexts) are corrupted by a multiple access channel attack $W_{Y|X_1X_2}$. A sufficient condition (in single-letter form) under which $(S_1, S_2)$ can be successfully embedded into $U$ under $W_{Y|X_1X_2}$ is established.**

## I. Introduction

Information hiding is the means to embed a secret message into a host message (covertext) so that the information hider is able to transmit the information even though the transmission is subject to manipulation by an attacker attempting to render the hidden information undetectable. A large body of literature including theoretical studies as well as various practical applications have recently been devoted to this new area (see, e.g., [1]–[8] and the references therein).

In the literature, the information-hiding scenario is usually modeled as a constrained channel coding problem. The secret messages, assumed to be uniformly distributed over a given message set, are one-by-one embedded into the host messages. Since the secret messages should not interfere perceptually with the host messages, a distortion constraint is placed between the encoder output and the original host messages. From an information-theoretic point of view, the problem is to find the largest embedding rate (known as embedding capacity) for which, at the encoder, the distortion between the host messages and the output (stegotexts) does not exceed a preset threshold, and at the decoder, the secret messages can be reproduced with an arbitrarily small probability of error.

In practical situations (e.g., instant (online) data-hiding), in order to reduce the complexity of coding, we may need to directly hide an information source (or correlated sources) with a nonuniform distribution. In this work we extend the point-to-point information-hiding model to a multi-user setting. Our model is depicted in Fig. 1. Instead of embedding uniformly distributed indices, two encoders independently embed two (arbitrarily distributed) discrete memoryless correlated sources $(S_1, S_2)$ into a common memoryless host source $U$, and transmit the resulting sequences to a common destination in the presence of discrete memoryless multiple access channel (MAC) attacks. One possible application of this scenario is that two agents separately embed noisy observations of the same source, and transmit the hidden information over a MAC attack channel. Throughout the paper, we deal with private information hiding; i.e., we assume that the decoder has perfect knowledge of $U$.

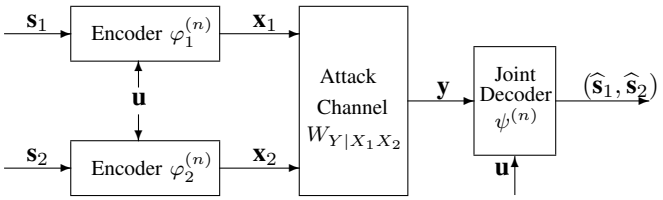


Fig. 1. A joint source coding and embedding model for multi-user information hiding.

Given the secret sources $(S_1, S_2)$, a MAC $W_{Y|X_1X_2}$, the host source $U$, and a distortion level pair $(D_1, D_2)$, one may ask whether there exists a coding scheme, such that $(S_1, S_2)$ can be embedded in $U$ within distortion levels $(D_1, D_2)$, and transmitted over $W_{Y|X_1X_2}$ with an arbitrarily small probability of error. To begin, we note that, especially in a multi-user system, jointly source coding and embedding the sources $(S_1, S_2)$ into $U$ might perform better than the traditional separate coding (i.e., concatenating lossless data compression and embedding). In this paper we investigate whether $(S_1, S_2)$ can be successfully transmitted under the MAC attacks by joint source coding and embedding codes. In particular, we establish a sufficient condition for successfully embedding $(S_1, S_2)$ into $U$ under the MAC $W_{Y|X_1X_2}$; see Theorem 1. Since our model jointly deals with information embedding (privacy protection) and compression of (nonuniform) correlated secret sources in a multiuser setting, it adds a new dimension to the traditional point-to-point information hiding problem. Note also that our problem can be viewed as a generalization of the problem of transmitting correlated sources over ordinary MAC channels [9]–[12].

## II. Preliminaries

Throughout, random variables (RV's) are denoted by capital letters, e.g., $X$, specific values are denoted by lower case letters, e.g., $x$, and their alphabets are denoted by calligraphic

*This work was supported in part by NSERC of Canada.

letters, e.g., $\mathcal{X}$. Similarly, random vectors are denoted by capital letters superscripted by their lengths, e.g., $X^n$, their alphabets are denoted by calligraphic letters superscripted by their lengths, e.g., $\mathcal{X}^n$, and their realizations are denoted by boldface lower case letters, e.g., $\mathbf{x} \triangleq (x_1, x_2, ..., x_n) \in \mathcal{X}^n$. The cardinality of a finite set $\mathcal{X}$ is denoted by $|\mathcal{X}|$. $\mathbb{E}(X)$ denotes the expectation of $X$. For any RV $X$, $P_X(x)$ denotes the probability that $X = x$. For jointly distributed RV's $X$ and $U$, $P_{X|U}(x|u)$ denotes the conditional probability of $X = x$ given that $U = u$. The probability of an independent and identically distributed (i.i.d.) sequence $\mathbf{x} \in \mathcal{X}^n$ is given by $P_{X^n}(\mathbf{x}) \triangleq \prod_{j=1}^n P_X(x_j)$. All alphabets are finite, and all logarithms are in natural base.

Let $V \triangleq (X_1, X_2, ..., X_m)$ be a superletter (an ordered collection of RV's) taking values in a finite set $\mathcal{V} \triangleq \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_m$ with joint distribution $P_V(x_1, ..., x_m)$, which for simplicity we also denote by $P_V(v)$. Denote by $T_\epsilon^{(n)}(V)$ or $T_\epsilon^{(n)}$ the set of all strongly $\epsilon$-typical sequences $(\mathbf{x}_1, \ldots, \mathbf{x}_m)$ [13] with respect to the joint distribution $P_V(v)$. Let $I_V \triangleq \{1, 2, ..., m\}$, and $I_G \subseteq I_V$. We then let $G = (X_{g_1}, X_{g_2}, ..., X_{g_{|I_G|}}) \in \mathcal{G}$ be a "sub-superletter" corresponding to $I_G$ such that $g_i \in I_G$. Let $G$, $K$, and $L$ be sub-superletters of $V$ such that $I_G$, $I_K$, $I_L$ are disjoint, and let $P_G$, $P_K$ and $P_{G|K}$ be the marginal and conditional distributions induced by $P_V$, respectively. Denote by $T_\epsilon^{(n)}(G)$ the restriction of $T_\epsilon^{(n)}(V)$ to the coordinates of $G$ [13]. Given $\mathbf{k} \in T_\epsilon^{(n)}(K)$, denote $T_\epsilon^{(n)}(G|\mathbf{k}) \triangleq \Big\{\mathbf{g} \in \mathcal{G}^n : (\mathbf{g}, \mathbf{k}) \in T_\epsilon^{(n)}(G, K)\Big\}$. $T_\epsilon^{(n)}(G|\mathbf{k})$ is sometimes referred to as $T_\epsilon^{(n)}(\cdot|\mathbf{k})$ if $G$ is clear from the context.

*Lemma 1:* [13, pp. 342-343] Let $G^n$, $K^n$ and $V^n$ be i.i.d. drawn according to $P_{G^n}$, $P_{K^n}$ and $P_{V^n}$. The following properties hold for sufficiently large $n$.

1) $P_{V^n}\Big(\{V^n \in T_\epsilon^{(n)}\}\Big) \geq 1 - \epsilon$, and $P_{K^n}\Big(\{K^n \in T_\epsilon^{(n)}(K)\}\Big) \geq 1 - \epsilon$.
2) For any $\mathbf{k} \in T_\epsilon^{(n)}(K)$, $|(1/n)\log P_{K^n}(\mathbf{k}) + H(K)| \leq \epsilon$.
3) $(1-\epsilon)e^{n(H(K)-\epsilon)} \leq \Big|T_\epsilon^{(n)}(K)\Big| \leq e^{n(H(K)+\epsilon)}$.
4) For any $\mathbf{k} \in T_\epsilon^{(n)}(K)$, $P_{G^n}\Big(\{G^n \in T_\epsilon^{(n)}(G|\mathbf{k})\}\Big) \geq 1 - \epsilon$.
5) For any $\mathbf{k} \in T_\epsilon^{(n)}(K)$, $(1-\epsilon)e^{n(H(G|K)-2\epsilon)} \leq \Big|T_\epsilon^{(n)}(G|\mathbf{k})\Big| \leq e^{n(H(G|K)+2\epsilon)}$.
6) For $(\mathbf{g}, \mathbf{k}) \in T_\epsilon^{(n)}(G, K)$, $\big|(1/n)\log P_{G^n|K^n}(\mathbf{g}|\mathbf{k}) + H(G|K)\big| \leq 2\epsilon$.

*Lemma 2:* (Markov Lemma [13, p. 579]) Let $G \to K \to L$ form a Markov chain in this order. For $0 < \epsilon < 1$ and $(\mathbf{g}, \mathbf{k}) \in T_\epsilon^{(n)}(G, K)$, if $L^n$ is i.i.d. drawn given $k$ according to $\prod_{j=1}^n P_{L|K}(l_j|k_j)$, then $\Pr\Big\{(\mathbf{g}, \mathbf{k}, L^n) \in T_\epsilon^{(n)}(G, K, L)\Big\} > 1 - \epsilon$ for $n$ sufficiently large.

## III. Problem Formulation and Main Results

Let the pair of memoryless correlated secret sources $\{(S_{1j}, S_{2j})\}_{j=1}^\infty$ have marginal distribution $P_{S_1S_2}$ and denote the marginal distribution of the host source $\{U_j\}_{j=1}^\infty$ by $P_U$. Assume $(S_1, S_2)$ and $U$ are independent. The attack channel is modeled as a two-sender one-receiver discrete memoryless MAC $W_{Y|X_1X_2}$ having input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$, output alphabet $\mathcal{Y}$, and a transition probability distribution $W_{Y|X_1X_2}(y|x_1, x_2)$. The probability of receiving $\mathbf{y} \in \mathcal{Y}^n$ conditioned on sending $\mathbf{x}_1 \in \mathcal{X}_1^n$ and $\mathbf{x}_2 \in \mathcal{X}_2^n$ is hence given by $W_{Y^n|X_1^nX_2^n}(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2) = \prod_{j=1}^n W_{Y|X_1X_2}(y_j|x_{1j}, x_{2j})$. Let $d_i : \mathcal{U} \times \mathcal{X}_i \to [0, \infty)$ be single-letter distortion measures and define $d_i^{max} \triangleq \max_{u, x_i} d_i(u, x_i)$ for $i = 1, 2$. For $\mathbf{u} \in \mathcal{U}^n$ and $\mathbf{x}_i \in \mathcal{X}_i^n$, let $d_i(\mathbf{u}, \mathbf{x}_i) = \sum_{j=1}^n d_i(u_j, x_{ij})$.

A joint source coding and embedding (JSCE) code $(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})$ with block length $n$ consists of two encoders $\varphi_1^{(n)} : \mathcal{S}_1^n \times \mathcal{U}^n \to \mathcal{X}_1^n$ and $\varphi_2^{(n)} : \mathcal{S}_2^n \times \mathcal{U}^n \to \mathcal{X}_2^n$ and a decoder $\psi^{(n)} : \mathcal{Y}^n \times \mathcal{U}^n \to \mathcal{S}_1^n \times \mathcal{S}_2^n$; see Fig. 1. The probability of error in reproducing the secret sources is given by

$$P_e^{(n)} = \sum_{\mathcal{S}_1^n \times \mathcal{S}_2^n \times \mathcal{U}^n} P_{S_1^nS_2^n}(\mathbf{s}_1, \mathbf{s}_2) P_{U^n}(\mathbf{u}) \sum_{\mathbf{y}:\psi^{(n)}(\mathbf{y},\mathbf{u}) \neq (\mathbf{s}_1, \mathbf{s}_2)} W_{Y^n|X_1^nX_2^n}(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2)$$

where $\mathbf{x}_i \triangleq \varphi_i^{(n)}(\mathbf{s}_i, \mathbf{u})$ $(i = 1, 2)$.

*Definition 1:* Given $P_U$ and distortion levels $D_1 > 0$ and $D_2 > 0$, we say that the secret sources $\{(S_{1j}, S_{2j})\}$ are $(D_1, D_2)$-*admissible* with respect to the MAC $W_{Y|X_1X_2}$, if there exists a sequence of codes $(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})$ such that $P_e^{(n)} \to 0$ as $n \to \infty$ and for any $\delta > 0$, $\frac{1}{n}\mathbb{E}[d_i(U^n, X_i^n)] \leq D_i + \delta$, $i = 1, 2$, for $n$ sufficiently large.

*Theorem 1:* $\{(S_{1j}, S_{2j})\}$ are $(D_1, D_2)$-*admissible* with respect to the MAC $W_{Y|X_1X_2}$ if there exist some RV $Q$ and a pair of conditional distributions $(P_{X_1|S_1UQ}, P_{X_2|S_2UQ})$ such that

$$H(S_1|S_2) < I(X_1; Y|X_2, S_2, U, Q), \quad (1)$$
$$H(S_2|S_1) < I(X_2; Y|X_1, S_1, U, Q), \quad (2)$$
$$H(S_1, S_2) < I(X_1, X_2; Y|U, Q), \quad (3)$$
$$\mathbb{E}[d_i(U, X_i)] \leq D_i, \ i = 1, 2, \quad (4)$$

where the above entropies, mutual informations, and expectations are taken with respect to the joint distribution

$$P_Q P_{S_1S_2} P_U P_{X_1|S_1UQ} P_{X_2|S_2UQ} W_{Y|X_1X_2}. \quad (5)$$

We remark that the RV $Q$ serves as a time-sharing RV and the cardinality of its alphabet can be bounded by $|\mathcal{Q}| \leq 6$.

The proof of the theorem, which employs a joint strong typicality coding argument [9] under additional distortion constraints, is deferred to Section V. Note that if $U$ is removed in (1)–(3), then the inequalities reduce to the sufficient condition under which the sources $\{(S_{1j}, S_{2j})\}$ can be reliably transmitted over the MAC $W_{Y|X_1X_2}$ obtained in [9], [10].

## IV. Special Cases

*1) Uniform and Independent Sources:* Suppose that the sources are independent and uniform, i.e., $P_{S_1}(s_1) = 1/|\mathcal{S}_1|$,

$P_{S_2}(s_2) = 1/|\mathcal{S}_2|$ and $P_{S_1S_2}(s_1,s_2) = P_{S_1}(s_1)P_{S_2}(s_2)$ for any $(s_1,s_2) \in \mathcal{S}_1 \times \mathcal{S}_2$. Define $\widetilde{R}_1 = H(S_1) = \log|\mathcal{S}_1|$ and $\widetilde{R}_2 = H(S_2) = \log|\mathcal{S}_2|$ to be the rates of the sources. By Theorem 1, $\{(S_{1j}, S_{2j})\}$ are $(D_1, D_2)$-admissible with respect to the MAC $W_{Y|X_1X_2}$ if there exists some RV $Q$ with $|\mathcal{Q}| \le 6$, and a pair of distributions $(P_{X_1|UQ}, P_{X_2|UQ})$ such that

$$\widetilde{R}_1 < I(X_1;Y|X_2,U,Q), \tag{6}$$
$$\widetilde{R}_2 < I(X_2;Y|X_1,U,Q), \tag{7}$$
$$\widetilde{R}_1 + \widetilde{R}_2 < I(X_1,X_2;Y|U,Q), \tag{8}$$
$$\mathbb{E}[d_i(U,X_i)] \le D_i,\ i=1,2, \tag{9}$$

where the above mutual informations and expectations are taken with respect to the joint distribution $P_Q P_U P_{X_1|UQ} P_{X_2|UQ} W_{Y|X_1X_2}$. If we further set $D_1 \ge d_1^{max}$ and $D_2 \ge d_2^{max}$ and let $U$ be deterministic, inequalities (6)–(9) give the capacity region of the MAC [13].

*2) Parallel Attack Channels:* Assume that the attack MAC is composed of two independent discrete memoryless channels $W_{Y|X_1X_2}(y|x_1,x_2) = W_{Y_1|X_1}(y_1|x_1) \times W_{Y_2|X_2}(y_2|x_2)$ where $W_{Y_i|X_i}$ has input alphabet $\mathcal{X}_i$ and output alphabet $\mathcal{Y}_i$ such that $\mathcal{Y}_1 \times \mathcal{Y}_2 = \mathcal{Y}$, $i = 1,2$. This can be interpreted as two attackers separately attacking the stegotexts. In this case, the condition given by Theorem 1 for successful embedding is equivalent to the following (see the proof in the appendix): $\{(S_{1j}, S_{2j})\}$ are $(D_1, D_2)$-admissible with respect to the MAC $W_{Y|X_1X_2}$ if

$$H(S_1|S_2) < C(W^{(1)}, D_1), \tag{10}$$
$$H(S_2|S_1) < C(W^{(2)}, D_2), \tag{11}$$
$$H(S_1,S_2) < C(W^{(1)}, D_1) + C(W^{(2)}, D_2), \tag{12}$$

where $C(W^{(i)}, D_i) = \max_{\mathbb{E}[d_i(U,X_i)] \le D_i} I(X_i;Y_i|U)$, $i = 1,2$ is the information-hiding capacity of the attack channel $W_{Y_i|X_i}$ with distortion threshold $D_i$ [4].

*3) Attack-Free Channel:* Let $l : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{Y}$ be a bijection and let $Y = l(X_1, X_2)$. In this case, Theorem 1 implies that $\{(S_{1j}, S_{2j})\}$ are $(D_1, D_2)$-admissible with respect to the MAC $W_{Y|X_1X_2}$ if

$$H(S_1|S_2) < H(X_1|X_2,S_2,U), \tag{13}$$
$$H(S_2|S_1) < H(X_2|X_1,S_1,U), \tag{14}$$
$$H(S_1,S_2) < H(X_1,X_2|U), \tag{15}$$
$$\mathbb{E}[d_i(U,X_i)] \le D_i,\ i=1,2, \tag{16}$$

where the entropies are taken under the joint distribution $P_{S_1S_2} P_U P_{X_1|S_1U} P_{X_2|S_2U}$. Note also that conditions (13)–(16) give the Slepian-Wolf lossless data compression region [13], [11] if we set $D_1 \ge d_1^{max}$, $D_2 \ge d_2^{max}$, and let $U$ be deterministic.

## V. Proof of Theorem 1

We first give an outline of the proof. We need to show that for given $P_{S_1S_2}$, $P_U$, and $W_{Y|X_1X_2}$, there exists a sequence of JSCE codes $(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})$ such that $P_e^{(n)} \to 0$ as $n \to \infty$ and for any $\delta > 0$, $\frac{1}{n}\mathbb{E}[d_i(U^n, X_i^n)] \le D_i + \delta$, $i = 1,2$, for $n$ sufficiently large. Fix $(P_Q, P_{X_1|S_1UQ}, P_{X_2|S_2UQ})$ such that the following are satisfied for some $\epsilon > 0$,

$$H(S_1|S_2) < I(X_1;Y|X_2,S_2,U,Q) - 7\epsilon, \tag{17}$$
$$H(S_2|S_1) < I(X_2;Y|X_1,S_1,U,Q) - 7\epsilon, \tag{18}$$
$$H(S_1,S_2) < I(X_1,X_2;Y|U,Q) - 7\epsilon, \tag{19}$$
$$\mathbb{E}[d_i(U,X_i)] \le D_i,\ i=1,2. \tag{20}$$

Define $P_i^{(n)} \triangleq \Pr\{d_i(U^n, X_i^n) > n(D_i + \epsilon)\}$, $i = 1,2$. We will prove that for any $\epsilon_1 > 0$, the following probabilities, which are averaged over a family of random codes $(\mathcal{C}_1, \mathcal{C}_2)$, $i = 1,2$, satisfy

$$\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}[P_e^{(n)}] \le \epsilon_1, \quad \mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}[P_i^{(n)}] \le \epsilon_1, \quad \mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}[P_i^{(n)}] \le \epsilon_1$$

for $n$ sufficiently large. Then $\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\{P_e^{(n)} + P_1^{(n)} + P_2^{(n)}\} \le 3\epsilon_1$, which guarantees that there exists at least one pair $(\mathcal{C}_1, \mathcal{C}_2)$ such that $P_e^{(n)} + P_1^{(n)} + P_2^{(n)} \le 3\epsilon_1$ and hence $P_e^{(n)} \le 3\epsilon_1$, $P_1^{(n)} \le 3\epsilon_1$, $P_2^{(n)} \le 3\epsilon_1$ are simultaneously satisfied for $n$ sufficiently large. Finally, it can be easily shown that $P_i^{(n)} \le 3\epsilon_1$ implies for $n$ sufficiently large that

$$\frac{1}{n}\mathbb{E}[d_i(U^n, X_i^n)] \le D_i + \epsilon + d_i^{max} P_i^{(n)} \le D_i + \delta.$$

### A. Random Code Design

*Random Code Generation*. Let $i \in \{1,2\}$. Choose a typical sequence $\mathbf{q} = (q_1, q_2, ..., q_n)$ arbitrarily in $T_\epsilon^{(n)}(Q)$. The sequence serves as a time sharing sequence and it is known at both the encoders and the decoder. For any sequences $\mathbf{s}_i, \mathbf{u}$ and the fixed $\mathbf{q}$, generate one $\mathbf{x}_i(\mathbf{s}_i, \mathbf{u}, \mathbf{q})$ sequence according to $\prod_{j=1}^{n} P_{X_i|S_iUQ}(x_{ij}|s_{ij}, u_j, q_j)$. Define codebook $\mathcal{C}_i$ as $\mathcal{C}_i \triangleq \{\mathbf{x}_i(\mathbf{s}_i, \mathbf{u}, \mathbf{q}) : (\mathbf{s}_i, \mathbf{u}) \in \mathcal{S}_i^n \times \mathcal{U}^n\}$. Reveal the codebooks to both the encoders and the decoder.

*Encoding*. Given $(\mathbf{s}_i, \mathbf{u}) \in \mathcal{S}_i^n \times \mathcal{U}^n$, Encoder $i$ sends $\mathbf{x}_i(\mathbf{s}_i, \mathbf{u}, \mathbf{q})$.

*Decoding*. The decoder has full knowledge of $\mathbf{u}$ (and also the time sharing sequence $\mathbf{q}$). Upon receiving sequence $\mathbf{y}$, the decoder finds the only pair $(\widehat{\mathbf{s}}_1, \widehat{\mathbf{s}}_2) \in T_\epsilon^{(n)}(S_1, S_2)$, such that $\mathbf{y} \in T_\epsilon^{(n)}(Y|\widehat{\mathbf{s}}_1, \widehat{\mathbf{s}}_2, \mathbf{u}, \mathbf{q}, \widehat{\mathbf{x}}_1, \widehat{\mathbf{x}}_2)$, where $\widehat{\mathbf{x}}_1 = \mathbf{x}_1(\widehat{\mathbf{s}}_1, \mathbf{u}, \mathbf{q})$ and $\widehat{\mathbf{x}}_2 = \mathbf{x}_2(\widehat{\mathbf{s}}_2, \mathbf{u}, \mathbf{q})$. If there is no or more than one such pair of sequences $(\widehat{\mathbf{s}}_1, \widehat{\mathbf{s}}_2)$, an error is declared.

For the sake of convenience, define the events

$$A_0 \ : \ (\mathbf{s}_1, \mathbf{s}_2, \mathbf{u}) \in T_\epsilon^{(n)}(S_1, S_2, U|\mathbf{q})$$
$$A_1 \ : \ (\mathbf{s}_1, \mathbf{s}_2, \mathbf{u}, X_1^n(\mathbf{s}_1, \mathbf{u}, \mathbf{q}), X_2^n(\mathbf{s}_2, \mathbf{u}, \mathbf{q})) \in T_\epsilon^{(n)}(\cdot|\mathbf{q}).$$

The following result is a consequence of the Markov lemma (Lemma 2).

*Lemma 3:* For any $\epsilon, \epsilon_2 \in (0,1)$, $\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}[\Pr(A_1^c|A_0)] \le \epsilon_2$ for $n$ sufficiently large, where the expectation is taken with respect to the random codes $\mathcal{C}_1$ and $\mathcal{C}_2$.

*B. Bounding* $\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}[P_e^{(n)}]$

$$\begin{aligned}\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}[P_e^{(n)}] \leq & \sum_{(T_\epsilon^{(n)}(S_1,S_2,U|\mathbf{q}))^c} P_{S_1^nS_2^n}(\mathbf{s}_1,\mathbf{s}_2)P_{U^n}(\mathbf{u}) \\ + & \sum_{T_\epsilon^{(n)}(S_1,S_2,U|\mathbf{q})} P_{S_1^nS_2^n}(\mathbf{s}_1,\mathbf{s}_2)P_{U^n}(\mathbf{u}) \\ \times & \ \mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\left[\sum_{\mathbf{y}:\psi^{(n)}(\mathbf{y},\mathbf{u})\neq(\mathbf{s}_1,\mathbf{s}_2)} W_{Y^n|X_1^nX_2^n}(\mathbf{y}|\mathbf{x}_1,\mathbf{x}_2)\right].\end{aligned}$$

The first term vanishes for $n$ sufficiently large by Lemma 1. It suffices to bound the expectation in the second term. Given $(\mathbf{s}_1,\mathbf{s}_2,\mathbf{u}) \in T_\epsilon^{(n)}(S_1,S_2,U|\mathbf{q})$, we have the following four error events:

$E_0$ : $(\mathbf{s}_1,\mathbf{s}_2,\mathbf{u},X_1^n(\mathbf{s}_1,\mathbf{u},\mathbf{q}),X_2^n(\mathbf{s}_2,\mathbf{u},\mathbf{q}),Y^n) \notin T_\epsilon^{(n)}(\cdot|\mathbf{q})$,

$E_1$ : $\exists\ \widehat{\mathbf{s}}_1 \neq \mathbf{s}_1$ such that $(\widehat{\mathbf{s}}_1,\mathbf{s}_2,\mathbf{u},X_1^n(\widehat{\mathbf{s}}_1,\mathbf{u},\mathbf{q}),X_2^n(\mathbf{s}_2,\mathbf{u},\mathbf{q}),Y^n) \in T_\epsilon^{(n)}(\cdot|\mathbf{q})$,

$E_2$ : $\exists\ \widehat{\mathbf{s}}_2 \neq \mathbf{s}_2$ such that $(\mathbf{s}_1,\widehat{\mathbf{s}}_2,\mathbf{u},X_1^n(\mathbf{s}_1,\mathbf{u},\mathbf{q}),X_2^n(\widehat{\mathbf{s}}_2,\mathbf{u},\mathbf{q}),Y^n) \in T_\epsilon^{(n)}(\cdot|\mathbf{q})$,

$E_3$ : $\exists\ \widetilde{\mathbf{s}}_1 \neq \mathbf{s}_1, \widetilde{\mathbf{s}}_2 \neq \mathbf{s}_2$ such that $(\widetilde{\mathbf{s}}_1,\widetilde{\mathbf{s}}_2,\mathbf{u},X_1^n(\widetilde{\mathbf{s}}_1,\mathbf{u},\mathbf{q}),\mathbf{u}),X_2^n(\widetilde{\mathbf{s}}_2,\mathbf{u},\mathbf{q}),Y^n) \in T_\epsilon^{(n)}(\cdot|\mathbf{q})$.

It then immediately follows from the union bound that

$$\begin{aligned}&\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\left[\sum_{\mathbf{y}:\psi^{(n)}(\mathbf{y},\mathbf{u})\neq(\mathbf{s}_1,\mathbf{s}_2)} W_{Y^n|X_1^nX_2^n}(\mathbf{y}|\mathbf{x}_1,\mathbf{x}_2)\right] \\ &\quad\leq \sum_{j=0}^{3}\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\Big[\Pr\{E_j|\,A_0\}\Big]. \qquad (21)\end{aligned}$$

To bound $\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\Big[\Pr\{E_0|\,A_0\}\Big]$, it follows from Lemmas 2 and 3 that

$$\begin{aligned}&\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\Big[\Pr\{E_0|\,A_0\}\Big] \\ &\leq \mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\Big[\Pr(A_1^c|A_0)\Big] + \mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\Big[\Pr\{E_0|\,A_0,A_1\}\Big] \\ &\leq \frac{\epsilon_0}{2}+\frac{\epsilon_0}{2} = \epsilon_0 \qquad (22)\end{aligned}$$

if $n$ sufficiently large, where $\epsilon_0$ will be specified later.

To bound $\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\Big[\Pr\{E_1|\,A_0\}\Big]$, write

$$\begin{aligned}&\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\Big[\Pr\{E_1|A_0\}\Big] \\ &\leq \sum_{\widehat{\mathbf{s}}_1\neq\mathbf{s}_1:\widehat{\mathbf{s}}_1\in T_\epsilon^{(n)}(S_1|\mathbf{s}_2)} \mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\left[\Pr\left\{\mathbf{v}_1\in T_\epsilon^{(n)}\middle|\,A_0\right\}\right] \qquad (23)\end{aligned}$$

where $\mathbf{v}_1 = (\widehat{\mathbf{s}}_1,\mathbf{s}_2,\mathbf{u},\mathbf{q},X_1^n(\widehat{\mathbf{s}}_1,\mathbf{u},\mathbf{q}),X_2^n(\mathbf{s}_2,\mathbf{u},\mathbf{q}),Y^n)$ and the expectation can be upper bounded by

$$\begin{aligned}&\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\left[\Pr\left\{\mathbf{v}_1\in T_\epsilon^{(n)}\middle|\,A_0\right\}\right] \\ &\leq \sum_{\mathcal{X}_1^n\times\mathcal{X}_2^n} P_{X_1^n|S_1^nU^nQ^n}(\widehat{\mathbf{x}}_1|\widehat{\mathbf{s}}_1,\mathbf{u},\mathbf{q})P_{X_2^n|S_2^nU^n}(\mathbf{x}_2|\mathbf{s}_2,\mathbf{u},\mathbf{q}) \\ &\qquad \sum_{\mathbf{y}\in T_\epsilon^{(n)}(Y|\widehat{\mathbf{s}}_1,\mathbf{s}_2,\mathbf{u},\mathbf{q},\widehat{\mathbf{x}}_1,\mathbf{x}_2)} P_{Y^n|S_2^nU^nQ^nX_2^n}(\mathbf{y}|\mathbf{s}_2,\mathbf{u},\mathbf{q},\mathbf{x}_2) \\ &\leq \left|T_\epsilon^{(n)}(Y|\widehat{\mathbf{s}}_1,\mathbf{s}_2,\mathbf{u},\mathbf{q},\widehat{\mathbf{x}}_1,\mathbf{x}_2)\right| e^{-n(H(Y|S_2,U,Q,X_2)-2\epsilon)} \qquad (24) \\ &\leq e^{n(H(Y|X_1,X_2)+2\epsilon)}e^{-n(H(Y|S_2,U,Q,X_2)-2\epsilon)} \qquad (25) \\ &= e^{n(H(Y|X_1,X_2,S_2,U,Q)+2\epsilon)}e^{-n(H(Y|S_2,U,Q,X_2)-2\epsilon)} \\ &= e^{-n(I(X_1;Y|X_2,S_2,U,Q)-4\epsilon)}, \qquad (26)\end{aligned}$$

where $\widehat{\mathbf{x}}_1 = \mathbf{x}_1(\widehat{\mathbf{s}}_1,\mathbf{u},\mathbf{q})$, $\mathbf{x}_2 = \mathbf{x}_2(\mathbf{s}_2,\mathbf{u},\mathbf{q})$, and (24) and (25) follow from Lemma 1. It then follows from (23), Lemma 1 and (17) that

$$\begin{aligned}&\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\Big[\Pr\{E_1|\,A_0\}\Big] \\ &\leq \left|T_\epsilon^{(n)}(S_1|\mathbf{s}_2)\right| e^{-n(I(X_1;Y|X_2,S_2,U,Q)-4\epsilon)} \\ &\leq e^{n(H(S_1|S_2)+2\epsilon)}e^{-n(I(X_1;Y|X_2,S_2,U,Q)-4\epsilon)} \\ &= e^{-n(I(X_1;Y|X_2,S_2,U,Q)-H(S_1|S_2)-6\epsilon)} \\ &\leq e^{-n\epsilon} \leq \epsilon_0, \qquad (27)\end{aligned}$$

for $n$ sufficiently large. Similarly, we can bound using (18)

$$\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\Big[\Pr\{E_2|\,A_0\}\Big] \leq \epsilon_0, \qquad (28)$$

for $n$ sufficiently large.

It remains to bound $\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\Big[\Pr\{E_3|\,A_0\}\Big]$. Write

$$\begin{aligned}&\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\Big[\Pr\{E_3|A_0\}\Big] \leq \sum_{\widetilde{\mathbf{s}}_1\neq\mathbf{s}_1,\widetilde{\mathbf{s}}_2\neq\mathbf{s}_2:(\widetilde{\mathbf{s}}_1,\widetilde{\mathbf{s}}_2)\in T_\epsilon^{(n)}(S_1,S_2)} \\ &\qquad \mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\left[\Pr\left\{\mathbf{v}_2\in T_\epsilon^{(n)}\middle|\,A_0\right\}\right], \qquad (29)\end{aligned}$$

where $\mathbf{v}_2 = (\widetilde{\mathbf{s}}_1,\widetilde{\mathbf{s}}_2,\mathbf{u},\mathbf{q},X_1^n(\widetilde{\mathbf{s}}_1,\mathbf{u},\mathbf{q}),X_2^n(\widetilde{\mathbf{s}}_2,\mathbf{u},\mathbf{q}),Y^n)$ and

$$\begin{aligned}&\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\left[\Pr\left\{\mathbf{v}_2\in T_\epsilon^{(n)}\middle|\,A_0\right\}\right] \\ &\leq \sum_{\mathcal{X}_1^n\times\mathcal{X}_2^n} P_{X_1^n|S_1^nU^nQ^n}(\widetilde{\mathbf{x}}_1|\widetilde{\mathbf{s}}_1,\mathbf{u},\mathbf{q})P_{X_2^n|S_2^nU^nQ^n}(\widetilde{\mathbf{x}}_2|\widetilde{\mathbf{s}}_2,\mathbf{u},\mathbf{q}) \\ &\qquad \sum_{\mathbf{y}\in T_\epsilon^{(n)}(Y|\widetilde{\mathbf{s}}_1,\widetilde{\mathbf{s}}_2,\mathbf{u},\mathbf{q},\widetilde{\mathbf{x}}_1,\widetilde{\mathbf{x}}_2)} P_{Y^n|U^nQ^n}(\mathbf{y}|\mathbf{u},\mathbf{q}) \\ &\leq \left|T_\epsilon^{(n)}(Y|\widetilde{\mathbf{s}}_1,\widetilde{\mathbf{s}}_2,\mathbf{u},\mathbf{q},\widetilde{\mathbf{x}}_1,\widetilde{\mathbf{x}}_2)\right| e^{-n(H(Y|U,Q)-2\epsilon)} \qquad (30) \\ &\leq e^{n(H(Y|U,Q,X_1,X_2)+2\epsilon)}e^{-n(H(Y|U,Q)-2\epsilon)} \qquad (31) \\ &= e^{-n(I(X_1,X_2;Y|U,Q)-4\epsilon)}\end{aligned}$$

where $\widetilde{\mathbf{x}}_1 = \mathbf{x}_1(\widetilde{\mathbf{s}}_1,\mathbf{u},\mathbf{q})$ and $\widetilde{\mathbf{x}}_2 = \mathbf{x}_2(\widetilde{\mathbf{s}}_2,\mathbf{u},\mathbf{q})$, and (30) and (31) follow Lemma 1. It then follows that,

$$\begin{aligned}&\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\Big[\Pr\{E_3|A_0\}\Big] \\ &\leq \sum_{\widetilde{\mathbf{s}}_1\neq\mathbf{s}_1,\widetilde{\mathbf{s}}_2\neq\mathbf{s}_2:(\widetilde{\mathbf{s}}_1,\widetilde{\mathbf{s}}_2)\in T_\epsilon^{(n)}(S_1,S_2)} e^{-n(I(X_1,X_2;Y|U)-4\epsilon)} \\ &\leq \left|T_\epsilon^{(n)}(S_1,S_2)\right| e^{-n(I(X_1,X_2;Y|U)-4\epsilon)} \\ &\leq e^{n(H(S_1,S_2)+2\epsilon)}e^{-n(I(X_1;Y|X_2,S_2,U)-4\epsilon)} \\ &= e^{-n(I(X_1;Y|X_2,S_2,U)-H(S_1,S_2)-6\epsilon)} \\ &\leq e^{-n\epsilon} \leq \epsilon_0 \qquad (32)\end{aligned}$$

for $n$ sufficiently large. Now plugging (22), (27), (28), and (32) back into (21), and setting $\epsilon_0 = \frac{\epsilon_1}{4}$, we see that $\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}[P_e^{(n)}] \leq \epsilon_1$ for $n$ sufficiently large.

### C. Bounding $\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}[P_i^{(n)}]$

Since the encoding is separately performed, Encoder 1 is independent of $\mathcal{C}_2$. Thus it suffices to show that $\mathbb{E}_{\mathcal{C}_1}[P_1^{(n)}] \le \epsilon_1$ for $n$ sufficiently large.

Clearly, if $(\mathbf{s}_1, \mathbf{u}, \mathbf{x}_1) \in T_\epsilon^{(n)}(S_1, U, X_1|\mathbf{q})$, then

$$\frac{1}{n} d_1(\mathbf{u}, \mathbf{x}_1(\mathbf{s}_1, \mathbf{u}, \mathbf{q})) \le \mathbb{E}[d_1(U, X_1)] + \epsilon \le D_1 + \epsilon$$

for $n$ sufficiently large, where the first inequality follows from the definition of strong typicality and the second inequality follows from (20). According to Lemma 1,

$$\begin{aligned}
\mathbb{E}_{\mathcal{C}_1}[P_1^{(n)}] &\le \sum_{(T_\epsilon^{(n)}(S_1,U|\mathbf{q}))^c} P_{S_1^n U^n}(\mathbf{s}_1, \mathbf{u}) \\
&\quad + \sum_{T_\epsilon^{(n)}(S_1,U|\mathbf{q})} P_{S_1^n U^n}(\mathbf{s}_1, \mathbf{u}) \mathbb{E}_{\mathcal{C}_1} \Phi\left\{\mathbf{v}_3 \notin T_\epsilon^{(n)}\right\} \\
&\le \frac{\epsilon_1}{2} + \frac{\epsilon_1}{2} = \epsilon_1 \qquad (33)
\end{aligned}$$

for $n$ sufficiently large, where $\mathbf{v}_3 = (\mathbf{s}_1, \mathbf{u}, \mathbf{q}, X_1^n(\mathbf{s}_1, \mathbf{u}, \mathbf{q}))$ and $\Phi(A)$ is the indicator function of the event $A$.

### D. Completing the Proof

As we mentioned in the beginning of the section,

$$\mathbb{E}_{\mathcal{C}_1,\mathcal{C}_2}\{P_e^{(n)} + P_1^{(n)} + P_2^{(n)}\} \le 3\epsilon_1,$$

implies that there exists a pair of codes $(\mathcal{C}_1, \mathcal{C}_2)$ such that $P_e^{(n)} \le 3\epsilon_1$, $P_1^{(n)} \le 3\epsilon_1$, $P_2^{(n)} \le 3\epsilon_1$ are simultaneously satisfied for $n$ sufficiently large. Furthermore, if $P_i^{(n)} \le 3\epsilon_1$, we have

$$\frac{1}{n}\mathbb{E}[d_i(U^n, X_i^n)] \le (D_i + \epsilon) + d_i^{max} P_i^{(n)} \le D_i + \delta_i,$$

as $n \to \infty$, by setting $\delta_i = \epsilon + 3\epsilon_1 d_i^{max}$. Thus the distortion constraint is satisfied. This completes the proof of Theorem 1.

## VI. Concluding Remarks

In this work we derive a sufficient condition with single-letter characterizations for hiding correlated sources against MAC attacks. An uncomputable (and somewhat trivial) outer bound (converse condition) can be easily formulated by applying Fano's inequality in terms of a sequence of $n$-dimensional joint distributions. We are currently studying the embedding of correlated sources with joint embedding-compression rate constraints. Our next step is to answer the question: when $(S_1, S_2)$ are $(D_1, D_2)$-admissible with respect to $W_{Y|X_1X_2}$, what is the compression limit for the sources $(S_1, S_2)$ and $U$?

## Appendix

*Proof of the Case of Parallel Attack Channels*: When $W_{Y|X_1X_2} = W_{Y_1|X_1} \times W_{Y_2|X_2}$, we see that (10)–(12) imply (1)–(4). In fact, if the maximums in (10)–(12) are achieved by $P^*_{X_1|U}(x_1|u)$ and $P^*_{X_2|U}(x_2|u)$, then simply letting $|\mathcal{Q}| = 1$, $P_{X_1|S_1U}(x_1|s_1, u) = P^*_{X_1|U}(x_1|u)$ and $P_{X_2|S_2U}(x_2|s_2, u) = P^*_{X_2|U}(x_2|u)$, we see that with this choice,

$$\begin{aligned}
I(X_1; Y|X_2, S_2, U, Q) &= I(X_1; Y_1|S_2, U, Q) = I(X_1; Y_1|U) \\
&= \max_{\mathbb{E}[d_1(U,X_1)] \le D_1} I(X_1; Y_1|U).
\end{aligned}$$

Similarly,

$$I(X_2; Y|X_1, S_1, U, Q) = \max_{\mathbb{E}[d_2(U,X_2)] \le D_2} I(X_2; Y_2|U),$$

and

$$\begin{aligned}
&I(X_1, X_2; Y_1, Y_2|U, Q) \\
&= \max_{\mathbb{E}[d_1(U,X_1)] \le D_1} I(X_1; Y_1|U) + \max_{\mathbb{E}[d_2(U,X_2)] \le D_2} I(X_2; Y_2|U).
\end{aligned}$$

We next show that (1)–(4) imply (10)–(12). We only need to show that for any $P_{X_1|S_1UQ}$ satisfying $\mathbb{E}[d_1(U, X_1)] < D_1$, the right hand side of (1) is upper bounded by (10). Since $(Q, S_1, U) \to X_1 \to Y_1$ form a Markov chain in this order,

$$\begin{aligned}
I(X_1; Y_1|U) &= H(Y_1|U) - H(Y_1|X_1, U) \\
&\ge H(Y_1|S_2, U, Q) - H(Y_1|X_1, S_2, U, Q) \\
&= I(X_1; Y_1|S_2, U, Q).
\end{aligned}$$

For any $P_{X_1|S_1UQ}$ satisfying $\mathbb{E}[d_1(U, X_1)] < D_1$, set

$$\widehat{P}_{X_1|U}(x_1|u) = \sum_{\mathcal{S}_1 \times \mathcal{Q}} P_{S_1}(s_1) P_Q(q) P_{X_1|S_1UQ}(x_1|s_1, u, q).$$

Under the corresponding $\widehat{P}_{X_1|U}(x_1|u)$, we have

$$\begin{aligned}
I(X_1; Y_1|U, S_2, Q) &\le I(X_1; Y_1|U) \\
&\le \max_{\mathbb{E}[d_1(U,X_1)] \le D_1} I(X_1; Y_1|U).
\end{aligned}$$

We can similarly show that (2)–(3) imply (11)–(12). □

## References


[1] A. S. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1639-1667, Jun. 2002.

[2] D. Karakos and A. Papamarcou, "A relationship between quantization and watermarking rates in the presence of additive Gaussian attacks," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1970-1982, August 2003.

[3] A. Maor and N. Merhav, "On joint information embedding and lossy compression in the presence of a memoryless attack," *IEEE Trans. Inform. Theory*, vol. 51, pp. 3166-3175, Sep. 2005.

[4] P. Moulin and J. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Trans. Inform. Theory*, vol. 49, no. 3, pp. 563-593, Mar. 2003.

[5] F. A. P. Petitcolas, R. J. Anderson and M. G. Kuhn, "Information hiding - a survey," *Proc. IEEE*, vol. 87, no. 7, pp. 1062-1078, Jul. 1999.

[6] A. Somekh-Baruch and N. Merhav, "On the capacity game of public watermarking systems," *IEEE Trans. Inform. Theory*, vol. 50, no. 3, pp. 511-524, Mar. 2004.

[7] E.-H. Yang and W. Sun, "On watermarking and compression rates of joint compression and private watermarking systems with abstract alphabets," *Proc. of the 2005 Canadian Workshop on Information Theory*, Montreal, Quebec, Canada, June 5-8, pp. 296-299, 2005.

[8] E.-H. Yang and W. Sun, "On information embedding when watermarks and covertexts are correlated"*Proc. of IEEE IEEE Int'l. Symp. Inform. Theory*, pp. 346–350, Seattle, USA, 2006.

[9] T. M. Cover, A. El Gamal and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inform. Theory*, vol. 26, no. 6, Nov. 1980.

[10] R. Ahlswede and T. S. Han, "On source coding with side information via a multiple-access channel and related problems in multi-user information theory," *IEEE Trans. Inform. Theory*, vol. 29, May 1983.

[11] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, pp. 471-480, July 1973.

[12] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Syst. Tech. Journal*, vol. 52, pp. 1037-1076, 1973.

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd Ed., Wiley, 2006.