**Communicating Correlated Sources over MAC and Interference Channels I: Separation-based Schemes**

<table>
<thead>
<tr>
<th>Journal:</th>
<th><em>IEEE Transactions on Information Theory</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>IT-17-0863.R2</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Regular Manuscript</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>18-Aug-2019</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Padakandla, Arun; Purdue University, Dept. of Computer Science</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Shannon theory, Joint source-channel coding, Inner bound, Achievability, Sufficient conditions, Correlated sources, constant co</td>
</tr>
</tbody>
</table>

---

**SCHOLARONE™ Manuscripts**

https://mc.manuscriptcentral.com/t-it
Communicating Correlated Sources over MAC and Interference Channels I: Separation-based Schemes

Arun Padakandla
Center for Science of Information
Purdue University

Abstract

We consider the two scenarios of communicating a pair $S_1, S_2$ of distributed correlated sources over 2–user multiple access (MAC) and interference channels (IC), respectively. While in the MAC problem, the receiver intends to reconstruct both sources losslessly, in the IC problem, receiver $j$ intends to reconstruct $S_j$ losslessly. We undertake a Shannon theoretic study and focus on characterizing sufficient conditions. Building on Dueck’s findings (1981), we propose a coding scheme based on fixed block-length (B-L) codes. We characterize its information theoretic performance and characterize a new set of sufficient conditions. We identify examples of the MAC and IC problems for which the latter conditions are proven to be strictly weaker than the current known tightest.

Index Terms

Shannon theory, Joint source-channel coding, Inner bound, Achievability, Sufficient conditions, Correlated sources, constant composition codes, single-letter coding scheme.

I. INTRODUCTION

Our primary focus in this article is the Shannon-theoretic study of the two scenarios depicted in Figs. 1, 2. Fig. 1 depicts the MAC problem wherein a pair $S_1, S_2$ of correlated sources, observed at the transmitters (Txs) of a 2–user multiple access channel (MAC), have to be communicated to the receiver (Rx). The Rx intends to reconstruct both the sources losslessly. Given a (generic) MAC $W_{Y|X_1,X_2}$, the MAC problem concerns characterizing necessary and sufficient conditions under which $S_1, S_2$ can be communicated over the MAC. Fig. 2 depicts the IC problem wherein a pair $S_1, S_2$ of correlated sources have to be communicated over a 2–user interference channel (IC) $W_{Y_1,Y_2|X_1,X_2}$. Rx $j$ wishes to reconstruct $S_j$ losslessly. The IC problem concerns characterizing necessary and sufficient conditions under which $S_1, S_2$ can be communicated over the IC $W_{Y_1,Y_2|X_1,X_2}$. Throughout our work, we restrict attention to characterizing sufficient conditions.

Being the simplest network generalizations of the point-to-point (PTP) problem studied by Shannon [1], the MAC and IC problems have had a rich history. Slepian and Wolf [2] solved the MAC problem for the particular case with $S_1 = (M_0, M_1)$, $S_2 = (M_0, M_2)$ wherein $M_0, M_1, M_2$ are independent messages. Han [3] leveraged tools from matroid theory to solve a slightly generalized version of this problem. The general MAC problem however remained open. Cover, El Gamal and Salehi [4] devised an elegant coding scheme, henceforth referred to as the CES scheme, to exploit source correlation amidst $S_1, S_2$ to correlate and thereby coordinate channel inputs $X_1, X_2$. Characterizing its performance, they derived sufficient conditions for the MAC problem, henceforth referred to as CES conditions. In spite of continued interest, the MAC problem remains open and the CES conditions has remained to be the tightest known set of sufficient conditions since 1981.

This work was supported by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-0939370. This work was presented in part at the IEEE International Symposium on Information Theory held in Barcelona, Spain (July 2016) and Aachen, Germany (June 2017).
In contrast to the MAC problem, the IC problem [5] remains unsolved [6] even when \(S_1, S_2\) are independent. Considering the general case, Liu and Chen [7] incorporated all known coding techniques - random source partitioning [8], message splitting via superposition coding [5] and the technique of inducing source correlation onto channel inputs [4] - and devised the current known best coding scheme for the IC problem, henceforth referred to as the LC coding scheme. The corresponding set of sufficient conditions (Thm. 2), henceforth referred to as LC conditions, remain to be the tightest known for the IC problem.

Our work is aimed at deriving a new set of sufficient conditions for the MAC and IC problems that are subsumed within the CES and LC conditions, respectively, and strictly weaker for identified examples. In this article, we present the first part of our findings. Specifically, we present a new set of sufficient conditions for the MAC (Thms. 4, 8) and IC problems (Thms. 5, 6, 7) that are analytically proven to be strictly weaker for identified examples (Exs. 1, 2). In the second part [9], we build on our findings here and achieve the aim mentioned above.

Within barely a few months of [4] coming to light, Dueck [10] identified a rich example and devised an ingenious, though very specific, coding scheme for that example to prove sub-optimality of CES scheme. Through the use of codes of non-asymptotic block-length (B-L), Dueck’s coding scheme transgressed long established conventions in information theory. Hence, in spite of being well aware, Dueck’s work has remained isolated. Our work recognizes a fundamental reasoning behind Dueck’s use of codes of non-asymptotic B-L. We elaborate on this in the sequel.

It is long believed that efficiency of a coding scheme designed for any communication scenario only increases with increasing B-L. This is true of the basic PTP channel and source coding tasks. Witsenhausen [11] studied the problem of extracting correlation from distributed correlated sources \(S_1, S_2\) that evolved IID over time. In a remarkable finding, he proved that the correlation amidst the extracted bits, quantified in terms of the agreement probability, strictly reduces with increasing B-L. Coding schemes designed for the MAC and IC problems has to simultaneously extract correlation from distributed sources and perform source and channel coding. The former prefers short B-Ls and the latter favors large B-Ls resulting in a tension in the choice of optimal B-L. Dueck’s use of codes of non-asymptotic B-L, henceforth referred to as fixed B-L codes, is a symptom of this tension. These point to a broader role for fixed B-L codes and a theory behind Dueck’s coding scheme. Our work is aimed at identifying the basic building blocks - coding and analytical tools - necessary to exploit this phenomenon.

The tools and the coding scheme we propose in this work differ fundamentally from conventional information-theoretic coding strategies. We have therefore adopted a pedagogical exposition in this first part and a concise description for the second [9]. In Sec. III, we describe the core ideas in the context of an example (Ex. 1) and introduce steps (Sec. III-A) of the analysis that will reappear in the following sections. We present the new coding and analytical tools in Sec. IV-A and describe our coding scheme in two steps - Step I in Sec. IV, Step II in Sec. V.

We conclude this section by recognizing closely related recent findings that have inspired our work. Dueck’s work [10] demonstrates that the CES scheme is sub-optimal in exploiting correlation amidst distributed information sources. Recently, Wagner, Kelly and Altug [12], building on [13], demonstrate an analogous sub-optimality of the Berger-Tung [14] coding scheme for the distributed source coding (DSC) problem. While [12] does not propose a new coding scheme, Shirani and Pradhan [15], [16] build on [12] and design a new coding scheme for a general instance of the DSC problem. We leverage tools developed in [15] to build the channel coding scheme proposed herein. While we borrow the idea of matrix coding and interleaving from [15], the problem of multiplexing information streams of difference B-Ls through a (single-input) channel and its performance characterization requires further tools. The use of constant composition codes and the newer analysis techniques developed herein provides a general framework for fixed B-L channel and joint source-channel coding.

A. Notation

We supplement standard information theory notation - upper case for RVs, calligraphic letters such as \(\mathcal{A}, \mathcal{S}\) for finite sets etc. - with the following. We let an underline denote an appropriate aggregation of related objects. For example, \(\mathcal{S}\) will be used to represent a pair \(S_1, S_2\) of RVs. \(\mathcal{S}\) will be used to denote either the pair \(S_1, S_2\) or the Cartesian product \(S_1 \times S_2\), and will be clear from context. If we have \(m\) components, say \((A_0, A_1, A_2)\), then \(\mathcal{A}\) will denote the triple, and we do not use an underline to denote pairs in this case. We let \(p_U(\cdot)\) and \(\mathbb{W}_U(\cdot)\) denote PMFs on \(U\). We reserve the latter for fixed PMFs such as those specified in the problem statement. \(p_U^l = \prod_{i=1}^l p_{U_i}\) denotes the product PMF on \(U^l\). When \(j \in \{1, 2\}\), then \(\overline{j}\) will denote the complement index, i.e., \(\{j, \overline{j}\} = \{1, 2\}\).

For \(m \in \mathbb{N}\), \([m] = \{1, \ldots, m\}\). For \(x^n \in \mathcal{X}^n\) and \(a \in \mathcal{X}\), let \(N(a | x^n) := \sum_{i=1}^n 1_{\{x_i = a\}}\) denote the number of occurrences of \(a\) in \(x^n\).

https://mc.manuscriptcentral.com//t-it
$T^n_b(U) = \{ u^n \in U^n : |n^{-1} \cdot N(b | u^n) - p_U(b)| \leq \delta p_U(b) \ \forall b \in U \}$
is our typical set. $u^n \sim \prod_{i=1}^n p_y$ abbreviates “$u^n$ is typical with respect to PMF $\prod_{i=1}^n p_y$”, and $u^n \sim \prod_{i=1}^n p_y$ abbreviates its negation. For a PMF $p_U$ on $U$, $b^* \in U$ will denote a symbol with the least positive probability wrt $p_U$. The underlying PMF $p_U$ will be clear from context. We let $\tau_\delta(K) = 2|K| \exp\{-2\delta^2 p_K(a^*)\}$ denote an upper bound on $P(K^1 \notin T^n_\delta(K))$. This upper bound can be derived using Hoeffding’s inequality [17, Problem 3.18(b)].

We adopt the notation in [17, Definition 2.1] with regard to types. Specifically, $P_{e^n}$ denotes type of $x^n$, and the set of sequences in $X^n$ of type $p$ is denoted $T^n_p$. We have used similar notation for typical sequences ($T^n_b(\cdot)$) and sequences of type $p$ ($T^n_p(\cdot)$). The particular reference will be clear from context.

For a map $f : S^t \to \mathcal{K}$, we let $f^n : S^n \to \mathcal{K}^n$ denote its $n$–letter extension defined by $f^n(s^n) := (f(s_1), f(s_2), \ldots, f(s_n))$. Boldfaced calligraphic letters such as $\mathcal{A} := \mathcal{A}^{m \times l}$ denote the set of all $m \times l$ matrices over $\mathcal{A}$. Boldfaced letters such as $a, A$ denote matrices. For an $m \times l$ matrix $a$, (i) $a(t, i)$ denotes the entry in row $t$, column $i$, (ii) $a(1 : m, i)$ denotes the $i^{th}$ column, $a(t, 1 : l)$ denotes $t^{th}$ row. “with high probability”, “single-letter”, “long Markov chain”, “block-length” are abbreviated as whp, S-L, LMC, B-L, respectively. We will be employing codes of fixed B-L whose B-L does not depend on the desired probability of error. Codes whose B-L will be chosen arbitrarily large as a function of the desired probability of error will be informally referred to as $\infty$–B-L codes.

For a point-to-point channel (PTP) $(\mathcal{U}, \mathcal{Y}, \mathcal{W}_{Y|U})$, let $E_r(R, p_U, \mathcal{W}_{Y|U})$ denote the random coding error exponent [17, Thm. 10.2] for constant composition codes of type $p_U$ and rate $R$. Specifically,

$$E_r(R, p_U, \mathcal{W}_{Y|U}) := \min_{p_{\mathcal{V}|\mathcal{U}}} \{ D(V|U) | p_{\mathcal{W}_{Y|U}|p_U} + |I(p_U; V|Y) - R| \}.$$ 

In the above expression, $I(p_U; V|Y) = I(U; Y)$, where $U$ is distributed with PMF $p_U$ and $Y$ has conditional PMF $p_{V|Y|U}$ given $U$. For a finite set $\mathcal{B}$ and $\mu \in [0, 1]$, we let

$$\ell_1(\mu, |\mathcal{B}|) := \log|\mathcal{B}| \mu + |\mathcal{B}| \log|\mathcal{B}|$$ 

(1)

If $A_1 \subseteq \mathcal{A}$ and $A_2 \subseteq \mathcal{A}$ are $(l-\text{length})$ random vectors, we let $\xi[0](A) := P(A_1 \neq A_2)$, and $\xi(A) := \xi[1](A)$. If $(A_1, A_2) : t \in [l]$ are independent and identically distributed (IID), we note

$$\xi[0](A) = 1 - (1 - \xi(A))^l \leq l\xi(A).$$

B. Problem Statement

Consider a 2–user MAC with input alphabets $\mathcal{X}_1, \mathcal{X}_2$, output alphabet $\mathcal{Y}$ and channel transition probabilities $\mathcal{W}_{Y|X_1X_2}$ (Fig. 1). Let $\mathcal{S} := (S_1, S_2)$, taking values over $\mathcal{S} := S_1 \times S_2$ with PMF $\mathcal{W}_{S_1S_2}$, denote a pair of information sources. For $j \in [2]$, $T_x$ observes $S_j$. The Rx aims to reconstruct $\hat{S}$ with arbitrarily small probability of error. With regard to the MAC problem, our objective is to characterize sufficient conditions for transmissibility of sources $(\mathcal{S}, \mathcal{W}_S)$ over the MAC $(\mathcal{X}, \mathcal{Y}, \mathcal{W}_{Y|X})$. A formal definition follows.

Definition 1: A pair $(\mathcal{S}, \mathcal{W}_S)$ is transmissible over MAC $(\mathcal{X}, \mathcal{Y}, \mathcal{W}_{Y|X})$ if for every $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that, for every $n \geq N_\epsilon$, there exist encoder maps $e_j : \mathcal{S}_j^n \to \mathcal{X}_j^n : j \in [2]$ and decoder map $d : \mathcal{Y}^n \to \mathcal{S}^n$ such that

$$\sum_{\mathcal{S}^n} \mathcal{W}_S^n(\mathcal{S}^n) \sum_{y^n \in \mathcal{Y}^n} \mathcal{W}_{Y|X}(y^n | e_1(s_1^n), e_2(s_2^n)) I(d(y^n) \neq \mathcal{S}^n) \leq \epsilon.$$ 

Consider a 2–user IC with input alphabets $\mathcal{X}_1, \mathcal{X}_2$, output alphabets $\mathcal{Y}_1, \mathcal{Y}_2$, and transition probabilities $\mathcal{W}_{Y_1Y_2|X_1X_2}$ (Fig. 2). Let $\mathcal{S} := (S_1, S_2)$, taking values over $\mathcal{S} := S_1 \times S_2$ with PMF $\mathcal{W}_{S_1S_2}$, denote a pair of information sources. For $j \in [2]$, $T_x$ observes $S_j$, and Rx $j$ aims to reconstruct $S_j$ with arbitrarily small probability of error. If this is possible, we say $\mathcal{S}$ is transmissible over IC $\mathcal{W}_{Y_1Y_2}$ with a formal definition follows.

Definition 2: A pair $(\mathcal{S}, \mathcal{W}_S)$ is transmissible over IC $(\mathcal{X}, \mathcal{Y}, \mathcal{W}_{Y|X})$ if for every $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that, for every $n \geq N_\epsilon$, there exist encoder maps $e_j : \mathcal{S}_j^n \to \mathcal{Y}_j^n : j \in [2]$ and decoder maps $d_j : \mathcal{Y}_j^n \to \mathcal{S}_j^n : j \in [2]$ such that

$$\sum_{\mathcal{S}^n} \mathcal{W}_S^n(\mathcal{S}^n) \sum_{y^n \in \mathcal{Y}^n} \mathcal{W}_{Y_j|X_j}(y_1^n | y_2^n) I(d_1(y_1^n) \neq s_1^n \text{ or } d_2(y_2^n) \neq s_2^n) \leq \epsilon.$$ 

With regard to the IC problem, our objective is to characterize sufficient conditions under which $(\mathcal{S}, \mathcal{W}_S)$ is transmissible over IC $(\mathcal{X}, \mathcal{Y}, \mathcal{W}_{Y|X})$.

1The bound $2e^{-2\delta^2 k}$ stated in [17, Problem 3.18(b)] is incorrect and must be replaced by $2e^{-\frac{\delta^2}{2}}$.

2$(1 - x)^t \geq 1 - tx$ for $x \in [0, 1]$. 

https://mc.manuscriptcentral.com/t-it
C. Current known coding techniques and sufficient conditions

The current known best strategy for these problems is based on the CES strategy [4] proposed in the context of the MAC problem. One idea of the CES strategy is to induce the source correlation onto channel inputs via S-L test channels \( p_{X_j|S_j} : j \in [2] \). Specifically, the codeword assigned for the source block \( s_j^n \) is picked with PMF \( \prod_{i=1}^n p_{X_j|S_j}(s_{ji}) \). While this idea induces correlation across the input symbols \( X_1, X_2 \), their joint PMF is constrained by a S-L LMC \( X_1 - S_1 - S_2 - X_2 \). A second idea is to exploit the GKW part \( K = f_j(S_j) : j \in [2] \) of the sources, whenever present, to permit a richer class of PMFs for \( X_1, X_2 \). The GKW part is specially coded using a common codebook technique, henceforth referred to as GKW coding. This permits the distributed encoders to agree on an information carrying\(^3\) common RV \( U \) (with a generic PMF). Consequently, the input symbols \( X_1, X_2 \) are not constrained to a S-L LMC \( X_1 - S_1 - S_2 - X_2 \). These ideas lead to the following sufficient conditions, henceforth referred to as CES conditions.

**Theorem 1** (Cover, El Gamal and Salehi, [4]): A pair of sources \((S, W)\) is transmissible over a MAC \((\mathcal{X}, \mathcal{Y}, \mathcal{W}|\mathcal{X})\) if there exist (i) a finite set \( U \), (ii) a PMF \( \mathbb{W}_S p_U p_{X|U} S | p_{X_1|U} S_1, p_{X_2|U} S_2 \mathbb{W}_Y | X_1, X_2 \) on \( \mathcal{S} \times \mathcal{U} \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y} \) such that

\[
H(S_j | S_2) < I(X_j ; Y | X_2, S_2, U) : j \in [2], \quad H(S_1) < I(X_1 ; Y, K, U), \quad H(S_2) < I(X_2 ; Y).
\]

(2)

where \( K = f_j(S_j) : j \in [2] \) taking values in \( K \) is the GKW part of \( S_1, S_2 \).

The reader is referred to [18, Sec. 14.1.1] for a proof of Thm 1.

**Remark 1:** GKW coding crucially relies on the selection of identical codewords by the distributed encoders. This is brought about by the use of identical codes and maps at the two encoders. It may be noted that in general, ignoring the GKW part results in a strictly sub-optimal performance.

The current known best coding technique for the IC problem, henceforth referred to as the LC technique [7] incorporates the technique of random source partitioning designed by Han and Costa [8], Han-Kobayashi technique of message splitting via superposition coding, and the CES strategy. In the following, we provide a characterization of the LC conditions for the specific case when the sources do not possess a GKW part. The reader is referred to [7, Thm. 1] for the general case and the proof of the following theorem.

**Theorem 2** (Liu and Chen, [7]): A pair of sources \((S, W)\) is transmissible over an IC \((\mathcal{X}, \mathcal{Y}, \mathcal{W}|\mathcal{X})\) if there exist (i) finite sets \( W_1, W_2, Q \), (ii) a PMF \( \mathbb{W}_S p Q p_{W_1|Q} p_{W_2|Q} p_{X_1|Q} S_1, p_{X_2|Q} S_2 \mathbb{W}_Y | X_1, X_2 \) defined on \( \mathcal{S} \times \mathcal{Q} \times \mathcal{W} \times \mathcal{X} \times \mathcal{Y} \) such that

\[
H(S_j) < I(S_j, X_j; Y_j | Q, W_j) : j \in [2],
\]

\[
H(S_1) + H(S_2) < \min \left\{ I(S_j, X_j; Y_j | Q, W_j) + I(W_j, S_j, X_j; Y_j | Q) : j \in [2] \right\},
\]

(3)

\[
H(S_1) + H(S_2) < \sum_{j=1}^2 I(S_j, W_j, X_j; Y_j | Q, W_j),
\]

\[
2H(S_j) + H(S_2) < I(S_j, X_j; Y_j | Q, W) + I(S_j, W_j, X_j; Y_j | Q) + I(S_j, W_j, X_j; Y_j | Q, W_j) : j \in [2].
\]

D. Tools : Constant composition codes and the random coding error exponent

The material presented in this section is made use of only in proofs of the theorems in Secs. IV, V. Readers may refer to this material as and when needed in those sections.

The ensemble of constant composition codes studied by Csiszár and Körner [17], [19] prove to be a very useful tool in our study. The following theorem, due to Csiszár and Körner, guarantee the existence of constant composition codes with guaranteed number of codewords and exponentially small error probabilities. In the sequel, we let

\[
l^*(A, B, \rho) := \min \left\{ l \in \mathbb{N} : \exp \left( \frac{l \rho}{2} \right) \geq 2(l + 1)^2 |A| + 2|A||B| \right\},
\]

(4)

\[
= \min \left\{ l \in \mathbb{N} : l \rho \geq \log 4 + (4|A| + 4|A||B|) \log(l + 1) \right\}
\]

where \( A, B \) are finite sets and \( \rho > 0 \).

\(^3\)The reader may distinguish this from a time sharing common RV that the distributed encoders agree.
The capacities of the PTPs $Y$.

Note that in the above eqn. the memoryless PTP is upper bounded is not stated in [17, Thm. 10.2], but is evident from the proof.

In other coding theorems (Thms. 4, 8, 5, 6, 7). Throughout Sec. III, the two following facts in conjunction with the mean value theorem. Firstly, the space of PTP discrete memoryless channel transition probabilities $p_{Y|U}$, B-L $\Lambda \geq \max \{l^*(U, Y, \rho), \rho \} : j \in \{2\}$, a type $p_{Y}$ of sequences in $U^l$, there exist a code $(l, M_u, e_u, d_u)$ of B-L $l$, encoder map $e_u : [M_u] \to U^l$ with codewords $u^l(m) : e_u(m) : m \in [M_u]$ each of type $p_{Y}$, decoder map $d_u : Y^l \to [M_u]$ such that (i) the codebook contains at least $M_u \geq \exp\{l\alpha\}$ codewords, and (ii) probability of error of the code, when employed on the memoryless PTP $(U, Y, P_{Y|U})$, is at most

$$\sum_{y^l \in Y^l} p_{Y|U}(y^l | u^l(m)) I_{\{d(y^l) \neq m\}} \leq (l + 1)^{2|U|^{|Y|}} \exp\{-lE_r(\alpha + \rho, p_U, p_{Y|U})\}$$

for every $m \in [M_u]$.

**Proof:** Follows from [17, Thm. 10.2]. The lower bound of $l^*(U, Y, \rho)$ on $l$ can be traced back to the proof of [17, Thm. 10.1] which forms the main ingredient in the proof of [17, Thm. 10.2]. It maybe noted that $\alpha + \rho, \alpha$ in our statement is equivalent to $R, R - \delta$ in [17, Thm. 10.2]. Lastly, the fact that the maximal probability of error is upper bounded is not stated in [17, Thm. 10.2], but is evident from the proof.

**Corollary 1:** Given any $\alpha > 0, \rho > 0$, finite alphabets $U, Y_1, Y_2$, channel transition probabilities $p_{Y_1Y_2|U}$, B-L $l \geq \max \{l^*(U, Y_1, \rho), l^*(U, Y_2, \rho), Y_2 \} : j \in \{2\}$, a type $p_{Y}$ of sequences in $U^l$, there exist a code $(l, M_u, e_U, d_{U1}, d_{U2})$ with message index set $[M_u]$ encoder map $e_u : [M_u] \to U^l$ with codewords $u^l(m) : d_{U1} = e_U(m) : m \in [M_u]$ each of type $p_{Y}$, and decoder maps $d_j : Y_j \to [M_u]$ such that, (i) the number of codewords $M_u \geq \exp\{l\alpha\}$, and (ii) maximal probability of decoding error of decoder $j$ is at most

$$\sum_{y^l \in Y^l} p_{Y_j|U}(y^l | u^l(m)) I_{\{d_j(y^l) \neq m\}} \leq (l + 1)^{2|U|^{|Y|}} \exp\{-lE_r(\alpha + \rho, p_U, p_{Y_j|U})\}$$

for every $m \in [M_u]$ and for any channel transition probabilities $p_{Y_j|U}$.

**Proof:** Follows from the fact that the bound in [17, Thm. 10.2] applies to every DMC, and in particular the two DMCs $p_{Y_j|U}$ and $p_{Y_2|U}$.

### III. Fixed B-L Coding over Isolated Channels

This section is aimed at explaining the central ideas of this article in a simplified setting. We do this in the context of a MAC example (Ex. 1) obtained via a simple generalization of Dueck’s ingenious example [10]. For Ex. 1, we first reason why the CES scheme is unable to communicate the sources over the MAC. A proof is provided in Appendix A. We then propose a new decoding scheme that permits the Rx recover the sources. In presenting the latter coding scheme and analyzing its performance, we introduce notation and steps of the analysis that reappear in other coding theorems (Thms. 4, 8, 5, 6, 7). Throughout Sec. III, $\xi[\cdot] : = \xi[\cdot](\cdot), \xi : = \xi(\cdot)$ and $\tau_{1, \delta} : = \tau_{1, \delta}(S_1)$.

**Example 1:** Source alphabets $S_1 = S_2 = \{0, 1, \cdots, a - 1\}$. Let $\eta \geq 6$ be a positive even integer. The source PMF is

$$W_{S_1S_2}(c^k, d^k) = \begin{cases} \frac{k-1}{k} & \text{if } c^k = d^k = 0^k \\ \frac{a^k}{ka^{a^k}} & \text{if } c^k = d^k, c^k \neq 0^k \\ 0 & \text{if } c^k = 0^k, d^k \neq 0^k, \text{ and otherwise.} \end{cases}$$

Note that in the above eqn. $c^k, d^k \in S_1$ abbreviate the $k$ digits $c_1c_2 \cdots c_k$ and $d_1d_2 \cdots d_k$, respectively. Fig. 3 depicts the source PMF with $\eta = 6$.

The MAC is depicted in Fig. 4 and described below. The input alphabets are $X_1 = U \times Y_1$ and $X_2 = U \times Y_2$. The output alphabet is $Y_0 \times Y_1 \times Y_2$. $U = Y_0 = \{0, 1, \cdots, a - 1\}$. $X_j : = (U_j, Y_j) \in U \times Y_j$ denotes $X_j$’s input. Moreover, $W_{Y_j|X} = W_{Y_j|U} W_{Y_j|V_j} = W_{Y_0|U} W_{Y_1|V_1} W_{Y_2|V_2}$, where

$$W_{Y_0|U}(y_0 | u_1, u_2) = \begin{cases} 1 & \text{if } y_0 = u_1 = u_2 \text{ or } u_1 \neq u_2, y_0 = 0, \\ 0 & \text{otherwise.} \end{cases}$$

The capacities of the PTPs $(Y_j, Y_j, W_{Y_j|V_j}) : j = 1, 2$ are $C_M = h_k(\frac{2}{a^2}) + \frac{1}{4} \log a$ and $C_M = h_k(\frac{2}{a^2})$, respectively, where $C_M = a \log a + 2h_k(\frac{a}{2}) + \frac{1}{4} \log a, \alpha = \frac{8k^c}{a^2}$. It can be verified that, for sufficiently large $a, k$, the capacity of the satellite channel $W_{Y_j|V_j}$ is at most $\frac{3}{4} \log a$. For all such $a, k$, we choose satellite channels for which $|Y_j| \leq a^\frac{3}{4}$.

**Remark 2:** We have only specified the capacities of the satellite PTP channels $V_j \to Y_j : j \in \{2\}$ leaving its exact specification open. This is because all of our arguments depend only on the capacities of satellite PTP channels. Any PTP channel with the specified capacities can be employed. The existence of such PTP channels follows from two following facts in conjunction with the mean value theorem. Firstly, the space of PTP discrete memoryless
channels endowed with the Euclidean metric is a path-connected metric space [20, Sec. 3.1]. Secondly, the capacity [20, Prop. 2] defined on this path-connected metric space is a continuous function.

In Appendix A, we prove that, for sufficiently large \( a, k \) the S-L CES technique is incapable of communicating \( S_1, S_2 \) over the MAC \( W_{y|x} \). We describe the reasoning below.

Consider sufficiently large values for \( a, k \). In regards to the source, it can be verified that \( H(S_1), H(S_2), H(S) \sim \log a \). Secondly, in regards to the MAC, the sum of the capacities of the satellite channel \( \mathbb{W}_{Y_j|U_j} \in [2] \) is \( o(\frac{\log a}{k}) \). The bulk of the source entropy (log \( a \)) has to be communicated via the shared \( \mathbb{W}_{Y_0|U} \)-channel. Studying the transition matrix \( \mathbb{W}_{Y_0|U} \), we note that in order to communicate close to a sum rate of \( \log a \) bits via the latter, it is necessary that \( U_1 \) must equal \( U_2 \) whp and \( U_1 = U_2 \) must be “close to” uniform. Can we induce such a PMF \( p_{U_1,U_2} \)?

Observe that \( S_1, S_2 \) do not possess a GKW part. The S-L CES technique is therefore constrained by a S-L LMC \( U_1V_1 - S_1 - S_2 - U_2V_2 \). For \( j \in [2] \), \( p_{U_j|S_j} \) can equivalently be viewed as \( U_j = g_j(S_j, W_j) \), for some function \( g_j \) and RVs \( W_1, W_2 \) that are independent. Owing to independence, \( W_1 \) and/or \( W_2 \) being non-trivial RVs, reduces \( P(U_1 = U_2) \). If we let \( W_1, W_2 \) be deterministic, the only way to make \( U_j \) uniform is to pool less likely symbols. However, the source is “highly non-uniform”. Since \( P(S_j = 0^k) \geq 1 - \frac{1}{k} \) for \( j \in [2] \), we can gather a probability of at most \( \frac{1}{k} \) by pooling all the less likely symbols together. Consequently, any \( p_{U_1,U_2} \) induced via a S-L coding scheme is sufficiently far from a PMF that satisfies the twin requirements of \( U_1 = U_2 \) whp and \( U_1 = U_2 \) being close to uniform. Proof of Lemma 1 in Appendix A formalizes this intuition. We therefore conclude that the constraint of S-L LMC is debilitating, and when constrained to a S-L CES technique, the shared channel - the main communication resource in communicating \( S \) to the decoder - cannot be utilized efficiently leading to the incapability of communicating \( S \) over the MAC.

A. Fixed B-L coding over isolated noiseless channels

The CES scheme is able to overcome the S-L LMC constraint only when the sources \( S_1, S_2 \) possess a GKW part \( K = f_1(S_1) = f_2(S_2) \). It does so via GKW coding which involves a shared codebook. Suppose we attempt GKW coding for Ex. 1 with a shared codebook \((U^n(s^n) : s^n \in S^n)\), wherein each codeword is picked independently and uniformly from \( U^n \). As the B-L \( n \) increases, we note that the PMFs of \( U^n(S_1^n) \) and \( U^n(S_2^n) \), individually tend to the desired product PMF \( \prod_{i=1}^{n} P(S_1^n) = P(S_2^n) \). However, \( P(U^n(S_1^n) = U^n(S_2^n)) \) tends to 0 as \( P(S_1^n = S_2^n) = (1 - \xi)^n \to 0 \) as \( n \to \infty \). Since conventional information-theoretic approach insists on choosing \( n \) arbitrarily large, GKW coding completely fails in co-ordinating the codewords input by the Txs, lending GKW coding inapplicable when \( P(S_1 = S_2) = \xi > 0 \), as in Ex. 1.

We therefore propose a coding scheme wherein the B-L of this shared code is fixed to \( l \), irrespective of the desired probability of error. \( l \) is chosen (i) large enough so that the PMFs of \( U^l(S_1^l) \) and \( U^l(S_2^l) \), individually are “reasonably” close to uniform, and (ii) small enough, to ensure \( \xi[l] \) is reasonably small. We refer to these \( l \)-length blocks as sub-blocks. Since \( l \) is fixed, there is a non-vanishing probability \( \phi > 0 \) that these source sub-blocks will be decoded erroneously. If one were to communicate an arbitrarily large number \( m \) of these sub-blocks, the number of erroneously decoded blocks concentrates around \( m\phi \). An outer code, operating on an arbitrarily large number \( m \) of these sub-blocks, will carry information to correct for these “errors” and communicate rest of the necessary information. Specifically, the decoded sub-blocks (including the erroneously decoded ones) from the
GKW coding are treated as side information and the rest of the necessary information is communicated through a standard Slepian-Wolf binning technique via the satellite channels.

We now specify the above codes and develop notation (that we adopt throughout this article). We also analyze the proposed coding technique and these steps will reappear in the subsequent coding theorems. We begin by employing a simple fixed B-L (inner) code. Let $T_0^j(S_1)$ be the source codebook, and let $C_U = U^l$ be the channel codebook. Let $t \alpha = \lceil \log d' \rceil$ bits, of the $\lceil \log |T_0^j(S_1)| \rceil$ bits output by the source code, be mapped to $C_U$. Both encoders use the same source codebook, channel codebook and mapping. The source to channel codebook mapping can be arbitrary. We reiterate that encoder 2 also employs source codebook $T_0^j(S_1)$, and not $T_0^j(S_2)$.

Suppose we communicate an arbitrarily large number $m$ of these sub-blocks on $\mathcal{W}_{Y_0|U}$ as above. Moreover, suppose $\text{Tx} 1$ communicates the rest of the $l \beta = \lceil \log |T_0^j(S_1)| \rceil - t \alpha$ bits output by its source code to $\text{Rx}$ on its satellite channel $\mathcal{W}_{Y_1|V_j}$.

How much more information needs to be communicated to enable the $\text{Rx}$ reconstruct $\hat{S}^m$?

We employ a matrix notation in the sequel. View the $m$ sub-blocks of the source $S_j$ as the rows of the matrix $S_j(1 : m, 1 : l) \in \mathbf{S}_j$. Let $\hat{\mathbf{K}}(1 : m, 1 : l) \in \mathbf{S}_1$ denote $\text{Rx}$’s reconstruction.

The $m$ sub-blocks

$$\left\{ (S_j(t, 1 : l), \hat{\mathbf{K}}(t, 1 : l) : j = 1, 2) : t \in [m] \right\}$$

are IID with an $l$-length distribution $\mathcal{W}_{S_j|S_1} p_{\hat{K}^1|S_1|S_2} = \mathcal{W}_{S_1} p_{\hat{K}^1|S_1|S_2}$. This suggests that we can treat the $l$-length sub-blocks as super-symbols and employ a standard binning technique. It suffices for encoder $j : j \in [2]$ to send $H(S_j^l|\hat{K}_j^l, S_j^l)$ bits per source sub-block, so long as their sum rate is at least $H(S_j^l|\hat{K}_j^l)$. We do not have a characterization of $p_{\hat{K}^1|S_1|S_2}$ and we therefore derive an upper bound. We have

$$H(S_j^l|\hat{K}_j^l, S_j^l) \leq H(S_j^l, 1_{\{\hat{K}_j^l \neq S_j^l\}}|\hat{K}_j^l, S_j^l) \leq h_b(P(\hat{K}_j^l \neq S_j^l)) + H(S_j^l|\hat{K}_j^l, S_j^l, 1_{\{\hat{K}_j^l \neq S_j^l\}})$$

(6)

$$\leq h_b(P(\hat{K}_j^l \neq S_j^l)) + P(\hat{K}_j^l \neq S_j^l) \log |S_j^l| + P(\hat{K}_j^l = S_j^l) H(S_j^l|\hat{K}_j^l, S_j^l)$$

$$\leq l \mathcal{L}_1(P(\hat{K}_j^l \neq S_j^l), |S_j|) + l H(S_j|S_1, S_2), \text{ for } j \in [2], \text{ and }$$

(7)

$$\mathcal{L}_1(P(\hat{K}_j^l \neq S_j^l), |S_j|) \leq h_b(P(\hat{K}_j^l \neq S_j^l)) + P(\hat{K}_j^l \neq S_j^l) \log |S_j^l| + P(\hat{K}_j^l = S_j^l) H(S_j^l|\hat{K}_j^l, S_j^l)$$

$$\leq l \mathcal{L}_1(P(\hat{K}_j^l \neq S_j^l), |S_j|) + l H(S_j|S_1)$$

(8)

represents the additional source coding rate needed to compensate for the errors in the fixed B-L decoding. It suffices to verify that the above rates are supported by the satellite channels. Our first step is to derive upper bounds on the RHSs of (7) and (8). Since $\mathcal{L}_1(\mu, |K|)$ is non-decreasing in $\mu$ for $\mu \in [0, \frac{1}{2}]$, we bound $P(\hat{K}_j^l \neq S_j^l)$ by a quantity that is less than $\frac{1}{2}$, and substitute the same to derive an upper bound on $\mathcal{L}_1(P(\hat{K}_j^l \neq S_j^l), |S_j|)$. Towards that end, note that $\{S_j^l \neq \hat{K}_j^l\} \subseteq \{S_j^l \neq S_j^l\} = T_0^l(S_j) \cup \{A_1 = A_2, \hat{A} \neq A_1\}$, where $A_j$ is the index of the $C_U$-codeword chosen by encoder $j$ and $\hat{A}$ is the output of the $C_U$ decoder at the $\text{Rx}$. Indeed, $S_j^l = S_j^l \in T_0^l(S_1)$ implies both encoders input same $C_U$-codeword and agree on the $l \beta$ bits communicated by encoder 1, in which case the occurrence of $\{S_j^l \neq \hat{K}_j^l\}$ will imply the occurrence of $\{A_1 = A_2, \hat{A} \neq A_1\}$. Since the MAC $\mathcal{W}_{Y_1|U}$ is noiseless, the latter event does not occur and we have

$$P(S_j^l \neq \hat{K}_j^l) \leq \phi,$$

(9)

where $\phi = \xi[l] + t_1 \delta$. Therefore

$$H(S_j^l|\hat{K}_j^l, S_j^l) \overset{(a)}{\leq} l \mathcal{L}_1(\phi, |S_j|) + l H(S_j|S_1, S_2) \text{ and } H(S_j^l|\hat{K}_j^l) \overset{(b)}{\leq} l \mathcal{L}_1(\phi) + l H(S_2|S_1).$$

(10)

We are thus left to identify choices for $l$ and $\delta$ such that $\phi = \xi[l] + t_1 \delta$ is less than $\frac{1}{2}$, and moreover when substituted in the above upper bounds yields rates that are supported on the satellite channels. These steps being specific to Ex. 1 is provided in Appendix B. We conclude with two remarks. First, note that $p_{\hat{K}_j^1|S_1|S_2}$ can in principle be

\[a\]

Through our description, we assume communication over $\mathcal{W}_{Y_1|V_j}$ is noiseless. In the end, we prove that the rate we demand of $\mathcal{W}_{Y_1|V_j}$ is less than its capacity, justifying this assumption.

\[b\]

Encoder $j$ could input any arbitrary $C_U$-codeword when its $\text{sub-block } S_j^l \notin T_0^l(S_1)$, and decoder $j$ could declare an arbitrary reconstruction when it observes $Y_0^l = 0$. Our probability of error analysis handles these events.
computed, once the fixed B-L codes, encoding and decoding maps are chosen, $S^l_{jm}$ will be binned at rate $H(S^l_{j} | \hat{K}^i_j)$ and the decoder can employ a joint-typicality based decoder using the computed $p^{S^l_{j} | \hat{K}^i_j}$.

**Remark 3:** The coding technique proposed here crucially relies on the choice of $l$ being neither too big, nor too small. This is elegantly captured as follows. As $l$ increases, $\xi(l)(\mathcal{S}) \to 1$, $\tau_i, \tau_j \to 0$. As $l$ decreases, $\xi(l)(\mathcal{S}) \to \xi(\mathcal{S})$, and $\tau_i, \tau_j \to 1$. If $\phi \to 0.5$, $\mathcal{L}_1(\phi, |S_j|) \to 0.5 \log |S_j| = \frac{1}{2} \log a$. This tension in the choice of $l$, as we mentioned in Sec. I is fundamental to this problem, not just a symptom of the proposed coding scheme.

**IV. FIXED B-L CODING OVER ARBITRARY MAC AND IC**

**Step 1:** **SEPARATE DECODING**

To facilitate pedagogy, we present our generalization of the coding scheme proposed in Sec. III-A in two steps. In Step I, the fixed B-L and the outer $\infty$-B-L codes are decoded separately. Step II in Sec. V adopts joint decoding.

**A. MAC Problem**

We state our first set of sufficient conditions for the MAC problem. A proof is provided in Sec. VI. In this section, we provide an outline of the new coding and analytical tools that form a main component of our contribution.

**Theorem 4:** A pair of sources $(\mathcal{S}_1, \mathcal{Y}_1)$ is transmissible over a MAC $(\mathcal{X}, \mathcal{Y}, \mathcal{W}_{Y|X})$ if there exist

(i) finite sets $\mathcal{K}, \mathcal{U}, \mathcal{V}_1, \mathcal{V}_2$,

(ii) maps $f_j : \mathcal{S}_j \to \mathcal{K}$, with $K_j = f_j(S_j)$ for $j \in [2],$

(iii) $\alpha, \beta \geq 0$, $\rho > 0$, $\delta > 0$,

(iv) $l \in \mathbb{N}$, $l \geq l^*(\rho, \mathcal{U}, \mathcal{Y})$, where $l^*(\cdot, \cdot, \cdot)$ is defined in (4),

(v) PMF $p_U p_V_1 p_V_2 p_X_1 p_U_1 p_X_2 p_X Y_1 Y_2$ defined on $\mathcal{U} \times \mathcal{V}_1 \times \mathcal{X} \times \mathcal{Y}$, where $p_U$ is a type of sequences in $\mathcal{U}^l$, such that

\[
(1 + \delta)H(K_a) < \alpha + \beta,
\]

\[
H(S_j | S_2, K_a) + \mathcal{L}_1(\phi, |S_j|) \leq I(V_j; Y_2 | V_2) - \mathcal{L}(\phi, |V_j|) + \beta < I(V_j; Y - \mathcal{L}(\phi, |V_j|),
\]

\[
\phi \in [0, 0.5) \text{ where } \phi = g(\alpha + \rho, l) + \xi(l)(\mathcal{K}) + \tau_i(\mathcal{K}_a), \text{ } g(R, l) := (l + 1)^{2l|\mathcal{Y}|} \exp\{-lE_r(R, p_U, p_V Y|U)} \tag{13}
\]

for some $a \in [2]$, where $\mathcal{L}_1(\cdot, \cdot), \mathcal{L}(\cdot, \cdot)$ as defined in (1).

**Remark 4:** The characterization provided here and those in Thms. 5, 6, 8 is via S-L PMFs and S-L expressions. A proof of bounds on the cardinality of auxiliary alphabets for more general coding theorems are provided in [9].

**Remark 5:** The sufficient conditions stated above (Thm. 4) can be strictly weaker than the CES conditions (Thm. 1). In particular, we prove in Appendix C that Ex. 1 satisfies the above conditions.

**Remark 6:** In this Thm. 4 as in the rest of Thms. 5, 6, 8, $K_j = f_j(S_j)$ : $j \in [2]$ can be arbitrary functions of $S_j$ : $j \in [2]$. For comprehension, it helps to visualize $K_1, K_2 \in \mathcal{K}$ as the near GKW parts of the sources $\mathcal{S}_j$. We informally refer to $K_1, K_2$ here and in proof of Thm. 6 as near GKW parts of $\mathcal{S}_j$. This is only to aid intuition. The following outlines put forth the new elements through an informal description and serve as a high level explanation for the bounds in the theorem statement.

**Outline of the coding scheme:** The overall coding technique is based on that proposed in Sec. III-A - fixed B-L GKW coding for $K_1, K_2$, Slepian-Wolf binning for $S_1, S_2$ with resulting indices communicated via an $\infty$-B-L MAC channel code. We also adopt the notation - sub-blocks, matrix of source symbols $S_j$ etc.- developed therein. In this outline, we only elaborate on how the two information streams of different B-Ls are multiplexed. As in Sec. III-A, let $C_U$ denote the fixed B-L channel code (of B-L $l$) employed at both encoders. Codewords from the MAC channel code of B-L $lm$ have to be multiplexed with $m$ codewords from $C_U$. If a single codeword from the former code is multiplexed with $m$ codewords of $C_U$, it experiences a channel with $l$-length memory. Since we seek an efficient technique based on S-L codes and a S-L characterization, we seek sub-vectors of the $lm$ symbols that make up the $m$ codewords of $C_U$, that are IID. The idea is to multiplex codewords of the outer MAC code along these sub-vectors, so that these codewords experience a memoryless channel. We are led to the elegant technique of interleaving devised by Shirani and Pradhan [15] in the related work of distributed source coding. Let rows of $U_j$ denote interleavers of $C_U$ obtained by encoding corresponding rows of $K_j$ via the fixed B-L source encoder and $C_U$. Since the $m$ sub-blocks $K_j(t, 1 : l) : t \in [m]$, are separately and identically coded, the $m$ pairs $(U_1(t, 1 : l), U_2(t, 1 : l)) : t \in [m]$, that constitute rows of $U_1, U_2$, are IID with an $l$-letter PMF $p_{U_1, U_2}$. If one were to randomly, independently and uniformly choose column numbers $\Pi_1 \in [l], \Pi_2 \in [l], \cdots, \Pi_m \in [l]$ from each of the rows, then the $m$ pairs $(U_1(t, \Pi_1), U_2(t, \Pi_2)) : t \in [m]$ are IID $p_{\Pi_1, \Pi_2} : = \frac{1}{l} \sum_{i=1}^{l} p_{U_1, U_2}$ (Lemma 5).
This leads us to the following idea. Suppose \( \Pi_t : [l] \to [l] : t \in [m] \) is a collection of \( m \) random independent and uniformly chosen surjective maps, then for every \( i \in [l] \), the sub-vector \((U_1(t, \Pi_t(i)), U_2(t, \Pi_t(i)) : t \in [m])\) has PMF \( \prod_{t=1}^{m} p_{\theta_t} \). One can therefore multiplex codewords chosen from \( l \) different outer MAC codes with these \( l \) sub-vectors and guarantee that each codeword experiences a memoryless channel.

A second idea concerns the choice of \( C_U \). We will choose and fix a \( C_U \) that we employ for all the \( m \) sub-blocks. Moreover, we do not randomize over the choice of \( C_U \) for our performance characterization. We therefore choose a \( C_U \) with guaranteed good performance. As we mentioned above, we will be multiplexing codewords of the outer MAC channel code along interleaved columns of \( C_U \). Unless we know the \( l \)-letter PMF of the codewords and the distribution of the message indexing \( C_U \), we will not have a characterization for \( p_{\theta_t} = \frac{1}{l} \sum_{t=1}^{l} p_{U_t} \), defined above. In general, we do not have this information of a good code \( C_U \) chosen for an arbitrary channel. We therefore choose \( C_U \) to be a constant composition code of type \( p_U \). Irrespective of the distribution of the message indexing \( C_U \), one can show that a randomly chosen index has PMF \( p_U \) and hence \( p_{\theta_t} = p_U \) (Lemma 6). In addition, Csiszár and Körner have derived tight characterization on their performance which we leverage in our work.

Outline of the analysis: Our analysis too, at a conceptual level, mimics that provided in Sec. III-A. We first characterize (a lower bound on) how much information is communicated through the fixed B-L codes. This is quantified through an upper bound \( \phi \) on \( P \left( K(t, 1 : l) \neq K_1(t, 1 : l) \right) \), where, as in Sec. III-A, \( K \in K \) denote the decoder’s reconstruction of \( K_1 \) based on the GKW coding. In addition to the two terms in (9) that made up \( \phi \), a third term quantifying (the channel coding) probability of error of code \( C_U \) appears in the specification of \( \phi \). Since we employ a constant composition code with \( \exp \{ \alpha \} \) codewords, the reader can identify the term \( g(\alpha + p, l) \) that make up \( \phi \) in conjunction with Thm. 3.6 One point worth noting is that, since \( l \) is fixed, \( K_j \) is not uniform on \( T^l_{\beta}(K_1) \) and hence the codewords of \( C_U \) are not employed uniformly. The fact that Csiszár and Körner provide bounds on the maximal probability of error for constant composition code comes to our rescue.

Next, we derive upper bounds on \( H(S_j^t | K^l, S_j^l) \), \( H(S_j^l | K^l) \) - the amount of information that needs to be communicated via the MAC code. This proceeds in the same spirit as in (6) - (8), and the reader can identify the terms on the LHS of (11) and (12) with \( K_1 \) substituted for \( S_j \) in the RHS of (10)(a) and (10)(b), respectively. Indeed, \( \beta \) as defined in Sec. III-A and appears only in the sum rate bound.

Lastly, we dwell on how much information can be communicated through the MAC channel code. This constitutes the new elements of our analysis. Our MAC channel code operates on input alphabets \( \mathcal{V}_1, \mathcal{V}_2 \) and output alphabet \( \mathcal{Y} \). What is the transition matrix of this effective MAC channel \( \mathcal{V}_1 \times \mathcal{V}_2 \to \mathcal{Y} \)? To answer this, we investigate the joint distribution induced by the coding scheme on the Cartesian product \( U \times U \times \mathcal{V}_1 \times \mathcal{V}_2 \times S_j \). Suppose \( C_U \) is made of message index set \( [M_a] \) and codewords \( (u^l(m) : m \in [M_a]) \). At Tx \( j \), the fixed B-L typical set source code encodes the \( t \)-th sub-block \( K_j(t, 1 : l) \) into a message part, of which indexes \( C_U \). Let \( A_{jt} \) denote this latter part. We therefore have the chosen \( C_U \) codeword in the \( t \)-th sub-block to be \( U_j(t, 1 : l) = u^l(A_{jt}) \). We note that PMF of \( (A_{1t}, A_{2t}) \) is invariant with \( t \), and hence let \( (A_1, A_2) \in [M_a] \times [M_a] \) have the same PMF of \( (A_{1t}, A_{2t}) \). The rows of \( U_1, U_2 \) are IID with PMF

\[
p_{U_1|U_2}(u'_1, u'_2) = \sum_{(a_1, a_2) \in [M_a] \times [M_a]} p(A_1 = a_1, A_2 = a_2) \text{I} \{ u^l(a_j) = u'_j : j \in [2] \}
\]

By choosing codewords of the \( l \) outer codes IID with PMF \( \prod_{t=1}^{m} p_{\theta_t} \) and the mapping from \( U \times U \to X_j \) IID with PMF \( \prod_{t=1}^{m} p_{X_j|U, \mathcal{V}_j} \), we ensure that the PMF of the \( l \)-length sub-blocks on \( U \times U \times \mathcal{V}_1 \times \mathcal{V}_2 \times X_1 \times X_2 \times Y \) is

\[
p_{U|Y}(u^l, v^l, x^l, y^l) = \left[ \sum_{(a_1, a_2) \in [M_a] \times [M_a]} p(A_1 = a_1, A_2 = a_2, (u(a_j) = u_j^l : j \in [2])) \right] \times \left[ \prod_{j=1}^{2} \left\{ \prod_{i=1}^{l} p_{V_j|U, \mathcal{V}_j}(v_{ji} | u_{ji}, v_{ji}) \right\} \right] \times \left[ \prod_{i=1}^{l} p_{Y|X_1, X_2}(y_i | x_{i1}, x_{i2}) \right]. \tag{14}
\]

6The explanation of \( \rho \) can be gleaned from the Proof of Thm. 3.
In other words, our coding scheme of B-L $lm$ which maybe viewed as $m$ sub-blocks of length $l$ induces a PMF $\prod_{i=1}^{m} p_{U_i}^{(i)} Y_{1i}^{(i)}(u_i^{(i)}, l_i^{(i)}, y_i^{(i)})$ on $U_{1m} \times U_{2m} \times Y_{1m} \times Y_{2m} \times X_{1m}^l \times X_{2m}^l \times Y_{lm}$. Each of the $l$ outer codes, operating on interleaved columns of these $m$ sub-blocks will experience a MAC with transition probabilities $p_{\Sigma \mid Y, \Sigma}$, where

$$p_{\Sigma \mid Y, \Sigma}^{(i)} (a, b, c, d) = \frac{1}{l} \sum_{i=1}^{l} p_{U_i, Y_i, X_{1i}, X_{2i}} (a_1, a_2, b_1, b_2, c_1, d),$$

and $p_{U_i, Y_i, X_{1i}, X_{2i}}$ is the PMF of the $i$-th coordinate of the Cartesian product of vectors $U_i^l, V_i^l, X_1^l, X_2^l, Y_i^l$ which is distributed with PMF (14). The rates of the $i$th MAC outer code is therefore constrained to lie within the achievable region of the MAC $(V_1, V_2, Y, p_{\Sigma \mid Y, \Sigma})$ corresponding to the PMF $p_{\Sigma \mid Y, \Sigma}$.

We are left to quantify $I(\gamma_1^{(i)}, \gamma_2^{(i)})$ and $I(\gamma_1^{(i)}, \gamma_2^{(i)})$ in terms of the PMF $p_{U, Y, X_{1}, X_{2}}$ provided in the theorem statement. We derive lower bounds on the above quantities. The reader is referred to the material following (36) through (44) where we prove that if $\frac{1}{2} \geq \epsilon > P(A_1 \neq A_2)$ and $C_U$ is a constant composition code of type $p_U$, then (40), (44) are lower bounds on $I(\gamma_1^{(i)}, \gamma_2^{(i)})$ and $I(\gamma_1^{(i)}, \gamma_2^{(i)})$, respectively. Recognize that (40) is $I(V_j, Y \mid V_j) - L(\epsilon, |Y|)$ and (44) is $I(V_1, V_2, Y) - L(\epsilon, |Y|)$. Since $\epsilon = \pi_{\tau}(K_1) + \xi_{\gamma}(K) < \phi$ and $L(\mu, |A|)$ is increasing in $\mu$ for $\mu \in (0, \frac{1}{2}]$, we are led to the sufficient condition that (i) the RHS of (10)(a) must be less than $H(V_j ; Y \mid V_j) - L(\phi, |Y|)$ for $j \in [2]$, and the (ii) sum of $1/2$ and RHS of (10)(b) must be less than $H(V_1, V_2 ; Y) - L(\phi, |Y|)$. These are indeed the sufficient conditions characterized in Thm. 4.

**B. IC Problem**

Our results in this section are analogous to those presented in Sec. IV-A for the MAC problem. We provide a new set of sufficient conditions for the IC problem in Thm. 5 that can be strictly weaker than the LC conditions.

**Theorem 5:** A pair of sources $(S, \Sigma)$ is transmissible over an IC $(X, \Sigma, Y \mid X, Y)$ if there exist

(i) finite sets $K, U, V_1, V_2$,

(ii) maps $f_j : S_j \rightarrow K_j$, with $K_j = f_j(S_j)$ for $j \in [2],$

(iii) $\alpha, \beta > 0, p^*, \delta > 0$,

(iv) $l \in \mathbb{N}, l \geq \max \{l^* (\rho, U, \gamma_j) : j \in [2] \}$, where $l^*(\cdot, \cdot, \cdot)$ is defined in (4),

(v) PMF $p_{U, V_1, V_2} p_{X_1} p_{U, V_1, V_2} \mid Y \mid X$ defined on $U \times V \times X \times X$, where $p_U$ is a type of sequences in $U^l$, such that

$$(1 + \delta) H(K_a) \leq \alpha + \beta, \quad H(S_j \mid K_a) + \beta + L_l(\phi_j, S_j) - I(V_j ; Y_j) - L(\phi_j, |V_j|) \quad \text{for } j \in [2], \quad \phi_j \leq \frac{1}{2}$$

where $\phi_j = g_j(\alpha, \rho, \lambda) + \xi_{\gamma}(K) + \pi_{\tau}(K_a), \quad g_j(R, l) : = (l + 1)^{2l(l+1)} \exp\{-lE_l(R, P_U, P_{Y_j \mid U})\}$ for $j \in [2]$,

and some $\alpha \in [2]$, where $L_l(\cdot, \cdot, \cdot), L(\cdot, \cdot, \cdot)$ is as defined in (1).

The proof contains no new elements beyond those presented for Thm. 4. The only differences are those that arise due to the separation of the decoders and are handled in the standard way. For example, the $\beta$ bits of the GKW code has to be transmitted by both encoders, and one has to characterize $\phi_j : j = 1, 2$ taking into account the error probability $g_j(\alpha, \rho, l) : j = 1, 2$ of $C_U$ on the two PTP channels $V_j \rightarrow Y_j$ in accordance with Corollary 1. We therefore omit a proof of the above theorem. We conclude with a remark that highlights its significance.

**Remark 7:** The sufficient conditions stated above can be strictly weaker than the LC conditions (Thm. 2). We prove this in Appendix E.

**V. Fixed B-L Coding Over Arbitrary MAC and IC Step 2: Conditional Decoding**

We enhance the coding scheme presented in Sec. IV via conditional decoding. In Step 1, the fixed B-L and $\infty$-B-L information streams caused interference to each other. The interference from the former can be nullified by conditional coding of the latter. Step 2 is based on this approach.

$^7$It can be verified that marginal $p_{\Sigma \mid Y_j}$ corresponding to PMF (14) factors as $\prod_{i=1}^{m} p_{V_i, \Sigma}$ and hence $p_{\Sigma \mid Y_2} = p_{V_1, \Sigma}$. These and other properties of (14) can be found in Lemma 2 of Appendix D.

$^8$The necessary notation is provided therein and the arguments can be easily followed.
A. IC Problem Step II: Joint decoding of Fixed and $\infty$-B-L information streams

It is natural to expect the sufficient conditions to take the form of (16) with $I(V_j; Y_j)$ on the RHS replaced by $I(V_j; Y_j|U)$ ignoring the change in correction terms $L_i(\cdot, \cdot)$. Indeed, as we will see, all of the sufficient conditions presented for the IC will involve corresponding substitutions. We present our first set of sufficient conditions for the IC based on conditional decoding of the outer code.

Theorem 6: A pair of sources $(\mathcal{S}, \mathcal{W}_Y)$ is transmissible over an IC $(\mathcal{X}_1, \mathcal{Y}_1, \mathcal{W}_Y|\mathcal{X}_1)$ if there exist

(i) finite sets $\mathcal{K}, \mathcal{U}, \mathcal{V}_1, \mathcal{V}_2$,

(ii) maps $f_j : \mathcal{S}_j \rightarrow \mathcal{K}$, with $K_j = f_j(\mathcal{S}_j)$ for $j \in [2]$,

(iii) $\alpha, \beta \geq 0$, $\rho > 0$, $\delta > 0$,

(iv) $l \in \mathbb{N}, l \geq \max\{l^*(\rho, \mathcal{U}, \mathcal{Y}_j) : j \in [2]\}$, where $l^*(\cdot, \cdot, \cdot)$ is defined in (4).

(v) PMF $p_U p_V p_{X_1|U} p_{X_2|U} p_{X_1|V_1} p_{X_2|V_2} \mathbb{W}_Y|\mathcal{X}_1 \mathcal{X}_2$ defined on $\mathcal{U} \times \mathcal{V}_1 \times \mathcal{X}_1 \times \mathcal{V}_2 \times \mathcal{X}_2$, where $p_U$ is a type of sequences in $\mathcal{U}$, such that for some $a \in [2]$, we have

\[
(1 + \delta)H(K_0) \leq \alpha + \beta, \quad H(S_j|K_0) + \beta + L_i(\phi_j, |S_j|) < I(V_j; Y_j|U) - L(\phi_j, |V|) \quad \text{for } j \in [2], \quad \phi_j \leq \frac{1}{2}(18)
\]

where $\phi_j = g_j(\alpha + \rho, l) + \xi[j](\mathcal{K}) + \tau_j(\mathcal{K})$,

\[
g_j(R, l) := (l + 1)^2[l|V|] \exp\{-lE[\rho, p_U, p_{Y_j|U}]\}, \quad \text{for } j \in [2] (19)
\]

Remark 8: If the sources have a GKW part $K = K_1 = K_2$, then $\xi(K) = 0$. One can choose $l$ arbitrarily large such that $\phi_1, \phi_2$ can be made arbitrarily small. The resulting inner bound corresponds to a very simple separation based scheme involving a common message communicated over the IC.

In the interest of brevity, we only provide the two central elements of the proof. In Sec. VI-B, we characterize (i) the PMF of the collection of matrices comprising of the transmitted, received and decoded codewords, and (ii) lower bounds on the capacity of the IC experienced by the interleaved $V_j$–codewords when subject to conditional decoding. A full proof is available in [21].

B. IC problem: Conditional decoding via Han Kobayashi technique

In communicating the $\infty$-B-L information stream, we can employ the Han-Kobayashi technique of message splitting via superposition coding. Each encoder builds outer codes on $\mathcal{W}_j, \mathcal{V}_j$ with the former carrying the public part and the latter, the private part. The output of the Slepian Wolf binning code is split into two parts, each indexing one of the above codes. A conditional Han Kobayashi decoding technique utilizing the interleaved vectors of the decoded fixed B-L code is employed. We present the following set of sufficient conditions. Techniques developed in Sec. V-A, in conjunction with Han-Kobayashi technique are employed to prove achievability. The following characterization of the Han-Kobayashi region is from [22].

Definition 3: Let $\mathbb{D}(\mathbb{W}_Y|\mathcal{X})$ denote the collection of PMFs $p_{U V W X_1 X_2}$ defined on $\mathcal{U} \times \mathcal{V}_1 \times \mathcal{X}_1 \times \mathcal{X}_2$ such that $U, V_j, W_j : j \in [2]$ are finite sets. For $p_{U V W X_1 X_2} \in \mathbb{D}(\mathbb{W}_Y|\mathcal{X})$, let $\alpha_{CHK}(p_{U V W X_1 X_2})$ be defined as the set of pairs $(R_1, R_2)$ that satisfy

\[
\begin{align*}
R_j & \leq d_j \\
R_j & \leq a_j + f_j \\
R_1 + R_2 & \leq e_1 + e_2 \\
2R_1 + R_2 & \leq 2a_j + f_j + e_j \\
-2R_1 & \leq 0
\end{align*}
\]

for $j \in [2]$ (20)

where

\[
\begin{align*}
a_j & = I(Y_j; V_j|U, W_j) - L(\phi_j, |V_j|) \\
b_j & = I(Y_j; W_j|U, V_j, W_2) - L(\phi_j, |W_2|) \\
c_j & = I(Y_j; W_j|U, V_j, W_2) - L(\phi_j, |W_2|) \\
d_j & = I(Y_j; V_j, W_j|U, W_2) - L(\phi_j, |V_j|) \\
e_j & = I(Y_j; V_j, W_2|U, W_j) - L(\phi_j, |W_2|) \\
f_j & = I(Y_j; W_j, W_2|U, V_j) - L(\phi_j, |W_2|)
\end{align*}
\]

for $j \in [2]$. We let

\[
\alpha_{CHK}(\mathbb{W}_Y|\mathcal{X}) = \text{coel} \left( \bigcup_{p_{U V W X_1 X_2} \in \mathbb{D}(\mathbb{W}_Y|\mathcal{X})} \alpha(p_{U V W X_1 X_2}) \right)
\]

https://mc.manuscriptcentral.com/t-it
where \( \mathcal{L}(\mu, |A|) = h_b(\mu) + \mu \log |A| \) for any \( \mu \in (0, 0.5) \), finite set \( A \), \( \text{cocl}(A) \) denotes the convex closure of \( A \subseteq \mathbb{R}^2 \).

**Theorem 7:** A pair of sources \((S, W)\) is transmissible over an IC \((X, Y, W_{Y|X})\) if there exist

(i) finite sets \(K, U, V_1, V_2, W_1, W_2\),
(ii) maps \(f_j : S_j \rightarrow K\), with \(K_j = f_j(S_j)\) for \(j \in [2]\),
(iii) \(\alpha, \beta \geq 0, \rho > 0, \delta > 0\),
(iv) \(l \in \mathbb{N}, l \geq \max\{l^*(\rho, U, Y) : j \in [2]\}\), where \(l^*(\cdot, \cdot, \cdot)\) is defined in (4),
(v) \(\text{PMF} p_U p_V p_{W_{Y|X}} p_{W_X} p_{S_1} p_{S_2} p_{W_{Y|X}}\) defined on \(U \times V \times X \times Y\), where \(p_U\) is a type of sequences in \(U^l\), such that for some \(a \in [2]\), we have

\[
(1 + \delta) H(K_a) \leq \alpha + \beta, \quad (H(S_j|K_a) + \beta + \mathcal{L}_l(\phi_j, |S_j|) : j \in [2]) \in \mathcal{C}(K_W) \quad (22)
\]

where \(\phi_j : = g_j(\alpha + \rho, l) + \xi_j(K_a), g_j(R, l) : = (l + 1)^2 l \cdot l^{|Y|} \exp\{-l E_r(R, p_U, p_{W_{Y|U}})\}\), for \(j \in [2]\), \(\mathcal{L}_l(\cdot, \cdot, \cdot)\) is as defined in (1).

**C. MAC Problem**

We present our second and final coding theorem for the MAC problem, wherein we incorporate conditional decoding of the outer code.

**Theorem 8:** A pair of sources \((S, W)\) is transmissible over a MAC \((X, Y, W_{Y|X})\) if there exist

(i) finite sets \(K, U, V_2\),
(ii) maps \(f_j : S_j \rightarrow K\), with \(K_j = f_j(S_j)\) for \(j \in [2]\),
(iii) \(\alpha, \beta \geq 0, \rho > 0, \delta > 0\),
(iv) \(l \in \mathbb{N}, l \geq \max\{l^*(\rho, U, Y) : j \in [2]\}\), where \(l^*(\cdot, \cdot, \cdot)\) is defined in (4),
(v) \(\text{PMF} p_U p_V p_{W_{Y|X}} p_{W_X} p_{S_1} p_{S_2} p_{W_{Y|X}}\) defined on \(U \times V \times X \times Y\), where \(p_U\) is a type of sequences in \(U^l\), such that for some \(a \in [2]\), we have

\[
(1 + \delta) H(K_a) \leq \alpha + \beta, \quad (H(S_j|K_a) + \beta + \mathcal{L}_l(\phi_j, |S_j|) : j \in [2]) \in \mathcal{C}(K_W) \quad (24)
\]

\[
\beta + H(S_j|K_a) + \mathcal{L}_l(\phi_j, |S_j|) \leq I(V_j; Y|U, V_2) - \mathcal{L}(\phi, |V_2|) \quad (25)
\]

\[
\phi \in [0, 0.5] \text{ where } \phi : = g(\alpha + \rho, l) + \xi_j(K_a), g(R, l) : = (l + 1)^2 l \cdot l^{|Y|} \exp\{-l E_r(R, p_U, p_{Y|U})\}\quad (26)
\]

\(\mathcal{L}_l(\cdot, \cdot, \cdot)\) is as defined in (1).

**Remark 9:** If the sources have a GKW part \(K = K_1 = K_2\), then \(\xi(K) = 0\). One can choose \(l\) arbitrarily large such that \(\phi\) can be made arbitrarily small. The resulting inner bound corresponds to separation based scheme involving a common message communicated over the MAC.

There are no new elements beyond those presented in proofs of Thms. 4, 6. The reader is referred to [21] wherein the key error events have been analyzed from first principles.

**VI. PROOFS OF THE CODING THEOREMS**

**A. Proof of Thm. 4**

**Coding Scheme:** Let \(K, \ldots\) be provided as in theorem statement. We assume \(a = 1\). The case \(a = 2\) is treated analogously. For simplicity we assume \(\beta = 0\) and \(\alpha > (1 + \delta) H(K_1)\). The B-L of the coding scheme is \(lm\), where \(l\) remains fixed and \(m\) is chosen arbitrarily large to satisfy the target probability of error. Let \(S_j \in S_j\) denote the source symbols observed by encoder \(j\). Specifically, \(S_j(t, i)\) is the symbol observed by encoder \(j\) during symbol interval \((t - 1)l + i\) for \((t, i) \in [m] \times [l]\) and \(K_j \in K\) is defined as \(K_j(t, i) = f_j(S_j(t, i))\) for \((t, i) \in [m] \times [l]\).

We first specify the codes, maps and follow it up with the encoding, decoding rules.

**Base-layer codes:** (i) Source code : Message index set \([T_{\delta}^l(K_1)]\), encoder map \(e_K : K^l \rightarrow [T_{\delta}^l(K_1)]\) and decoder map \(d_K : [T_{\delta}^l(K_1)] \rightarrow K^l\) such that \(d_K(e_K(k^l)) = k^l\) for \(k^l \in T_{\delta}^l(K_1)\) define the source code. (ii) Channel code : A constant composition code \(C_U = (l, M_u, e_u, d_u)\) of B-L \(l\) built over \(U\) consisting of \(M_u \geq |T_{\delta}^l(K_1)|\) codewords each of type \(p_U\) serves as the base-layer channel code. \(C_U\) is characterized via encoder map \(e_u : [M_u] \rightarrow U^l\) and decoder map \(d_u : U^l \rightarrow [M_u]\). \(C_U\) is chosen such that \(M_u \geq |T_{\delta}^l(K_1)|\) and

\[
\max_{a \in [M_u]} \sum_{y^l : d_u(y^l) \neq a} p_{Y^l|U}(y^l|e_u(a)) \leq g(\alpha + \rho, l, \cdot).
\]

https://mc.manuscriptcentral.com//t-it
Since $l \geq l^*(\rho, \mathcal U, \mathcal V)$, $\alpha > (1 + \delta) H(K_1)$ and $|T^l(K_1)| \leq \exp\{l(1 + \delta) H(K_1)\}$, the existence of such a constant composition code $C_{\mathcal V}$ is guaranteed by Corollary 1.

**Encoding of base-layer codes:** Base-layer coding in illustrated in Fig. 5. Let $A_{jt} := e_K(K_j(t, 1 : l))$ denote the message output by the B-L $l$ source code $T^l(K_1)$ corresponding to the $t$-th row of $K_j$. Without pooling these messages, Tx $j$ communicates $A_{jt} : t \in [m]$ via channel code $C_{\mathcal V}$. We let $u^l(a) : e_u(a) : a \in [M_u]$, and for $a \in [M_u]^m$, we let $u(a) \in \mathcal U$ be defined through $u(a)(t, 1 : l) = u^l(a_t) : t \in [m]$. Base layer coding will result in encoder $j$ identifying $U_j = u(A_j)$.

**Satellite-layer codes:** The satellite-layer codes communicate the Slepian-Wolf bin index of the observed source matrix $S_j$. This bin index, as mentioned in the outline (Sec. IV-A), is split into $l$ indices and communicated through $l$ channel codes. We begin with the description of the Slepian-Wolf source code. (i) The Slepian-Wolf source code is characterized via a partition of $S_j^{lm}$ into $M_{V_j}$ bins. Let $S_j^{lm} \rightarrow [M_{V_j}]$ denote the partition map and let $B_j = (B_{j1}, \cdots, B_{jl}) \in [M_{V_j}]^l$ denote the bin index of $S_j$. (ii) $l$ channel codes - $C_{V_1}, \cdots, C_{V_l}$ - each built over $V_j$ communicate the bin index $(B_{j1}, \cdots, B_{jl})$. For $i \in [l]$, channel code $C_{V_j,i} = (m, M_{V_j}, e_{V_{ji}}, d_{V_{ji}})$ is of B-L $m$, has message index set $[M_{V_j}]$ and is characterized via encoder map $e_{V_{ji}} : [M_{V_j}] \rightarrow [V_j]^m$, decoder map $d_{V_{ji}} : V_j^m \rightarrow [M_{V_j}] \times [M_{V_j}]$. We let $v_{ji}^{m}(b) : b \in [M_{V_j}]$ denote the codewords of $C_{V_{ji}}$. Tx $j$ has to multiplex the collection $v_{ji}^{m}(B_{ji}) : i \in [l]$ of codewords with $U_j = u(A_j)$.

**Multiplexing unit:** The multiplexing unit (depicted in Fig. 6) consists of (i) $m$ surjective maps $\pi_t : [l] \rightarrow [l] : t \in [m]$, and (ii) predefined matrices $x_j(u, v_j) \in \mathcal X_j$ for $j \in [2], u \in \mathcal U, v_j \in \mathcal V_j$. The $m$ surjective maps enable us identify $l$ sub-vectors of $U_j$, along which the codewords chosen from $C_{V_1,1}, \cdots, C_{V_l,1}$ will be multiplexed. Specifically, $v_{ji}^{m}(B_{ji})$ will be multiplexed with the $i$-th interleaved column $(U_{jt}(t, \pi_t(i)) : t \in [m])$. We employ the following notation in the sequel which greatly simplifies exposition in relation to interleaving.

For $A \in \mathcal A$, and a collection $\lambda_t : [l] \rightarrow [l] : t \in [m]$ of surjective maps, we let $A^\lambda \in \mathcal A$ be such that $A^\lambda(t, i) := A(t, \lambda_t(i))$ for each $(t, i) \in [m] \times [l]$. To reduce clutter, we let $A^\lambda = A^{\lambda_t}$. If $A \in \mathcal A, B \in \mathcal B$, then $[AB]^\lambda(1 : m, i) := (A^\lambda(1 : m, i), B^\lambda(1 : m, i))$.

For $j \in [2], B_j \in [M_{V_j}]^l$, we let $v_j(B_j) \in \mathcal V_j$ be defined through $v_j(B_j)^m : (1 : m, i) = v_{ji}^{m}(B_{ji}) : i \in [l]$. Our last step in the encoding rule is to map $U_j := u(A_j), V_j := v(B_j)$ into channel inputs on $X_j$. Encoder $j$ maps $(A_j, B_j) \in [M_{V_j}]^m \times [M_{V_j}]^l$ into $x_j(u(A_j), v(B_j)) \in \mathcal X_j$. For $(t, i) \in [m] \times [l]$, the encoder inputs symbol $x_j(A_j, B_j)(t, i)$ on the channel during symbol interval $(t - 1)l + i$.

**Encoding rule:** Refer to Table I and Figs. 5, 7, 6. Tx $j$ observes $A_j = (A_{jt} : t \in [m]) \in [M_u]^m$ and $B_j \in [M_{V_j}]$. For

https://mc.manuscriptcentral.com//t-it
**Fig. 7. Satellite Layer Coding**

![Slepian-Wolf Binning Code](image)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_2^1(K_1)$</td>
<td>Source code employed at both encoders</td>
<td>$T_2^1(K_1)$ is used to encode rows of $K_1, K_2$.</td>
</tr>
<tr>
<td>$e_K : K^l \rightarrow [T_2^1(K_1)]$</td>
<td>Encoder map of (common) typical set source code</td>
<td>$e_K$ are fixed B-L $l$.</td>
</tr>
<tr>
<td>$A_{jl}$</td>
<td>Index output by $T_2^1(K_1)$ corresponding to $K_j(t, 1 : l)$</td>
<td>$A_{jl} = (A_{j1}, \cdots, A_{jm})$.</td>
</tr>
<tr>
<td>$d_K : [T_2^1(K_1)] \rightarrow T_3^1(K_1)$</td>
<td>Decoder map of (common) typical set source code</td>
<td>$d_K$ is constant composition of type $p_U$.</td>
</tr>
<tr>
<td>$C_U$</td>
<td>The common channel code over $U$ of fixed B-L $l$. Employed at both encoders.</td>
<td>$C_U$ is constant composition of type $p_U$ and $\beta \rightarrow 0$, we have $M_u \geq</td>
</tr>
<tr>
<td>$[M_u]$</td>
<td>Message Index set of $C_U$</td>
<td>$e_{V_{ji}} : [M_v] \rightarrow [V_u^\ell]$</td>
</tr>
<tr>
<td>$v_{ji} : [V_v] \rightarrow [V_u^\ell]$</td>
<td>Encoder map of $C_{V_{ji}}$.</td>
<td>Message output by this code lies in $[M_v]$.</td>
</tr>
<tr>
<td>$d_{V_j} : Y^m \rightarrow [M_v_1] \times [M_v_2]$</td>
<td>Decoder map of MAC channel code $C_{V_{1,i}, V_{2,i}}$.</td>
<td>$\beta_j : S_j^m \rightarrow [M_v]$</td>
</tr>
<tr>
<td>$\pi_j : [l] \rightarrow [l]$</td>
<td>Surjective maps employed for multiplexing</td>
<td>$\pi_j : [l] \rightarrow [l]$</td>
</tr>
<tr>
<td>$a_j(t, \pi_j(i))$</td>
<td>Predefined matrices employed for mapping</td>
<td>$a_j(t, \pi_j(i))$</td>
</tr>
<tr>
<td>$x_j(u, v) : U \times V_j \rightarrow X_j$</td>
<td>$x_j : U \times V_j \rightarrow X_j$</td>
<td>$x_j : U \times V_j \rightarrow X_j$</td>
</tr>
<tr>
<td>$\theta_j(u, v) : = u$ and $v_j(\cdot)$ &amp; $x_j(u, v)$</td>
<td>$\theta_j(u, v) : = u$ and $v_j(\cdot)$ &amp; $x_j(u, v)$</td>
<td>$\theta_j(u, v) : = u$ and $v_j(\cdot)$ &amp; $x_j(u, v)$</td>
</tr>
</tbody>
</table>

**TABLE I**

**DESCRIPTION OF ELEMENTS THAT CONSTITUTE THE CODING SCHEME**

$t \in [m]$, let $U_j(t, 1 : l) : = e_{u}(A_{jl}) = u^l(A_{jl})$. For $i \in [l]$, let $V_j^l(1 : m, i) = v_j^m(B_{ji})$. Let $X_j = x_j(U_j, V_j)$. $X_j(t, i)$ is input on the channel during symbol-interval $(t - 1)l + i$.  

**Decoding rule:** Let $Y \in Y$ denote the matrix of received symbols with $Y(t, i)$ being the symbol received during symbol interval $(t - 1)l + i$. The channel decoder attempts to recover $(A_1, B_1, B_2) \in [M_u]^{m} \times [M_v_1]^l \times [M_v_2]^l$.  

**Base-layer channel decoder:** $d_u$ - the decoder of $C_U$ - operates on rows of $Y$ separately to output $\hat{A}_l := d_u(Y(t, 1 : l)) : t \in [m]$ - the decoded messages corresponding to $A_1$.

---

9 Note $v_j^m(1 : m, i) = v_j(1, \pi_1(i)) \cdots v_j(m, \pi_m(i))$. 

[Link to manuscript](https://mc.manuscriptcentral.com//t-it)
\textbf{Satellite-layer channel decoder:} This unit aims to decode $\hat{B}_1, \hat{B}_2$. Let

\[ p_{U^l V^l X^l Y^l}(u^l, v^l, x^l, y^l) = \sum_{(a_1, a_2) \in \mathcal{M}_a \times \mathcal{M}_a} P(A_1 = a_1, A_2 = a_2) \mathbb{I}\{u^l(v_i) = \hat{u}^l_i, j \in [2]\} \times \prod_{j=1}^{l} \left\{ \prod_{i=1}^{2} p_{V_i}(v_{ij}) p_{X_i | U^l V^l}(x_{ij} | u_{ij}, v_{ij}) \right\} \times \prod_{i=1}^{l-1} \mathbb{W}_{Y_i}(y_{i} | x_{i1}, x_{i2}) \tag{28} \]

be a PMF on $\mathcal{U}^l \times \mathcal{V}^l \times \mathcal{X}^l \times \mathcal{Y}^l$ and

\[ p_{\hat{U}_i \hat{V}_i \hat{X}_i \hat{Y}_i}(\hat{u}_i, \hat{v}_i, \hat{x}_i, \hat{y}_i ; a, \hat{b}, \hat{c}, d) := \frac{1}{l} \sum_{i=1}^{l} p_{U_i, V_i, X_i, Y_i}(a, \hat{b}_i, \hat{c}_i, d) \tag{29} \]

be a PMF on $\mathcal{U} \times \mathcal{V} \times \mathcal{X} \times \mathcal{Y}$. In Appendix D, we list and prove certain simple properties of PMFs (28), (29) that we will have opportunity to leverage in the sequel. For $i \in [l]$, populate $D_i(Y) := \left\{ (\hat{b}_{i1}, \hat{b}_{i2}) : (v^m_{i1}(\hat{b}_{i1}), v^m_{i2}(\hat{b}_{i2}), Y^\pi(1 : m, i)) \text{ is jointly typical wrt } \prod_{t=1}^{m} p_{\hat{V}_t \hat{Y}_t} \right\}$. For $i \in [l]$ such that $D_i(Y)$ is empty, set $\hat{B}_{i1} = \hat{B}_{i2} = (1, 1)$. For $i \in [l]$ such that $D_i(Y)$ is not empty, choose one among the pairs in $D_i(Y)$ uniformly at random, and set $(\hat{B}_{i1}, \hat{B}_{i2})$ to be that pair. Note that if $D_i(Y)$ is a singleton for each $i \in [l]$, there is a unique choice for $\hat{B}_1, \hat{B}_2$. The channel code decoder forwards $\hat{A}, \hat{B}_1, \hat{B}_2$ to the source code decoder.

\textbf{Base-layer source decoder:} The decoded messages $\hat{A}_i : t \in [m]$ is mapped to the corresponding typical sequences in $T^1_\beta(K_1)$. Let $\hat{K}(t, 1 : l) : t \in [m]$ denote the corresponding sequences. The map from $\hat{A}$ to $\hat{K}$ is via the decoder of the fixed B-L typical set source code $T^1_\beta(K_1)$.

\textbf{Satellite-layer source decoder:} In the second step, the Slepian Wolf decoder performs a standard joint-typicality decoding within the indexed pair $\hat{B}_1, \hat{B}_2$ of bins, treating the rows of $\hat{K}$ as $m$ super-symbols of side-information. Specifically,

\[ \mathcal{D}(\hat{K}, \hat{B}_1, \hat{B}_2) := \left\{ (s_1, s_2) : \beta(s_j) = \hat{B}_j : j \in [2], \text{ and } (\hat{K}, s_1, s_2) \text{ is jointly typical wrt } \prod_{t=1}^{m} p_{K^t \hat{S}_1^t \hat{S}_2^t} \right\}, \]

where, for $s_1^t = s_{11} s_{12} \cdots s_{1t}, s_2^t = s_{21} s_{22} \cdots s_{2t}$,

\[ p_{K^t \hat{S}_1^t \hat{S}_2^t}(k_1^t, s_1^t, s_2^t) = p_{K^t | \hat{S}_1^t \hat{S}_2^t} \left( k_1^t | f_1(s_{11}), \ldots, f_1(s_{1t}) \right) \prod_{i=1}^{l} \mathbb{W}_{S_i S_2}(s_{1i}, s_{2i}), \tag{30} \]

and $p_{Y^\pi(U^l V^l)}$ is given by the corresponding conditional marginal in (28). If $\mathcal{D}(\hat{K}, \hat{B}_1, \hat{B}_2)$ is empty, set $(\hat{S}_1, \hat{S}_2)$ to a predefined pair in $\mathcal{S}_1 \times \mathcal{S}_2$ that is arbitrarily fixed upfront. Otherwise, choose one among the pairs in $\mathcal{D}(\hat{K}, \hat{B}_1, \hat{B}_2)$ uniformly at random and set $(\hat{S}_1, \hat{S}_2)$ to be that pair. Declare $(\hat{S}_1, \hat{S}_2)$ as the decoded matrix of source symbols.

\textbf{Error event:} Let us characterize the error event $\mathcal{E}$. Suppose

\[ \mathcal{E}_1 := \bigcup_{i=1}^{l} \left\{ (B_{i1}, B_{i2}) \neq (\hat{B}_{i1}, \hat{B}_{i2}) \right\}, \mathcal{E}_2 := \left\{ (\hat{K}, S_1, S_2) \text{ is not typical wrt } \prod_{t=1}^{m} p_{K^t \hat{S}_1^t \hat{S}_2^t} \right\}, \mathcal{E}_3 := \bigcup_{(s_1, s_2) \in \mathcal{S}_1 \times \mathcal{S}_2} \left\{ (S_1, S_2) \neq (\hat{S}_1, \hat{S}_2), \beta(s_j) = \hat{B}_j : j \in [2] \right\}, \text{ then note that } \mathcal{E} \subseteq \mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3. \]

In the following, we derive upper bounds on $P(\mathcal{E}_1), P(\mathcal{E}_2), P(\mathcal{E}_3).$
**Probability of Error Analysis:** We analyze probability of a random code. Towards that end, let us describe its distribution.

**Distribution of the random code:** As we mentioned, we do not randomize over the choice of fixed B-L typical set source code and the constant composition code $C_U$. This leaves us with having to specify the distribution of random coding scheme lends this analysis some complexity. For channel $p$, there exists $\xi > 0$ such that

\[
\lim_{m \to \infty} P \left( \left. \| X \|_{\mathcal{B}} \leq \xi \right| \pi_t \right) = 0.
\]

To begin with, these four elements are mutually independent. With regard to the bin indices, the collections $(\beta_l(s_{1m}^l) : s_{1m}^l \in S_{1m}^l)$ and $(\beta_2(s_{2m}^l) : s_{2m}^l \in S_{2m}^l)$ are mutually independent. Moreover, for $j \in [2]$, the bin indices $\beta_j(s_{jm}^l) : s_{jm}^l \in S_{jm}^l$ are uniformly and independently chosen from $[M_{V_j}]$. The $m$ surjective maps $\Pi_t : t \in [m]$ are mutually independent and uniformly distributed over the entire collection of surjective maps over $[l]$. Each codeword in the collection $(V_{ij}^m(b_{ij}) : b_{ij} \in [M_{V_j}], i \in [l], j \in [2])$ is mutually independent of the others and $V_{ij}^m(b_{ij}) \sim \prod_{t=1}^m p_{V_j}(|v(t, i), v_j(t, i)|$. This defines the distribution of our random code. We employ an analogous notation for our random code. For example, given $b_j = (b_{ij} : i \in [l])$, we let $V_j(b_j) \in V_j$ be defined through $V_j(b_j)_{1 : m, i} = V_{ji}^m(b_{ij}) : i \in [l]$, and similarly $X_j(u, v_j) = X_j(u, v_j)$.

**Upper bound on $P(\xi_2, P(\xi_3))$:** $\xi_2, \xi_3$ constitute the error events with regard to the Slepian-Wolf code. If we can prove that the rows of $(\hat{K}, S_1, S_2)$ are IID with PMF $p_{K, S_1, S_2}$, then using standard techniques one can verify that there exists $\xi > 0$ such that

\[
\max \{ P(\xi_2), P(\xi_3) \} \leq \exp\{-m\xi\} \text{ if } \frac{\log M_{V_1}^l_{V_2}}{m} > H(S_1^{l} | \hat{K}_l, \hat{S}_2^{l}) : j \in [2], \quad \frac{\log M_{V_1}^l_{V_2}}{m} > H(S_1^{l} | \hat{S}_2^{l} | \hat{K}_l). \tag{32}
\]

We prove this in Appendix F. (89) establishes truth of (32).

**Upper bound on $P(\xi_1)$:** We now derive an upper bound on $\sum_{i=1}^l P((B_{1i}, B_{2i}) \neq (\hat{B}_{1i}, \hat{B}_{2i}))$. Towards that end, let us focus on one of the terms in the latter sum. Note that

\[
P((B_{1i}, B_{2i}) \neq (\hat{B}_{1i}, \hat{B}_{2i})) \leq P((V_{1i}^m(B_{1i}), V_{2i}^m(B_{2i}), Y^{\Pi}(1:m,i)) \approx \prod_{t=1}^m p_{r_1, r_2, y}) + P \left( \bigcup_{b_{1i}, b_{2i} \in M_{V_1} \times M_{V_2}} \left( (B_{1i}, B_{2i}) \neq (\hat{B}_{1i}, \hat{B}_{2i}), (V_{1i}^m(b_{1i}), V_{2i}^m(b_{2i}), Y^{\Pi}(1:m,i)) \approx \prod_{t=1}^m p_{r_1, r_2, y} \right) \right) \tag{33}
\]

With regard to the first term in (33), it suffices to prove

\[
(V_{1i}^m(B_{1i}), V_{2i}^m(B_{2i}), Y^{\Pi}(1:m,i)) \text{ is distributed with PMF } \prod_{t=1}^m p_{r_1, r_2, y}. \tag{34}
\]

This is established in Appendix G. We therefore conclude existence of a $\xi > 0$ such that the first term in the RHS of (33) is

\[
P((V_{1i}^m(B_{1i}), V_{2i}^m(B_{2i}), Y^{\Pi}(1:m,i)) \approx \prod_{t=1}^m p_{r_1, r_2, y}) \leq \exp\{-m\xi\} \tag{35}
\]

and hence can be made arbitrarily small by choosing $m$ sufficiently large.

We are left to analyze the second term in (33). Having proved that codes $C_{V_1,i}, C_{V_2,i}$ experience a MAC channel $p_{r_1, r_2, y}$, the latter analysis will rely on the codewords of $C_{V_1,i}, C_{V_2,i}$ being mutually independent. Though conceptually straightforward, the involved nature of the coding scheme lends this analysis some complexity. For the sake of completeness, we provide the same in Appendix H, and present the new elements in the sequel.

---

10Refer to [18, Problem 10.9].
We summarize our proof thus far. In conjunction with Appendix H, we have proved that if

\[ H(S_j^l|\hat{K}^l, S_2^l) < \frac{\log M_1^j}{m} < \mathbb{I}(\mathcal{Y}_j; \mathcal{Y}_2^j) : j \in [2], \text{ and } H(S_1^l, S_2^l|\hat{K}^l) < \frac{\log M_1^j M_2^j}{m} < \mathbb{I}(\mathcal{Y}_1^l \mathcal{Y}_2^l; \mathcal{Y}) \]  

(36)

where \( S_1^l, S_2^l, \hat{K}^l \) and \( \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y} \) are distributed as in (30) and (28), respectively, then the proposed coding scheme can enable the decoder recover \( S_1, S_2 \) with arbitrarily high reliability by choosing \( m \) sufficiently large. Our last step involves characterizing the upper and lower bounds in (36) in terms of the PMF \( p_{UVXY} \) provided in the Theorem statement. We begin with the channel coding bounds.

**Lower bounds on \( I(\mathcal{Y}_j; \mathcal{Y}_2^j) \) and \( I(\mathcal{Y}_1^l \mathcal{Y}_2^l; \mathcal{Y}) \):** Suppose \((U_1^l, \hat{V}_1^l, X_1^l, Y_1^l) = (U_1^l, U_2^l, V_2^l, X_1^l, X_2^l, Y_1^l)\) is distributed with PMF (28), and \( I \in \{1, \cdots, l\} \) is a random index independent of the collection \((U_1^l, V_1^l, X_1^l, Y_1^l)\), then \( U_{1I}, U_{2I}, V_{1I}, V_{2I}, X_{1I}, X_{2I}, Y_I \) is distributed with PMF (29). Hence we study \( I(V_{2I}; Y_I, V_{2I}) = I(\mathcal{Y}_j; \mathcal{Y}_2^j) \) and \( I(V_{1I}, V_{2I}; Y_I) = I(\mathcal{Y}_1^l \mathcal{Y}_2^l; \mathcal{Y}) \). We begin by recognizing that

\[ \{A_{1l} \neq A_{2l}\} \subseteq \{K_1(t, 1 : l) \neq T_0^l(K_1)\} \cup \{K_1(t, 1 : l) = K_2(t, 1 : l)\}, \text{ and hence} \]

\[ P(A_{1l} \neq A_{2l}) \leq \epsilon = \xi(0(K)) + \tau_{\delta_1}(K_1). \]  

(37)

where the last inequality follows from the theorem hypothesis. We therefore have \( \frac{1}{2} \geq \epsilon \geq P(A_1 \neq A_2) \geq P(U_1^l \neq U_2^l) \), and hence

\[ I(V_{1I}; V_{2I}) = H(V_{1I}) - H(V_{1I}|V_{2I}) \geq H(V_{1I}) - H(V_{1I}|V_{2I}, Y_I) \]

(38)

\[ \geq H(V_{1I}) - (H(V_{1I}|V_{2I}) - \mathbb{I}|U_1^l = U_2^l) - h_b(\epsilon) \]

\[ = H(V_{1I}) - P(U_1^l = U_2^l) \mathbb{I}|U_1^l = U_2^l = 1) - H(V_{1I}|V_{2I}, Y_I) = h_b(\epsilon) \]

(39)

\[ \geq H(V_{1I}) - H(V_{1I}, Y_I) - \epsilon \log |V| - h_b(\epsilon) = I(V_{1I}; V_{1I}) - \epsilon \log |V| - h_b(\epsilon) \]

(40)

where (38) follows from \( p_{V_1} = p_{V_2} = p_{V} \) (Lemma 2) and \( \frac{1}{2} \geq \epsilon \geq P(U_1^l \neq U_2^l) \), (40) follows from Lemma 3 in Appendix D and \( \frac{1}{2} \geq \epsilon \geq P(U_1^l \neq U_2^l) \). Indeed, note that Lemma 3 states

\[ P(U_{1I} = u, V_{2I} = u, V_{1I} = v_1, V_{2I} = v_2, X_{1I} = x_1, X_{2I} = x_2, Y_I = y| U_1^l = U_2^l = 1) = p_{UVXY}(u, v, x, y) \]  

(41)

for every \( u, v, x, y \in \mathcal{U} \times \mathcal{V} \times \mathcal{X} \times \mathcal{Y} \) and hence, any functional of the PMF on the LHS of (41) is equal to any functional of the PMF on the RHS of (41), in particular the entropy functional. Following an analogous sequence of steps, we have

\[ I(V_{1I}, V_{2I}; Y_I) = H(V_{1I}, V_{2I}) - H(V_{1I}, V_{2I}|Y_I) \geq H(V_{1I}, V_{2I}) - H(V_{1I}, V_{2I}, Y_I) = \mathbb{I}|U_1^l = U_2^l = 1) - h_b(\epsilon) \]

(42)

\[ = H(V_{1I}, V_{2I}) - P(U_1^l = U_2^l) \mathbb{I}|U_1^l = U_2^l = 1) - H(V_{1I}, V_{2I}, Y_I) = h_b(\epsilon) \]

\[ \geq H(V_{1I}, V_{2I}) - H(V_{1I}, V_{2I}, Y_I) - \epsilon \log |V| - h_b(\epsilon) = I(V_{1I}, V_{1I}) - \epsilon \log |V| - h_b(\epsilon) \]

(43)

\[ \geq H(V_{1I}, V_{2I}) - [H(V_{1I}, V_{2I}, Y_I) - H(Y_I)] - \epsilon \log |V| - h_b(\epsilon) = I(V_{1I}, Y_I, V_{2I}) - \epsilon \log |V| - h_b(\epsilon) \]

(44)

**Upper bounds on \( H(S_j^l|\hat{K}^l, S_2^l) \), \( H(S_1^l, S_2^l|\hat{K}^l) \):** Recall from (89) that \( p_{S_1, S_2, \hat{K}} \) is the PMF of any row of the triplet \( S_1, S_2, \hat{K} \) of matrices. Appealing to the sequence of steps from (6) through (8) we recognize that it suffices to characterize an upper bound \( \phi \) on \( P(K(t, 1 : l) \neq K_1(t, 1 : l)) \), that is at most \( \frac{1}{2} \). Towards that end, recall that our typical set source code ensures \( d_k(e_k(k^l)) = k_1^l \) for every \( k_1 \in T_0^l(K_1) \). This guarantees \( \{K(t, 1 : l) \neq K_1(t, 1 : l)\} \subseteq \{A_{1l} \neq A_l\} \). In order to derive an upper bound on the latter event, we are required to characterize the channel \( p_{Y|U}|U \) experienced by codewords of \( C_U \). In particular, since

\[ P(A_{1l} \neq A_l) \leq P(A_{1l} \neq A_{2l}) + P(\hat{A}_l \neq A_{1l}, A_{1l} = A_{2l}) \leq \epsilon + P(\hat{A}_l \neq A_{1l}, A_{1l} = A_{2l}), \]

(45)

it suffices to prove that the second term in (45) is at most \( g(\alpha + \rho, l) \). Indeed, our hypothesis guarantees \( \epsilon + g(\alpha + \rho, l) = \phi \leq \frac{1}{2} \). We are therefore required to prove that if the two transmitters choose a common \( C_U \)-codeword, then the latter experiences a memoryless \( \prod_{i=1}^l p_{Y|U} \) channel. We can then appeal to our choice of the constant composition code, Thm. 3 and conclude that RHS of (45) is at most \( \phi \leq \frac{1}{2} \). We provide this in Appendix I.
B. Two elements in the Proof of Thm. 6

We provide our arguments for the case \( a = 1, \beta = 0 \) which implies \( \alpha > (1 + \delta)H(K_1) \). We refer the reader to the coding scheme presented in Sec. VI-A. The two layers of codes, the multiplexing unit and the encoding rule is identical to that presented there, and is hence not repeated. This gets us to the decoding rule.

**Decoding Rule:** Let \( Y_j \in Y_{j}^{m \times t} \) denote the matrix of symbols received by decoder \( j \) with \( Y_j(t, i) \) being the symbol received during symbol interval \((t-1)t + i\). The channel-code decoder attempts to recover \( (\hat{A}_1, \hat{B}_j) \in [M_u]_i \times [M_V^t]' \).

- **Base-layer channel decoder:** \( d_u - \) the \( C_U \) - decoder - decodes rows of \( Y_j \) separately and identically into \( \hat{A}_{jt} : = d_{u,j}(Y_j(t, 1 : l)) : t \in [m] \) and reconstructs \( \hat{U}_j : = u\{\hat{A}_j\} \).
- **Satellite-layer channel decoder:** For each \( i \in [l] \), the \( C_{V_j,i} \) - decoder looks for all messages \( \hat{b}_{ji} \in [M_{V_j}] \) such that the corresponding codeword is jointly typical with \( Y_j\{1 : m, i\}, u\{\hat{A}_j\} \{1 : m, i\} \). Specifically for \( i \in [l] \), populate

\[
\mathcal{D}_i(Y_j, \hat{A}_j) : = \left\{ \hat{b}_{ji} \in [M_{V_j}] : \left( v_{ji}^m(\hat{b}_{ji}), Y_j^p(1 : m, i), u\{\hat{A}_j\} \{1 : m, i\} \right) \text{ is jointly typical wrt } \prod_{l=1}^{m} p_{Y_j | Y_j^p | u\{\hat{A}_j\} \{1 : m, i\}} \right\}, \tag{46}
\]

where

\[
p_{Y_j, U_j, X_j | Y_j^p, u\{\hat{A}_j\}}(y_j, u_j, x_j) = \sum_{(a_1, a_2) \in [M_u]_i \times [M_u]} P(\frac{A_1 = a_1}{A_2 = a_2}) \mathbf{1}_{\{u_j^{(a_1)}(a_2) = u_j^{(a_1)}(a_2)\}} \times \prod_{l=1}^{2} \prod_{i=1}^{l} p_{v_{ji}(v_{ji})} p_{x_{ji}(x_{ji} | u_{ji}, v_{ji})} \mathbf{1}_{\{u_{ji} = d_{u,j}(y_{ji})\}}. \tag{47}
\]

If \( \mathcal{D}_i(Y_j, \hat{A}_j) \) is a singleton for each \( i \in [l] \), denote \( \hat{B}_{ji} \) to be the element in \( \mathcal{D}_i(Y_j, \hat{A}_j) \). In this case, we have a unique choice for \( \hat{B}_{ji} \). In the alternate case, i.e., when at least one of \( \mathcal{D}_i(Y_j, \hat{A}_j) : i \in [l] \) is not a singleton, we declare an error.

- **Base layer source-code decoder:** Let \( \hat{K}_j(t, 1 : l) = d_K(\hat{A}_{jt}) \) be the reconstructions output by the fixed B-L typical set decoder.
- **Satellite layer source-code decoder:** The decoder of the Slepian-Wolf code looks for

\[
\mathcal{D}(\hat{K}_j, \hat{B}_j) : = \left\{ \hat{s}_j \in \mathcal{S}_j : \beta_j(\hat{s}_j) = \hat{B}_j \text{ and } (s_j, \hat{K}_j) \text{ is jointly typical wrt } \prod_{l=1}^{m} p_{s_j | \hat{K}_j} \right\}. \tag{49}
\]

If \( \mathcal{D}(\hat{K}_j, \hat{B}_j) \) contains a unique matrix \( \hat{S}_j \), declare the same as the decoded matrix of source symbols. Otherwise, an error is declared.

The first element we present here will establish that the rows of

\[
U_j : = u\{\hat{A}_j\}, V_j : = V_j\{\hat{B}_j\}, X_j : = X_j\{\hat{A}_j, \hat{B}_j\}, Y_j, \hat{U}_j : = u\{\hat{A}_j\} : j \in [2]
\]

are IID with PMF \( p_{U_j, V_j, X_j, Y_j} \) defined in (48). Establishing this will take us through steps analogous to those that took us from (91) to (100). We provide these steps in Appendix J.

The second element we present derives lower bounds on the capacity of the IC experienced by the satellite-layer codes. It can be proved that one can communicate at rate \( I(V_j; Y_j, Z_j) \) via the satellite-layer codes, where \( V_j, Y_j, Z_j \) is the corresponding marginal of the PMF defined in (47). We now derive lower bounds on \( I(V_j; Z_j) \) in terms of the PMF \( p_{U_j, V_j, X_j} \). **Lower Bounds on** \( I(V_j; Y_j, Z_j) \): Suppose \( (U_j^l, V_j^l, X_j^l, Y_j^l) = (U_j^l, U_j^l, V_j^l, V_j^l, X_j^l, X_j^l, Y_j^l, Y_j^l, \hat{U}_j^l, \hat{U}_j^l) \) is distributed with PMF (47), and \( j \in \{1, \cdots, l\} \) is a random index independent of the collection \( U_j^l, V_j^l, X_j^l, Y_j^l \), then
\( U_{13}, U_{23}, V_{13}, V_{23}, X_{13}, X_{23}, Y_{13}, Y_{23}, \bar{U}_{12}, \bar{U}_{23} \) is distributed with PMF (48). Hence we study \( I(V_{j3}; Y_3, \bar{U}_{j3}) = I(\gamma_j; v_j, \bar{\gamma}_j) \). Suppose \( \frac{1}{2} \geq \phi \geq P(\{U_1^l \neq U_2^l\} \cup \{U_1^l \neq \bar{U}_1^l\}) \), then
\[
I(V_{j3}; Y_3, \bar{U}_{j3}) = H(V_{j3}) - H(V_{j3}| Y_3, \bar{U}_{j3}, I_{\{U_1^l=U_2^l=\bar{U}_1^l\}}) - h_b(\phi)
\]
\[
\geq H(V_3) - H(V_{j3}| Y_3, \bar{U}_{j3}, I_{\{U_1^l=U_2^l=\bar{U}_1^l\}}) - h_b(\phi)
\]
\[
= H(V_3) - P(U_1^l = U_2^l = \bar{U}_1^l) \left[ H(V_{j3}| Y_3, \bar{U}_{j3}, I_{\{U_1^l=U_2^l=\bar{U}_1^l\}}) = 1 \right] - h_b(\phi)
\]
\[
= H(V_3) - P(U_1^l = U_2^l = \bar{U}_1^l) \left( H(V_{j3}| Y_3, \bar{U}_{j3}, I_{\{U_1^l=U_2^l=\bar{U}_1^l\}} = 1) \right) - h_b(\phi)
\]
where (50) follows from \( p_{V_{j3}} = p_{V_{j1}} = p_{V_{j2}} \) (Lemma 2) and \( \frac{1}{2} \geq \phi \geq P(\{U_1^l \neq U_2^l\} \cup \{U_1^l \neq \bar{U}_1^l\}) \), (52) follows from Lemma 3 in Appendix D and from \( \frac{1}{2} \geq \epsilon \geq P(U_1^l \neq U_2^l) \).

ACKNOWLEDGEMENT

The author is thankful to (i) Prof. Sandeep Pradhan, Farhad Shirani for sharing their insights on [15], (ii) Deepanshu Vasal for several technical discussions and (iii) Prof. P R Kumar for his support and encouragement. The author is particularly very grateful to Prof. Wojciech Szpankowski for the inspiration, his support and encouragement.

APPENDIX A

CES STRATEGY IS INCAPABLE OF COMMUNICATING \( S \) OVER \( W_{Y|X} \) IN EX. 1

The following proof is based on Dueck’s arguments [10].

Lemma 1: Consider Ex. 1 with any \( \eta \in \mathbb{N} \). There exist \( a_\eta \in \mathbb{N}, k_\eta \in \mathbb{N} \), such that for any \( a \geq a_\eta \) and any \( k \geq k_\eta \), the sources and the MAC described in Ex. 1 do not satisfy CES conditions that are stated in [4, Thm. 1].

Proof: Given any set \( Q \) and any PMF \( W_{SPQPX_1|S_1QPX_2|S_2QW_{Y|X}} \), where \( X_j = U_j, V_j : j \in [2] \), we will prove
\[
I(X_j; Y_j| Q) = I(U_j; Y_j| Q) < H(S),
\]
thereby contradicting (2). Towards that end, we derive a lower bound on \( H(S) \). By simple substitution, verify that
\[
H(S) \geq H(S_2) = h_b(\frac{1}{k}) + 1 \log(a^k - 1) \geq 1 \log\left(\frac{ka^k}{2}\right)
\]
\[
\geq \log a + \frac{1}{k} \log\left(\frac{ka^k}{2}\right) \geq \log a,
\]
whenever \( a^k \geq 2, k \geq 2 \). We now consider the LHS of (54).

Let \( R = 1_{\{S_1 = S_2 \}} \).
\[
I(X_j; Y_j| Q) = I(U_j; Y_j| Q) \leq I(U_j, R; Y_j| Q) \leq \log 2 + I(V_j; Y_j| Q, R) \leq \log 2 + \frac{\log|\gamma_0| \times \gamma_2| \times 2^{r(2)}}{k} + (1 - \frac{1}{k})I(U_j; Y_j| Q, R = 1)
\]
\[
= \log 2 + \frac{\log|\gamma_0| \times \gamma_2| \times 2^{r(2)}}{k} + (1 - \frac{1}{k})I(U_j; Y_j| Q, R = 1)
\]
We focus on the third term in the above sum. Conditioned on \( R = 1 \), the sources are equal to \( (0^k, 0^k) \). It can be verified that \( V_1 U_1 - S_1 Q - S_2 Q - V_2 U_2 \). Given \( Q = q, R = 1, X_1 = (V_1, U_1) \) is independent of \( X_2 = (V_2, U_2) \) and hence
\[
I(U_j; Y_j| Q, R = 1) = \max_{p_{V_1 U_1}p_{V_2 U_2}} I(U_j; Y_j) \leq 2C_M + h_b\left(\frac{2}{k}\right) + \frac{1}{k} \log a + h_b\left(\frac{2}{ka^k}\right) + \max_{p_{V_1 U_1}} H(Y_0)
\]
We now evaluate an upper bound on the maximum value of \( H(Y_0) \) subject to \( U_1, U_2 \) being independent. We evaluate the following three possible cases.
Case 1a: For some \( u \in \mathcal{U} \), \( P(U_1 = u) \geq \frac{1}{2} \) and \( P(U_2 = u) \geq \frac{1}{2} \). Then \( P(Y_0 = u) \geq \frac{1}{4} \) (independence of \( U_1, U_2 \)) and hence \( H(Y_0) \leq \log 2 + \frac{3}{4} \log a \).

Case 1b: For some \( u \in \mathcal{U} \), \( P(U_1 = u) \geq \frac{1}{2} \) and \( P(U_2 = u) \leq \frac{1}{2} \). Then \( P(U_2 \neq u) \geq \frac{1}{2} \) and hence \( P(Y_0 = 0) \geq \frac{1}{4} \) and hence \( H(Y_0) \leq \log 2 + \frac{3}{4} \log a \).

Case 2a: For every \( u \in \mathcal{U} \), \( P(U_1 = u) \leq \frac{1}{2} \). Then for any \( u \in \mathcal{U} \), \( P(U_2 \neq u) = \sum_u \sum_{z \neq u} P(U_2 = u)P(U_1 = z) \geq \frac{1}{2} \sum_u P(U_2 = u) = \frac{1}{2} \), implying \( P(Y_0 = 0) \geq \frac{1}{2} \) and hence \( H(Y_0) \leq \log 2 + \frac{3}{4} \log a \).

In all cases, we have \( H(Y_0) \leq \log 2 + \frac{3}{4} \log a \). Substituting this in (57) and then back in (56), we conclude

\[
I(X; Y|Q) \leq 2 \log 2 + 2C_M + h_b\left(\frac{2}{k}\right) + \frac{1}{k} \log a + h_b\left(\frac{2}{ka^{\frac{2}{k}}}\right) + \frac{3}{4} \log a + \frac{\log |Y|}{k} < \log a
\]

for sufficiently large \( k, a \). In (58), we have used the fact that for sufficiently large \( a, k \) the satellite channels are chosen such that \( |Y|^2 \leq a^{\frac{k}{2}} \).

\[\Box\]

Appendix B

Satellite Channels in Ex. 1 Support Slepian-Wolf binning rates

Our goal here is to identify values for \( l, \delta \) such that \( \phi = \xi[l] + \tau_{l, \delta} \leq \frac{1}{2} \), and when substituted in (10) yields rates that are supported on the satellite channels. Specifically, our choice for \( l, \delta \) must satisfy

\[
\mathcal{L}_l(\phi, |S_j|) + \beta 1_{|j|=1} + H(S_j|S_1)1_{|j|=2} \leq C_M + h_b\left(\frac{2}{ka^{\frac{2}{k}}}|S_j|\right) + \frac{1}{k} \log a \leq 0
\]

(59)

\[
\mathcal{L}_l(\phi, |S|) + \beta + H(S_2|S_1) \leq 2C_M + h_b\left(\frac{2}{ka^{\frac{2}{k}}}|S|\right) + \frac{1}{k} \log a
\]

(60)

where the RHSs in (59), (60) are the capacities of \( \mathbb{W}_{Y_j|X_j} \) and the MAC \( \mathbb{W}_{Y_i|X_j} \) comprised of the two satellite channels \( \mathbb{W}_{Y_1|X_1}, \mathbb{W}_{Y_2|X_2} \), respectively. We note that

\[
\tau_{l, \delta} = 2a^k \exp\left(-\frac{\delta^2 l}{2ka^{\frac{2}{k}}}\right)\] and \( \xi[l] \leq \frac{k^3}{a^{\frac{k}{2}}} \).

Choose \( l = k^4 a^{\frac{2k}{3}}, \delta = \frac{1}{k} \), substitute in (61) and verify

\[
\tau_{l, \delta} \leq 2a^k \exp\left(-\frac{1}{2}\left(\frac{3}{2} - 2\right)k\right), \xi[l] \leq \frac{k^3}{a^{\frac{k}{2}}} \]. Since \( \eta \geq 6 \), we have \( \phi = \xi[l] + \tau_{l, \delta} \leq 2k^3 a^{\frac{k}{2}} \frac{2k}{3} < \frac{1}{2} \) for sufficiently large \( a, k \). Verifying

\[
\mathcal{L}_l^s(2k^3 a^{\frac{k}{2}}|S_j|) \leq \alpha(h_b(\alpha + \log a)), \text{ with } \alpha = \frac{8k^4}{a^{\frac{k}{2}}}
\]

(63)

for sufficiently large \( a, k \). Substituting \( \delta = \frac{1}{k} \), verify\(^{11}\)

\[
\beta \leq (2/l) + (1/k) \log a + (1+(1/k))h_b(1/k).
\]

(64)

Since \( h_b(\frac{2}{k}) - (1 + \frac{1}{k})h_b(\frac{1}{k}) \geq 0 \) for \( k \) satisfying \((k-2)^2 \geq 4k\sqrt{k} \), we have \( \beta \leq h_b(\frac{2}{k}) + \frac{5}{4k} \log a \) for \( a, k \) satisfying \((k-2)^2 \geq 4k\sqrt{k} \) and \( k^3 a^{\frac{2k}{3}} \log a \geq 8 \). Lastly, note that \( H(S_2|S_1) \leq h_b(\frac{1}{k(k-1) a^{\frac{2k}{3}}}) + \frac{2a^k}{k} \log a \leq h_b(\frac{2}{ka^{\frac{2}{k}}}) + \frac{1}{k} \log a \) for \( a, k \) that satisfy \( k > 2 \) and \( a^{\frac{2k}{3}} \geq 8k \). The validity of (59), (60) for \( a, k \) that satisfy (i) \((k-2)^2 \geq 4k\sqrt{k} \), (ii) \( k^3 a^{\frac{2k}{3}} \log a \geq 8 \) and (iii) \( a^{\frac{2k}{3}} \geq 8k \) can now be verified by substituting (63) and the above derived bounds.

Appendix C

Fixed B-L Coding can Strictly Outperform CES

In here, we consider Ex. 1 and prove existence \( a^* \in \mathbb{N} \) and \( k^* \in \mathbb{N} \) such that for any \( a \geq a^* \) and any \( k \geq k^* \), \( S \) and MAC \( \mathbb{W}_{Y_i|UX} \) satisfy conditions stated in Thm. 4. In conjunction with Lemma 1, we would thus establish truth of Remark 5.

Consider the following assignment for the auxiliary parameters in Thm. 4. Let \( \mathcal{K} = S_1 = S_2, \mathcal{U} \) be the input alphabet of the shared channel \( \mathbb{W}_{Y_1|UX} \), \( Y_j = X_j : j \in [2] \) be the input alphabet of the satellite channels \( \mathbb{W}_{Y_j|X_j} : j \in [2] \), respectively. Let \( f_j(s) = s \) for \( s \in S_j \) be the identity map, and hence \( K_j = S_j \) for \( j \in [2] \). Let \( \alpha = (1 - \frac{1}{3k}) \log a, \beta = \frac{5}{4k} \log a + (1 + \frac{1}{k}) h_b(\frac{1}{k}), \rho = \frac{1}{4k} \log a, \delta = \frac{1}{k} \). Let \( l = k^3 a^{\frac{2k}{3}} \). Let \( p_U \) be the uniform

\(^{11}\) Use \( H(S_1) \leq \log a + h_b(\frac{1}{k}) \) and \( |T_i| = \frac{1}{k} \log(1+\delta)H(S_1) \).

https://mc.manuscriptcentral.com/t-it
PMF on $\mathcal{U} = \{0, \ldots, a - 1\}$. Let $p_{V_j} : j \in [2]$ be the capacity achieving distribution on satellite channels $\mathcal{W}_{Y|X_j} : j \in [2]$, respectively. Note that, for any $u \in \mathcal{U}$, $l_{P_U}(u) = k^4a^{\frac{a}{2} - 1}$ is a natural number since $\eta \geq 6$ is an even integer. For the above assignment, note that (4) is

$$l^*(\frac{1}{4k} \log \frac{a}{4} \mathcal{U}, \mathcal{Y}) = \min \left\{ l : \frac{l}{4k} \log \frac{a}{4} \geq \log 4 + 4a(1 + a|\mathcal{Y}_1||\mathcal{Y}_2|) \log (l + 1) \right\} \leq \min \left\{ l : \frac{l}{4k} \log \frac{a}{4} \geq \log 4 + 4a(1 + a|\mathcal{Y}_1||\mathcal{Y}_2|) \log 2l \right\}. \quad(65)$$

Recall that for sufficiently large $a, k$, the satellite channels defined in Ex. 1 have $|\mathcal{Y}_j| \leq a^{\frac{3}{2}}$. It can be verified that the RHS of (65) is lesser than or equal to $k^4a^{\frac{a}{2}}$ for sufficiently large $a, k$. Therefore, the assignment $l = k^4a^{\frac{a}{2}} \geq l^*(\frac{1}{4k} \log \frac{a}{4} \mathcal{U}, \mathcal{Y})$ for sufficiently large $a, k$.

Verify that $(1 + \delta)H(K_1) = (1 + \delta)H(S_1) < (1 + \frac{1}{k}) \log a + (1 + \frac{1}{k}) h_b(\frac{1}{k}) = \alpha + \beta$. Since $p_U$ is uniform and $p_{V_0|U}$ induced by the chosen PMF is deterministic, it can be verified that $E_r(\alpha + \rho, p_U, p_{V_0|U}) = \log a - (\alpha + \rho) = \frac{1}{k^4} \log 4$. Hence

$$g(\alpha + \rho, l) = (l + 1)^{2l} \log 4 \exp \left\{ -\frac{l}{4k} \log 4 \right\} \leq 4^{-\frac{1}{4l}} (l + 1)^{2a^\frac{a}{2} + 2} \leq 4^{-\frac{1}{4l}} (2l)^{2a^\frac{a}{2}} \leq \frac{k^3}{a^\frac{a}{2}} \quad(66)$$

for sufficiently large $a, k$. Since our choice of $\delta = \frac{1}{k}, l = k^4a^{\frac{a}{2}}$ are identical to that in Sec. III-A, we appeal to (61), (62) and conclude

$$\tau_{l,\delta}(\mathcal{K}) + \xi[l](\mathcal{K}) \leq \frac{2k^3}{a^\frac{a}{2}}, \quad \text{and in conjunction with (66) we have}, \quad \phi \leq \frac{3k^3}{a^\frac{a}{2}} \leq \frac{1}{2}. \quad(67)$$

for sufficiently large $a, k$. Substituting this upper bound in $\mathcal{L}_l(\cdot, \cdot)$ and $\mathcal{L}(\cdot, \cdot)$, it can be verified that

$$\mathcal{L}_l(\phi, |\mathcal{S}_j|) \leq \mathcal{L}_l(\phi, |\mathcal{S}|) \leq \frac{1}{l} h_b(\alpha) + \frac{\alpha}{2} \log a \quad(68)$$

where we have used that fact that for large $a, k$, we have $|\mathcal{Y}_j| \leq a^{\frac{3}{2}}$. We are now set to prove the remaining inequalities (11), (12). This follows by simple substitution of $\beta = \frac{5}{4k} \log a + (1 + \frac{1}{k}) h_b(\frac{1}{k})$, upper bound of $h_b(\frac{\frac{2}{k^4}}{\frac{a}{4l}}) + \frac{2 \log a}{a^\frac{a}{2}}$ on $H(S_2|S_1)$, capacities of $\mathcal{W}_{Y|X_j}$ for $I(V_j; Y|V_2)$, the sum of these capacities for $I(V_j; Y)$, (68) and is left to the reader.

**APPENDIX D**

**PROPERTIES OF PMFS (28), (29) EMPLOYED IN DECODING RULE**

Let us recall

$$p_{U_1U_2V_1V_2}(u_1, u_2, v_1, v_2, y') = \sum_{\{a_1, a_2\} \in \mathcal{M}_a \times \mathcal{M}_a} P(\mathcal{A}_1 = a_1) \mathbb{I}_{\{u_1' = a_1\}} \prod_{i=1}^l \left[ \prod_{j=1}^l p_{V_j}(v_{ji}) p_{X_j|U_j}(x_{ji}|u_{ji}, v_{ji}) \right] \prod_{i=1}^l \mathcal{W}_{Y|X_i} x_i(y_i|x_{i1}, x_{i2}) \quad(69)$$

be a PMF$^{12}$ on $\mathcal{U}^l \times \mathcal{Y}^l \times \mathcal{A}^l \times \mathcal{Y}^l$, and

$$p_{U_1U_2V_1V_2}(u_1, u_2, v_1, v_2, x_1, x_2, y_{1,2}, a, b, c, d) := \frac{1}{l} \sum_{i=1}^l p_{U_1, U_2, V_1, V_2, X_1, X_2, Y_{1,2}}(a, b, c, d) \quad(70)$$

**Lemma 2:** Let $U_1^l, U_2^l, V_1^l, V_2^l, X_1^l, X_2^l, Y^l$ take values in $\mathcal{U}^l \times \mathcal{Y}^l \times \mathcal{X}^2 \times \mathcal{Y}^l$ with PMF (69) and consider the PMF (70) defined on $\mathcal{U} \times \mathcal{Y} \times \mathcal{X} \times \mathcal{Y}$, Suppose $I$ is a random index taking values in $\{1, \ldots, l\}$ that is uniformly distributed and independent of $U_1^l, U_2^l, V_1^l, V_2^l, X_1^l, X_2^l, Y^l$. The following are true.

$^{12}$In (69), $U_1U_2V_1V_2X_1X_2Y'$ abbreviates $U_1U_2V_1V_2X_1X_2Y'$ and similarly $y, y', x, y$ abbreviates $u_1, u_2, v_1, v_2, x_1, x_2, y'$. https://mc.manuscriptcentral.com/t-it
1) \( p_{V_i[V_j]} = \prod_{i=1}^l p_{V_i} p_{V_j} \).
2) \( U_{11}, U_{21}, V_{11}, V_{21}, X_{11}, X_{21}, Y_1 \) has PMF (70).
3) \( p_{Y_j} = p_{V_j} \) for \( j \in [2] \),
4) The marginals

\[
\begin{align*}
p_{V_i[U_j]}(u_1, u_2) &= \sum_{(a_1, a_2) \in [M_a] \times [M_a]} P (A_1 = a_1, A_2 = a_2) \mathbb{1}_{u'(a_i) = u'_i, j \in [2]}, 
\end{align*}
\]

(71)

\[
\begin{align*}
p_{V_i[X_i|U_j]}(v'_i, u_i | u_j) &= \left[ \prod_{i=1}^l \left\{ \prod_{j=1}^{2} p_{V_j} (v_{ji}) p_{X_j|U_j} (x_{ji} | u_{ji}) \right\} \mathbb{W}_{Y_i|X_i} (y_i | x_{1i}, x_{2i}) \right], \text{ and in particular (72)}
\end{align*}
\]

(72)

\[
\begin{align*}
p_{V_i[X_i|U_j]}(v'_i, u_i | u_j) &= \left[ \prod_{i=1}^l \left\{ \prod_{j=1}^{2} p_{V_j} (v_{ji}) p_{X_j|U_j} (x_{ji} | u_{ji}) \right\} \mathbb{W}_{Y_i|X_i} (y_i | x_{1i}, x_{2i}) \right]
\end{align*}
\]

(73)

\[
\begin{align*}
p_{V_i[X_i|U_j]}(v'_i, u_i | u_j) &= \prod_{i=1}^l p_{V_i} (v_{i} x_{1i}, x_{2i}) \mathbb{W}_{Y_i|X_i} (y_i | x_{1i}, x_{2i}) , \text{ and hence (74)}
\end{align*}
\]

(74)

\[
\begin{align*}
p_{V_i[U_j]}(v'_i, u_i | u_j) &= \prod_{i=1}^l p_{V_i} (v_{i} | u_i)
\end{align*}
\]

(75)

**Proof:** 1) Follows by just computing the marginal \( p_{V_i[V_j]} \) wrt (69). 2) Straightforward to verify. 3) Follows from previous two assertions. 4) Follows by just evaluating the LHSs wrt to (69). 

**Lemma 3:** Given \( l \in \mathbb{N} \), finite alphabet sets \( A, B_1, B_2, C \) and a PMF \( p_{A B_1, B_2 C} = p_{A B_1} p_{B_2} p_{A C} \) on \( A \times B_1 \times B_2 \times C \) such that \( p_A \) is a type of sequences in \( A^l \). Suppose \( (A_1^l, A_2^l, B_1^l, B_2^l, C^l) \) take values in \( A^l \times B_1^l \times B_2^l \times C^l \) with PMF \( p_{A_1^l, A_2^l, B_1^l, B_2^l, C^l} \) given by

\[
\begin{align*}
p_{A_1^l, A_2^l, B_1^l, B_2^l, C^l}(a_1^l, a_2^l, b_1^l, b_2^l, c^l) &= \left[ \sum_{m_1, m_2} P(M_1 = m_1, M_2 = m_2) \mathbb{1}_{u'(m_1) = a_1^l, \ u'(m_2) = a_2^l} \right] \prod_{i=1}^l \left\{ \frac{p_{B_1}(b_{i1}|a_{i1}) p_{B_2}(b_{i2}|a_{i2})}{p_{C}(c_i|b_{i1}, b_{i2})} \right\},
\end{align*}
\]

where (i) \( u^l : [M_a] \rightarrow A^l \) is a map such that \( u^l(m) \in A^l \) is of type \( p_A \) for every \( m \), (ii) \( (M_1, M_2) \in [M_a] \times [M_a] \) are a pair of (message) random variables with PMF \( P(M_1 = \cdot, M_2 = \cdot) \). Suppose \( I \) is a random index taking values in \( \{1, \cdots , l\} \) that is uniformly distributed and independent of \( A_1^l, A_2^l, B_1^l, B_2^l, C^l \), then

\[
\begin{align*}
P (A_{1I} = x, A_{2I} = x, B_{1I} = y_1, B_{2I} = y_2, C_I = z | A_1^l = A_2^l) = p_{A B_1, B_2 C}(x, y_1, y_2, z).
\end{align*}
\]

(76)

**Proof:** Let \( J = \mathbb{1}_{A_1^l = A_2^l} \). It can be verified by summing over \( b_1^l, b_2^l, c^l \) that

\[
\begin{align*}
p_{A_1^l A_2^l}(a_1^l, a_2^l) &= \left[ \sum_{m_1, m_2} P(M_1 = m_1, M_2 = m_2) \mathbb{1}_{u'(m_1) = a_1^l, \ u'(m_2) = a_2^l} \right],
\end{align*}
\]

and hence \( p_{B_1^l B_2^l C^l|A_1^l A_2^l}(b_1^l, b_2^l, c^l|a_1^l, a_2^l) = \prod_{i=1}^l p_{B_1 B_2 C}(b_{i1}, b_{i2}, c_i | a_i) \). Since \( B_1^l B_2^l C^l = A_1^l A_2^l - J \) forms a Markov chain, we have

\[
\begin{align*}
p_{A_1^l A_2^l B_1^l B_2^l C^l|J}(a_1^l, a_2^l, b_1^l, b_2^l, c^l|1) &= p_{A_1^l A_2^l|J}(a_1^l, a_2^l|1) \prod_{i=1}^l p_{B_1 B_2 C}(b_{i1}, b_{i2}, c_i | a_i).
\end{align*}
\]
The PMF of a randomly chosen coordinate is given by

\[
P_{A_1,A_2,B_1,B_2,C_1}(x, x_1, y_1, y_2, z | 1) = \frac{1}{l} \sum_{i=1}^{l} p_{A_i,A_2,B_1,B_2,C_1}(x, x_1, y_1, y_2, z | 1)
\]

\[
= \frac{1}{l} \sum_{i=1}^{l} \sum_{s_i \in \mathcal{A}_1} \sum_{t_i \in \mathcal{B}_1} \sum_{v_i \in \mathcal{C}_1} p_{A_i,A_2,B_1,B_2,C_1}(s_i^{-1}x s_i^{-1}, t_i^{-1}y_1 t_i^{-1}, u_i^{-1}y_2 u_i^{-1}, v_i^{-1}z v_i^{-1} | 1)
\]

\[
= \frac{1}{l} \sum_{i=1}^{l} \sum_{s_i \in \mathcal{A}_1} \sum_{t_i \in \mathcal{B}_1} \sum_{v_i \in \mathcal{C}_1} p_{A_i,A_2,B_1,B_2,C_1}(s_i^{-1}x s_i^{-1}, t_i^{-1}y_1 t_i^{-1}, u_i^{-1}y_2 u_i^{-1}, v_i^{-1}z v_i^{-1} | 1)p_{B_i,B_2}(y_1, y_2, z | x) \prod_{j=1}^{l} p_{B_j,B_2}(t_j, u_j, v_j | s_j)
\]

\[
= \frac{1}{l} \sum_{i=1}^{l} \sum_{s_i \in \mathcal{A}_1} p_{A_i,A_2,B_1,B_2,C_1}(s_i^{-1}x s_i^{-1}, t_i^{-1}y_1 t_i^{-1}, u_i^{-1}y_2 u_i^{-1}, v_i^{-1}z v_i^{-1} | 1)p_{B_i,B_2}(y_1, y_2, z | x)
\]

\[
= p_{B_1,B_2}(y_1, y_2, z | x) \sum_{i=1}^{l} \sum_{s_i \in \mathcal{A}_1} p_{A_i,A_2}(x, x_1, y_1, y_2, z | 1) = p_{B_1,B_2}(y_1, y_2, z | x)p_{A_1,A_2}(x) = p_{A_1,A_2,B_1,B_2,C_1}(x, y_1, y_2, z),
\]  

(77)

where, \( \sum_{i=1}^{l} p_{A_i,A_2}(x, x_1, y_1, y_2, z | 1) = l p_{A}(x) \) is argued as follows. Since \( u_i^l(m) \in \mathcal{A}_1 \) is of type \( p_A \) for every \( m \in [M_u] \),

\[
\sum_{a_i^l,a_2^l \in \mathcal{A}_1} p_{A_i,A_2}(a_i^l, a_2^l) = 1, \quad \text{and hence} \quad \sum_{a_i \in \mathcal{A}_1} p_{A_i}(a_i | 1) = 1, \quad \text{which implies}
\]

\[
\sum_{i=1}^{l} p_{A_i,A_2}(x, x_1, y_1, y_2, z | 1) = \sum_{i=1}^{l} \sum_{a_i \in \mathcal{A}_1} p_{A_i}(a_i | 1) \mathbf{1}_{\{a_i = x\}} = \sum_{a_i \in \mathcal{A}_1} \sum_{i=1}^{l} p_{A_i}(a_i | 1) \mathbf{1}_{\{a_i = x\}}
\]

\[
= \sum_{a_i \in \mathcal{A}_1} \sum_{a_i \in \mathcal{A}_1} p_{A_i}(a_i | 1) l p_{A}(x) = l p_{A}(x).
\]

\[
\text{APPENDIX E}
\]

**Fixed B-L Coding Can Strictly Outperform LC**

We introduce an example analogous to Ex. 1 and prove that while it does not satisfy LC conditions, it satisfies those stated in Thm. 5.

**Example 2:** Let \((S, \mathbb{W}_S)\) be the source described in Ex. 1. The IC is described below. The input alphabets are \(U \times X_1\) and \(U \times X_2\). The output alphabets are \(X_1 \times X_1 \) and \(X_2 \times X_2\). The symbols received by Rx \(j\)'s \(Y_j\) denotes \(Y_j \in Y_0 \times X_j\) denotes symbols received by Rx \(j\). The symbols \(Y_0\) received at both Rxs agree with probability 1. \(\mathbb{W}_{Y_0} = \mathbb{W}_{Y_1} \times U_1 \times U_2 = \mathbb{W}_{Y_2} \times X_2 \times U_1 \times U_2\), where

\[
\mathbb{W}_{Y_0}(y_0 | u_1, u_2) = \begin{cases} 1 & \text{if } y_0 = u_1 = u_2 \text{ or } u_1 \neq u_2, y_0 = 0, \\ 0 & \text{otherwise}. \end{cases}
\]

The capacities of PTP channels \(\mathbb{W}_{Y_j} : j = 1, 2\) are \(C_1 = h_b(\frac{2}{7}) + \frac{2}{7} \log a\) and \(C_2 = h_b(\frac{2}{7a^2})\), respectively. Just as in Ex. 1, it can be verified that, for sufficiently large \(a\), the capacity of satellite channel \(\mathbb{W}_{Y_j} \times X_j\) is at most \(\frac{5}{27} \log a\). For all such \(a, k\), we choose satellite channels for which \(|Y_j| \leq \frac{2}{5}\).

**Lemma 4:** Consider Ex. 2 with any \(\eta \in \mathbb{N}\). There exist \(a_0 \in \mathbb{N}, k_0 \in \mathbb{N}\), such that for any \(a \geq a_0\) and any \(k \geq k_0\), the sources and the IC described in Ex. 2 do not satisfy LC conditions that are stated in [7, Thm. 1].

**Proof:** Since the sources do not have a GKW part, it suffices to prove that Ex. 2 does not satisfy conditions stated in Thm. 2. Let \(Q \subseteq \mathbb{W}_{Q\times Q}\) be any collection of RVs whose PMF factorizes as \(\mathbb{W}_{Q\times Q\times Q\times Q\times Q}\). We prove

\[
H(S) \geq I(S_1 X_1 U_1; Y_1 Y_0 Q W) + I(W_1 S_2 X_2 U_2; Y_2 Y_0 Q) - I(S_1 S_2)
\]

(78)
and thereby contradicting (3). The lower bound on \(H(S)\) follows from (55). Secondly, the RHS of (78) can be bounded above by

\[
I(S_1 X_1 U_1; Y_1 Y_0 Q W) + I(W_1 S_2 X_2 U_2; Y_2 Y_0 | Q) - I(S_1; S_2)
\]

\[
\leq I(S_1 X_1 U_1; Y_1 Y_0 | Q W) + I(W S_2 X_2 U_2; Y | Q) - I(S_1; S_2)
\]

\[
\leq I(W S_2 X_2 U_2; Y | Q) + I(S_1 X_1 U_1; Y Q W) - I(S_1 X_1 U_1; Y W S_2 X_2 U_2) - I(S_1; S_2)
\]

\[
= I(X U; Y | Q) + I(S_1 X_1 U_1; Y Q W) - I(S_1 X_1 U_1; Y W S_2 X_2 U_2) - I(S_1; S_2)
\]

\[
= I(X U; Y | Q) + I(S_1 X_1 U_1; Y Q W) - I(S_1 X_1 U_1; Y W S_2 X_2 U_2 | Q W) \leq I(X U; Y | Q)
\]

Following the steps identical to proof of Lemma 1, it can be verified that

\[
I(X U; Y | Q) \leq 2 \log 2 + 2 C_t + h_b(\frac{2}{k a^q}) + \frac{3}{4} \log a + \frac{\log |Y|}{k} < \log a
\]

for sufficiently large \(k, a\). In view of the lower bound on \(H(S)\) (55) and the upper bound on the RHS of (78) via (80) and (81), we are done.

We now provide a choice for parameters in Thm. 5 to prove Ex. 2 satisfies those conditions. Let \(K = S_1 = S_2\), \(U\) be the input alphabet of the shared channel \(W Y_j | U\), \(V_j = X_j : j \in [2]\) be the input alphabet of the satellite channels \(W Y_j | X_j : j \in [2]\), respectively. Let \(f_j(s) = s\) for \(s \in S_j\) be the identity map, and hence \(K_j = S_j\) for \(j \in [2]\). Let \(\alpha = (1 - \frac{1}{k} \log a + \beta = \frac{5}{4k} \log a + (1 + \frac{1}{k}) h_b(1) + \frac{1}{k} \log a, \rho = \frac{1}{k} \log a, \delta = \frac{1}{k}\). Let \(l = k^4 a^\frac{ab}{2}\). Let \(p_U\) be the uniform PMF on \(U = \{0, \ldots, a - 1\}\). Let \(p_{V_j} : j \in [2]\) be the capacity achieving distribution on satellite channels \(W Y_j | X_j : j \in [2]\), respectively. Note that, for any \(u \in U\), \(p_{V_j}(u) = k^4 a^\frac{ab}{2} + 1\) is a natural number since \(a \geq 6\) is an even integer. We refer to the arguments in Appendix C that proves the choice \(l = k^4 a^\frac{ab}{2} \leq l^{*} (\frac{1}{4k} \log \frac{a}{k}, U, V)\) for sufficiently large \(a, k\).

Note that

\[
\beta + H(S_j | S_l) + \mathcal{L}(\phi, | S_j |) + \mathcal{L}(\phi, | V_j |) \leq \frac{5}{4k} \log a + (1 + \frac{1}{k}) h_b(1) + 1 \eta_k(2) h_b(2 \frac{a}{k a^n}) + \phi + (1 + \frac{1}{k}) \log a
\]

\[
+ (1 + \frac{1}{k}) h_b(\phi) \leq \frac{2}{k} \log a + h_b(2 \frac{a}{k a^n}) + 1 \eta_k(2) h_b(2 \frac{a}{k a^n})
\]

for sufficiently large \(a, k\) because with the above choice for \(\delta, l, \alpha, \rho, \beta\), we have from (67) \(\phi \leq \frac{8 k^4}{a^2}\) for sufficiently large \(a, k\). The RHS of (82) is \(I(V_j; Y_j)\) and we have therefore proved (22) for the choice of \(p_{V_j}\) being the capacity achieving PMF.

**APPENDIX F**

**ROWS OF (S_1, S_2, K) ARE IID WITH PMF P_{k^4 S_1 S_2}**

We verify this by just characterizing the distribution of the associated random variables and the use of law of total probability. Suppose for \(j \in [2], t \in [m], i \in [l]\), we have \(k_j(t, i) = f_j(s_j(t, i))\) and \(a_{jt} = e_k(k_j(t, 1 : l))\), then

\[
P\left(S_j = s_j, X_j, V_j | a_{jt}, b_{jt} = x_j, B_j = b_j \right) = P\left(S_j = s_j, S_2 = s_2, M_{V_1}, M_{V_2} \prod_{t=1}^{m} \left( \prod_{i=1}^{l} W Y | (y(t, i) x_1(t, i) x_2(t, i)) \right) \right)
\]

\[
\left\{ \left( \prod_{j=1}^{2} p_{V_j}(y_j(t, t_i))^p_{X_j} y_j(t, i) e_{u}(a_{jt}) v_j(t, i) \right) \right\} \left\{ \hat{k}(t, 1 : l) = d_k \left( d_u(y(t, 1 : l)) \right) \right\}
\]

wherein \(e_u(a_{jt})\) denotes the \(i\)-th symbol in \(e_u(a_{jt}) = u'(a_{jt}) \in U^l\). Expression (84) is equal to

\[
P\left(S_1 = s_1, S_2 = s_2, M_{V_1}, M_{V_2} \prod_{t=1}^{m} \left( \prod_{i=1}^{l} W Y | x_1(t, i) x_2(t, i) \right) \prod_{j=1}^{2} p_{V_j}(y_j(t, i)) p_{X_j} y_j(t, i) e_u(a_{jt}) v_j(t, i) \right)
\]

\[
\left\{ \hat{k}(t, 1 : l) = d_k \left( d_u(y(t, 1 : l)) \right) \right\}
\]

\[
P\left(S_1 = s_1, S_2 = s_2, M_{V_1}, M_{V_2} \prod_{t=1}^{m} p_{V_1} Y_1 | U^l \left( v_1(t, 1 : l), v_2(t, 1 : l), x_1(t, 1 : l) e_u(a_{1t}) x_2(t, 1 : l), y(t, 1 : l) e_u(a_{2t}) \right) \right)
\]

\[
\left\{ \hat{k}(t, 1 : l) = d_k \left( d_u(y(t, 1 : l)) \right) \right\}
\]
where (i) (85) is obtained by re-ordering the product $\prod_{i=1}^m p_{v_i}(v_j(t, \Pi_i(i))) = \prod_{i=1}^m p_{v_i}(v_j(t, i))$, and
(iv) (86) follows from noting that the marginal $p_{u_{1,2}}$ wrt PMF in (28) is given by

$$p_{u_{1,2}}(u_1^t, u_2^t) = \sum_{(a_1, a_2) \in [M_s] \times [M_s]} P(\ A_1 = a_1, A_2 = a_2 \ ) \mathbb{1}\{u_1(a_1) = u_2^t, j \in [2]\}$$

and hence

$$p_{v_{1,2}|Y|U(t)}(u_1^t, u_2^t | y^t) = \left[ \prod_{j=1}^l \left( \prod_{i=1}^l p_{v_j}(v_{ji}) p_{X_j|U, V_j}(x_{ji} | u_{ji}, v_{ji}) \right) \right] \left[ \prod_{i=1}^l \mathbb{W}_{Y|X_1, X_2, y_i \mid x_{1i}, x_{2i}} \right].$$

Following from (84) to (86), we conclude that if

$$k_j(t, i) = f_j(s_j(t, i)) \quad \text{and} \quad a_{ji} = e_k(k_j(t, 1 : l))$$

for $j \in [2], t \in [m], i \in [l]$, we have

$$P(S_1 = s_1, S_2 = s_2) \prod_{i=1}^l \mathbb{W}_{Y|U(t)}(y(t, 1 : l) | e_u(e_k(k_1(t, 1 : l))), e_u(e_k(k_2(t, 1 : l)))) \mathbb{1}\{\hat{k}(t, 1 : l) = d_k(d_u(y(t, 1 : l)))\}.$$
APPENDIX G

PROOF OF (34)

The reader will note that, in proving of (34), we would be asserting that the codes $C_{V_1,i}, C_{V_2,i}$ experience a MAC channel $p_{U|Y_2}$. The steps till (35) are aimed at establishing (34).

Since $V_{1i}^m(B_1i), V_{2i}^m(B_2i), Y_i^m(1 : m, i) = [V_1(B_1) \ V_2(B_2) \ Y_i^m](1 : m, i)$, (34) holds if

$$(V_1(B_1), V_2(B_2), Y)$$

is distributed with PMF $\prod_{t=1}^{m} p_{V_1|V_2|Y; \Pi}$ and $\Pi_1, \ldots, \Pi_m$ is independent of $V_1(B_1), V_2(B_2), Y$. (90)

Indeed, sufficiency of (90) follows from Lemma 6 (Appendix K). Note that

$$P \left( u \{A_j\} = u_j, V_j \{B_j\} = v_j \ \middle| \ A_j = a_j, B_j = b_j \ : j \in [2], \Pi_t = \pi_t : t \in [m] \right) = \sum_{\substack{a_j, a_j \ \in \ B_j, B_j \ : j \in [2], \Pi_t = \pi_t : t \in [m]}} P \left( u \{a_j\} = u_j, V_j \{b_j\} = v_j \ \middle| \ A_j = a_j, B_j = b_j \ : j \in [2] \right).$$

(91)

We now break up the second factor of a generic term in the sum above. In particular,

$$P(\Pi_t = \pi_t : t \in [m]|A_j = a_j, B_j = b_j : j \in [2]) = \frac{1}{m}$$

(92)

$$P \left( V_j \{b_j\} = v_j \ \middle| \ u \{a_j\} = u_j, A_j = a_j, B_j = b_j \ : j \in [2], \Pi_t = \pi_t : t \in [m] \right) = P \left( V_j^m(b_j) = v_j^m(1 : m, i) : i \in [2], j \in [2], u \{a_j\} = u_j : j \in [2] \right)$$

(93)

$$= \prod_{t=1}^{m} \left\{ \begin{array}{l} u_j(t,1:l) \ \\ \prod_{i=1}^{l} p_{V_j}(v_j(t,\pi_t(i))) \end{array} \right\}.$$ (94)

$$P \left( X_j \{a_j\} = x_j \ \middle| \ u \{a_j\} = u_j, A_j = a_j, B_j = b_j \ : j \in [2], \Pi_t = \pi_t : t \in [m] \right) = \prod_{t=1}^{m} \prod_{i=1}^{l} \left\{ \begin{array}{l} u_j(t,1:l) \ \\ \prod_{i=1}^{l} p_{X_j|U_j}(x_j(t,i)|u_j(t,i)v_j(t,i)) \end{array} \right\}.$$ (95)

Substituting (93) - (95), rewriting $\prod_{t=1}^{m} \prod_{i=1}^{l} p_{V_j}(v_j(t,\pi_t(i)))$ as $\prod_{t=1}^{m} \prod_{i=1}^{l} p_{V_j}(v_j(t,i))$, (91) is given by

$$P \left( u \{A_j\} = u_j, V_j \{B_j\} = v_j \ \middle| \ X_j \{a_j\} = x_j \ : j \in [2], Y = y, \Pi_t = \pi_t : t \in [m] \right) = \sum_{\substack{a_j, a_j \ \in \ B_j, B_j \ : j \in [2] \ \text{such that} \ \sum_{t=1}^{m} \prod_{i=1}^{l} p_{V_j}(v_j(t,i))p_{X_j|U_j}(x_j(t,i)|u_j(t,i)v_j(t,i))}} \left( \frac{1}{m} \right)^{m}$$

(96)

$$\sum_{\substack{a_j, a_j \ \in \ B_j \ : j \in [2] \ \text{such that} \ \sum_{t=1}^{m} \prod_{i=1}^{l} p_{V_j}(v_j(t,i))p_{X_j|U_j}(x_j(t,i)|u_j(t,i)v_j(t,i))}} \left( \frac{1}{m} \right)^{m}$$

(97)

$$= \sum_{\substack{a_j, a_j \ \in \ B_j \ : j \in [2] \ \text{such that} \ \sum_{t=1}^{m} \prod_{i=1}^{l} p_{V_j}(v_j(t,i))p_{X_j|U_j}(x_j(t,i)|u_j(t,i)v_j(t,i))}} \left( \frac{1}{m} \right)^{m}$$

(98)
To begin with, we derive an upper bound on 

\[
\sum_{a_1, a_2} P(A_j = a_j) \prod_{j \in [2]} P\left(u_j(t, 1 : l) : u_j(a_j), j \in [2]\right) \left[\prod_{i=1}^l \left[\prod_{j=1}^2 p_{V_j}(v_j(t, i)) p_{X_j}(x_j(t, i) \mid u_j(t, i))\right]\right]
\]

where (98) follows from the invariance of the distribution of \(A_j \sim eK(k_j(t, 1 : l))\) with \(t \in [m]\). Recall that \((A_1, A_2)\) is identically distributed as \((A_{1t}, A_{2t})\) for any \(t \in [m]\). This was stated prior to (28). In arriving at (99), we (i) leveraged the sum over \(a_1, a_2\) being over all of \([M_a]^m \times [M_a]^m\), (ii) the rest of the terms not depending on \(a_1, a_2\), and (iii) \(u_j(\cdot)\) being invariant with \(t\). Finally, (100) follows from definition of (28). We conclude

\[
P \left(\bigcup_{a_1, a_2} \bigcup_{\theta, \varphi} \bigcup_{\phi_{1}, \phi_{2}} \bigcup_{\eta_{1}, \eta_{2}} \bigcup_{\xi_{1}, \xi_{2}} \{B_{1i} = b_{1i}, B_{2i} = b_{2i}, (V_{1i}^m(b_{1i}), V_{2i}^m(b_{2i}), Y^\Pi(1 : m, i)) \in T^m_{\beta}(p_{Y \mid Y})\} \right)
\]

and in particular

\[
P \left(\bigcup_{a_1, a_2} \bigcup_{\theta, \varphi} \bigcup_{\phi_{1}, \phi_{2}} \bigcup_{\eta_{1}, \eta_{2}} \bigcup_{\xi_{1}, \xi_{2}} \{B_{1i} = b_{1i}, B_{2i} = b_{2i}, (V_{1i}^m(b_{1i}), V_{2i}^m(b_{2i}), Y^\Pi(1 : m, i)) \in T^m_{\beta}(p_{Y \mid Y})\} \right)
\]

(103) proves (90).

**Appendix H**

**Upper Bound on \(P(\mathcal{E}_3)\)**

To begin with, we derive an upper bound on

\[
P \left(\bigcup_{b_1, b_2} \bigcup_{\hat{b}_1, \neq b_1, \hat{b}_2, \neq b_2} \{B_{1i} = b_{1i}, B_{2i} = b_{2i}, (V_{1i}^m(\hat{b}_{1i}), V_{2i}^m(\hat{b}_{2i}), Y^\Pi(1 : m, i)) \in T^m_{\beta}(p_{Y \mid Y})\} \right)
\]

By the union bound and the law of total probability, the above quantity is bounded on the above by

\[
\sum_{a_1, a_2} \sum_{b_1} \sum_{b_2} \sum_{\hat{b}_1} \sum_{\neq b_1} \sum_{\hat{b}_2} \sum_{\neq b_2} P \left(\bigcup_{x_{1}^{m}, x_{2}^{m}} \bigcup_{\hat{x}_{1}^{m}, \neq x_{1}^{m}} \bigcup_{\hat{x}_{2}^{m}, \neq x_{2}^{m}} P \left(A_j = a_j, B_j = b_j, V_{ji}^m(\hat{b}_{ji}) = \hat{y}_{ji}^m, V_j(\{b_j\}^m(1 : m, i) = y_{j}^m, X_j(\{a_j, b_j\})^m(1 : m, i) = x_{j}^m, j \in [2], Y^\Pi(1 : m, i) = y_{j}^m\right) \right)
\]

Consider a generic term in the above sum. Firstly, the triple \(A_j, B_j, V_j(\{b_j\}^m(1 : m, i) = j \in [2]\) is independent of \(V_{ji}^m(\hat{b}_{ji}) : j \in [2]\). This is because (i) the codebook generation process is independent of the messages, and (ii) \(V_j(\{b_j\})^m(1 : m, i) : j \in [2]\) is a function of \(V_j(b_j(t)) : j \in [2], i \in [\ell]\) and \(\Pi_t : t \in [m]\), and these random objects are mutually independent of \(V_{ji}^m(\hat{b}_{ji}) : j \in [2]\). Secondly, \(X_j(\{a_j, b_j\})^m(1 : m, i) : j \in [2]\) is conditionally independent of \(V_{ji}^m(\hat{b}_{ji}) : j \in [2]\) given \(A_j, B_j, V_j(\{b_j\}^m(1 : m, i) = j \in [2]\). This is true because (i) \(X_j(\{a_j, b_j\})^m(1 : m, i) : j \in [2]\) is conditionally independent of the rest of the random objects, given \(V_j(\{b_j\}^m(1 : m, i), u_j(\{A_j\})^m(1 : m, i) : j \in [2]\), where \(u_j(\{A_j\})\) is a deterministic function\(^{13}\) of \(A_j\), and (ii) \(\Pi_t : t \in [m]\) is independent of \(V_{ji}^m(\hat{b}_{ji}) : j \in [2]\). Finally,

\(^{13}\)We do not randomize over the fixed B-L code \(C_U\).
The article continues with mathematical expressions and proofs related to the previous ones. The text involves multiple summations, integrals, and inequalities. The text is dense and technical, typical of a research paper in information theory.
Substituting (110) in (108) and summing over \(a_j,b_j,x^m_j : j \in [2], v^m_i\), we obtain
\[
\sum_{b_1: \, (\bar{b}_1, \tilde{b}_1, \bar{g}_1, \tilde{g}_1) \in T_3^m(p_{\bar{g}_1}, p_{\tilde{g}_1})} P(\{V_2 (B_2) \in T_2^m(p_{B_2}) \}) \sum_{t=1}^{m} P_Y(v_t) \sum_{t=1}^{m} p_{\tilde{v}_t} (\tilde{v}_t) p_{\bar{v}_t} (\bar{v}_t) p_{Y}(v_{2t}, y_{2t}) (111)
\]
as an upper bound on (108), where the last equality follows from arguments identical to those that established truth of (107). Once again, based on standard typicality argument, for example lemma [18, Lemma 3.1], it can be proved that given any \(\eta > 0\), there exist a choice for \(\beta > 0\) and \(m_{\beta, \eta} \in \mathbb{N}\) such that for all \(m \geq m_{\beta, \eta}\), (111) is at most \(\eta\) if \(\frac{\log M_{\eta}}{m} < I(Y_1; Y_2)\).

**APPENDIX I**

**COMMON CODEWORD EXPERIENCES PTP \(\prod_{t=1}^{T} P_Y[U]\)**

We are required to prove that whenever the two Txs choose a common \(C_U\) codeword, the latter experiences a PTP channel with transition probabilities \(\prod_{t=1}^{T} P_Y[U]\). Towards that end, we note that
\[
P \left( \{u_{j}^{m} \mid j \in [2], Y = y \} \right) = \sum_{b_1, \tilde{b}_1, \bar{a}_1, \tilde{a}_1} P(\{a_j = \bar{a}_j, b_j = \tilde{b}_j : j \in [2]\} \mid \{u_j^{m} \mid j \in [2], Y = y \}) P(\{a_j = \bar{a}_j, b_j = \tilde{b}_j : j \in [2]\}) \tag{112}
\]
where (112) is identical to (91) except for the range of the summation. Using (93), (94), (95) and following a sequence of steps analogous to those that took us from (91) to (97), we have (112) equal to
\[
P \left( \{u_{j}^{m} \mid j \in [2], Y = y \} \right) = \prod_{t=1}^{T} p_Y(v_t) p_X(v_t) p_{X,Y}(y_t | u_t^{m}) \left( u_1(t, 1 : l), u_2(t, 1 : l) \right) \tag{113}
\]
where \(u_1(t, 1 : l)\) is the \(i\)-th symbol of \(u_1(t, 1 : l)\) and (113) follows from (72). Since LHS of (112) is equal to the RHS of (113), summing the latter (113), we have
\[
\sum_{u_1, u_2, v_1, v_2, x_1, x_2} P \left( \{u_{j}^{m} \mid j \in [2], Y = y \} \right) = \prod_{t=1}^{T} p_Y(v_t) p_X(v_t) p_{X,Y}(y_t | u_t^{m}) \left( u_1(t, 1 : l), u_2(t, 1 : l) \right) \tag{114}
\]
and hence
\[
P(Y = y \mid A_1 = a_1, A_2 = a_2) = \prod_{t=1}^{T} p_Y(v_t) p_X(v_t) p_{X,Y}(y_t | u_t^{m}) \left( u_1(t, 1 : l), u_2(t, 1 : l) \right) \tag{115}
\]
where (114) follows from (75).

**APPENDIX J**

**ROWS OF TRANSMITTED AND DECODED MATRICES ARE IID**

We choose the PMF of the random code identical to that in Sec. VI-A. Note that
\[
P \left( \{u_{j}^{m} \mid j \in [2], Y = y \} \right) = \prod_{t=1}^{T} p_Y(v_t) p_X(v_t) p_{X,Y}(y_t | u_t^{m}) \left( u_1(t, 1 : l), u_2(t, 1 : l) \right) \tag{116}
\]

We break down the second factor in a generic term above just as we did for the analogous term in (91). Essentially (93), (94), (95), and in addition
\[
P(\begin{align*}
  u_{\tilde{A}_j} &= \tilde{u}_j, \\
  Y_j &= y_j, \\
  X_j &= x_j : j \in [2]
\end{align*})
\]
leads us to breaking down the second factor in a generic term of (115) as
\[
P\left(\begin{align*}
  u_{\tilde{A}_j} &= u_j, \\
  Y_j &= y_j, \\
  X_j &= x_j : j \in [2]
\end{align*}\right)
\]
Substituting (117) in (115) and following a sequence of steps that took us from (93) to (100), we have
\[
\text{where (118) follows from the fact that } \prod_{t=1}^{l} p_{\alpha_j}(v_j(t, \Pi_{t}(i))) = \prod_{t=1}^{l} p_{\alpha_j}(v_j(t, i))\text{ and (119) follows from the invariance of the distribution of } A_{\alpha_j} = d_{\alpha_j}(y_j(t, 1 : l)) \text{ with } t \leq |m|.
\]


**APPENDIX K**

**INTERLEAVING RESULTS IN IID DISTRIBUTIONS**

**Lemma 5:** Let $\mathcal{A}$ be a finite set and $p_{A^t}$ be a PMF on $\mathcal{A}$. Let $A(1, 1 : l), A(2, 1 : l), \ldots, A(m, 1 : l) \in \mathcal{A}^l$ be independent and identically distributed vectors with PMF $p_{A^t}$. Let $\Lambda_1, \ldots, \Lambda_m$ be independent and uniformly distributed indices taking values in $\{1, \ldots, l\}$. Moreover, $\Lambda_1, \ldots, \Lambda_m$ is independent of the collection $A(1, 1 : l), A(2, 1 : l), \ldots, A(m, 1 : l)$. Then the components $A(t, \Lambda_t) : t \in [m]$ are independent and identically distributed with PMF $\frac{1}{l} \sum_{i=1}^l p_{A_i}$, where $p_{A_i}$ is the PMF of $A(t, i)$.

**Proof:** Note that

$$P(A(t, \Lambda_t) = a_t : t \in [m]) = \sum_{j_1 \in [l]} \cdots \sum_{j_m \in [l]} P(A(t, j_t) = a_t, \Lambda_t = j_t : t \in [m])$$

$$= \frac{1}{l^m} \sum_{j_1 \in [l]} \cdots \sum_{j_m \in [l]} P(A(t, j_t) = a_t : t \in [m])$$

$$= \frac{1}{l^m} \sum_{j_1 \in [l]} \cdots \sum_{j_m \in [l]} P(A(t, j_t) = a_t)$$

$$= \frac{1}{l^m} \left( \frac{1}{l} \sum_{j_1 \in [l]} P(A(t, j_t) = a_t) \right) = \frac{1}{l^m} \left( \frac{1}{l} \sum_{i=1}^l p_{A_i}(a_t) \right),$$

where (i) (120) follows from independence of $(\Lambda_1, \ldots, \Lambda_m)$ and $A(1 : m, 1 : l)$, (ii) (121) follows from the independence of the vectors $A(1, 1 : l), A(2, 1 : l), \ldots, A(m, 1 : l) \in \mathcal{A}^l$, (iii) (122) follows from $A(1, 1 : l), A(2, 1 : l), \ldots, A(m, 1 : l) \in \mathcal{A}^l$ being identically distributed, and moreover, $p_{A_i}(a) = P(A(t, i) = a)$.

**Lemma 6:** Let $\mathcal{A}$ be a finite set and $p_{A^t}$ be a PMF on $\mathcal{A}$. Let $A(1, 1 : l), A(2, 1 : l), \ldots, A(m, 1 : l) \in \mathcal{A}^l$ be independent and identically distributed vectors with PMF $p_{A^t}$. Let $\Theta_l$ be the set of all surjective maps on the set $\{1, 2, \ldots, l\}$. Let surjective maps $\Lambda_1, \Lambda_2, \ldots, \Lambda_m$ be chosen uniformly and independently from $\Theta_l$. For $i = 1, 2, \ldots, l$, let

$$B(t, i) = A(t, \Lambda_i(i)) : t \in [m], i \in [l].$$

The $l$ vectors $B(1 : m, i) : i = 1, 2, \ldots, l$ are identically distributed with PMF $\prod_{i=1}^m \frac{1}{l} \sum_{j=1}^l p_{A_i}$, where

$$p_{A_i}(a) = \sum_{a_1 \in A} \cdots \sum_{a_{i-1} \in A} \sum_{a_i \in A} p_{A}^{(1)}(a_1, \ldots, a_{i-1}, a, a_{i+1}, \ldots, a_l).$$

**Proof:** For any $i \in [l]$, note that

$$P(B(t, i) = a_t : t \in [m]) = \sum_{j_1 \in [l]} \cdots \sum_{j_m \in [l]} P(A(t, j_t) = a_t, \Lambda_t(i) = j_t : t \in [m])$$

$$= \frac{1}{l^m} \sum_{j_1 \in [l]} \cdots \sum_{j_m \in [l]} P(A(t, j_t) = a_t : t \in [m])$$

$$= \frac{1}{l^m} \sum_{j_1 \in [l]} \cdots \sum_{j_m \in [l]} P(A(t, j_t) = a_t)$$

$$= \frac{1}{l^m} \left( \frac{1}{l} \sum_{j_1 \in [l]} P(A(t, j_t) = a_t) \right) = \frac{1}{l^m} \left( \frac{1}{l} \sum_{i=1}^l p_{A_i}(a_t) \right),$$

where (i) (123) follows from independence of the surjective maps $(\Lambda_1, \ldots, \Lambda_m)$ and $A(1 : m, 1 : l)$, (ii) (124) follows from the independence of the vectors $A(1, 1 : l), A(2, 1 : l), \ldots, A(m, 1 : l) \in \mathcal{A}^l$, (iii) (125) follows from $A(1, 1 : l), A(2, 1 : l), \ldots, A(m, 1 : l) \in \mathcal{A}^l$ being identically distributed, and moreover, $p_{A_i}(a) = P(A(t, i) = a)$.

**Lemma 7:** Let $C_U$ be a constant composition code of type $p_U$ with message index set $[M^m]$, encoder map $e_a : [M^m] \rightarrow U^l$ with codewords $u^l(a) : a \in [M^m]$. Let $A_t \in [M^m]$ be a (random) message and $U^l = u^l(A_t)$.
denote the corresponding codeword. Suppose $I\{1, \cdots, l\}$ is uniformly distributed and independent of $U^l$, then $p_{u_i} = \frac{1}{l} \sum_{i=1}^l p_{u_i} = p_U$.

Proof: Finally, let us identify $\prod_{i=1}^m \frac{1}{l} \sum_{i=1}^l p_{u_{ji}}$, the PMF of these sub-vectors. Observe that $C_U$ is a constant composition code of type $p_U$. Irrespective of the PMF of the messages $A_{jt}$ indexing this codebook, the indexed codeword $u^l(A_{jt})$ has type $p_U$. A uniformly chosen symbol from $u^l(A_{jt})$ will therefore have PMF $p_U$. We make this formal through the following identities. Note that

$$\frac{1}{l} \sum_{i=1}^l p_{u_{ji}}(c) = \frac{1}{l} \sum_{i=1}^l \sum_{u^l \in U^l} p_{u^l}(u^l) \mathbb{I}_{\{u_i = c\}} = \frac{1}{l} \sum_{u^l \in U^l} \sum_{i=1}^l p_{u^l}(u^l) \mathbb{I}_{\{u_i = c\}}$$

$$= \frac{1}{l} \sum_{u^l \in U^l} \sum_{i=1}^l P(u(A_{ji}) = u^l) \mathbb{I}_{\{u_i = c\}} = \frac{1}{l} \sum_{u^l \in U^l} \sum_{i=1}^l P(u(A_{ji}) = u^l) \mathbb{I}_{\{u_i \text{ has type } p_U\}} \mathbb{I}_{\{u_i = c\}}$$

$$= \frac{1}{l} \sum_{u^l \in U^l} P(u(A_{ji}) = u^l) \mathbb{I}_{\{u_i \text{ has type } p_U\}} p_U(c) \sum_{u^l \in U^l} P(u(A_{ji}) = u^l) \mathbb{I}_{\{u_i \text{ has type } p_U\}} = p_U(c),$$

and hence conclude sub-vector $(U_j(t, \Lambda_1(t)) : t \in [m])$ has PMF $\prod_{t=1}^m \frac{1}{l} \sum_{i=1}^l p_{u_{ji}} = \prod_{t=1}^m p_U$.

APPENDIX L
THE INTERLEAVING CONSTRUCT

Let $p_{A^l B^l}$ be a PMF on $A^l \times B^l$. We will prove

$$\sum_{a \in A} \sum_{b \in B} \prod_{t=1}^m p_{A^l B^l}(a(t, 1 : l), b(t, 1 : l)) \frac{1}{l^{m-l}} \sum_{\pi_1(i)=1}^l \cdots \sum_{\pi_m(i)=1}^l \left\{ \begin{array}{ll} [ab]^\pi (1 : m, i) \end{array} \right\} = \sum_{a \in A^m} \sum_{b \in B^m} \prod_{t=1}^m \left\{ \begin{array}{ll} [ab]^\pi (1 : m, i) \end{array} \right\} = \prod_{t=1}^m p_{AB}(a_t, b_t)$$

where

$$p_{AB}(u, v) = \frac{1}{l} \sum_{i=1}^l p_{A_i B_i}(u, v)$$

$p_{A_i B_i}$ is the PMF of the $i$-th component of $A^l, B^l$. Observe that

$$\sum_{a \in A^m} \sum_{b \in B^m} \prod_{t=1}^m p_{A^l B^l}(a(t, 1 : l), b(t, 1 : l)) \frac{1}{l^{m-l}} \sum_{\pi_1(i)=1}^l \cdots \sum_{\pi_m(i)=1}^l \left\{ \begin{array}{ll} [ab]^\pi (1 : m, i) \end{array} \right\} = \sum_{a \in A^m} \sum_{b \in B^m} \prod_{t=1}^m [ab]^\pi (1 : m, i) \prod_{t=1}^m \left\{ \begin{array}{ll} [ab]^\pi (1 : m, i) \end{array} \right\} = \prod_{t=1}^m p_{AB}(a_t, b_t)$$

\[
= \sum_{x^m \in \mathcal{A}^m} \sum_{y^m \in \mathcal{B}^m} \sum_{l \in \pi_1(i)=1} \sum \prod_{i=1}^{l} \mathbb{1}_{\left\{\left(x^m, y^m\right) \notin T^m_{\delta}(p_{\hat{\mathcal{A}}})\right\}} \prod_{t=1}^{m} \frac{1}{l_{m}} \prod_{t=1}^{m} p_{A_{x(l)} B_{x(l)}(x_t, y_t)} \mathbb{1}_{\left\{\left(x^m, y^m\right) \notin T^m_{\delta}(p_{\hat{\mathcal{A}}})\right\}} = \sum_{x^m \in \mathcal{A}^m} \sum_{y^m \in \mathcal{B}^m} \prod_{t=1}^{m} p_{A_{x(t)} B_{x(t)}(x_t, y_t)} \mathbb{1}_{\left\{\left(x^m, y^m\right) \notin T^m_{\delta}(p_{\hat{\mathcal{A}}})\right\}}
\]

which is what we sought out to prove.

REFERENCES