Communicating Correlated Sources over Multi-user Channels

Arun Padakandla
Center for Science of Information
Purdue University

Abstract—A new coding technique, based on fixed block-length codes, is proposed for the problems of communicating a pair of correlated sources over (i) a 2-user multiple access channel (MAC) and (ii) a 2-user interference channel. Its performance is analyzed to derive a new set of sufficient conditions for both problems. The derived conditions are proven to be strictly less binding than the current known best. Our findings are inspired by Dueck’s example [1].

I. INTRODUCTION

The presence of correlation amongst distributed information sources can be exploited via intelligent information processing techniques to enable efficient communication. Transferring source correlation via test channels [2] or binning [3] enables one to exploit soft or probabilistic correlation. When the sources possess Gács-Körner-Witsenhausen common (GKW) part the technique of conditional coding provides further benefits. In this article, we propose a coding technique to exploit the presence of near GKW part amongst distributed sources.

We are concerned with the problem of communicating distributed correlated sources over multi-user channels. Our primary focus are the two scenarios depicted in Figs. 1, 2. Consider a pair \( S_1, S_2 \) of correlated sources. Transmitter (Tx) \( j \) observes \( S_j \). In the MAC problem (Fig. 1), the receiver (Rx) \( j \) wishes to reconstruct both \( S_1, S_2 \) losslessly. In the IC problem (Fig. 2), Rx \( j \) wishes to reconstruct \( S_j \) losslessly. We undertake a Shannon-theoretic study and restrict attention to characterizing sufficient conditions for transmissibility of \( S_1, S_2 \) over the corresponding channels.

Cover, El Gamal and Salehi [2] proved that separate source and channel coding is sub-optimal for the MAC problem. By transferring soft or probabilistic correlation via single letter (S-L) test channels, they enabled better co-ordination. This was further enhanced via conditional coding in the presence of GKW part amongst the sources. Their elegant technique [2], henceforth referred to as CES, is fundamental to joint source-channel coding and has been employed in the context of broadcast [4] and interference channels [5] too. In particular, Liu and Chen [5] bring to bear all current known techniques - CES, random source partitioning [4] and message splitting via superposition coding [6] - and propose a coding (LC) technique for the IC problem. They derive a set of sufficient (LC) conditions [5, Thm. 1] that remains to be the current known best for the IC problem.

It is important to note that ignoring GKW part, i.e., treating it as soft correlation is strictly sub-optimal. In other words, conditional coding yields strictly enhanced throughput. We consider the case when the sources have a near but not perfect, GKW part. While the CES and LC techniques are constrained to treating this as soft correlation, we propose ‘approximate conditional coding’ via fixed block-length (B-L) codes, and harness a fraction of the gains promised by (perfect) conditional coding. This enables us outperform the CES and LC techniques, and we therefore prove that the CES and LC conditions are, in general, not necessary for the respective problems. We undertake an information-theoretic analysis of the proposed coding technique and derive a set of sufficient conditions for both problems. Through examples, we prove that the derived conditions can be strictly less binding than those of CES and LC.

Presence of GKW part enables encoders co-ordinate their inputs, and thereby eliminate interference for the corresponding component of the channel input. Moreover, GKW part enables co-ordination even while enjoying separation. In other words, one can design a channel code corresponding to an optimizing input pmf, unconstrained by the source pmf. If \( S_1, S_2 \) do not possess a GKW part, a single-letter (S-L) technique is constrained by the S-L long Markov chain (LMC) \( X_1 - S_1 - S_2 - X_2 \). The S-L LMC can, in general, severely constrain the set of achievable input pmfs (Ex. 1, Rem. 1). If \( S_1, S_2 \) possess a near GKW part, i.e., \( K_j = f_j(S_j) : j \in [2] \)

This work was supported by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-0939370.

Fig. 1. Transmission of correlated sources over MAC.

Fig. 2. Transmission of correlated sources over 2-IC.
such that $\xi = P(K_1 \neq K_2)$ is ‘quite’ small, (relatively) large sub-blocks of length $l$ could agree with high probability. Indeed, $\xi^{[l]} = P(K_1^{[l]} \neq K_2^{[l]}) = 1 - (1 - \xi)^l \leq l\xi$ can be held small by appropriately choosing $l$. If the encoders employ conditional coding, i.e., identical source to channel mappings, restricted to sub-blocks of fixed length $l$, then the encoders can enjoy the benefits of separation and co-ordination on a good fraction (at least $\sim (1 - l\xi)$) of the $l$-length sub-blocks. Indeed, we prove in Section III that the latter technique outperforms CES and LC coding techniques. In Sections IV, V, we build on this idea to propose a general coding technique for an arbitrary problem instance.

Joint source-channel coding over multi-user channels have been studied from Shannon-theoretic viewpoint in [7]–[9]. In particular, [9] studies necessary conditions for the MAC problem. In the interest of developing feasible strategies based on the separation approach, [10] proposes hybrid coding. Fundamental performance limits for communicating Gaussian sources over Gaussian channels have been studied in [11], [12] and the latter considers communication over on IC.

Our findings highlight the sub-optimality of (current known) S-L joint source-channel coding techniques (Rem. 2). Notwithstanding this, we derive a S-L characterization (Rem. 6) of a new inner bound that strictly enlarges the current known best (CES and LC bound). Indeed, the fixed B-L coding technique is an $l$-letter technique. An important second contribution is therefore, a framework - codes and tools (interleaving) - for stitching together S-L techniques in a way that permits performance analysis of the resulting $l$-letter technique and characterization via S-L expressions. Stepping beyond performance characterization, our third contribution is a new coding technique for communicating correlated sources over an IC.

This is part of an evolving work [14]–[16] on joint source-channel coding, and is inspired by Dueck’s novel example [1] and his very specific, yet ingenious, fixed B-L coding. Here, we restrict attention to separation based schemes\(^1\) and focus on providing a clear step-by-step description of the ideas. Unifying fixed B-L coding and inducing source correlation onto channel inputs [2] involves additional challenges, and is dealt in a concurrent submission [15, Sec. V].

II. PRELIMINARIES: NOTATION, PROBLEM STATEMENT

We supplement standard information theory notation - upper case for RVs, calligraphic for sets etc. - with the following. We let an underline denote an appropriate aggregation of related objects. For example, $\underline{S}$ will be used to represent a pair $S_1, S_2$ of RVs. $\underline{S}$ 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. When $j \in \{1, 2\}$, then $j$ will denote the complement index, i.e., $\{j, \overline{j}\} = \{1, 2\}$. For $m \in \mathbb{N}$, $[m] = \{1, \cdots, m\}$.

$$T_3^n(U) = \{u^n \in U^n : \frac{N(b|u^n)}{n} - p_U(b) \leq \delta p_U(b) \ \forall b \in U\}$$

is our typical set. For a pmf $p_U$ on $U$, $b^* \in U$ will denote a symbol with the least positive probability wrt $p_U$.\(^2\) Boldfaces letters such as $\mathbf{A}$ denote matrices. For a $m \times l$ matrix $\mathbf{A}$, (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 whp, S-L, LMC, B-L respectively.

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

$$\min \{D(\mathcal{W}_{Y|U} || \mathcal{W}_{Y|U}(p_U)) + I(p_U; \mathcal{W}_{Y|U}) - R^+\}.$$  

For RVs $A_1, A_2$, we let $\xi^{[l]}(A) = P(A_1^{[l]} \neq A_2^{[l]})$, and $\xi(A) = \xi^{[1]}(A)$. Throughout Sec. III, $\xi^{[l]} = \xi^{[l]}(\mathcal{S})$ and $\xi = \xi(\mathcal{S})$. If $A$ is IID, we note $\xi^{[l]} = 1 - (1 - \xi)^l \leq l\xi$. We let $\rho_{\delta}(K) = 2|K| \exp\{-2\delta^2 \rho_{\delta}^2(\alpha | l)\}$ denote an upper bound on $P(K \notin T_3^\delta(K))$ where $T_3^\delta(K)$ denotes our typical set.

Consider a 2-user MAC with input alphabets $X_1, X_2$, output alphabet $Y$ and channel transition probabilities $\mathcal{W}_{Y|X_1 X_2}$ (Fig. 1). Let $S = (S_1, S_2)$, taking values over $\mathcal{S} = S_1 \times S_2$ with pmf $\mathcal{W}_{S_1 S_2}$, denote a pair of information sources. For $j \in \{2\}$, Tx $j$ observes $S_j$. The Rx aims to reconstruct $S$ 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 $(\mathcal{A}, Y, \mathcal{W}_{Y|X})$.

Consider a 2-user IC with input alphabets $X_1, X_2$, output alphabets $Y_1, Y_2$, and transition probabilities $\mathcal{W}_{Y_1 Y_2|X_1 X_2}$ (Fig. 2). Let $S = (S_1, S_2)$, taking values over $\mathcal{S} = S_1 \times S_2$ with pmf $\mathcal{W}_{S_1 S_2}$, denote a pair of information sources. For $j \in \{2\}$, Tx $j$ observes $S_j$. The Rx $j$ aims to reconstruct $S_j$ with arbitrarily small probability of error. If this is possible, we say $S$ is transmissible over IC $\mathcal{W}_{Y_j|X}$. With regard to the IC problem, our objective is to characterize sufficient conditions under which $(\mathcal{S}, \mathcal{W}_S)$ is transmissible over IC $\mathcal{W}_{Y_j|X}$.

III. FIXED B-L CODING OVER ISOLATED CHANNELS

We consider simple generalizations (Ex. 2 for MAC problem, Ex. 1 for IC problem) of Dueck’s example [1] and propose a coding technique that enables transmissibility of the sources over the corresponding channels. We also prove all current known joint source-channel coding techniques, in particular CES and LC techniques, are incapable of the same. On the one hand, this proves strict sub-optimality of the latter\(^3\) and on the other, highlights the need for fixed B-L coding.

A. Fixed B-L coding for IC problem

Example 1: Source alphabets $S_1 = S_2 = \{0, 1, \cdots, a - 1\}^k$. Let $\eta \geq 8$ be a positive even integer. The source PMF is

$$\mathcal{W}_{S_1 S_2}(c^k, d^k) = \begin{cases} \frac{k}{k} & \text{if } c^k = d^k = 0^k \\ \frac{k}{k a^{\eta} (a - 1)} & \text{if } c^k = d^k, c^k \neq 0^k, \\ \frac{k}{k a^{\eta} (a - 1)} & \text{if } c^k = 0^k, d^k \neq 0^k, \end{cases}$$

\(^1\)As was done in [14].

\(^2\)The underlying pmf $p_U$ will be clear from context.

\(^3\)Strict sub-optimality of CES conditions was proven by Dueck [1]. Strict sub-optimality of LC technique can be inferred from [1]. To verify this, modify the MAC therein, to an IC with identical outputs, and use the arguments presented in proof of Lemma 1. Surprisingly, this has not been documented in [5].
the decoders can reconstruct $S_1$ and $S_2$ if each of them is provided $Y_1$ and $Y_2$ (and $Y_0$). We then prove that this is not permissible, by following an argument similar to [1, Sec. III.c].

Lemma 1: Consider Ex. 1 with any $\eta \in \mathbb{N}$. There exists an $a_* \in \mathbb{N}, k_* \in \mathbb{N}$, such that for any $a \geq a_*$ and any $k \geq k_*$, the sources and the IC described in Ex. 1 do not satisfy LC conditions that are stated in [5, Thm. 1].

Proof: Since the sources do not have a GKW part, it suffices to prove that Ex. 1 does not satisfy conditions stated in [5, Corollary 1]. Let $S\subseteq\{Q,W\} \subseteq U \times X$ be any collection of RVs whose pmf factorizes as $\mathbb{P}(Q,W) = \mathbb{P}(Q)\mathbb{P}(W)$.

We prove:

$$H(S) = \max_{\pi \in \mathcal{P}(Q,W)} H(\pi) \leq \max_{\pi \in \mathcal{P}(Q,W)} H(\pi | Q,W) \leq H(Q,W) \leq H(Q) + H(W)$$

whenever $\eta \geq 2$. Secondly, the RHS of (1) can be bounded above by

$$H(S) \leq \log(a^k - 1) + \frac{1}{k} \log(k) - \log a \geq \log a + \frac{1}{k} \log(k) - \log a$$

Towards that end, we first note

$$\frac{1}{k} \log(k) - \log a \geq \frac{1}{k} \log(k) - \log a$$

whenever $a^k \geq 2$. Secondly, the RHS of (1) can be bounded above by

$$\log a + \frac{1}{k} \log(k) \geq \log a$$

Figure 3. On the left, the source pmf is depicted through a bipartite graph. Larger probabilities are depicted through edges with thicker lines. On the right, we depict the probability matrix.

Figure 4. Source channel setup of Example 1.

0 otherwise. Note that in the above eqn. $a^k, d^k \in S_1$ abbreviate the $k$ ‘digits’ $c_1, c_2, \ldots, c_k$ and $d_1, d_2, \ldots, d_k$ respectively. Fig. 3 depicts the source pmf with $\eta = 6$.

The IC is depicted in Fig. 4 and described below. The input alphabets are $\mathcal{U} \times \mathcal{X}_1$ and $\mathcal{U} \times \mathcal{X}_2$. The output alphabets are $\mathcal{Y}_1 \times \mathcal{Y}_1 \times \mathcal{Y}_1$ and $\mathcal{Y}_0 \times \mathcal{Y}_1 \times \mathcal{Y}_1$. $U = (0, 1, \ldots, a - 1)$ and $(Y_0, Y_j) \in \mathcal{Y}_0 \times \mathcal{Y}_1$ denote encoder $j$’s input and $(Y_0, Y_j) \in \mathcal{Y}_0 \times \mathcal{Y}_1$ denotes symbols received by decoder $j$. The symbols $Y_0$ received at both decoders agree with probability 1. $\mathbb{P}(S_1, S_2 | X_1, X_2) = \mathbb{P}(S_1 | X_1) \mathbb{P}(S_2 | X_2) \mathbb{P}(Y_0 | U_1, U_2)$, where $\mathbb{P}(Y_0 | U_1, U_2) = \begin{cases} 1 & \text{if } y_0 = u_1 = u_2 \\ 0 & \text{if } u_1 \neq u_2 \text{ and } y_0 = 0 \end{cases}$.

0 otherwise. The capacities of PTP channels $\mathbb{P}(S \mid j = 1, 2) = C = \log(a^k) + \log(a^k - 1)$ and $C = \log(a^k) + \log(a^k - 1)$ respectively. We identify key aspects of Ex. 1. Let $a, k$ be chosen sufficiently/quiet large. While $S$ does not possess a GKW part, $S_1$ and $S_2$ agree on most, but not all, realizations. Indeed, $\xi(S) = \frac{1}{k^2 - a^k}$ is very small. We also have $H(S_1, S_2), H(S) = \log(a^k) + \log(a^k - 1)$.

Each decoder benefits a lot by decoding either source, or any function thereof. Secondly, $\mathbb{P}(S \mid j = 1, 2)$ is ‘very far’ from the uniform pmf, and hence any S-L function $g_j(S_j)$ will remain ‘considerably’ non-uniform.

The IC supports a sum capacity of at most $\log(a^k + 1) + \log(a^k - 1)$, since $\mathbb{P}(X_1, X_2) = 2C + \log(a^k)$. The $\mathbb{P}(Y_0 | U_1, U_2)$-channel must carry bulk of the information (for large $k$).

The latter channel carries very little information when $U_1 \neq U_2$, and moreover, it is necessary that $U_1 = U_2$ and $U_1 = U_2$ be close to uniform, in order to communicate $\sim \log a$ bits over $\mathbb{P}(Y_0 | U_1, U_2)$.

We first prove Ex. 1 does not satisfy LC conditions. The proof is based on the following argument. Suppose LC tech-
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 $u \in U$, $P(U_1 = u) \geq \frac{1}{2}$ 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 $H(Y_0) \leq \log 2 + 2 \log a$.

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

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

In all cases, we have $H(Y_0) \leq \log 2 + \frac{1}{2} \log a$. Substituting through (4) and above, we conclude

$$I(X; U; Y | Q) \leq 2 \log 2 + 2C + h_b \left( \frac{2}{ka\gamma_k} \right) + \frac{3}{4} \log a + \frac{\log |Y|}{k} < \log a$$

for sufficiently large $k, a$.

Remark 1: Why is the LC technique incapable of communicating $S_0^3$? Any valid pmf $p_{U_1, U_2}$ induced by a S-L coding scheme is constrained to the LMC $U_1 - S_1 - S_2 - U_2$. For $j \in [2]$, $p_{U_1 (S_j)}$ can equivalently be viewed as $U_j = g_j(S_j, W_j)$, for some function $g_j$ and RV $W_j$, that satisfy $W_1 \perp W_2$. Owing to the latter, $W_1$ 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, and even by pooling all the less likely symbols, we can gather a probability, of at most, $\frac{1}{k}$. Consequently, any $p_{U_1, U_2}$ induced via a S-L coding scheme is sufficiently far from any pmf that satisfies $U_1 = U_2$ w.h.p and $U_1 = U_2$ close to uniform.

Remark 2: An $l$-letter (multi-letter with $l > 1$) coding scheme is constrained by an $l$-letter LMC $U_1 - S_1 - S_2 - U_2$. Suppose we choose $l$ reasonably large such that $1) f^{|f|}(S)$ is not high, and 2) $S_1$ is reasonably uniform on its typical set $T_j^3(S_1)$, and define $U_j : j \in [2]$ through identical functions $U_1^j = g(S_1^j) : j \in [2]$, then one can easily visualize the existence of $g$ such that $p_{U_1 (U_2)}$, satisfies the twin objectives of $U_1 = U_2$ w.h.p and $U_1 = U_2$ close to uniform. Our coding scheme, will in fact, identify such $g$ maps. This portrays the sub-optimality of S-L schemes for joint source-channel coding.

We now present our fixed B-L coding for the IC problem. In order to input codewords on the $\mathbb{W}_{Y_0|U_2}$-channel, that agree, we employ the same source code, same channel code and same mapping, each of fixed B-L$L$ $l$, at both encoders. $l$ is chosen large enough such that the source can be reasonably efficiently compressed, and yet small enough, to ensure $f^{|f|}(S)$ is reasonably small. We refer to these $l$-length blocks as sub-blocks. Since $l$ is fixed, there is a non-vanishing probability that these source sub-blocks will be decoded erroneously. An outer code, operating on an arbitrarily large number $m$ of these sub-blocks, will carry information to correct for these ‘errors’. The outer code will operate over satellite channel $\mathbb{W}_{Y_1|X_j}$. We begin with a description of the fixed B-L codes.

We employ a simple fixed B-L (inner) code. Let $T_j^3(S_1)$ be the source code, and let $C_U = U^l$ be the channel code. Let $lA = [\log d']$ bits, of the $[\log |T_j^3(S_1)|]$ bits output by the source code, be mapped to $C_U$. Both encoders use the same source code, channel code and mapping.

Suppose we communicate an arbitrarily large number $m$ of these sub-blocks on $\mathbb{W}_{Y_1|X_j}$ as above. Moreover, suppose encoder $j$ communicates the rest of the $lB = [\log |T_j^3(S_1)|] - lA$ bits output by its source code to decoder $j$ on its satellite channel $\mathbb{W}_{Y_1|X_j}$. How much more information needs to be communicated to decoder $j$, to enable it reconstruct $S_j^m$? We do a simple analysis that suggests a natural coding technique.

View the $m$ sub-blocks of $S_j$ as the rows of the matrix $S_j(1 : m, 1 : l) \in S_j^{m \times l}$. Let $K_j(1 : m, 1 : l) \in S_j^{m \times l}$ denote decoder $j$’s reconstruction of $S_j(1 : m, 1 : l)$

$\{ (S_j(t, 1 : l), K_j(t, 1 : l) : j = 1, 2, \ldots : t \in [m] \}$

are iid with an $l-$length distribution $\mathbb{W}_{S_j^l, S_j^l} P_{K_j} K_j^l S_j^l S_j^l S_j^l S_j^l$.

Since, in principle, we can operate by treating these $l-$length sub-blocks as a *super-symbol*, and employ standard binning technique over these $m$ super-symbols, decoder $j$ needs only $H(S_j^l | K_j^l)$ bits per source sub-block. We have no characterization of $P_{K_j} K_j^l S_j^l S_j^l S_j^l S_j^l$ and hence we derive an upper bound.

$$H(S_j^l | K_j^l) \leq H(S_j^l, I_{K_j^l S_j^l} K_j^l) \leq h_b (P(K_j^l \neq S_j^l)) + \frac{1}{2} \log (P(K_j^l \neq S_j^l))$$

represents the additional source coding rate needed to compensate for the errors in the fixed B-L decoding. It suffices to prove $L^p_j (P(K_j^l \neq S_j^l)) = 0$ and $H(S_j^l | K_j^l) \leq \frac{1}{2} \log (P(K_j^l \neq S_j^l))$ by a quantity that is less than $\frac{1}{2}$. Towards that end, note that $S_j^l \neq K_j^l$ implies both encoders input same $C_U$-codeword and agree on the $lB$ bits communicated to their respective decoders. Therefore $P(S_j^l \neq K_j^l) \leq \phi$, where $\phi = \tau_{1,l}(S_1) + \xi(S)$,

$$\tau_{1,l}(S_1) \leq 2a^k \exp\left(-\frac{\delta^2 l}{2k^2 a^2 k}\right) + \xi(S) \leq \frac{1}{3} \log k,$$
verify \( \phi \leq 2^{k^2} a^{-\frac{ab}{k^2}} < \frac{1}{2} \) for sufficiently large \( a, k \). Verify
\[
L^S(2^{k^2} a^{-\frac{ab}{k^2}}, |S_j|) \leq \frac{1}{4k} \log a \tag{9}
\]
for sufficiently large \( a, k \). Recall \( IB = \lceil \log |T^S_1(S_1)| \rceil - lA \). Substituting \( \delta = \frac{1}{4k} \), verify
\[
B \leq (2/l) + (1/k) \log a + (1 + (1/k)) h_2(1/k). \tag{10}
\]
Since \( h_2(\frac{2}{k}) - (1 + 1/k) h_2(\frac{2}{k}) \geq \frac{1}{2} \log \frac{k}{\delta a n^a} \) for large enough \( k \), RHS of (16), (17) sum to at most \( \frac{2}{k} + h_2(\frac{2}{k}) + \frac{a}{2k} \log a \) for large enough \( a, k \). Furthermore, \( H(S_2|S_1) \leq h_2(\frac{1}{k} + 1) a^2 + \frac{1}{k} \log a \leq h_2(\frac{2}{k^2 a^2}) + \frac{1}{k} \log a \) for sufficiently large \( a, k \). It can now be easily verified that the satellite channels support these rates for large enough \( a, k \).

A few details with respect to the above coding technique are worth mentioning. \( p_{S_j^1|S_j^2} \in \mathcal{S} \) can in principle be computed, once the fixed block-length codes, encoding and decoding maps are chosen. \( S^m \) will be binned at rate \( H(S_j^1|K_j^l) \) and the decoder can employ a joint-typicality based decoder using the computed \( p_{S_j^1|K_j^l} \).

We conclude the following.

**Theorem 1:** The LC conditions stated in [5, Thm. 1] are not necessary. Refer to Ex. 1. There exists \( a^* \in \mathbb{N} \) and \( k^* \in \mathbb{N} \) such that for any \( a \geq a^* \) and any \( k \geq k^* \), \( S_1, S_2 \) and the IC \( \mathcal{W}_{Y_j|X_j} \) do not satisfy LC conditions, and yet, \( S \) is transmissible over IC \( \mathcal{W}_{Y_j|U_j} \).

**Remark 3:** The above scheme crucially relies on the choice of \( l \) - neither too big, nor too small. This is elegantly captured as follows. As \( l \) increases, \( \xi(l) \to 1 \), \( \tau(l) \), and \( (g_{p,l}) \to 0 \). As \( l \) decreases, \( \xi(l) \to \xi(\infty) \), \( \tau(l) \), and \( (g_{p,l}) \to 1 \). If \( \phi \to 0.5 \), \( L^0(\phi, |S_j|) \to 0.5 \log |S_j| = \frac{1}{2} \log a \).

### B. Fixed B-L coding for the MAC problem

**Example 2:** Consider the source of Ex. 1 \( (S, \mathcal{W}_S) \) with \( \eta = 10 \). The source is depicted in Fig. 5.

The MAC is depicted in Fig. 6 and described below. The input alphabets are \( U \times X_1 \) and \( U \times X_2 \). The output alphabet is \( Y_0 \times Y_1 \times Y_2 \). \( U = Y_0 = \{0, 1, \cdots, a - 1\} \). \((U_j, X_j) \in U \times X_j \) denotes encoder \( j \)'s input. Moreover, \( \mathcal{W}_{Y_j|U_j} = \mathcal{W}_{Y_0|U_0} \mathcal{W}_{Y_1|X_1} \mathcal{W}_{Y_2|X_2} \), where
\[
\begin{align*}
\mathcal{W}_{Y_0|U_0} &= \{1 \mid y_0 = u_1 = u_2 \} \\
\mathcal{W}_{Y_1|U_1} &= \{1 \mid u_1 \neq u_2, y_0 = 0, 0 \text{ otherwise} \}
\end{align*}
\]

0 otherwise. The capacities of the PTPs \( \mathcal{W}_{Y_j|X_j} \) : \( j = 1, 2 \) are \( C + h_2(\frac{2}{k a^2}) + \frac{1}{2} \log a \) and \( C + h_2(\frac{2}{k a^2}) \) respectively, where \( C = \log a + \log h_2(\frac{2}{k a^2}) \).

**Theorem 2:** Refer to Ex. 1. There exists \( a^* \in \mathbb{N} \) and \( k^* \in \mathbb{N} \) such that for any \( a \geq a^* \) and any \( k \geq k^* \), 1) \( S \) and MAC \( \mathcal{W}_{U_j|X_j} \) do not satisfy CES conditions [2, Thm. 1], and yet, 2) \( S \) is transmissible over MAC \( \mathcal{W}_{U_j|X_j} \).

**Proof:** To establish the first assertion, it suffices to prove
\[
H(S) > I(XU; Y|Q)
\]

13Use \( H(S_1) \leq \log a + h_2(\frac{2}{k}) \) and \( |T_0(S_1)| \leq \exp((1 + \delta)H(S_1)) \).

14\( g_{p,l} \) is defined in the sequel.

Note that encoder 2 also employs source code \( T^0_1(S_1) \), (not \( T^0_2(S_2) \)).
$C_U = \mathcal{U}$ be the channel code employed at both encoders. $lA = \lfloor \log a \rfloor$ bits, of the $\lfloor \log |T^T_\delta(S_1)| \rfloor$ bits output by the source code is mapped to $C_U$. Both encoders employ the same map.\textsuperscript{10}

The rest of the $IB = \lfloor \log |T^T_\delta(S_1)| \rfloor - lA$ bits is communicated by Encoder 1\textsuperscript{17} on the satellite channel using an arbitrary large B-L code (of B-L lm). After observing the $\mathcal{W}_{Y_0|U_i}$ channel output for an arbitrarily large number $m$ of sub-blocks and having decoded the rest $lB$ bits for each sub-block, how much more information does the decoder need?

We employ a matrix notation in the sequel. View the $m$ sub-blocks of the source $S_j$ as the rows of the matrix $S_j(1:m,1:l) \in S_j^{m \times l}$. Let $\hat{K}(t, 1:m, 1:l) \in S_1^{m \times l}$ denote decoder’s reconstruction. The $m$ sub-blocks

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

are iid with an $l$-length distribution $\mathcal{W}_{S_1|S_2 |P_{\hat{K}^{1}|S_1} S_2}$.\textsuperscript{18} This suggests that we can treat the $l$-length sub-blocks as supersymbols and employ a standard binning technique. It suffices for encoder $j : j \in [2]$ to send $H(S_j^l |K^l)$ bits per source sub-block. We do not have a characterization of $p_{\hat{K}^{1}|S_1} S_2$ and we therefore derive an upper bound. We have

$$H(S_j^l |\hat{K}^l) \leq H(S_j^l | 1_{K_j \neq S_1^l} | \hat{K}^l) \leq h_b(P(\hat{K}^l \neq S_1^l)) +$$

$$+ P(\hat{K}^l \neq S_1^l) \log |S_1^l| + P(\hat{K}^l = S_1^l) H(S_j^l | S_1^l)$$

$$\leq l\mathcal{L}_S(\phi, |\hat{K}^l \neq S_1^l), |S_1^l|) + lH(\hat{S}_j^l | S_1^l).$$

where the above RHS is the capacity of $\mathcal{W}_{Y_j|X_j}$. Since $\mathcal{L}_S(\phi, |\hat{K}^l \neq S_1^l) \neq 0$-non-decreasing in $\phi$ if $\phi \leq \frac{1}{2}$, we bound $P(\hat{K}^l \neq S_1^l)$ by a quantity that is less than $\frac{1}{2}$. Towards that end, note that $\{S_1^l \neq \hat{K}^l \} \subseteq \{S_1^l \neq S_2^l \} \cup \{S_1^l \neq T^T_\delta(S_1) \}$. Indeed, $S_1^l = S_2^l \in T^T_\delta(S_1)$ implies both encoders input same $C_U$-codeword and agree on the $l$ bits communicated by encoder 1. Therefore $P(\hat{K}^l \neq S_1^l) \leq \phi$, where $\eta^l \phi = \xi^l + \eta^l \delta$, $\tau^l \delta \leq 2a^k \exp\left(-\frac{\delta^2}{2a^k} \right)$ and $\xi^l \leq \frac{1}{a^k 2^{a^k}}$.\textsuperscript{15}

Choose $l = k^4 a^5 k^2$, $\delta = \frac{1}{2}$, substitute in (15) and verify $\phi \leq \frac{2k^3}{a^k 2^{a^k}} < \frac{1}{2}$ for sufficiently large $a, k$. Verify\textsuperscript{20}

$$B \leq (2/l + (1/k) \log a + (1 + (1/k)) h_b(1/k)$$

$$\leq h_b(\frac{1}{k}) - \left(1 + (1/k) h_b(1/k) \right) \geq \frac{1}{2k} \log a$$

for sufficiently large $k$, verify\textsuperscript{20}

$$B \leq h_b(\frac{1}{\alpha}) + \frac{1}{4k} \log a$$

for sufficiently large $a, k$. Substituting $\delta = \frac{1}{2}$, verify\textsuperscript{20}

$$B \leq \left(1 + (\log a) + (1 + (1/k)) h_b(1/k)$$

Since $h_b(\frac{1}{\alpha}) \geq \frac{1}{2k} \log a$ for large enough $k$, we have $B \leq h_b(\frac{1}{k}) + \frac{1}{4k} \log a$ for sufficiently large $a, k$.

\textsuperscript{16}When $S_1^l \notin T^T_\delta(S_1)$, some chosen codeword is input on $\mathcal{W}_{Y_0|U}$. We handle this case.

\textsuperscript{17}Encoder 2 does not communicate the rest of the $lB$ bits.

\textsuperscript{18}$\mathcal{W}_{S_1|S_2, \mathcal{P}_{\hat{K}^{1}|S_1} S_2}$ does not necessarily factor, owing to the $l$-length encoder and decoder mappings.

\textsuperscript{19}Recall $\xi^l = \xi^l(S_1), \xi = \xi(S_1), \tau^l \delta = \tau^l(S_1)$ in this section.

\textsuperscript{20}Use $H(S_1^l) \leq \log a + h_b(\frac{1}{k})$ and $|T_\delta(S_1)| \leq \exp(l(1 + \delta) H(S_1^l))$.

Lastly, note that $H(S_2|S_1) \leq h_b(\frac{1}{k(1 - \log a)}) + \frac{1}{4k} \log a$ for sufficiently large $a, k$. The validity of (14) for sufficiently large $a, k$, can now be verified by substituting (16) and the above derived bounds.

\textbf{Remark 4:} $\mathcal{W}_{Y_0|U_i}$ is precious and must be utilized efficiently. This requires extraction of common bits.\textsuperscript{21} $\mathcal{W}_S$ is ‘very far’ from uniform requiring significant compression. If you seek very efficient compression, by choosing $l$ too large, $\xi^l \rightarrow 1$ resulting in very low probability of co-ordination. On the other hand, setting $l = 1$, results in high co-ordination ($\xi^l = \xi$), but least compression. The optimal value for $l$ is neither too big, nor too small. This is elegantly captured as follows. If $l$ increases, $\xi^l \rightarrow 1$, $\tau^l \delta \rightarrow 0$. If $l$ decreases, $\xi^l \rightarrow \xi$, and $\tau^l \delta \rightarrow 1$. If $\phi \rightarrow 0.5$, $\mathcal{L}_S(\phi, |S_1|) \rightarrow 0.5 \log |S_1|.$

\textbf{Remark 5:} The fixed B-L codes enabled decoder reconstruct $K$ that was highly correlated to $S$. The rest of the necessary information was communicated via separation using an $l$-letter source code operating over sub-blocks (supersymbols). Yet, 1) the isolation between $\mathcal{W}_{Y_j|X_j}$ and $\mathcal{W}_{Y_0|U}$ channels and 2) the separation technique permitted us to communicate this via a S-L channel code over $\mathcal{W}_{Y_j|X_j}$.

\textbf{IV. FIXED BL CODES OVER AN ARBITRARY IC}

Our analysis (Sec. III-A) focused on proving

$$\mathcal{L}_S(\phi, |S_1|) + B + H(S_1|S_1) \leq I(X_j; Y_j)$$

where $\phi < \frac{1}{2}$ was an upper bound on $P(\hat{K}^l \neq S_1^l)$. All our sufficient conditions will take this form. The lack of isolation between channels carrying fixed B-L and infinite B-L codes will throw primarily two challenges.\textsuperscript{22} We present our generalization in three pedagogical steps.

In general, $P(S_1^l \neq K^l) \leq \tau^l \delta + \xi^l \phi + g_{\rho,l}$, where the first two terms are as in (8), and $g_{\rho,l}$ is the probability that any of the decoders incorrectly decodes the $C_U$-codeword, conditioned on both encoders choosing the same $C_U$ codeword.\textsuperscript{23} Our fixed B-L code $C_U$ will be a constant composition code, and in the statements of all theorems, $g_{\rho,l}$ is defined as

$$g_{\rho,l} = \sum_{j=1}^{2} \exp\left(-l \mathcal{E}_r(A + \rho, p_1 p_2, p_{Y_j, U} - \rho) \right).$$

In all our theorems, $\mathcal{L}_S(\phi, |S_1|)$ is defined as in (7), $g_{\rho,l}$ as above, $\phi = \tau^l \delta (K^l) + \xi^l \phi + g_{\rho,l}$ will serve as an upper bound on $P(S_1^l \neq K^l)$ that is less than $\frac{1}{2}$.

\textbf{A. Designing independent streams ignoring self-interference}

The main challenges in generalizing pertain to 1) multiplexing a fixed B-L code with an infinite B-L code through a single channel input, and 2) the effect of erroneous conditional coding on the outer code. We adapt tools developed by Shirani and Pradhan [18], [19] in the context of distributed source coding. The following very simple generalization is chosen to illustrate our ideas. In particular, we live with self-interference between the two streams.

\textsuperscript{21}$U_1 \neq U_2$ results in very low information transfer.

\textsuperscript{22}And an additional loss in the channel rate, denoted $\mathcal{L}_C (\cdot, \cdot)$.

\textsuperscript{23}For Ex.1, $g_{\rho,l} = 0$, and we ignored it. For general IC, $g_{\rho,l}$ is non-zero.
Theorem 3: \( (S, W) \) is transmissible over IC 
\( (X, Y, W) \) if there exists (i) a finite set \( \mathcal{K} \), maps \( f_j : S_j \rightarrow \mathcal{K} \), with \( K_j = f_j(S_j) \) for \( j \in [2] \), (ii) \( l \in \mathbb{N}, \delta > 0 \), (iii) finite set \( U, V_1, V_2 \) and pmf \( p_{UV_1V_2}p_{XU_1V_2|UV_1V_2} \) defined on \( U \times Y \times X \), where \( p_U \) is a type of size sequences\(^{24}\) in \( \mathcal{U} \), (iv) \( A, B \geq 0, \rho \in (0, A) \) such that \( \phi \in [0, 0.5) \),

\[
A + B \geq (1 + \delta)H(K_1), \quad \text{and for } j \in [2], \quad B + H(S_j|X_j) + L^2(\phi, |U|) < (I(V_j; Y_j) - L_j^2(\phi, |V|)),
\]

where, \( \phi := g_{\rho, l} + e^{l(0)}(K) + \tau_{1, \delta}(K_1) \), \( L_j^2(\phi, |U|) = h_b(\phi) + \phi \log |U| + \|\phi\| |U| \phi \log \frac{1}{\phi} \).

Remark 6: The characterization provided in Thm. 7 (and those in Thms. 4, 5) is via S-L PMFs and S-L expressions.

Remark 7: In the above, the fixed B-L code operates over \( K_j \) instead of \( S_j, L_j^2(\phi, |U|) \) quantifies the loss in rate of the \( \mathcal{U} \) due to erroneous channel coding. Note that, in Ex. 1, satellite channel remained unaffected when the encoders placed different \( C \) codewords. The latter events imply, the \( V_j - Y_j \) channel is not \( p_{UV_j|Y_j} \), \( L_j^2(\phi, |U|) \) is a bound on the difference in the mutual information between \( p_{UV_j|Y_j} \) and the actual channel. Note that \( L_j^2(\phi, |U|) \to 0 \) as \( \phi \to 0 \).

Proof: We elaborate on the new elements. The rest follows from standard arguments \([16]\). The source-coding module, and the mappings to the channel-coding module are identical to Section III-A. We describe the structure of \( C \) and how it is multiplexed with the outer code built on \( V_j \). If we build a single code \( C \) of B-L \( mLm \) and multiplex it with \( m \) blocks of \( C \), then \( C \) does not experience an IID memoryless channel.

Let \( U_j(t, 1 : l) \) denote encoder \( j \)'s chosen codeword from \( C \) corresponding to the \( l \)-th sub-block of \( K_j(t, 1 : l) \). We seek to identify sub-vectors of \( U_j \) that are IID, and whose pmf we know. Then we can multiplex the outer code along with these sub-vectors. Interleaving \([19]\) enables us do this.

Suppose, for \( t \in [m] \), \( A(t, 1 : l) \sim p_{UV} \) and the \( m \) vectors \( A(1, 1 : l), \ldots, A(m, 1 : l) \) are iid \( p_{A} \). Let \( \pi_i : [l] \to [l] : t \in [m] \) be \( m \) surjective maps, that are independently and uniformly chosen among the collection of surjective maps (permuters). Then, for each \( i \in [l] \), the \( m \)-length vector

\[
A(1, \pi_i(1)), A(2, \pi_i(2)), \ldots, A(m, \pi_i(m)) \sim \prod_{i=1}^{l} \left\{ \frac{1}{l} \prod_{i=1}^{l} P_{A_i} \right\}.
\]

The above fact can be therefore be stated as \( A^* \) is \( \sim \prod_{i=1}^{l} (\frac{1}{l} \sum_{i=1}^{m} P_{A_i}) \).

One can now easily prove that, if \( C \) is a constant composition code of type \( p_U \), and \( m \) codewords are independently chosen\(^{25}\) from \( C \) and placed as rows of \( U_j \), then for any \( i \in [l] \), the interleaved vector \( U^* \) is \( \sim \prod_{i=1}^{l} P_{A_i} \). We now build \( l \)-codebooks (independently drawn), one for each of these interleaved vectors.

Following is our channel code structure. \( C \) is constant composition code of type \( p_U \) and B-L \( l \). Encoder \( j \) picks \( l \) independent codes of type \( p_U \), \( j \in [2] \), each iid \( \sim \prod_{i=1}^{m} p_{V_i} \), and pmf \( p_{UV_1V_2}p_{XU_1V_2|UV_1V_2} \) defined on \( U \times Y \times X \), where \( p_U \) is a type of sequences\(^{24}\) in \( \mathcal{U} \). (iv) \( A, B \geq 0, \rho \in (0, A) \) such that \( \phi \in [0, 0.5) \),

\[
A + B \geq (1 + \delta)H(K_1), \quad \text{and for } j \in [2],
\]

\[
B = \mathcal{L}_j(\phi, |U|) = h_b(\phi) + 5\phi \log |U| + \|\phi\| |U| \phi \log \frac{1}{\phi},
\]

where, \( \phi \), \( L_j^2(\phi, |U|) \) are as defined in Thm 7.

\[\text{[26]}\]

\[\text{We are unaware of the transition probabilities of this IID PTP.}\]
Remark 8: For simplicity and compact description, we derive a uniform upper bound on all the mutual-information quantities involved in the description of the Han-Kobayashi region. This explains the large constant multiplying $\phi \log \frac{1}{\phi}$. Our third step is to use the decoded fixed B-L channel codewords towards conditional decoding of the outer code. The outer code is built on $X_j$ and is superimposed on (interleaved vectors of) $C_V$. The challenge is that a fraction $\phi$ of the decoded codewords are erroneous. The approach is to treat the interleaved columns of the decoded $U$ as a noisy state-side information. Interleaving ensures that these sub-vectors have a S-L IID pmf. Proof is similar to [15, Proof of Thm. 1].

Theorem 5: $(\mathcal{S}, \mathcal{W}_S)$ is transmissible over an IC $\mathbb{W}_{Y|X}$ if there exists (i) a finite set $K$, maps $f_j : S_j \rightarrow K$, with $K_j = f_j(S_j)$ for $j \in [2]$, (ii) $l \in \mathbb{N}, \delta > 0$, (iii) finite set $\mathcal{U}$ and pmf $p_{V_U | U} p_X | V_U = U$ defined on $U \times X$, where $p_{V_U}$ is a type of sequences in $U^l$, (iv) $A, B \geq 0$, $\rho \in (0, A)$ such that $\phi \in [0, 0.5)$, where

$$A + B \geq (1 + \delta)H(K_1), \text{ and for } j \in [2],$$

$$B + H(S_j | K_1) + \mathcal{L}^S(\phi, |K|) = I(V_j; Y_j | U) - \mathcal{L}^C_2(\phi, |U|),$$

$$\mathcal{L}^C_2(\phi, |U|) = h_{\phi}(\phi) + \phi \log |U| + |X_j| |U| (1 + |X_j|) \phi \log \frac{1}{\phi}$$

and $\phi$, $\mathcal{L}^S(\phi, |K|)$ are as defined in Thm 7.

The final step in our generalization will combine the techniques of Thms. 4, 5. In particular, we employ Han-Kobayashi technique in the superposition layer. The message to be communicated through the outer code is split into private and public parts and coded using separate codebooks. Decoder $j$ uses the decoded fixed B-L channel codeword and employs a conditional Han-Kobayashi decoder. We omit a characterization in the interest of brevity.

V. FIXED BL CODES OVER AN ARBITRARY MAC

We present our generalization for the MAC problem in two pedagogical steps. In the first step (Sec. V-A), we decode the fixed B-L code and the infinite B-L codes using independent decoders. Since each stream of information cases interference to the other, this is sub-optimal. We enhance this technique via joint decoding in our second step (V-B). In the interest of brevity, we omit a proof of the theorems.

Throughout this section, we let

$$g_{\rho, l} := \exp\{-l(E_r(A, \rho, p_U, p_Y | U) - \rho)\},$$

$\mathcal{L}^S(\phi, |S_j|)$ be defined as in (7), and $\phi = \tau_{l, \delta}(K_1) + \xi[|K|]$ and $g_{\rho, l} < \frac{1}{2}$ serve as an upper bound on $P(S_i^l \neq K_j^l)$.

A. Designing independent streams of information ignoring self-interference

Theorem 6: $(\mathcal{S}, \mathcal{W}_S)$ is transmissible over MAC $(\mathcal{X}, \mathcal{Y}, \mathbb{W}_{Y|X})$ if there exists (i) a finite set $K$, maps $f_j : S_j \rightarrow K$, with $K_j = f_j(S_j)$ for $j \in [2]$, (ii) $l \in \mathbb{N}, \delta > 0$, (iii) finite set $\mathcal{U}$, $V_j$, and pmf $p_{V_U} p_{V_j} p_{X_j} | V_U = U p_{X_j} | V_U = U$ defined on $U \times X \times X_j$, where $p_{V_U}$ is a type of sequences in $U^l$, (iv) $A, B \geq 0$, $\rho \in (0, A)$ such that $\phi \in [0, 0.5)$,

$$A + B \geq (1 + \delta)H(K_1), \text{ and for } j \in [2],$$

$$H(S_j | K_1, S_j) + \mathcal{L}^S(\phi, |S_j|) < I(V_j; Y_j | V_j) - \mathcal{L}^C_2(\phi, |V|),$$

$$B + H(S_j | K_1) + \mathcal{L}^C_2(\phi, |S_j|) < I(V_j; Y_j) - \mathcal{L}^C_2(\phi, |V|),$$

where,

$$\phi := g_{\rho, l} + \xi[|K|] + \tau_l, \delta(\phi), \mathcal{L}^C_2(\phi, |U|) = h_{\phi}(\phi) + \phi \log |U| + |Y| |U| \phi \log \frac{1}{\phi}.$$


