Communicating Correlated Sources over an Interference Channel

Arun Padakandla

Abstract—A new coding technique, based on fixed block-length codes, is proposed for the problem of communicating a pair of correlated sources over a 2-user interference channel. Its performance is analyzed to derive a new set of sufficient conditions. The latter is proven to be strictly less binding than the current known best, which is due to Liu and Chen [1]. Our findings are inspired by Dueck’s example [2].

I. INTRODUCTION

Network information theory has provided us with elegant techniques to exploit correlation amongst distributed information sources. Such a correlation is handled at two levels. Probabilistic (soft) correlation is exploited via binning or transferring them via test channels [3]. When the sources possess common bits - Gács-Körner-Witsenhausen (GKW) common part-, conditional coding provides enhanced benefits. In this article, we propose a new coding technique to exploit the presence of near GKW parts, amongst distributed sources.

Our primary focus is the scenario depicted in Fig. 1. A pair $S_1, S_2$ of correlated sources have to be communicated over a 2-user interference channel (IC). Receiver (Rx) $j$ wishes to reconstruct $S_j$ losslessly. We undertake a Shannon-theoretic study and restrict attention to characterizing sufficient conditions under which $S_j$ can be reconstructed at Rx $j$.

The current known best set of sufficient conditions (LC conditions) for this problem is due to Liu and Chen [1] Thm. 1 and are proven to be optimal for a class of deterministic ICs [1] Thm. 2. In this article, we propose a new coding technique based on fixed block-length (B-L) codes and derive a new set of sufficient conditions. Through an example (Ex. 1), we prove (Lem. 2) the latter conditions are strictly less binding.

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., $K_j = f_j(S_j) : j \in [2]$ such that $\xi = P(K_1 \neq K_2)$ is ‘quite’ small, (relatively) large sub-blocks of length $l$ can 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 these $l$-length sub-blocks. Indeed, we prove in Section III-A that the latter technique outperforms Liu and Chen’s coding technique (LC technique). In Section IV, we build on this idea to propose a general coding technique for an arbitrary problem instance.

Joint source-channel coding over multi-user channels has received considerable attention with regard to characterizing fundamental limits [4]–[6] and designing feasible strategies [7]. Fundamental performance limits for communicating Gaussian sources over Gaussian channels have been studied in [8], [9] [10] and the latter considers communication over 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. 3) of a new inner bound that strictly enlarges the current known best (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 [11]–[13] on joint source-channel coding, and is inspired by Dueck’s novel example [2] and his very specific, yet ingenious, fixed B-L coding. Here, we restrict attention to separation based schemes 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 [14] involves additional challenges, and is dealt in a concurrent submission [12] Sec. V.

II. PRELIMINARIES: NOTATION, PROBLEM STATEMENT

We let an underline denote an appropriate aggregation of related objects. For example, $S$ will be used to represent a pair $S_1, S_2$ of RVs. $\underline{S}$ will be used to denote either the

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

1As was done in [1].
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 $\overline{j}$ will denote the complement index, i.e., $\{j, \overline{j}\} = \{1, 2\}$. For $m \in \mathbb{N}$, $\{m\} := \{1, \ldots, m\}$. For a pmf $p_{U}$ on $U$, $b^* \in U$ will denote a symbol with the least positive probability wrt $p_{U}$. Boldfaced letters such as $A$ denote matrices. For a $m \times l$ 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 wph, S-L, LMC, B-L respectively.

For a point-to-point channel (PTP) $(U, Y, \mathbb{W}_{Y|U})$, let $E_{r}(R, p_{U}, \mathbb{W}_{Y|U})$ denote the random coding exponent for constant composition codes of type $p_{U}$ and rate $R$ [13] Thm 10.2]. Specifically, $E_{r}(R, p_{U}, \mathbb{W}_{Y|U})$ is defined as

$$
\min_{\mathbb{W}_{Y|U}} \left\{ D(V_{Y|U}||\mathbb{W}_{Y|U}|p_{U}) + |I(p_{U}; V_{Y|U}) - R|^{+} \right\}.
$$

For RVs $A_1, A_2$, we let $\xi(\xi[A]) : = P(A_1 \neq A_2)$, and $\xi(A) : = \xi(\xi[A])$. Throughout Sec. [11] $\xi(\xi[A])$ and $\xi(A)$, if $A$ is i.i.d., we note $\xi(\xi[A]) = 1 - (1 - \xi)^{l} \leq \xi$. We let $\eta_\delta(K) = 2|K| \exp(-2\delta^2 p_{U}(a))$ denote an upper bound on $P(K \notin T_n^\delta(K))$ where $T_n^\delta(K)$ denotes our typical set.

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 $\mathbb{W}_{Y_1Y_2|X_1X_2}$. Let $\mathcal{S}_j(\mathcal{S}_2) := (S_1, S_2)$, taking values over $\mathcal{S}_1 \times \mathcal{S}_2$ with pmf $\mathbb{W}_{S_1S_2}$, denote a pair of information sources. For $j \in \{2\}$, encoder $j$ observes $S_j$ and decoder $j$ aims to reconstruct $S_j$ with arbitrarily small probability of error (Fig. 1). If this is possible, we say $\mathcal{S}$ is transmissible over IC $\mathbb{W}_{Y_j|X_j}$. In this article, our objective is to characterize sufficient conditions under which $(\mathcal{S}, \mathcal{S}_2)$ is transmissible over IC $\mathbb{W}_{Y_j|X_j}$.

III. FIXED B-L CODING OVER ISOLATED CHANNELS

We consider a simple generalization (Ex. 1) of Dueck’s example [2] and propose a coding technique that enables transmissibility of the sources over the corresponding IC. We also prove all current known joint source-channel coding techniques, and in particular, LC is incapable of the same. On the one hand, this proves strict sub-optimality of the latter in the other hand, highlights the need for fixed BL coding.

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

$$
\mathbb{W}_{S_1S_2}(c^k, d^k) = \begin{cases} 
\frac{k-1}{k} & \text{if } c^k = d^k = 0^k \\
\frac{k}{k(a^k-1)} & \text{if } c^k = d^k, c^k \neq 0^k, \\
\frac{1}{k(a^k-1)} & \text{if } c^k = 0^k, d^k \neq 0^k, \text{and}
\end{cases}
$$

0 otherwise. 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. 2 depicts the source pmf with $\eta = 6$.

![Fig. 2. 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.](image)

The IC is depicted in Fig. 2 and described below. The input alphabets are $U \times X_1$ and $U \times X_2$. The output alphabets are $Y_0 \times Y_1 \times Y_2$, $U = Y_0 = \{0, 1, \ldots, a - 1\}$. $(U_j, X_j) \in U \times X_j$ denotes encoder $j$’s input and $(Y_0, Y_j) \in Y_0 \times Y_j$ denotes symbols received by decoder $j$. The symbols $Y_0$ received at both decoders agree with probability 1. $\mathbb{W}_{Y_0Y_1Y_2|X_1X_2U_1U_2} = \mathbb{W}_{Y_0|U_1U_2}(y_0|u_1u_2) \equiv 1$ if $y_0 = u_1 = u_2$ \[1\] if $u_1 \neq u_2, y_0 = 0$, and 0 otherwise. The capacities of PTP channels $\mathbb{W}_{Y_j|X_j} : j = 1, 2$ are $C := h_b(\frac{1}{2}) + \frac{1}{2} \log a$ and $C + h_b(\frac{2}{a\eta_{\text{min}}})$ 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, $\frac{\xi(\xi[A])}{k^{\gamma/2}}$ is very small. We also have $H(S_1), H(S_2), H(S)$ and $\frac{1}{k^{\gamma/2}} \sim \log a$ and naturally $H(S_j|S_2^*) : j = 2$ is very small. Each decoder benefits a lot by decoding either source, or any function thereof. Secondly, $\mathbb{W}_{S2}$ 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 $\log a + 2C + h_b(\frac{2}{a\eta_{\text{min}}})$. Since $\sup_{p_{X_1X_2}} I(X_1Y_1) = 2C + h_b(\frac{2}{a\eta_{\text{min}}})$, the $\mathbb{W}_{Y_0|U_1U_2}$-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 $\mathbb{W}_{Y_0|U_1U_2}$.

We first prove Ex. 1 does not satisfy LC conditions. The proof is based on the following argument. Suppose LC technique enables decoder $j$ reconstruct $S_j$ for $j \in \{2\}$, then both 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 [2] Sec. III.c.

In fact, $(1 - \frac{1}{a\eta_{\text{min}}}) \log a - \frac{\log 2}{a} \leq H(S_1), H(S_2), \leq \log a + h_b(\frac{1}{2}) + \frac{\log 2}{a}$.
Lemma 1: Consider Ex. 1 with any η ∈ N. There exists an \( a_\ast \in \mathbb{N}, k_\ast \in \mathbb{N} \), such that for any \( a \geq a_\ast \) and any \( k \geq k_\ast \), the sources and the IC described in Ex. 1 do not satisfy LC conditions that are stated in [1] Thm. 1.

The proof is detailed in [13].

Proof: Since the sources do not have a GKW part, it suffices to prove that Ex. 1 does not satisfy conditions stated in [1] Corollary 1. Let \( S|QWXYU \) be any collection of RVs whose pmf factorizes as \( \mathbb{P}(S|QWXYU) = \mathbb{P}(S|Q) \mathbb{P}(W|QXU) \mathbb{P}(Y|QWXU) \mathbb{P}(XU|Y) \mathbb{P}(Z|WZU) \). We prove

\[
H(S) > I(S_1X_1U_1;Y_1|Q,W) + I(W_1S_2X_2U_2;Y_2|Q,W) - I(S_1;S_2)
\]

and thereby contradicting [1] Eqn. (44), Corollary 2).Towards that end, we first note

\[
H(S) \geq H(S_2) = h\left( \frac{1}{k} \right) + \frac{1}{k} \log(a^k - 1) \geq \frac{1}{k} \log \frac{ka^k}{2}
\]

whenever \( a^k \geq 2 \). Secondly, the RHS of (4) can be bounded above by

\[
\begin{align*}
I(S_1X_1U_1;Y_1|Q,W) + I(W_1S_2X_2U_2;Y_2|Q,W) - I(S_1;S_2) & \leq \frac{I(W_1S_2X_2U_2;Y_2|Q)}{I(S_1;S_2)} - I(S_1;S_2) \\
& \leq \frac{I(W_1S_2X_2U_2;Y_2|Q)}{I(S_1;S_2)} - I(S_1;S_2)
\end{align*}
\]

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

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

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

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

\[
I(XU;Y|Q,W) \leq 2 \log 2 + 2C + h_b\left( \frac{2}{ka^k} \right) + \frac{3}{4} \log a + \frac{\log |Y|}{k} < \log a
\]

for sufficiently large \( k, a \).

Remark 1: Why is the LC technique incapable of communicating \( S^2 \)? Any valid pmf \( p_{U_1U_2} \) induced by a S-L coding scheme is constrained to the LMC \( U_1 - S_1 - S_2 \rightarrow U_2 \). For \( j \in \{2\} \), \( p_{U_1U_2} \) can equivalently be viewed as \( U_1 = g_j(S_1,W_j) \), for some function \( g_j \) and RV \( W_j \), that satisfy \( W_j \perp W_2 \). 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, \( \frac{1}{3} \). Consequently, any \( p_{U_1U_2} \) induced via a S-L coding scheme is sufficiently far from any pmf that satisfies \( U_1 = U_2 \) 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 \( U_l^1 - S_l^1 - S_l^2 \rightarrow U_l^2 \). Suppose we choose \( l \) reasonably large such that 1) \( \varepsilon(||S||) \) is not high, and 2) \( S_l^1 \) is reasonably uniform on its typical set \( T_l^1(S_l) \), and define \( U_j : j \in \{l\} \) through identical functions \( U_l^1 = g(S_l^1) \) : \( j \in \{l\} \), then one can easily visualize the existence of \( g \) such that \( p_{U_1U_2} \) satisfies the twin objectives of \( U_l^1 = U_l^2 \) whp and \( U_l^1 = S_l^1 \) 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.

A. Fixed Block-length Coding over a Noiseless Channel

In order to input codewords on the \( W_{Y_0|U} \)-channel, that agree, we employ the same source code, same channel code and same mapping, each of fixed B-L[1] \( l \), at both encoders. \( l \) is chosen large enough such that the source can be reasonably efficiently compressed, and yet small enough, to ensure \( \varepsilon(||S||) \) 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 \( W_{Y_0|X_1} \). We begin with a description of the fixed B-L codes.
We employ a simple fixed B-L (inner) code. Let \( T_{A}^{S}(S_{j}) \) be the source code, and let \( C_{U} = U^{l} \) be the channel code. Let \( lA = \lceil \log d' \rceil \) bits, of the \( \lceil \log |T_{A}^{S}(S_{j})| \rceil \) bits output by the source code, be mapped to \( C_{U} \). Both encoders use the same source code \( T_{A}^{S}(S_{j}) \) channel code and mapping.

Suppose we communicate an arbitrarily large number \( m \) of these sub-blocks on \( W_{Y_{j|x_j}} \) as above. Moreover, suppose encoder \( j \) communicates the rest of the \( IB = \lceil \log |T_{A}^{S}(S_{j})| \rceil - lA \) bits output by its source code to decoder \( j \) on its satellite channel \( W_{Y_{j|x_j}} \). How much more information needs to be communicated to decoder \( j \), to enable it reconstruct \( S_{j}^{lm} \)? 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 \( S_{j}(1 : m, 1 : l) \in S_{m \times l}^{m} \). Let \( K_{j}(1 : m, 1 : l) \in S_{m \times l}^{m} \) denote decoder \( j \)'s reconstruction of \( S_{j}(1 : m, 1 : l) \). The \( m \) sub-blocks

\[
\left\{ (S_{j}(t, 1 : l), K_{j}(t, 1 : l)) : j = 1, 2, \ldots, t \in [m] \right\}
\]

are iid \(^8\) with an \( l \)-length distribution \( W_{S_{j}^{m}|S_{j}^{m}P_{K_{j}^{m}|S_{j}^{m}P_{S_{j}^{m}}}^{m} \). 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 \( p_{K_{j}^{l}|K_{j}^{l}|S_{j}^{m}|S_{j}^{m}} \), and hence we derive an upper bound.

\[
H(S_{j}^{l}|K_{j}^{l}) \leq H(S_{j}^{l}, \mathbf{1}_{(K_{j}^{l} \neq S_{j}^{l})}) \leq h_{b}(P(K_{j}^{l} \neq S_{j}^{l})) + \log |I_{j}| + \log |I_{j}|^{l} + H(S_{j}|S_{j}). (5)
\]

\[
L_{j}^{S}(\phi, |S_{j}|) := \frac{1}{l} h_{b}(\phi) + \phi \log |K|,
\]

represents the additional source coding rate needed to compensate for the costs in the fixed B-L decoding. It suffices to prove \( L_{j}^{S}(\phi, |S_{j}|) + B + H(S_{j}|S_{j}) \leq |C_{U, j}| \) of \( W_{Y_{j|x_j}} \). Since \( L_{j}^{S}(\phi, |K|) \) is non-decreasing in \( \phi \) if \( \phi < \frac{1}{2} \), we bound \( P(K_{j}^{l} \neq S_{j}^{l}) \) by a quantity that is less than \( \frac{1}{2} \). Towards that end, note that \( \{S_{j}^{l} \neq K_{j}^{l}\} \subseteq \{S_{j}^{l} \neq S_{j}^{l}\} \cup \{S_{j}^{l} \neq T_{A}^{S}(S_{j})\} \).

Indeed, \( S_{j}^{l} \neq S_{j}^{l} \) 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 \( \phi = \tau_{l,\delta}(S_{j}) + \tau(S_{j}) \).

\[
\tau_{l,\delta}(S_{j}) \leq 2a^{2} \exp\left(-\frac{\delta^{2}l}{2k^{2}a^{2}k} \right) \quad \text{and} \quad \tau(S_{j}) \leq \frac{l}{ka^{2}k}. (7)
\]

Choose \( l = k^{4}a^{2\frac{3}{2}}, \delta = \frac{1}{2} \), substitute in (7). Since \( \eta \geq 6 \), verify \( \phi \leq 2k^{3}a^{2\frac{3}{2}} \). For sufficiently large \( a, k \).

\[
L_{j}^{S}(2k^{3}a^{-\frac{3}{2}}, |S_{j}|) \leq \frac{1}{4k} \log a \quad (8)
\]

for sufficiently large \( a, k \). Recall \( IB = \lceil \log |T_{A}^{S}(S_{j})| \rceil - lA \). Substituting \( \delta = \frac{1}{2} \), verify (5).

\[
B \leq \frac{2}{l} + \frac{1}{l} \log a + \frac{1}{l} \log a \quad (9)
\]

Since \( h_{b}(\frac{1}{l}) - \frac{1}{l} \) \( \log a \) \( \geq \frac{1}{l} \log \frac{1}{a^2} \) \( \) for large enough \( k, \) RHS of (8), (9) sum to at most \( \frac{3}{l} + \frac{1}{l} \log a \) \( \) for large enough \( a, k \). Furthermore, \( H(S_{j}^{l}|S_{j}^{l}) \leq h_{b}(\frac{2}{k^{2}a^{2}k}) + \frac{1}{l} \log a \) \( \) 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 regard to the above coding technique is worth mentioning. \( p_{K_{j}^{m}|S_{j}^{m}|S_{j}^{m}} \) can in principle be computed, once the fixed block-length codes, encoding and decoding maps are chosen. \( S_{j}^{lm} \) will be binned at rate \( H(S_{j}^{l}|K_{j}^{l}) \) and the decoder can employ a joint-type based typical decoder using the computed \( p_{S_{j}^{m}|K_{j}^{m}} \). We conclude the following.

**Theorem 1:** The LC conditions stated in \( W_{Y_{j|x_j}} \) 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 \( k \geq k^{*} \), \( S_{j}^{l} \) and the IC \( W_{Y_{j|x_j}} \) do not satisfy LC conditions, and yet, \( S_{j}^{l} \) is transmissible over IC \( W_{Y_{j|x_j}} \).

**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, \( \tau(l,\delta) \rightarrow 1 \), \( \tau(l,\delta) \) (and \( g_{p,l} \)) \( \to 0 \). As \( l \) decreases, \( \tau(l,\delta) \rightarrow \tau(S_{j}) \), \( \tau(l,\delta) \rightarrow 1 \). If \( \phi \to 0.5 \), \( L_{j}^{S}(\phi, |S_{j}|) \to 0.5 \log |S_{j}| \to \frac{k}{2} \log a \).

**IV. FIXED BL CODES OVER AN ARBITRARY IC**

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

\[
L_{j}^{S}(\phi, |S_{j}|) + B + H(S_{j}|S_{j}) \leq 1(X_{j};Y_{j}) \quad (10)
\]

where \( \phi < \frac{1}{2} \) was an upper bound on \( P(K_{j}^{l} \neq S_{j}^{l}) \). 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. We present our generalization in three pedagogical steps.

In general, \( P(S_{j}^{l} \neq K_{j}^{l}) \leq \tau_{l,\delta} + \tau(l,\delta) + g_{p,l} \), where the first two terms are as in (7), and \( g_{p,l} \) is the probability that any of the decoders incorrectly decodes the \( C_{U} \)-codeword, conditioned on both encoders choosing the same \( C_{U} \) codeword.\(^{12}\)

Our fixed B-L code \( C_{U} \) will be a constant composition code, and in the statements of all theorems, \( g_{p,l} \) is defined as

\[
g_{p,l} := \sum_{j=1}^{2} \exp\left(-l(E_{r}(A + p_{U}p_{Y_{j}|U} - \rho))\right). \quad (11)
\]

In all our theorems, \( L_{j}^{S}(\phi, |S_{j}|) \) is defined as \( g_{p,l}, g_{p,l} \) above, \( \phi = \tau_{l,\delta}(K_{j}) + \tau(l,\delta) + g_{p,l} \) will serve as an upper bound on \( P(S_{j}^{l} \neq K_{j}^{l}) \) that is less than \( \frac{1}{2} \).

\(^{13}\)Use \( H(S_{j}) \) \( \leq \log a + h_{b}(\frac{1}{l}) \) and \( |T_{A}(S_{j})| \leq \exp\{1 + \frac{1}{2}H(S_{j})\} \).

\(^{14}\)Two cases are defined in the sequel.

\(^{15}\)And an additional loss in the channel rate, denoted \( L_{C}(\gamma) \).

\(^{16}\)For Ex. 1, \( g_{p,l} = 0 \), and we ignored it. For general IC, \( g_{p,l} \) is non-zero.
A. Designing independent streams ignoring self-interference

The main challenges in generalizing 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 [16, 17] 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.

**Theorem 2:** $(\mathcal{S}, \mathcal{W}, \mathcal{G})$ is transmissible over IC $(\mathcal{X}, \mathcal{Y}, \mathcal{W}, \mathcal{X}, \mathcal{Y})$ if there exists $(i)$ a finite set $\mathcal{K}$, maps $f_j : \mathcal{S}_j \to \mathcal{K}$, with $K_j = f_j(\mathcal{S}_j)$ for $j \in [2]$, $(ii)$ $l \in \mathbb{N}, \delta > 0$, $(iii)$ finite set $\mathcal{U}$, $\mathcal{V}_1, \mathcal{V}_2$ and pmf $p_{\mathcal{U}l}\mathcal{U}_1, \mathcal{P}_l, \mathcal{P}_l, \mathcal{U}_1, \mathcal{V}_1, \mathcal{V}_2$ defined on $\mathcal{U} \times \mathcal{Y} \times \mathcal{X}$, where $p_{\mathcal{U}l}$ is a type of sequence $\mathcal{T}$ in $\mathcal{U}$, $(iv)$ $A, B \geq 0$, $\rho \in (0, A)$ such that $\phi \in [0, 0.5)$,

$$A + B \geq 1 + \delta H(K_1),$$
and for $j \in [2]$, $B + H(S_j|K_1) + L^S_j(\phi, |S_j|) < I(V_j; Y_j) - L^C_j(\phi, |V_j|)$,

where, $\phi = g_{\rho l} + \epsilon^l(K_j) + \tau_1, \delta(K_j)$, $L^S_j(\phi, |\mathcal{U}|) = h_0(\phi) + \phi log |\mathcal{U}| + |\mathcal{U}| log \phi 1/\phi$.

**Remark 4:** The characterization provided in Thm. 2 (and those in Thm. 3 [2]) is via S-L PMFs and S-L expressions.

**Remark 5:** In the above, the fixed B-L code operates over $K_j$ instead of $S_j^\phi(\phi, |\mathcal{U}|)$ quantifies the loss in rate of the outer code due to erroneous conditional coding. Note that, in Ex. 1 satellite channel remained unaffected when the encoders placed different C.U. codewords. The latter events imply, the $V_j - Y_j$ channel is not $p_{\mathcal{Y}l}|\mathcal{Y}_j$, $L^C_j(\phi, |\mathcal{U}|)$ is a bound on the difference in the mutual information between $p_{\mathcal{Y}l}|\mathcal{Y}_j$ and the actual channel. Note that $L^C_j(\phi, |\mathcal{U}|) \to 0$ as $\phi \to 0$.

**Proof:** We elaborate on the new elements. The rest follows from standard arguments [13]. The source-coding module, and the mappings to the channel-coding module are identical to Section III-A. We describe the structure of $C_{U1}$ and how it is multiplexed with the outer code built on $V_j$.

If we build a single code $C_{Vj}$ of B-L lm and multiplex it with $m$ blocks of $C_{U1}$, then $C_{Vj}$ does not experience an IID memoryless channel.

Let $U_j(t, 1 : l)$ denote encoder $i$’s chosen codeword from $C_{U1}$ corresponding to the $i$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. We can then multiplex the outer code along these sub-vectors. **Interleaving** enables us to do this.

Suppose, for $t \in [m]$, $A(t, 1 : l)$ is $p_{\mathcal{A}l}$, and the $m$ vectors $A(1, 1 : l), \cdots, A(m, 1 : l)$ are iid $p_{\mathcal{A}l}$. Let $\pi_1 : [l] \to [l] : t \in [m]$ be $m$ surjective maps, that are independent and uniformly chosen among the collection of surjective maps (permuters). Then, for each $i \in [l]$, the $m$-length vector

$$A(1, \pi_1(i)), A(2, \pi_2(i)), \cdots, A(m, \pi_m(i)) \sim \prod_{i=1}^{m} \left( \frac{1}{l} \sum_{i=1}^{l} p_{A,i} \right).$$

**[12] Appendix A** contains a proof. The following notation will ease exposition. For $A \in \mathbb{A}^{m \times l}$, and a collection $\pi : [l] \to [l] : t \in [m]$ of surjective maps, we let $A^\pi \in \mathbb{A}^{m \times l}$ be such that $A^\pi(t, i) = A(t, \pi_t(i))$ for each $(t, i) \in [m] \times [l]$.

The above fact can be therefore be stated as $A^\pi(1 : m, i) \sim \prod_{i=1}^{m} \left( \frac{1}{l} \sum_{i=1}^{l} p_{A,i} \right)$.

One can now easily prove that, if $C_{U1}$ is a constant composition code of type $p_{\mathcal{U}l}$, and $m$ codewords are independently chosen from $C_{U1}$ and placed as rows of $U_j$, then for any $i \in [l]$, the interleaved vector $U_i^\pi(1 : m, i) \sim \prod_{i=1}^{m} U_i$.

We now build codebooks (independently drawn), one for each of these interleaved vectors.

Following is our channel code structure. $C_{U1}$ is constant composition code of type $p_{\mathcal{U}l}$ and B-L l. Encoder $j$ picks $l$ independent codes $C_{Vj,i} : i \in [l]$, each iid $\prod_{i=1}^{m} p_{\mathcal{P}_i}$, each B-L $m$. $C_{Vj,i}$ is multiplexed along with sub-vector $U_j^\pi(1 : m, i)$. Outer code message is split into $l$ equal parts ($M_{j1}, \cdots, M_{jl}$). $V_j \in \mathbb{Y}^{m \times l}$ is defined as $V_j^\pi(1 : m, i) = C_{Vj,i}(M_{ij})$. For each $(i, t) \in [m] \times [l]$, $X_j(t, i)$ is chosen IID wrt $p_{\mathcal{X}j}\mathcal{X}(\mathcal{U}(t, i), \mathcal{V}(t, i))$. Symbols in $X_j \in \mathbb{X}^{m \times l}$ are input on the channel. It can be verified that (1) the codewords of $C_{U1}$ and $C_{Vj,i}$ pass through an IID memoryless channels whose transition probabilities are ‘characterized’ in the sequel.

Since each codebook $C_{Vj,i}$ and each codeword is IID, $U_j(t, 1 : l) \to Y_j(t, 1 : l)$ is IID $p_{\mathcal{V}l}^\mathcal{V}$-channel. Interleaving ensures $V_j^\pi(1 : m, i) \to Y_j^\pi(1 : m, i)$ is IID. But, unless $U_1 = U_2$, we are not guaranteed the latter channel is $\prod_{j=1}^{m} p_{Vj}|\mathcal{V}_j$.

In fact, we only know certain marginals of $V_j^\pi(1 : m, i), V_j^\pi(1 : m, i), V_j^\pi(1 : m, i)$.

Let $(V_j^\pi(1 : m, i), V_j^\pi(1 : m, i)) \sim \prod_{j=1}^{m} p_{Vj}|\mathcal{V}_j$, where $p_{Vj} = p_{Vj}$.

We wish to bound the difference $I(V_j; Y_j) - I(V_j; Y_j)$ from above. Using the relations $p_{Uj} = p_{Uj}$, $p_{Uj} = p_{Uj}v_y(u, u)$, where $p_{Uj}v_y(u, u)$ is the pmf of the interleaved vector $U_j^\pi(1 : m, i)U_j^\pi(1 : m, i)$, $\sum_{v \neq u} p_{Uj}v_y(u, u)$, we can prove $p_{Vj}Y_j(v, y) \leq \phi$. These steps are analogous to those in Appendix B. Using [15] Proof of Lemma 2.7, we conclude $I(V_j; Y_j) - I(V_j; Y_j) \leq L_C(\phi, |\mathcal{U}|)$. We refer to reader to [13] for the rest of the proof which is quite standard.

**Lemma 2:** The conditions stated in Thm. 2 are strictly weaker than those stated in Thm. 1.

**Proof:** Ex. 1 with $a, k$ chosen sufficiently large, satisfies the conditions stated in Thm. 2. In particular, choose $\delta = \frac{k}{l}, \rho = 1, A = (1 - \frac{1}{l}) log a, B = H(S_l) - A, l = kz^l, K = S_j, f_j = \text{identity}, V_j = A_j, p_{\mathcal{V}l}$ uniform, $p_{\mathcal{V}l}$ capacity achieving. The result now follows from Lemma 1.

B. Additional information via Message-Splitting

We now employ Han-Kobayashi technique to communicate the rest of the information (LHS of [10]). Towards that end, let $H \mathcal{H}(p_{V_1W_1X_1}, p_{V_2W_2X_2})$ be the Han-Kobayashi inner bound defined in [4] Proposition 3.

**Theorem 3:** $(\mathcal{S}, \mathcal{W}, \mathcal{G})$ is transmissible over IC $\mathcal{W}$ if there exists (i) a finite set $\mathcal{K}$, maps $f_j : S_j \to K_j = f_j(S_j)$ for $j \in [2]$, (ii) $l \in \mathbb{N}, \delta > 0$, (iii) finite set $\mathcal{U}$, $\forall j \neq j$, $W_j : j \in [2]$ and pmf $p_{\mathcal{U}l}\mathcal{U}_1, \mathcal{P}_l, \mathcal{P}_l, \mathcal{U}_1, \mathcal{V}_1, \mathcal{V}_2$ defined

18Not necessarily uniformly. In fact the index output from the source code, owing to the fixed block-length $l$ is not necessarily uniform

19We are unaware of the transition probabilities of this IID PTP.
on $\mathcal{U} \times \mathcal{Y} \times \mathcal{W} \times \mathcal{X}$, where $p_U$ is a type of sequences in $\mathcal{U}$, (iv) $A, B \geq 0, \rho \in (0, A)$ such that $\phi \in [0, 0.5)$.

Remark 6: 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 $\log \frac{1}{\phi}$. 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) $C_U$. The challenge is that a fraction $\phi$ of the decoded codewords are erroneous. The approach is to treat the interleaved columns of the decoded $\bar{U}$ as a noisy state/side information. Interleaving ensures that these sub-vectors have a S-L IID pmf. Proof is similar to [12] Proof of Thm. 1.

Theorem 4: $(S, W_S)$ is transmissible over an IC $\mathcal{W}_{Y|X}$ 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 $p_U p_{X_j|U} p_{X_j|U}$ defined on $\mathcal{U} \times \mathcal{X}$, where $p_U$ is a type of sequences in $\mathcal{U}$, (iv) $A, B \geq 0, \rho \in (0, A)$ such that $\phi \in [0, 0.5)$, where

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

$$B + H(S_j|K_1) + {\mathcal{L}^C}(\phi, |K|) \leq I(X_j; Y_j|U) - {\mathcal{L}^C}(\phi, |U|)$$

and $\phi, \mathcal{L}^C(\phi, |K|)$ are as defined in Thm 2.

The final step in our generalization will combine the techniques of Thm. 3-4. 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.

REFERENCES

[1] W. Liu and B. Chen, “Interference Channels with Arbitrarily Correlated Sources,” Trans. on Info. Th., vol. 57, no. 12, pp. 8027–8037, Dec 2011.

[2] G. Dueck, “A note on the multiple access channel with correlated sources (corresp.),” Trans. on Info. Th., vol. 27, no. 2, pp. 232 – 235, mar 1981.

[3] T. Cover, A. Gamal, and M. Salehi, “Multiple access channels with arbitrarily correlated sources,” Trans. on Info. Th., pp. 648 – 657, Nov 1980.

[4] K. Khezeli and J. Chen, “A Source-channel Separation theorem with application to the source broadcast problem,” IEEE Trans. Info. Th., vol. 62, no. 4, pp. 1764–1781, 2016.

[5] Y. Murin, R. Dabora, and D. Gunduz, “On Joint Source-Channel Coding for Correlated Sources over Multiple-Access Relay Channels,” Trans. on Info. Th., vol. 60, no. 10, pp. 6231–6253, Oct 2014.

[6] A. Lapidoth and M. Wigger, “A necessary condition for the transmissibility of correlated sources over a MAC,” in 2016 ISIT, July 2016.

[7] P. Minero, S. H. Lim, and Y.-H. Kim, “A unified approach to Hybrid Coding,” Trans. on Info. Th., vol. 61, no. 4, pp. 1509–1523, April 2015.

[8] A. Lapidoth and S. Tinguely, “Sending a Bivariate Gaussian over a Gaussian MAC,” Trans. on Info. Th., vol. 56, pp. 2714–2752, 2010.