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) 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

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

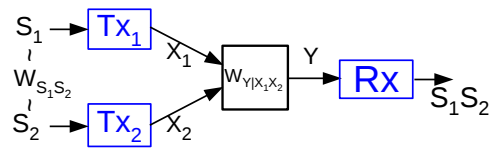

Fig. 1. Transmission of correlated sources over MAC.

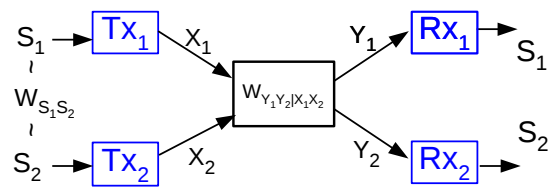

Fig. 2. Transmission of correlated sources over 2-IC.

(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., $K_j = f_j(S_j) : j \in [2]$

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 these $l-$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] [13] 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 <u>underline</u> denote an appropriate aggregation of related objects. For example, $\underline{S}$ will be used to represent a pair S_1, S_2 of RVs. $\underline{\mathcal{S}}$ will be used to denote either the pair $\mathcal{S}_1, \mathcal{S}_2$ or the Cartesian product $\mathcal{S}_1 \times \mathcal{S}_2$, and will be clear from context. When $j \in \{1, 2\}$, then $\not{j}$ will denote the complement index, i.e., $\{j, \not{j}\} = \{1, 2\}$. For $m \in \mathbb{N}$, $[m] := \{1, \cdots, m\}$.

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

is our typical set. For a pmf p_U on $\mathcal{U}$, $b^* \in \mathcal{U}$ 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) $\mathbf{A}(t, i)$ denotes the entry in row t, column i, (ii) $\mathbf{A}(1 : m, i)$ denotes the ith column, $\mathbf{A}(t, 1 : l)$ 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) $(\mathcal{U}, \mathcal{Y}, \mathbb{W}_{Y|U})$, 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 [17, Thm 10.2]. Specifically, $E_r(R, p_U, \mathbb{W}_{Y|U})$ is defined as

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

For RVs A_1, A_2, we let $\xi^{[l]}(\underline{A}) := P(A_1^l \neq A_2^l)$, and $\xi(\underline{A}) := \xi^{[1]}(\underline{A})$. Throughout Sec. III, $\xi^{[l]} = \xi^{[l]}(\underline{S})$ and $\xi = \xi(\underline{S})$. If $\underline{A}$ is IID, we note[3] $\xi^{[l]} = 1 - (1-\xi)^l \leq l\xi$. We let $\tau_{l,\delta}(K) = 2|\mathcal{K}| \exp\{-2\delta^2 p_K^2(a^*) l\}$ denote an upper bound on $P(K^l \notin T_\delta^l(K))$ where $T_\delta^l(K)$ denotes our typical set.

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

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 $\mathbb{W}_{Y_1Y_2|X_1X_2}$ (Fig. 2). Let $\underline{S} := (S_1, S_2)$, taking values over $\underline{\mathcal{S}} := \mathcal{S}_1 \times \mathcal{S}_2$ with pmf $\mathbb{W}_{S_1S_2}$, denote a pair of information sources. For $j \in [2]$, Tx j observes S_j, and Rx j aims to reconstruct S_j with arbitrarily small probability of error. If this is possible, we say $\underline{S}$ *is transmissible over IC* $\mathbb{W}_{\underline{Y}|\underline{X}}$. With regard to the IC problem, our objective is to characterize sufficient conditions under which $(\underline{\mathcal{S}}, \mathbb{W}_{\underline{S}})$ is transmissible over IC $\mathbb{W}_{\underline{Y}|\underline{X}}$.

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[4], and on the other, highlights the need for fixed B-L coding.

A. Fixed B-L coding for IC problem

Example 1: Source alphabets $\mathcal{S}_1 = \mathcal{S}_2 = \{0, 1, \cdots, a-1\}^k$. Let $\eta \geq 8$ be a positive even integer. The source PMF is

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

[1] As was done in [14].

[2] The underlying pmf p_U will be clear from context.

[3] $(1-x)^l \geq 1 - xl$ for $x \in [0, 1]$.

[4] 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].

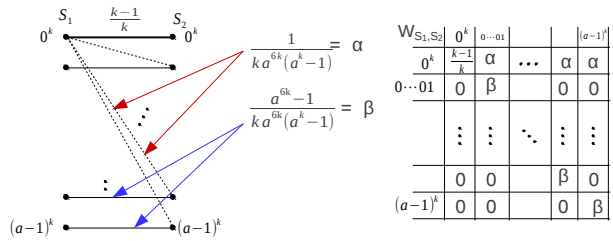

Fig. 3. On the left, the source pmf is depicted through a bipartite graph. Larger probabilities are depicted through edges with thicker lines. On the right, we depict the probability matrix.

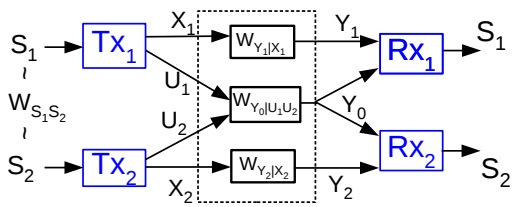

Fig. 4. Source channel setup of Example 1.

0 otherwise. Note that in the above eqn. $c^k, d^k \in \mathcal{S}_1$ 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 IC is depicted in Fig. 6 and described below. The input alphabets are $\mathcal{U}\times\mathcal{X}_1$ and $\mathcal{U}\times\mathcal{X}_2$. The output alphabets are $\mathcal{Y}_0\times\mathcal{Y}_1$ and $\mathcal{Y}_0\times\mathcal{Y}_2$. $\mathcal{U} = \mathcal{Y}_0 = \{0, 1, \cdots, a-1\}$. $(U_j, X_j) \in \mathcal{U}\times\mathcal{X}_j$ denotes encoder j's input and $(Y_0, Y_j) \in \mathcal{Y}_0\times\mathcal{Y}_j$ denotes symbols received by decoder j. The symbols Y_0 received at both decoders agree with probability 1. $\mathbb{W}_{Y_0Y_1Y_2|X_1U_1X_2U_2} = \mathbb{W}_{Y_1|X_1}\mathbb{W}_{Y_2|X_2}\mathbb{W}_{Y_0|U_1U_2}$, where

$$\mathbb{W}_{Y_0|U_1U_2}(y_0|u_1,u_2) = \begin{cases} 1 & \text{if } y_0 = u_1 = u_2 \\ 1 & \text{if } u_1 \neq u_2, y_0 = 0, \text{ and} \end{cases}$$

0 otherwise. The capacities of PTP channels $\mathbb{W}_{Y_j|X_j} : j = 1, 2$ are $\mathcal{C} := h_b(\frac{2}{k}) + \frac{2}{k}\log a$ and $\mathcal{C} + h_b(\frac{2}{ka^{\eta k}})$ respectively.

We identify key aspects of Ex. 1. Let a, k be chosen sufficiently/quiet large. While $\underline{S}$ does *not* possess a GKW part, S_1 and S_2 agree on most, but not all, realizations. Indeed, $\xi(\underline{S}) = \frac{1}{ka^{\eta k}}$ is very small. We also have $H(S_1), H(S_2), H(\underline{S})$ are[5] $\sim \log a$ and naturally $H(S_j|S_{\not j}) : j \in [2]$ is very small. Each decoder benefits a lot by decoding *either* source, or any function thereof. Secondly, $\mathbb{W}_{\underline{S}}$ 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 $\log a + 2\mathcal{C} + h_b(\frac{2}{ka^{\eta k}})$. Since $\sup_{p_{X_1X_2}} I(\underline{X};\underline{Y}) = 2\mathcal{C} + h_b(\frac{2}{ka^{\eta k}})$, the $\mathbb{W}_{Y_0|\underline{U}}$−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 $\mathbb{W}_{Y_0|\underline{U}}$.

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 [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 $\underline{SQWXUY}$ be *any* collection of RVs whose pmf factorizes as $\mathbb{W}_{\underline{S}}p_Q p_{W_1|Q}p_{W_2|Q}p_{X_1U_1|S_1W_1Q}p_{X_2U_2|S_2W_2Q}\mathbb{W}_{\underline{Y}|\underline{XU}}$. We prove

$$H(\underline{S}) > I(S_1X_1U_1;Y_1Y_0|Q\underline{W}) + I(W_1S_2X_2U_2;Y_2Y_0|Q) - I(S_1;S_2) \quad (1)$$

and thereby contradicting [5, Eqn. (44), Corollary 2]. Towards that end, we first note

$$\begin{aligned} H(\underline{S}) &\geq H(S_2) = h_b(\frac{1}{k}) + \frac{1}{k}\log(a^k-1) \geq \frac{1}{k}\log\frac{ka^k}{2} \\ &\geq \log a + \frac{1}{k}\log(\frac{k}{2}) \geq \log a, \end{aligned} \quad (2)$$

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

$$\begin{aligned} &I(S_1X_1U_1;Y_1Y_0|Q\underline{W}) + I(W_1S_2X_2U_2;Y_2Y_0|Q) - I(S_1;S_2) \\ &\leq I(S_1X_1U_1;Y_1Y_0|Q\underline{W}) + I(\underline{W}S_2X_2U_2;\underline{Y}|Q) - I(S_1;S_2) \\ &\leq I(\underline{WSXU};\underline{Y}|Q) + I(S_1X_1U_1;\underline{Y}|Q\underline{W}) \\ &\quad -I(S_1X_1U_1;\underline{Y}|Q\underline{W}S_2X_2U_2) - I(S_1;S_2) \\ &= I(\underline{XU};\underline{Y}|Q) + I(S_1X_1U_1;\underline{Y}|Q\underline{W}) \\ &\quad -I(S_1X_1U_1;\underline{Y}X_2U_2|Q\underline{W}S_2) - I(S_1;S_2) \quad (3) \\ &= I(\underline{XU};\underline{Y}|Q) + I(S_1X_1U_1;\underline{Y}|Q\underline{W}) \\ &\quad -I(S_1X_1U_1;\underline{Y}X_2U_2S_2|Q\underline{W}) \\ &\leq I(\underline{XU};\underline{Y}|Q) \end{aligned}$$

where (3) follows from Markov chains $Q\underline{SW} - \underline{XU} - \underline{Y}$ and $S_1X_1U_1 - S_2Q\underline{W} - X_2U_2$ and (4) follows from the independence of $\underline{S}$ and $\underline{W}Q$.

We now argue the RHS above is strictly lesser than $\log a$. Our argument pretty much follows Dueck [1, Sec. III.c] verbatim and is provided here for the sake of completeness. Let $R = \mathbb{1}_{\{(S_1,S_2)=(0^k,0^k)\}}$.

$$\begin{aligned} &I(\underline{XU};\underline{Y}|Q) \leq I(\underline{XU}R;\underline{Y}|Q) \leq \log 2 + I(\underline{XU};\underline{Y}|Q,R) \\ &\leq \log 2 + \frac{\log|\mathcal{Y}_0\times\mathcal{Y}_1\times\mathcal{Y}_2|}{k} + (1-\frac{1}{k})I(\underline{XU};\underline{Y}|Q,R=1) \end{aligned}$$

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 $X_1U_1 - S_1Q - S_2Q - X_2U_2$. Given $Q = q, R = 1$, (X_1, U_1) is independent of (X_2, U_2) and hence

$$\begin{aligned} I(\underline{XU};\underline{Y}|Q,R=1) &\leq \max_{p_{X_1U_1}p_{X_2U_2}} I(\underline{XU};\underline{Y}|Q,R=1) \\ &\leq 2\mathcal{C} + h_b(\frac{2}{ka^{\eta k}}) + \max_{p_{U_1}p_{U_2}\mathbb{W}_{Y_0|\underline{U}}} H(Y_0) \end{aligned} \quad (4)$$

[5] In fact, $(1-\frac{1}{a^{\eta k}})\log a - \frac{\log 2}{k} \leq \begin{smallmatrix} H(S_1), H(S_2), \\ H(\underline{S}) \end{smallmatrix} \leq \log a + h_b(\frac{1}{k}) + \frac{\log 2}{k}$

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 $u \in \mathcal{U}$, $P(U_2 \neq U_1) = \sum_u \sum_{z \neq u} P(U_2 = u) P(U_1 = z) \geq \frac{1}{2} \sum_u P(U_2 = u) = \frac{1}{2}$, 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 through (4) and above, we conclude

$$I(\underline{X}\underline{U}; \underline{Y}|Q) \leq 2 \log 2 + 2\mathcal{C} + h_b(\frac{2}{ka^{\eta k}}) + \frac{3}{4} \log a + \frac{\log |\underline{\mathcal{Y}}|}{k} < \log a$$

for sufficiently large k, a. ■

Remark 1: Why is the LC technique incapable of communicating $\underline{S}$? 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 $j \in [2]$, $p_{U_j|S_j}$ can equivalently be viewed as $U_j = g_j(S_j, W_j)$, for some function g_j and RV W_j, that satisfy $W_1 \perp\!\!\!\perp W_2$. 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 $U_1 = U_2$ whp *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 $U_1^l - S_1^l - S_2^l - U_2^l$.[6] Suppose we choose l *reasonably* large such that 1) $\xi^{[l]}(\underline{S})$ is not high, and 2) S_j^l is *reasonably* uniform on its typical set $T_\delta^l(S_j)$, and define $U_j : j \in [2]$ through *identical* functions $U_j^l = g(S_j^l) : j \in [2]$, then one can easily visualize the existence of g such that $p_{U_1^l U_2^l}$ satisfies the twin objectives of $U_1^l = U_2^l$ whp and $U_1^l = U_2^l$ is 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 $\mathbb{W}_{Y_0|\underline{U}}-$channel, that agree, we employ the same source code, same channel code and same mapping, each of *fixed B-L*[7] l, at both encoders. l is chosen large enough such that the source can be reasonably efficiently compressed, and yet small enough, to ensure $\xi^{[l]}(\underline{S})$ 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 $\mathbb{W}_{Y_j|X_j}$. We begin with a description of the fixed B-L codes.

[6]Constraint will *not* be via a S-L LMC.

[7]irrespective of the desired prob. of error

We employ a simple fixed B-L (inner) code. Let $T_\delta^l(S_1)$ be the source code, and let $C_U = \mathcal{U}^l$ be the channel code. Let $lA = \lfloor \log a^l \rfloor$ bits, of the $\lceil \log |T_\delta^l(S_1)| \rceil$ bits output by the source code, be mapped to C_U. Both encoders use the same source code[8], channel code and mapping.

Suppose we communicate an *arbitrarily large* number m of these sub-blocks on $\mathbb{W}_{Y_0|\underline{U}}$ as above. Moreover, suppose encoder j communicates the rest of the $lB = \lceil \log |T_\delta^l(S_1)| \rceil - lA$ bits output by its source code to decoder j on its satellite channel $\mathbb{W}_{Y_j|X_j}$.[9] 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 $\mathbf{S}_j(1:m, 1:l) \in \mathcal{S}_j^{m \times l}$. Let $\hat{\mathbf{K}}_j(1:m, 1:l) \in \mathcal{S}_1^{m \times l}$ denote decoder j's reconstruction of $\mathbf{S}_1(1:m, 1:l)$[10]. The m sub-blocks

$$\left\{ \left(\mathbf{S}_j(t, 1:l), \hat{\mathbf{K}}_j(t, 1:l) : j = 1, 2 \right) : t \in [m] \right\} \quad (5)$$

are iid[11] with an $l-$length distribution $\mathbb{W}_{S_1^l S_2^l} p_{\hat{K}_1^l \hat{K}_2^l | S_1^l S_2^l}$[12]. 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 $H(S_j^l | \hat{K}_j^l)$ bits per source sub-block. We have no characterization of $p_{\hat{K}_1^l \hat{K}_2^l | S_1^l S_2^l}$, and hence we derive an upper bound.

$$\begin{aligned} H(S_j^l | \hat{K}_j^l) &\leq H(S_j^l, \mathbb{1}_{\{\hat{K}_j^l \neq S_1^l\}} | \hat{K}_j^l) \leq h_b(P(\hat{K}_j^l \neq S_1^l)) + \\ &+ P(\hat{K}_j^l \neq S_1^l) \log |\mathcal{S}_j^l| + P(\hat{K}_j^l = S_1^l) H(S_j^l | S_1^l). \quad (6) \\ &\leq l \mathcal{L}_l^S(P(\hat{K}_j^l \neq S_1^l), |\mathcal{S}_j|) + l H(S_j | S_1), \text{ where} \\ &\mathcal{L}_l^S(\phi, |\mathcal{K}|) := \frac{1}{l} h_b(\phi) + \phi \log |\mathcal{K}|, \quad (7) \end{aligned}$$

represents the additional source coding rate needed to compensate for the errors in the fixed B-L decoding. It suffices to prove $\mathcal{L}_l^S(P(\hat{K}_j^l \neq S_1^l), |\mathcal{S}_j|) + B + H(S_j|S_1) <$ capacity of $\mathbb{W}_{Y_j|X_j}$. Since $\mathcal{L}_l^S(\phi, |\mathcal{K}|)$ is non-decreasing in ϕ if $\phi < \frac{1}{2}$, we bound $P(\hat{K}_j^l \neq S_1^l)$ by a quantity that is less than $\frac{1}{2}$. Towards that end, note that $\{S_1^l \neq \hat{K}_j^l\} \subseteq \{S_1^l \neq S_2^l\} \cup \{S_1^l \notin T_\delta^l(S_1)\}$. Indeed, $S_1^l = S_2^l \in T_\delta^l(S_1)$ implies both encoders input same C_U-codeword and agree on the lB bits communicated to their respective decoders. Therefore $P(S_1^l \neq \hat{K}_j^l) \leq \phi$, where $\phi = \tau_{l,\delta}(S_1) + \xi(\underline{S})$,

$$\tau_{l,\delta}(S_1) \leq \quad 2a^k \exp\{-\frac{\delta^2 l}{2k^2 a^{2k}}\} \text{ and } \xi^{[l]}(\underline{S}) \leq \frac{l}{ka^{\eta k}}. \quad (8)$$

Choose $l = k^4 a^{\frac{\eta k}{2}}, \delta = \frac{1}{k}$, substitute in (8). Since $\eta \geq 6$,

[8]Encoder 2 also employs source code $T_\delta^l(S_1)$, (and not $T_\delta^l(S_2)$).

[9]Through our description, we assume communication over $\mathbb{W}_{Y_j|X_j}$ is noiseless. In the end, we prove that the rate we demand of $\mathbb{W}_{Y_j|X_j}$ is lesser than its capacity, justifying this assumption.

[10]1) Encoder j could input any arbitrary C_U-codeword when its sub-block $S_j^l \notin T_\delta^l(S_1)$, and decoder j could declare an arbitrary reconstruction when it observes $Y_0^l = 0^l$. 2) $\hat{\mathbf{K}}_2(1:m, 1:l)$ is also viewed as reconstruction of $\mathbf{S}_1(1:m, 1:l)$.

[11]The $l-$length mappings from the source sub-blocks to the $l-$length channel codewords input on the $\mathbb{W}_{Y_0|\underline{U}}$ are identical across the sub-blocks and the source sub-blocks are mutually independent.

[12]$p_{\hat{K}_1^l \hat{K}_2^l | S_1^l S_2^l}$ does not necessarily factor, owing to the $l-$length maps.

verify $\phi \leq 2k^3 a^{-\frac{\eta k}{2}} < \frac{1}{2}$ for sufficiently large a, k. Verify

$$\mathcal{L}_l^S(2k^3 a^{-\frac{\eta k}{2}}, |\mathcal{S}_j|) \leq \frac{1}{4k} \log a \tag{9}$$

for sufficiently large a, k. Recall $lB = \lceil \log |T_\delta^l(S_1)| \rceil - lA$. Substituting $\delta = \frac{1}{k}$, verify[13]

$$B \leq (2/l) + (1/k) \log a + (1 + (1/k)) h_b(1/k). \tag{10}$$

Since $h_b(\frac{2}{k}) - (1 + \frac{1}{k}) h_b(\frac{1}{k}) \geq \frac{1}{2k} \log \frac{k}{256}$ for large enough k, RHS of (16), (17) sum to at most $\frac{2}{l} + h_b(\frac{2}{k}) + \frac{5}{4k} \log a$ for large enough a, k. Furthermore, $H(S_2|S_1) \leq h_b(\frac{1}{(k-1)a^{\eta k}}) + \frac{2}{a^{\eta k}} \log a \leq h_b(\frac{2}{k a^{\eta k}}) + \frac{1}{4k} \log a$ 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. $p_{\hat{K}_1^l \hat{K}_2^l | S_1^l S_2^l}$ 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 $H(S_j^l | \hat{K}_j^l)$ and the decoder can employ a joint-typicality based decoder using the computed $p_{S_j^l | \hat{K}_j^l}$. 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 $\mathbb{W}_{\underline{Y}|\underline{XU}}$ do not satisfy LC conditions, and yet, $\underline{S}$ is transmissible over IC $\mathbb{W}_{\underline{Y}|\underline{XU}}$.

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, $\xi^{[l]}(\underline{S}) \to 1$, $\tau_{l,\delta}$ (and $g_{\rho,l}$)[14] $\to 0$. As l decreases, $\xi^{[l]}(\underline{S}) \to \xi(\underline{S})$, and $\tau_{l,\delta}$ (and $g_{\rho,l}$) $\to 1$. If $\phi \to 0.5$, $\mathcal{L}_l^S(\phi, |\mathcal{S}_j|) \to 0.5 \log |\mathcal{S}_j| = \frac{k}{2} \log a$.

B. Fixed B-L coding for the MAC problem

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

The MAC is depicted in Fig. 6 and described below. The input alphabets are $\mathcal{U} \times \mathcal{X}_1$ and $\mathcal{U} \times \mathcal{X}_2$. The output alphabet is $\mathcal{Y}_0 \times \mathcal{Y}_1 \times \mathcal{Y}_2$. $\mathcal{U} = \mathcal{Y}_0 = \{0, 1, \cdots, a-1\}$. $(U_j, X_j) \in \mathcal{U} \times \mathcal{X}_j$ denotes encoder j's input. Moreover, $\mathbb{W}_{\underline{Y}|\underline{UX}} = \mathbb{W}_{Y_0|\underline{U}} \mathbb{W}_{Y_1|X_1} \mathbb{W}_{Y_2|X_2}$,where

$$\mathbb{W}_{Y_0|U_1U_2}(y_0|u_1, u_2) = \begin{cases} 1 & \text{if } y_0 = u_1 = u_2 \\ 1 & \text{if } u_1 \neq u_2, y_0 = 0, \text{ and} \end{cases}$$

0 otherwise. The capacities of the PTPs $(\mathcal{Y}_j, \mathcal{X}_j, \mathbb{W}_{Y_j|X_j})$: $j = 1, 2$ are $\mathcal{C} + h_b(\frac{2}{k}) + \frac{1}{k} \log a$ and $\mathcal{C} + h_b(\frac{2}{k a^{10k}})$ respectively, where $\mathcal{C} = \alpha \log a + \alpha h_b(\alpha) + \frac{1}{4k} \log a$, $\alpha = \frac{8k^4}{a^{5k}}$.

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) $\underline{S}$ and MAC $\mathbb{W}_{\underline{Y}|\underline{UX}}$ do not satisfy CES conditions [2, Thm. 1], and yet, 2) $\underline{S}$ is transmissible over MAC $\mathbb{W}_{\underline{Y}|\underline{UX}}$.

Proof: To establish the first assertion, it suffices to prove

$$H(\underline{S}) > I(\underline{XU}; Y|Q)$$

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

[14]$g_{\rho,l}$ is defined in the sequel.

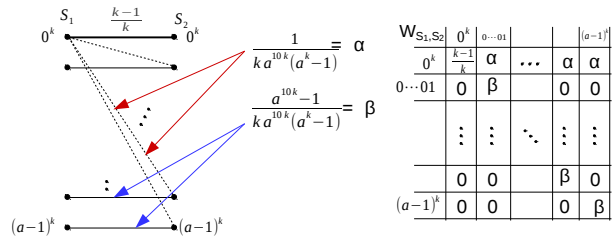

Fig. 5. On the left, the source pmf is depicted through a bipartite graph. Larger probabilities are depicted through edges with thicker lines. On the right, we depict the probability matrix.

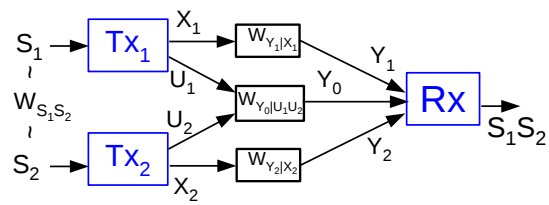

Fig. 6. MAC of Example 1.

for every pmf $p_Q p_{X_1|S_1Q} p_{X_2|S_2Q}$. An upper bound on $H(\underline{S})$ was derived in (2). Following arguments identical to proof of Thm. 1, we obtain the following,

$$\begin{aligned} I(\underline{XU}; Y|Q, R=1) &\leq \max_{p_{X_1U_1} p_{X_2U_2}} I(\underline{XU}; Y|Q, R=1) \\ &\leq 2\mathcal{C} + h_b(\frac{2}{k a^{10k}}) + h_b(\frac{2}{k}) + \frac{1}{k} \log a \\ &\quad + \max_{p_{U_1} p_{U_2} \mathbb{W}_{Y_0|\underline{U}}} H(Y_0) \end{aligned} \tag{11}$$

where $R = \mathbb{1}_{\{(S_1,S_2)=(0^k,0^k)\}}$. Indeed, (11) is analogous to (4). Following the arguments in proof of Thm. 1 verbatim, the last term on the RHS of (11) can be bounded on the above to yield

$$\begin{aligned} I(\underline{XU}; Y|Q) &\leq 2\log 2 + 2\mathcal{C} + h_b(\frac{2}{k a^{10k}}) + h_b(\frac{2}{k}) \\ &\quad + \frac{1}{k} \log a + \frac{3}{4} \log a + \frac{\log |\mathcal{Y}|}{k} \\ &< \log a \end{aligned}$$

for sufficiently large k, a. We now propose a technique based on fixed B-L codes to prove the second assertion.

The goal is to co-ordinate ($U_1 = U_2$) and communicate efficiently ($U_1 = U_2$ is close to uniform) over $\mathbb{W}_{Y_0|\underline{U}}$−channel. The key idea is to let every input symbol (on $\mathbb{W}_{Y_0|\underline{U}}$) be determined by a fixed number l of source symbols, henceforth referred to as *sub-block*. l is *reasonably* large so that, the (corresponding) sub-block of source symbols is approximately uniform on its typical set, and *reasonably* small so that S_1^l, S_2^l agree with *relatively* high probability. Since l is fixed, we need an *outer code* to correct for sub-block 'errors' and communicate rest of the information. This outer code operates over satellite channels $\mathbb{W}_{Y_j|X_j}$, and codes over *arbitrarily large* number m of sub-blocks.

$l-$length mapping from S_j^l to U_j^l is via a simple separation. Employ $T_\delta^l(S_1)$ as the source code at both encoders.[15] Let

[15]Note that encoder 2 also employs source code $T_\delta^l(S_1)$, (and not $T_\delta^l(S_2)$).

$C_U = \mathcal{U}^l$ be the channel code employed at both encoders. $lA = \lfloor \log a^l \rfloor$ bits, of the $\lceil \log |T_\delta^l(S_1)| \rceil$ bits output by the source code is mapped to C_U. Both encoders employ the same map.[16]

The rest of the $lB = \lceil \log |T_\delta^l(S_1)| \rceil - lA$ bits is communicated by Encoder 1[17] on the satellite channel using an arbitrary large B-L code (of B-L lm). After observing the $\mathbb{W}_{Y_0|\underline{U}}-$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 $\mathbf{S}_j(1:m, 1:l) \in \mathcal{S}_j^{m\times l}$. Let $\hat{\mathbf{K}}(1:m,1:l) \in \mathcal{S}_1^{m\times l}$ denote decoder's reconstruction. The m sub-blocks

$$\left\{ \left(\mathbf{S}_j(t,1:l), \hat{\mathbf{K}}(t,1:l) : j = 1,2 \right) : t \in [m] \right\} \tag{12}$$

are iid with an $l-$length distribution $\mathbb{W}_{S_1^l S_2^l} p_{\hat{K}^l|S_1^l S_2^l}$[18]. 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 $H(S_j^l|\hat{K}^l)$ bits *per* source *sub-block*. We do not have a characterization of $p_{\hat{K}^l|S_1^l S_2^l}$ and we therefore derive an upper bound. We have

$$\begin{aligned} H(S_j^l|\hat{K}^l) &\le H(S_j^l, \mathbb{1}_{\{\hat{K}^l \neq S_1^l\}}|\hat{K}^l) \le h_b(P(\hat{K}^l \neq S_1^l)) + \\ &+ P(\hat{K}^l \neq S_1^l)\log|\mathcal{S}_j^l| + P(\hat{K}^l = S_1^l) H(S_j^l|S_1^l) \\ &\le l\mathcal{L}_l^S(P(\hat{K}^l \neq S_1^l), |\mathcal{S}_j|) + lH(S_j|S_1), \text{ where,} \end{aligned} \tag{13}$$

$\mathcal{L}_l^S(\phi, |\mathcal{K}|)$ is as defined in (7). It suffices to prove

$$\begin{array}{l} \mathcal{L}_l^S(P(\hat{K}^l \neq S_1^l), |\mathcal{S}_j|) \\ +B\mathbb{1}_{\{j=1\}} + H(S_j|S_1)\mathbb{1}_{\{j=2\}} \end{array} \le \begin{array}{l} \mathcal{C} + h_b(\frac{2}{ka^{\eta k}})\mathbb{1}_{\{j=2\}} \\ +[h_b(\frac{2}{k}) + \frac{1}{k}\log a]\mathbb{1}_{\{j=1\}}, \end{array} \tag{14}$$

where the above RHS is the capacity of $\mathbb{W}_{Y_j|X_j}$. Since $\mathcal{L}_l^S(\phi, |\mathcal{K}|)$ is non-decreasing in ϕ if $\phi < \frac{1}{2}$, 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 $\{S_1^l \neq \hat{K}^l\} \subseteq \{S_1^l \neq S_2^l\} \cup \{S_1^l \notin T_\delta^l(S_1)\}$. Indeed, $S_1^l = S_2^l \in T_\delta^l(S_1)$ implies both encoders input same C_U-codeword and agree on the lB bits communicated by encoder 1. Therefore $P(S_1^l \neq \hat{K}^l) \le \phi$, where[19] $\phi = \xi^{[l]} + \tau_{l,\delta}$,

$$\tau_{l,\delta} \le 2a^k \exp\{-\frac{\delta^2 l}{2k^2 a^{2k}}\} \text{ and } \xi^{[l]} \le \frac{l}{ka^{10k}}. \tag{15}$$

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

$$\mathcal{L}_l^S(2k^3 a^{-5k}, |\mathcal{S}_j|) \le \mathcal{C} - \frac{1}{4k}\log a \tag{16}$$

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

$$B \le (2/l) + (1/k)\log a + (1 + (1/k))h_b(1/k). \tag{17}$$

Since $h_b(\frac{2}{k}) - (1 + \frac{1}{k})h_b(\frac{1}{k}) \ge \frac{1}{2k}\log\frac{k}{256}$ for large enough k, we have $B \le h_b(\frac{2}{k}) + \frac{5}{4k}\log a$ for sufficiently large a, k. Lastly, note that $H(S_2|S_1) \le h_b(\frac{1}{(k-1)a^{\eta k}}) + \frac{2}{a^{\eta k}}\log a \le h_b(\frac{2}{ka^{\eta k}}) + \frac{1}{4k}\log a$ 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. ■

Remark 4: $\mathbb{W}_{Y_0|\underline{U}}$ is precious and must be utilized efficiently. This requires extraction of common bits.[21] $\mathbb{W}_{\underline{S}}$ is 'very far' from uniform requiring significant compression. If you seek very efficient compression, by choosing l too large, $\xi^{[l]} \to 1$ 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, $\xi^{[l]} \to 1$, $\tau_{l,\delta} \to 0$. If l decreases, $\xi^{[l]} \to \xi$, and $\tau_{l,\delta} \to 1$. If $\phi \to 0.5$, $\mathcal{L}^S(\phi, |\mathcal{S}|) \to 0.5\log|\mathcal{S}|$.

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

IV. Fixed BL codes over an arbitrary IC

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

$$\mathcal{L}_l^S(\phi, |\mathcal{S}_j|) + B + H(S_j|S_1) \le I(X_j; Y_j) \tag{18}$$

where $\phi < \frac{1}{2}$ was an upper bound on $P(\hat{K}_j^l \neq S_1^l)$. 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.[22] We present our generalization in three pedagogical steps.

In general, $P(S_1^l \neq \hat{K}_j^l) \le \tau_{l,\delta} + \xi^{[l]} + g_{\rho,l}$, 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.[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

$$g_{\rho,l} := \sum_{j=1}^{2} \exp\{-l(E_r(A+\rho, p_U, p_{Y_j|U}) - \rho)\}.$$

In all our theorems, $\mathcal{L}_l^S(\phi, |\mathcal{S}_j|)$ is defined as in (7), $g_{\rho,l}$ as above, $\phi = \tau_{l,\delta}(K_1) + \xi^{[l]}(\underline{K}) + g_{\rho,l}$ will serve as an upper bound on $P(S_1^l \neq \hat{K}_j^l)$ that is less than $\frac{1}{2}$.

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.

[16] When $S_j^l \notin T_\delta^l(S_1)$, some chosen codeword is input on $\mathbb{W}_{Y_0|\underline{U}}$. We handle this event.

[17] Encoder 2 does *not* communicate the rest of its lB bits.

[18] $\mathbb{W}_{S_1^l S_2^l} = \prod_{i=1}^l \mathbb{W}_{S_1 S_2}$, but $p_{\hat{K}_1^l \hat{K}_2^l|S_1^l S_2^l}$ does not necessarily factor, owing to the $l-$length encoder and decoder mappings.

[19] Recall $\xi^{[l]} = \xi^{[l]}(\underline{S}), \xi = \xi(\underline{S}), \tau_{l,\delta} = \tau_{l,\delta}(S_1)$ in this section.

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

[21] $U_1 \neq U_2$ results in very low information transfer.

[22] And an additional loss in the channel rate, denoted $\mathcal{L}^C(\cdot,\cdot)$.

[23] For Ex.1, $g_{\rho,l} = 0$, and we ignored it. For general IC, $g_{\rho,l}$ is non-zero.

Theorem 3: $(\underline{\mathcal{S}}, \mathbb{W}_{\underline{S}})$ is transmissible over IC $(\underline{\mathcal{X}}, \underline{\mathcal{Y}}, \mathbb{W}_{\underline{Y}|\underline{X}})$ if there exists (i) a finite set $\mathcal{K}$, maps $f_j : \mathcal{S}_j \to \mathcal{K}$, with $K_j = f_j(S_j)$ for $j \in [2]$, (ii) $l \in \mathbb{N}, \delta > 0$, (iii) finite set $\mathcal{U}, \mathcal{V}_1, \mathcal{V}_2$ and pmf $p_U p_{V_1} p_{V_2} p_{X_1|UV_1} p_{X_2|UV_2}$ defined on $\mathcal{U} \times \underline{\mathcal{V}} \times \underline{\mathcal{X}}$, where p_U is a type of sequences[24] in $\mathcal{U}^l$, (iv) $A, B \geq 0$, $\rho \in (0, A)$ such that $\phi \in [0, 0.5)$,

$$A + B \geq (1+\delta) H(K_1), \text{ and for } j \in [2],$$
$$B + H(S_j|K_1) + \mathcal{L}_l^S(\phi, |\mathcal{S}_j|) < I(V_j; Y_j) - \mathcal{L}_j^C(\phi, |\mathcal{V}|),$$

where, $\phi := g_{\rho,l} + \xi^{[l]}(\underline{K}) + \tau_{l,\delta}(K_1)$, $\mathcal{L}_j^C(\phi, |\mathcal{U}|) = h_b(\phi) + \phi \log |\mathcal{U}| + |\mathcal{Y}||\mathcal{U}| \phi \log \frac{1}{\phi}$.

Remark 6: The characterization provided in Thm. 7 (and those in Thm. 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^l instead of S_j^l. $\mathcal{L}_j^C(\phi, |\mathcal{U}|)$ 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 $p_{Y_j|V_j}$. $\mathcal{L}_j^C(\phi, |\mathcal{U}|)$ is a bound on the difference in the mutual information between $p_{Y_j|V_j}$ and the actual channel. Note that $\mathcal{L}_j^C(\phi, |\mathcal{U}|) \to 0$ 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_U and how it is multiplexed with the outer code built on $\mathcal{V}_j$. If we build a single code C_{V_j} of B-L lm and multiplex it with m blocks of C_U, then C_{V_j} does *not* experience an IID memoryless channel.

Let $\boldsymbol{U}_j(t, 1:l)$ denote encoder j's chosen codeword from C_U corresponding to the tth sub-block of $\boldsymbol{K}_j(t, 1:l)$. We seek to identify sub-vectors of $\boldsymbol{U}_j$ that are IID, and whose pmf we know. We can then multiplex the outer code along these sub-vectors. *Interleaving* [19] enables us do this.

Suppose, for $t \in [m]$, $\mathbf{A}(t, 1:l) \sim p_{A^l}$ and the m vectors $\mathbf{A}(1, 1:l), \cdots, \mathbf{A}(m, 1:l)$ are iid p_{A^l}. Let $\pi_t : [l] \to [l] : t \in [m]$ 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

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

[15, Appendix A] contains a proof. The following notation will ease exposition. For $\mathbf{A} \in \mathcal{A}^{m \times l}$, and a collection $\pi_t : [l] \to [l] : t \in [m]$ of surjective maps, we let $\mathbf{A}^\pi \in \mathcal{A}^{m \times l}$ be such that $\mathbf{A}^\pi(t, i) := \mathbf{A}(t, \pi_t(i))$ for each $(t, i) \in [m] \times [l]$. The above fact can be therefore be stated as $\mathbf{A}^\pi(1:m, i) \sim \prod_{t=1}^{m} \{\frac{1}{l} \sum_{i=1}^{l} p_{A_i}\}$.

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

[24]Please refer to [17, Defn 2.1]

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

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

Since each codebook $C_{V_{ji}}$ and each codeword is IID, $\boldsymbol{U}_j(t, 1:l) \to \boldsymbol{Y}_j(t, 1:l)$ is IID $p_{Y_j|U}-$channel. Interleaving ensures $\boldsymbol{V}_j^\pi(1:m, i) \to \boldsymbol{Y}_j^\pi(1:m, i)$ is IID. But, unless $\boldsymbol{U}_1 = \boldsymbol{U}_2$, we are not guaranteed the latter channel is $\prod p_{Y_j|V_j}$.[26] In fact, we only know certain marginals of

$$\boldsymbol{U}_j^\pi(1:m, i), \boldsymbol{V}_j^\pi(1:m, i), \boldsymbol{X}_j^\pi(1:m, i), \boldsymbol{Y}_j^\pi(1:m, i) \ : \ j \in [2]$$

Let $(\boldsymbol{V}_j^\pi(1:m, i), \boldsymbol{Y}_j^\pi(1:m, i)) \sim \prod p_{\hat{V}_j \hat{Y}_j}$, where $p_{\hat{V}_j} = p_{V_j}$. We wish to bound the difference $I(V_j; Y_j) - I(\hat{V}_j; \hat{Y}_j)$ from above. Using the relations $p_{\hat{U}_j} = p_U$, $p_{\hat{V}_j \hat{Y}_j | \hat{U}_1 \hat{U}_2}(v, y|u, u) = p_{V_j Y_j | U}(v, y|u)$, where $\prod p_{\hat{U}_1 \hat{U}_2}$ is the pmf of the interleaved vector $\boldsymbol{U}_1^\pi(1:m, i) \boldsymbol{U}_2^\pi(1:m, i)$, $\sum_{u_1 \neq u_2} p_{\hat{U}_1 \hat{U}_2} \leq \phi$, we can prove $|p_{V_j Y_j}(v, y) - p_{\hat{V}_j \hat{Y}_j}(v, y)| \leq \phi$. These steps are analogous to those in [15, Appendix B]. Using [17, Proof of Lemma 2.7], we conclude $I(V_j; Y_j) - I(\hat{V}_j; \hat{Y}_j) \leq \mathcal{L}_C(\phi, |\mathcal{U}|)$. We refer to reader to [16] for the rest of the proof which is quite standard. ■

Lemma 2: The conditions stated in Thm. 7 are strictly weaker than those stated in [5, Thm. 1].

Proof: Ex. 1, with a, k chosen sufficiently large, satisfies conditions stated in Thm 7. In particular, choose $\delta = \frac{1}{k}, \rho = 1, A = (1 - \frac{1}{k^3}) \log a, B = H(S_1) - A, l = k^4 a^{\frac{\eta k}{2}}, \mathcal{K} = \mathcal{S}_j$, $f_j =$ identity, $\mathcal{V}_j = \mathcal{X}_j, p_U$ uniform, p_{V_j} 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 (18)). Towards that end, let $\mathscr{HK}(p_{V_1 W_1 X_1} p_{V_2 W_2 X_2})$ be the Han-Kobayashi inner bound defined in [5, Proposition 3].

Theorem 4: $(\underline{\mathcal{S}}, \mathbb{W}_{\underline{S}})$ is transmissible over IC $\mathbb{W}_{\underline{Y}|\underline{X}}$ if there exists (i) a finite set $\mathcal{K}$, maps $f_j : \mathcal{S}_j \to \mathcal{K}$, with $K_j = f_j(S_j)$ for $j \in [2]$, (ii) $l \in \mathbb{N}, \delta > 0$, (iii) finite set $\mathcal{U}, \mathcal{V}_j, \mathcal{W}_j : j \in [2]$ and pmf $p_U p_{V_1 W_1} p_{V_2 W_2} p_{X_1|UV_1W_1} p_{X_2|UV_2W_2}$ defined on $\mathcal{U} \times \underline{\mathcal{V}} \times \underline{\mathcal{W}} \times \underline{\mathcal{X}}$, where p_U is a type of sequences in $\mathcal{U}^l$, (iv) $A, B \geq 0$, $\rho \in (0, A)$ such that $\phi \in [0, 0.5)$,

$$A + B \geq (1+\delta) H(K_1), \text{ and for } j \in [2],$$
$$\begin{pmatrix} B + \mathcal{L}^C(\phi, |\underline{\mathcal{UVW}}|) \\ H(S_j|K_1) + \mathcal{L}_l^S(\phi, |\mathcal{S}_j|) \end{pmatrix} : j \in [2] \Big) \in \mathscr{HK}(p_{V_1 W_1 X_1} p_{V_2 W_2 X_2}),$$
$$\mathcal{L}_j^C(\phi, |\mathcal{U}|) = h_b(\phi) + 5\phi \log |\underline{\mathcal{UVW}}| + ||\underline{\mathcal{UVW}}|^3 \phi \log \frac{1}{\phi}.$$

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

[26]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 $\mathcal{X}_j$ and is superimposed on (interleaved vectors of) C_U. The challenge is that a fraction ϕ of the decoded codewords are erroneous. The approach is to treat the interleaved columns of the decoded $\hat{\boldsymbol{U}}$ 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: $(\underline{\mathcal{S}}, \mathbb{W}_{\underline{S}})$ is transmissible over an IC $\mathbb{W}_{\underline{Y}|\underline{X}}$ if there exists (i) a finite set $\mathcal{K}$, maps $f_j : \mathcal{S}_j \to \mathcal{K}$, 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 $p_U p_{X_1|U} p_{X_2|U}$ defined on $\mathcal{U} \times \underline{\mathcal{X}}$, where p_U is a type of sequences in $\mathcal{U}^l$, (iv) $A, B \geq 0$, $\rho \in (0, A)$ such that $\phi \in [0, 0.5)$, where

$$A + B \geq (1+\delta) H(K_1), \text{ and for } j \in [2],$$
$$B + H(S_j|K_1) + \mathcal{L}^S(\phi, |\mathcal{K}|) < I(X_j; Y_j|U) - \mathcal{L}_j^C(\phi, |\mathcal{U}|)$$
$$\mathcal{L}_j^C(\phi, |\mathcal{U}|) = h_b(\phi) + \phi \log |\mathcal{U}| + |\mathcal{X}_j||\mathcal{Y}||\mathcal{U}|(1 + |\mathcal{X}_{\not j}|)\phi \log \frac{1}{\phi}$$

and ϕ, $\mathcal{L}^S(\phi, |\mathcal{K}|)$ are as defined in Thm 7.
The final step in our generalization will combine the techniques of Thm. 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

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

$\mathcal{L}_l^S(\phi, |\mathcal{S}_j|)$ be defined as in (7), and $\phi = \tau_{l,\delta}(K_1) + \xi^{[l]}(\underline{K}) + g_{\rho,l} < \frac{1}{2}$ serve as an upper bound on $P(S_1^l \neq \hat{K}_j^l)$.

A. Designing independent streams of information ignoring self-interference

Theorem 6: $(\underline{\mathcal{S}}, \mathbb{W}_{\underline{S}})$ is transmissible over MAC $(\underline{\mathcal{X}}, \mathcal{Y}, \mathbb{W}_{Y|\underline{X}})$ if there exists (i) a finite set $\mathcal{K}$, maps $f_j : \mathcal{S}_j \to \mathcal{K}$, with $K_j = f_j(S_j)$ for $j \in [2]$, (ii) $l \in \mathbb{N}, \delta > 0$, (iii) finite set $\mathcal{U}, \mathcal{V}_1, \mathcal{V}_2$ and pmf $p_U p_{V_1} p_{V_2} p_{X_1|UV_1} p_{X_2|UV_2}$ defined on $\mathcal{U} \times \underline{\mathcal{V}} \times \underline{\mathcal{X}}$, where p_U is a type of sequences[27] in $\mathcal{U}^l$, (iv) $A, B \geq 0$, $\rho \in (0, A)$ such that $\phi \in [0, 0.5)$,

$$A + B \geq (1+\delta) H(K_1), \text{ and for } j \in [2],$$
$$H(S_j|K_1, S_{\not j}) + \mathcal{L}_l^S(\phi, |\mathcal{S}_j|) < I(V_j; Y_j|V_{\not j}) - \mathcal{L}_j^C(\phi, |\mathcal{V}|),$$
$$B + H(\underline{S}|K_1) + \mathcal{L}_l^S(\phi, |\mathcal{S}_j|) < I(\underline{V}; Y_j) - \mathcal{L}_j^C(\phi, |\mathcal{V}|),$$

where, $\phi := g_{\rho,l} + \xi^{[l]}(\underline{K}) + \tau_{l,\delta}(K_1)$, $\mathcal{L}_j^C(\phi, |\mathcal{U}|) = h_b(\phi) + \phi \log |\mathcal{U}| + |\mathcal{Y}||\mathcal{U}|\phi \log \frac{1}{\phi}$.

B. Joint decoding of fixed and infinite B-L codes

Theorem 7: A pair of sources $(\underline{\mathcal{S}}, \mathbb{W}_{\underline{S}})$ is transmissible over a MAC $(\underline{\mathcal{X}}, \mathcal{Y}, \mathbb{W}_{Y|\underline{X}})$ if there exists (i) a finite set $\mathcal{K}$, maps $f_j : \mathcal{S}_j \to \mathcal{K}$, with $K_j = f_j(S_j)$ for $j = 1, 2$, (ii) $l \in \mathbb{N}, \delta > 0$, (iii) finite set $\mathcal{U}$ and pmf $p_U p_{X_1|U} p_{X_2|U}$ defined on $\mathcal{U} \times \underline{\mathcal{X}}$, where p_U is a type of sequences[28] in $\mathcal{U}^l$, (iv) $A, B \geq 0$, $\rho \in (0, A)$ such that $\phi \in [0, 0.5)$, where

$$\phi := g_{\rho,l} + \xi^{[l]} + \tau_{l,\delta}, \xi^{[l]} := 1 - (1-\xi)^l,$$
$$\xi = P(K_1 \neq K_2), \tau_{l,\delta} := 2|\mathcal{K}| \exp\{-2\delta^2 p_{K_1}^2(a^*) l\},$$
$$g_{\rho,l} = \exp\{-l[E_r(A+\rho, p_U, p_{Y|U}) - \rho]\},$$
$$A + B \geq (1+\delta) H(K_1), \text{ and for } j \in [2],$$
$$H(S_j|S_{\not j}, \underline{K}) + \theta < I(X_j; Y, X_{\not j}|U) - h_b(\phi) - \phi \log |\mathcal{U}|$$
$$-|\mathcal{X}_{\not j}||\mathcal{Y}||\mathcal{U}|(1 + |\mathcal{X}_j|)\phi \log \frac{1}{\phi}, \quad (19)$$
$$B + H(\underline{S}|\underline{K}) + \theta < H(X_1|U) + H(X_2|U) - H(\underline{X}|Y, U)$$
$$-2h_b(\phi) - 2\phi \log |\mathcal{U}| - (1 + |\underline{\mathcal{X}}|)|\mathcal{Y}||\mathcal{U}|\phi \log \frac{1}{\phi}, \quad (20)$$

where $\theta = h_b(\phi) + 2\phi \log |\mathcal{K}|$.

Acknowledgements

The author is thankful to (i) Sandeep Pradhan and Farhad Shirani of Univ. of Michigan for discussions and sharing their insights on their findings [18], [19], (ii) Deepanshu Vasal of Univ. of Wisconsin at Madison for several technical discussions. The author is very grateful to Prof. P R Kumar of Texas A & M University and Prof. Wojciech Szpankowski of Purdue University for all the support, encouragement and technical inputs. The author is also thankful to Robert (Bob) Brown of CSoI for all the logistic support.

This work was supported by the Center for Science of Information at Purdue University, and in particular NSF grants NSF CCF-0939370.

References

[1] 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.

[2] T. Cover, A. Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," Information Theory, IEEE Transactions on, vol. 26, no. 6, pp. 648 – 657, nov 1980.

[3] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," Information Theory, IEEE Transactions on, vol. 19, no. 4, pp. 471 – 480, July 1973.

[4] T. Han and M. H. M. Costa, "Broadcast channels with arbitrarily correlated sources," IEEE Trans. on Info. Th., pp. 641–650, Sep 1987.

[5] W. Liu and B. Chen, "Interference channels with arbitrarily correlated sources," Trans. on Info. Th., vol. 57, no. 12, pp. 8027–8037, Dec 2011.

[27] Please refer to [17, Defn 2.1]

[28] Please refer to [20, Defn 2.1]

[6] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," Information Theory, IEEE Transactions on, vol. 27, no. 1, pp. 49 – 60, jan 1981.

[7] 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.

[8] 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.

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

[10] 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.

[11] A. Lapidoth and S. Tinguely, "Sending a bivariate gaussian over a gaussian mac," Trans. on Info. Th., vol. 56, no. 6, pp. 2714–2752, 2010.

[12] C. Tian, S. Diggavi, and S. Shamai, "The achievable distortion region of sending a bivariate gaussian source on the gaussian broadcast channel," Trans. on Info. Th., vol. 57, no. 10, pp. 6419–6427, Oct 2011.

[13] I. E. Aguerri and D. Gündüz, "Correlated gaussian sources over gaussian weak interference channels," in 2015 IEEE ITW, Oct 2015, pp. 84–88.

[14] A. Padakandla, "Communicating correlated sources over a MAC in the absence of a Gács-Körner common part," in Proc. 2016 ISIT, pp. 2014–18.

[15] ——, "Communicating Correlated Sources Over a MAC in the absence of a Gács-Körner Common Part," available at http://arxiv.org/abs/1601.07091 .

[16] ——, "Communicating Correlated Sources Over an Interference Channel," available at https://arxiv.org/abs/1701.06977 .

[17] I. Csiszár and J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems, 2nd ed. Cambridge Univ. Press, 2011.

[18] F. Shirani and S. Pradhan, "Finite block-length gains in distributed source coding," in Proc. IEEE 2014 ISIT, June 2014, pp. 1702–1706.

[19] F. Chaharsooghi, A. Sahebi, and S. Pradhan, "Distributed source coding in absence of common components," in Proc. 2013 ISIT, pp. 1362–66.

[20] I. Csiszár and J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems, 2nd ed. Budapest: Cambridge University Press, June 2011.