Abstract—We consider a set of $S$ independent encoders that must transmit a set of correlated sources through a network of noisy, independent broadcast channels to $T$ receivers, with no interference at the receivers. For the general problem of sending correlated sources through broadcast networks, it is known that the source–channel separation theorem breaks down and the achievable rate region as well as the proper method of coding are unknown. For our scenario, however, we establish the optimal rate region using a form of joint source–channel coding. When the optimal channel input distribution from transmitter $i$ to receiver $j$ is independent of $j$, our result has a max-flow/min-cut interpretation. Specifically, in this case our result implies that if it is possible to send the sources to each receiver separately while ignoring the others, then it is possible to send to all receivers simultaneously.

I. INTRODUCTION

While many point-to-point communication problems are relatively well understood, the more general problem of how to transmit correlated sources to multiple receivers is much harder to analyze. Two of the key features that complicate network information theory problems include the lack of source–channel separation and the conflicting goals that arise when multiple destinations receive the same transmission (broadcast) or when multiple transmitters send to the same destination (multiple access).

In order to understand how these issues effect communication system design, we consider the problem of sending $S$ correlated sources through a noisy broadcast network with $T$ receivers as illustrated in Figure 1. Specifically, we imagine that transmitter $i$ observes a source $U_i$ which consists of a set of $n$ independent and identically distributed samples, $U_{i}[1], U_{i}[2], \ldots, U_{i}[n]$. The sources are correlated (across the transmitters) in the sense that the joint distribution (across $i$) is

$$P(U_1, U_2, \ldots, U_S) = \prod_{m=1}^{n} P(U_{1}[m], U_{2}[m], \ldots, U_{S}[m]).$$

The channels are independent broadcast channels in the sense that if each transmitter chooses a block of channel inputs $X_{i}[1], X_{i}[2], \ldots, X_{i}[n]$ then each receiver $j$ observes the $i$ channel output sequences $Y_{i,j}[1], Y_{i,j}[2], \ldots, Y_{i,j}[n]$ without interference. The aggregate channel law is thus assumed to be

$$P(Y_{1,1} \ldots Y_{S,T}|X_{1} \ldots X_{S}) = \prod_{m=1}^{n} \prod_{i=1}^{S} \prod_{j=1}^{T} P_{i,j}(Y_{i,j}[m]|X_{i}[m])$$

for some set of discrete memoryless channels $\{P_{i,j}(y|x)\}$. Of interest are communication schemes, possibly involving joint source–channel coding, that allow each receiver to decode all of the source sequences $U_1, U_2, \ldots, U_S$ with high probability.

Ho et. al considered Slepian-Wolf [1] coding over networks within the context of network coding [2], with implicit use of joint source–channel coding schemes. They do not, however, consider noisy channels within their framework. Ramamoorthy et. al considered separating Slepian-Wolf coding from network coding and developed a scenario where separate source–channel coding is insufficient [3]. Other work that illustrates max–flow/min–cut interpretations of communication across networks of noisy channels includes [4]. Cover, El Gamal, and Salehi [5] considered the problem of transmitting correlated sources over multiple access networks. In their model, there is only a single receiver (i.e., $T = 1$) and the multiple access interference does not factor as in (2). As a result, [5] shows that the source–channel separation theorem for point-to-point communication [6, Sec. 14.10] breaks down and a joint source–channel coding strategy is required.

In contrast, as in the case of [7], a key feature of our network model is that each receiver observes the output of the channel from each transmitter independently and so there is no multiple access interference. Consequently, effects such as coherent or incoherent interference, beam-forming, or transmitter cooperation are not possible. Note, however, that there is a type of “broadcast interference” in the sense that transmitter $i$ cannot choose to send independent signals to each receiver. Instead, transmitter $i$ must choose a channel input that simultaneously conveys the desired message to all receivers. Similarly, there is also a kind of “distribution interference” in the sense that transmitter $i$ cannot choose a different input

Fig. 1. Transmitting correlated sources over a broadcast network: Two correlated sources $U_1$ and $U_2$ are observed by independent transmitters and both sources must be communicated to each receiver.

Todd P. Coleman is with the University of Illinois, Urbana-Champaign, Urbana, IL 61801, coleman@uiuc.edu. Emin Martinian is with Bain Capital, Boston, MA, emin@alum.mit.edu. Erik Ordentlich is with Hewlett-Packard Laboratories, 1501 Page Mill Rd., Palo Alto, CA 94304, email: eord@hpl.hp.com.
random variables $X_1,\ldots,X_n$ is the result of passing $X_1$ through a binary symmetric channel (BSC) while $Y_{1,2}$ is the result of passing $X_1$ through a $Z$-channel, then it is not possible to simultaneously transmit at the point-to-point capacity of the individual channels.

Another closely related paper is that of Tuncel [8], which considers the problem of transmitting a single source over a broadcast channel to receivers that also observe side-information correlated with the source. A single-letter characterization of the source-channel pairs for which reliable communication is possible is derived. In contrast, our setting involves multiple sources and requires that each source be correctly decoded by each receiver. We assume no receiver side-information, although this could be easily incorporated, as can requiring each receiver to correctly decode only a prespecified, receiver-dependent subset of the sources. Either generalization would then include the setting of [8]. The achievability proof of [8] relies on a joint source-channel coding scheme using a kind of virtual binning in which a bin consists of all source sequences whose corresponding channel codewords are jointly-typical with the received signal. Our achievability proof is based on a similar approach.

II. DEFINITIONS

- We denote a length $n$ sequence $(x[1],x[2],\ldots,x[n])$ as $x^n$.
- We use the notation $S$ to denote the set $\{1,2,\ldots,S\}$ and $T$ to denote the set $\{1,2,\ldots,T\}$.
- Source $i$, the channel input from transmitter $i$, and the output from transmitter $i$ to transmitter $j$ shall be assumed to be over the respective finite alphabets $U_i$, $A_i$, and $Y_{i,j}$. Sequences over these alphabets shall respectively be denoted as $u^n_i$, $x^n_i$, and $y^n_{i,j}$.
- For any $A \subseteq S$ we define $U_A$ and $u^n_A$ as
  \[ U_A = \{ u_i \}_{i \in A}, \quad u^n_A = \{ u^n_i \}_{i \in A}. \]

We denote by $A^n_r(P_{U_S})$ the jointly typical set associated with the discrete probability distribution $P_{U_S}$ on $U_S$ and $\epsilon > 0$:

\[ A^n_r(P_{U_S}) = \left\{ u^n_S \mid \forall A \subseteq S, \left| H(P_{U_A}) - \frac{1}{n} \log \frac{1}{P(u^n_A)} \right| < \epsilon \right\}. \]

- For any $A \subseteq S$ we define $W_A(u^n_S)$, the set of all sequences $\tilde{u}^n_A$ such that $\tilde{u}^n_A$ differs from $u^n_A$ for each $i \in A$ and that, when combined with $u^n_{A^c}$, are jointly typical, as
  \[ W_A(u^n_S) = \left\{ \tilde{u}^n_A : \{ \tilde{u}^n_A \neq u^n_A \}_{i \in A}, \ (\tilde{u}^n_{A^c}, u^n_{A^c}) \in A^n_r(P_{U_S}) \right\} \]

We note the following properties [6] for a collection of i.i.d. random variables $\{(X(i),Y(i))\}$ with probability distribution $P_{X,Y}$:

- We denote a pair of random variables sequences $(X^n,Y^n)$ as jointly typical with respect to the distribution $P_{X,Y}$ if $(X^n,Y^n) \in A^n_r(P_{X,Y})$.

- For any $\epsilon > 0$ and for large enough $n$:
  \[ P(A^n_r(P_{X,Y})) > 1 - \epsilon. \]

- If $(X^n,Y^n)$ are i.i.d. distributed according to $P_X P_Y$ and $P_X$ and $P_Y$ agree with the marginals of $P_{X,Y}$, then
  \[ P((X^n,Y^n) \in A^n_r(P_{X,Y})) \leq 2^{-n(I(X;Y) - 3 \epsilon)}. \]

The following is also a well known consequence of joint typicality.

- If $u^n_{S,j} \in A^n_r(P_{U_S})$, then
  \[ |W_A(u^n_S)| \leq 2^{n[H(U_A|U_A^c)+\epsilon]}. \]

III. STATEMENT OF RESULTS

Let $C_{i,j}$ denote the point-to-point capacity of the channel connecting source $i$ to receiver $j$:

\[ C_{i,j} = \max_{P(x_i)} I(X_i;Y_{i,j}). \]

Specifically, $C_{i,j}$ is the maximum rate of information that can be sent from transmitter $i$ to receiver $j$, if all other receivers are ignored. Thus, if we only cared about receiver $j$ it is trivial to determine whether all the sources could be successfully communicated to the receiver with no bandwidth expansion/compression:1

**Proposition 1:** If we only consider receiver $j$, the sources can be reliably communicated to receiver $j$ if for some $\epsilon_1$ and $\epsilon_2$

\[ (C_{1,j} - \epsilon_1, C_{2,j} - \epsilon_2, \ldots, C_{S,j} - \epsilon) \in R_{SW}(U_1, U_2, \ldots, U_S), \]

where $R_{SW}(U_1, U_2, \ldots, U_S)$ is the Slepian-Wolf [11] rate region for the sources $U_1, U_2, \ldots, U_S$. Conversely, if the sources can be reliably communicated to receiver $j$ then

\[ (C_{1,j}, C_{2,j}, \ldots, C_{S,j}) \in R_{SW}(U_1, U_2, \ldots, U_S). \]

While (8) is sufficient if we only want to communicate to receiver $j$, it is not sufficient if we wish to communicate to all receivers simultaneously, in a source–channel separated fashion. Specifically, even if (8) is satisfied for each $j$, it may be impossible to successfully communicate the sources to all the receivers with separate source and channel decoding.

![Fig. 2. Strict suboptimality of source-channel separation at both encoders and decoders.](image)

For example, as illustrated in Figure 2, suppose each channel is a BSC where $C_{1,1} = 0.9$, $C_{1,2} = 0.6$, $C_{2,1} = 0.6$, $C_{2,2} = 0.9$ and suppose we would like to communicate the correlated sources $U_1$ and $U_2$ where $H(U_1) = H(U_2) = 1$, $H(U_1, U_2) = 1.5$. Then from the cut-set bounds [6], if we

1With no bandwidth expansion/compression, if $n$ source samples are observed for each $U_i$, then exactly $n$ channel samples are used. We consider the more general case of differing numbers of samples in Section V.
were to perform complete source-channel separation at the encoders and decoders, then message 1 can be decoded by receiver 2 only if $R_1 \leq 0.6$. Similarly, message 2 can be decoded by receiver 1 only if $R_2 \leq 0.6$. This implies that $R_1 + R_2 \leq 1.2 < H(U, V)$. So performing Slepian-Wolf encoding and decoding in a modular fashion with channel coding does not suffice even though (8) is satisfied for all $j$.

Any scheme based on Slepian-Wolf source coding followed by channel encoding for the above example would require transmitting across some of the BSCs at rates greater than their capacities. Nevertheless, the following theorem, which is our main result, implies that the use of a certain joint source–channel coding procedure does permit the reliable transmission of the sources to each of the receivers.

**Theorem 1:** The sources $U_1, \ldots, U_S$ can be reliably communicated to all receivers simultaneously if there exist random variables $X_1, \ldots, X_S$ satisfying

$$H(U_A|U_{A^c}) \leq \sum_{i\in A} I(X_i; Y_{i,j})$$

for all $j \in T$ and $A \subseteq S$. Conversely, if the sources $U_1, \ldots, U_S$ can be reliably communicated to all receivers simultaneously then there exist random variables $X_1, \ldots, X_S$ satisfying

$$H(U_A|U_{A^c}) \leq \sum_{i\in A} I(X_i; Y_{i,j})$$

for all $j \in T$ and $A \subseteq S$.

Consider the case where for each $i$, there is an optimal input distribution $P(x_i^n)$ that simultaneously achieves the point-to-point capacity from transmitter $i$ to receiver $j$:

$$\forall i, \exists P(x_i^n) \text{ such that } \forall j, I(X_i^n; Y_{i,j}) = C_{i,j}. \quad (12)$$

When this condition is satisfied (e.g., when each channel from $i$ to $j$ is a BSC), Theorem 1 has a max-flow/min-cut interpretation. Specifically, if we consider a cut-set separating receiver $j$ from all the transmitters we can easily determine whether receiver $j$ can decode all the sources using capacity arguments and the Slepian-Wolf theorem via Proposition 1. Evidently when (12) is true, Theorem 1 says that we can determine whether a given communication problem is feasible (i.e., whether all the receivers can successfully decode all the sources) simply by considering the minimum cut. It seems that in this scenario, information acts like a fluid in the sense that analyzing flows is sufficient and worrying about “broadcast interference” is not required.

**IV. PROOF OF THE THEOREM 1**

A. Achievability

1) **Encoder:** Select at random, independently for each transmitter $i$, a source-sequence-to-channel-codeword encoder $x_i^n(\cdot) : U_i^n \rightarrow X_i^n$ by letting the symbols of $x_i^n(u_i^n)$ for all $u_i^n \in U_i^n$ be independently and identically distributed (within and across codewords) according to $P_{X_i}$.  

2) **Decoder:** At the $j$th decoder, decode to a source tuple $(\tilde{u}_1^n, \ldots, \tilde{u}_S^n)$ in $A^n$ $(P_{U_S})$ Also satisfying $(x_{i}^n(\tilde{u}_i^n), y_{i,j}) \in A^n(P_{X_i, Y_{i,j}})$ for all $i$.

We shall use the upper case symbols $U_i^n$, $X_i^n(U_i^n)$, and $Y_{i,j}$ to denote the random values of the source sequence, channel codewords, and channel outputs, respectively, with a joint distribution induced by the source distribution, the above random encoder selection, and the above broadcast network channel model.

3) **Error Probability Analysis:** To be concise, we shall let $A^n_{c,i,j}$ denote

$$A^n_{c,i,j} \triangleq A^n_e(P_{X_i, Y_{i,j}}).$$

The error probability $P_e$ can be bounded by

$$P_e \leq \sum_{j \in T} [P(E_{j,0}) + \sum_{A \subseteq S} P(E_{j,A})] \quad (13)$$

where:

- $E_{j,0} \triangleq \{ U^n_S \notin A^n_e(P_{U_S}) \}$
- $E_{j,A} \triangleq \{ U^n_S \in A^n_e(P_{U_S}) \} \cap \{ \exists \forall u^n_A \in \mathcal{W}_A(U^n_S) \text{ s.t. } (X^n_i(u^n_i), Y_{i,j}) \in A^n_{c,i,j} \forall i \in A \}$.  

We now bound $P(E_{j,A})$ which is $P(E_{j,A})$ averaged over the random encoding.

For each $i \in A$, and any $u^n_i \neq \tilde{u}_i^n$,

$$P\left( (X^n_i(u^n_i), Y^n_i) \in A^n_{c,i,j} | U^n_i = \tilde{u}_i^n \right) \leq 2^{-n(I(X_i; Y_{i,j}) - 3e)} \quad (14)$$

which follows from the independence of $X^n_i(u^n_i)$ and $Y^n_i$ and (5).

Letting

$$E(u^n, i, j, e) \triangleq \{ (X^n_i(u^n_i), Y^n_i) \in A^n_{c,i,j} \},$$

it follows that

$$\mathcal{P}(E_{j,A}) = \sum_{\tilde{u}^n_S \in A^n_{c}(P_{U_S})} P_{U^n_S}(\tilde{u}^n_S) \times$$

$$P\left( \bigcup_{u^n_A \in \mathcal{W}_A(\tilde{u}^n_A)} \bigcap_{i \in A} E(u^n_i, i, j, e) | U^n_S = \tilde{u}^n_S \right) \leq \sum_{\tilde{u}^n_S \in A^n_{c}(P_{U_S})} 2^{n(H(U_A|U_{A^c}) + e)} P_{U^n_S}(\tilde{u}^n_S) \times$$

$$\max_{\{u^n_i \neq \tilde{u}^n_i\} \in A} \sum_{i \in A} P\left( \bigcap_{i \in A} E(u^n_i, i, j, e) | U^n_S = \tilde{u}^n_S \right) \leq 2^{n(H(U_A|U_{A^c}) + e)} \prod_{i \in A} 2^{-n(I(X_i; Y_{i,j}) - 3e)} \quad (16)$$

$$= 2^{n[H(U_A|U_{A^c}) - (\sum_{i \in A} I(X_i; Y_{i,j}) - 4e)]} \rightarrow 0 \quad (17)$$
where (15) follows from the union bound, (3) and (6); (16) follows from (14) and the independence of the encoder selection across transmitters; and (17) follows by (10).

B. Converse

Let us denote the error probability at receiver $j \in T$ as $P_e^j$. For each $A \subseteq S$, define

$$Y_{A,j}^n \triangleq \{Y_{i,j}^n\}_{i \in A}$$
$$X_A^n \triangleq \{X_i^n\}_{i \in A}.$$ 

WLOG, assume that $A = \{1, \ldots, |A|\}$. Fano’s inequality tells us that if $P_e^j \to 0$ as $n \to \infty$, then:

$$H(U_A^n|Y_{A,j}^n, Y_{A'}^n) \leq n\epsilon_n \text{ where } \epsilon_n \to 0. \quad (18)$$

Thus

$$H(U_A^n|Y_{A,j}^n, Y_{A'}^n) = H(U_A^n|Y_{A,j}^n, X_A^n) + H(U_A^n|Y_{A,j}^n, X_A^n) \leq I(U_A^n; Y_{A,j}^n, U_A^n) + n\epsilon_n \quad (19)$$

$$\leq I(X_{A,j}^n; Y_{A,j}^n, U_A^n) + n\epsilon_n$$

$$= H(Y_{A,j}^n) - H(Y_{A,j}^n|X_{A,j}^n, U_A^n) + n\epsilon_n$$

$$\leq H(Y_{A,j}^n) - \sum_{a \in A} H(Y_{a,j}^n|\{Y_{a',j}^n\}_{a' < a}, X_{A,j}^n, U_A^n) + n\epsilon_n$$

$$\leq \sum_{a \in A} n I(X_a[t]; Y_{a,j}[t]) + n\epsilon_n$$

$$\leq n \sum_{a \in A} I(X_a[t]; Y_{a,j}) + n\epsilon_n \quad (20)$$

where (19) is due to (18), (20) follows because the channels are independent, (21) follows because the channels are memoryless and because conditioning reduces entropy, and (22) follows by defining

$$P_{X_a^n} = \frac{1}{n} \sum_{t=1}^n P_{X_a[t]}$$

and noting the concavity of mutual information.

V. BANDWIDTH EXPANSION AND COMPRESSION

Theorem 1 assumes that $n$ source symbols are encoded into $n$ channel symbols and characterizes the source–channel pairs for which this can be done reliably. Additional flexibility can be introduced by allowing encoder $i$ to map $n$ source symbols to $\lfloor \kappa_i n \rfloor$ channel symbols, where the bandwidth factor $\kappa_i$ can be interpreted as a suitably normalized bandwidth of the channel at sender $i$. By slightly modifying the proof of Theorem 1, we can obtain the following generalization that characterizes the set of reliably achievable bandwidth factors $(\kappa_1, \ldots, \kappa_S)$ for a given source and channel pair.

**Theorem 2**: The sources $U_1, \ldots, U_S$ can be reliably communicated to all receivers simultaneously, using bandwidth factors $\kappa_1, \ldots, \kappa_S$, if there exist random variables $X_1, \ldots, X_S$ satisfying

$$H(U_A^n|U_{A'}^n) \leq \sum_{i \in A} \kappa_i I(X_i; Y_{i,j}) \quad (23)$$

for all $j \in T$ and $A \subseteq S$. Conversely, if the sources $U_1, \ldots, U_S$ can be reliably communicated to all receivers simultaneously, using bandwidth factors $\kappa_1, \ldots, \kappa_S$, then there exist random variables $X_1, \ldots, X_S$ satisfying

$$H(U_A^n|U_{A'}^n) \leq \sum_{i \in A} \kappa_i I(X_i; Y_{i,j}) \quad (24)$$

for all $j \in T$ and $A \subseteq S$.

VI. CONCLUDING REMARKS

To summarize, we illustrated that for broadcast networks, performing source–channel separation for encoding and decoding is in general suboptimal. Next, we characterized the optimal performance for this setting via Theorem 1, the proof of which shows the existence of a joint source–channel coding scheme attaining optimal performance. In the (random) construction of this scheme (Section IV-A.1), channel codewords are assigned to source sequences in an independent, identically distributed fashion. The proof can thus be interpreted as demonstrating the existence of a set of channel codewords such that almost all assignments of these codewords to source sequences yields a good source–channel code, when decoded using the joint source–channel decoder of Section IV-A.2. This is in contrast to multiple access networks exhibiting multiple access interference, where a careful assignment of channel codewords to source sequences appears to be needed to optimally match the correlation in the sources to the structure of the interference. Thus, in a sense, a kind of source–channel separation exists on the encoding side of our problem. A in-depth discussion of various degrees of source–channel separation can be found in the related work [8] mentioned above.

Our result may have practical relevance as a robust architecture - for example, in sensor networks with fusion centers, processing power is usually not nearly as constrained of a resource for the fusion center as is the case for the individual sensors. Consequently making the signaling methods for the individual sensor nodes to require the least possible cooperation and computation is generally desirable. Note that these properties fit well with the achievability methods we proposed here, and they still attain asymptotically optimal performance.

REFERENCES


