Linear Network Coding Capacity Region of 2-Receiver MIMO Broadcast Packet Erasure Channels with Feedback

Chih-Chun Wang, David J. Love; {chihw,djlove}@purdue.edu
Center of Wireless Systems and Applications (CWSA)
School of Electrical and Computer Engineering, Purdue University, USA

Abstract*—This work studies the capacity of the 2-receiver multiple-input/multiple-output (MIMO) broadcast packet erasure channels (PECs) with channel output feedback, which is in contrast with the single-input/single-output setting of the existing works. Motivated by the immense success of linear network coding (LNC) in theory and in practice, this work focuses exclusively on LNC schemes and characterizes the LNC feedback capacity region (R_1^*, R_2^*) of 2-receiver MIMO broadcast PECs. A new linear-space-based approach is proposed, which unifies the problems of finding a capacity outer bound and devising the achievability scheme into a single linear programming (LP) problem. Specifically, *an LP solver is used to exhaustively search for the LNC scheme(s) with the best possible throughput, the result of which is thus guaranteed to attain the LNC feedback capacity.

I. Introduction

It is well known that feedback enlarges the capacity region of broadcast channels [2], [8]. Recently the feedback capacity region of K-receiver broadcast *packet erasure channels* (PECs) has been fully characterized for $K \leq 3$ and for the special case of perfectly-fair capacity with spatially independent erasure events [4], [9]. Specifically, a 1-to-K PEC takes an input symbol (also known as a packet) W from some sufficiently large finite field $\mathsf{GF}(q)$ and each destination d_k (out of K destinations) receives either the input packet $Z_k = W$ or an erasure $Z_k = *$, depending on whether the packet W has successfully arrived at d_k. The capacity results in [4], [9] capture closely the network coding capacity for the downlink scenario from a single access point to multiple clients with one antenna and simple modulation schemes [6].

On the other hand, such a single-input/single-output (SISO) broadcast PEC model does not take into account several commonly used modern communication schemes. For example, 2 antennas may be used at both the source s and the destinations d_k, which corresponds to a multiple-input/multiple-output (MIMO) PEC [3] that takes $(W^{[1]}, W^{[2]}) \in (\mathsf{GF}(q))^2$ as input, and each d_k may receive one of the four possible outcomes $(W^{[1]}, *)$, $(*, W^{[2]})$, $(W^{[1]}, W^{[2]})$, and $(*, *)$ depending on whether the packet $W^{[m]}$ sent by antenna m, $m = 1, 2$, is decodable or not. Even when only a single antenna is used, source s may use Orthogonal Frequency Division Multiple Access (OFDMA), which, in each time slot, can send out multiple streams of packets over different sub-carriers. Each sub-carrier may experience different erasure events. Each d_k constantly scans all subcarriers and records any overheard packets. The multiple sub-carriers in OFDMA can again be modelled as a MIMO broadcast PEC.

This works considers the MIMO broadcast PEC with 2 receivers and *channel output feedback*. Motivated by the immense success of linear network coding (LNC) [7], this work focuses exclusively on LNC schemes and characterizes the full LNC feedback capacity region (R_1^*, R_2^*). A new framework is proposed that unifies the problems of finding a capacity outer bound and designing the corresponding bound-achieving solution into a single linear-programming (LP) problem. Specifically we use an LP solver to search exhaustively over all possible LNC design choices and find the LNC solution that achieves the highest throughput. The exhaustiveness guarantees that the resulting LNC scheme is throughput optimal (among all LNC solutions) and thus achieves the LNC capacity.

Remark: Such a constructive optimality proof has been widely used in the networking community but not in the information theory community. For example, in the networking society, *the optimal multi-path routing scheme is found by simply searching over all possible routing decisions that obey the flow-conservation law,* which is in contrast with the information-theoretic approach that first finds a cut and an achievability scheme and later proves that they meet. This exhaustive-search-based approach was previously not possible since there are too many LNC design choices. With a new framework that leverages upon the underlying linear space structure of LNC, we can greatly reduce the number of design choices and are thus able to design provably optimal LNC schemes without the need of finding any cut condition!

II. Problem Formulation

The M-input 2-receiver MIMO broadcast PEC is defined as follows. For any time slot, source s sends M symbols $\mathbf{W} \triangleq (W^{[1]}, W^{[2]}, \cdots, W^{[M]}) \in (\mathsf{GF}(q))^M$ and each d_i receives a random subset $\mathsf{rx}_i \subset \{1, 2, \cdots, M\}$ of the $W^{[\cdot]}$ symbols for $i = 1, 2$. The MIMO broadcast PEC can be described by the joint reception probability $p_{\mathsf{rx}_1,\mathsf{rx}_2}$ such that $\sum_{\forall \mathsf{rx}_1,\mathsf{rx}_2} p_{\mathsf{rx}_1,\mathsf{rx}_2} = 1$. For example, when $M = 3$, $p_{\{1,2\},\{2,3\}}$ is the probability that d_1 receives $\mathbf{Z}_1 \triangleq (W^{[1]}, W^{[2]}, *)$ and d_2 receives $\mathbf{Z}_2 \triangleq (*, W^{[2]}, W^{[3]})$. We consider only stationary and memoryless channels, i.e., $\{p_{\mathsf{rx}_1,\mathsf{rx}_2}\}$ does not change with respect to time and the reception events for any distinct time slots $t_1, t_2, \cdots$ are independent. The above setting is a strict generalization of the $(M, 2)$ *erasure broadcast channel* in [3], which requires the independence across different receivers and across all M sub-channels.

We use $p_{a_1a_2}^{[m]}$ to denote the (marginal) reception probabilities for the m-th symbol $W^{[m]}$ where each bit a_i indicates whether d_i receives $W^{[m]}$ or not for $i=1,2$. For example,

$$p_{10}^{[m]} = \sum_{\forall \mathsf{rx}_1,\mathsf{rx}_2 \text{ s.t. } m\in\mathsf{rx}_1, m\notin\mathsf{rx}_2} p_{\mathsf{rx}_1,\mathsf{rx}_2}.$$

Consider the following communication problem. For any rate vector (R_1,R_2), within n time slots source s would like to send two independent packet streams $\mathbf{X}_i \triangleq (X_{i,1},X_{i,2},\cdots,X_{i,nR_i}) \in (\mathsf{GF}(q))^{nR_i}$ to destination d_i for $i=1,2$. At the end of each time slot, each d_i reports back to s which subset of symbols it has received (the rx_i value) through the use of ACK or NACK. This channel output feedback setting was not considered in the existing work [3].

If we use the input argument "(t)", $t=1,2,\cdots,n$, to distinguish the n channel usages, a network code can be described by n encoding functions: for all $t=1,\cdots,n$,

$$\mathbf{W}(t) = f_t(\mathbf{X}_1,\mathbf{X}_2,[\mathsf{rx}_1,\mathsf{rx}_2]_1^{t-1}) \quad (1)$$

and two decoding functions: for all $i=1,2$,

$$\hat{\mathbf{X}}_i = g_i\left([\mathbf{Z}_i]_1^n, \{f_t(\cdot,\cdot,[\mathsf{rx}_1,\mathsf{rx}_2]_1^{t-1}) : t=1,\cdots,n\}\right), \quad (2)$$

where $[\mathbf{Z}_i]_1^n$ denotes what d_i has received from time 1 to n, and $[\mathsf{rx}_1,\mathsf{rx}_2]_1^{t-1}$ denotes the channel output information from time 1 to $(t-1)$. In (2) we assume that each d_i knows[1] how the coded symbols are generated (the functions $f_t(\cdot,\cdot,[\mathsf{rx}_1,\mathsf{rx}_2]_1^{t-1})$) but does not know the actual information symbols $\mathbf{X}_1$ and $\mathbf{X}_2$ used to generate $\mathbf{W}(t)$.

A network code is linear if the encoders f_t are linear with respect to $\mathbf{X}_1$ and $\mathbf{X}_2$, i.e., (1) can be written as

$$\mathbf{W}(t) = \overline{\mathbf{X}} \cdot \mathbf{C}_t, \quad (3)$$

where $\mathbf{C}_t$ is an $(nR_1+nR_2)\times M$ matrix in $\mathsf{GF}(q)$ and $\overline{\mathbf{X}} \triangleq (\mathbf{X}_1,\mathbf{X}_2)$ is an $n(R_1+R_2)$-dimensional row vector consisting of all information symbols. The choice of $\mathbf{C}_t$ depends on $[\mathsf{rx}_1,\mathsf{rx}_2]_1^{t-1}$. In practice [1], the *coding coefficients* $\mathbf{C}_t$ are often embedded in the header of the packets so that a decoder of the form of (2) can be used.

Definition 1: A rate vector (R_1,R_2) is achievable by LNC if for any $\epsilon>0$ there exists a linear network code of length n and finite field $\mathsf{GF}(q)$ such that $\mathsf{Prob}(\hat{\mathbf{X}}_i \neq \mathbf{X}_i) < \epsilon$ for all $i=1,2$. The LNC capacity region is defined as the closure of all (R_1,R_2) that are achievable by LNC.

III. An Illustrative Example

A simple inner bound on the above capacity problem is to perform LNC over the M different SISO broadcast PECs (one for each of the M inputs) separately. We thus have

Lemma 1: A rate vector (R_1,R_2) is achievable by LNC if there exist $2M$ non-negative variables $R_1^{[m]}$ and $R_2^{[m]}$, for all $m=1,\cdots,M$, such that the following conditions are satisfied.

$$\forall i=1,2, \quad \sum_{m=1}^{M} R_i^{[m]} = R_i \quad (4)$$

$$\forall m=1,\cdots,M, \quad \begin{cases} \frac{R_1^{[m]}}{p_{10}^{[m]}+p_{11}^{[m]}} + \frac{R_2^{[m]}}{p_{10}^{[m]}+p_{01}^{[m]}+p_{11}^{[m]}} < 1 \\ \frac{R_1^{[m]}}{p_{10}^{[m]}+p_{01}^{[m]}+p_{11}^{[m]}} + \frac{R_2^{[m]}}{p_{01}^{[m]}+p_{11}^{[m]}} < 1 \end{cases}. \quad (5)$$

Proof: Eq. (4) follows from summing up the per-input LNC rates and (5) follows from the feedback capacity region results for 1-input 2-receiver broadcast PECs [5]. ■

If we restrict ourselves to consider only $(M,2)$ *erasure channels* in [3], for which all sub-channels are independent, then we can follow the capacity outer bound construction in [8] and use the zero-feedback capacity results in [3] to construct the following capacity outer bound.

Lemma 2: Consider $(M,2)$ erasure channels. Even when we allow the use of non-linear codes (see (1)), a rate vector (R_1,R_2) is achievable only if there exist $4M$ non-negative variables $R_i^{[m,k]}$ for all $i,k\in\{1,2\}$ and $m\in\{1,\cdots,M\}$ such that the following conditions are satisfied.

$$\forall i,k\in\{1,2\}, \quad \sum_{m=1}^{M} R_i^{[m,k]} = R_i$$

$$\forall m\in\{1,\cdots,M\}, \quad \begin{cases} \frac{R_1^{[m,1]}}{p_{10}^{[m]}+p_{11}^{[m]}} + \frac{R_2^{[m,1]}}{p_{10}^{[m]}+p_{01}^{[m]}+p_{11}^{[m]}} \leq 1 \\ \frac{R_1^{[m,2]}}{p_{10}^{[m]}+p_{01}^{[m]}+p_{11}^{[m]}} + \frac{R_2^{[m,2]}}{p_{01}^{[m]}+p_{11}^{[m]}} \leq 1 \end{cases}.$$

For some special choices of the parameters $p_{a_1a_2}^{[m]}$, either (i) the inner and outer bounds do meet; or (ii) we can design simple LNC solutions that follow the ideas of [5] and attain the outer bound. The capacity region is thus determined for those values of $p_{a_1a_2}^{[m]}$. However, the capacity characterization problem remains open for general $p_{a_1a_2}^{[m]}$. Take the following 2-input, 2-receiver MIMO broadcast PEC for example. Suppose the joint reception probability $\{p_{\mathsf{rx}_1,\mathsf{rx}_2}\}$ satisfies

$$p_{00}^{[1]}=0, \quad p_{01}^{[1]}=0.125, \quad p_{10}^{[1]}=0, \quad p_{11}^{[1]}=0.875;$$
$$p_{00}^{[2]}=0.04, \quad p_{01}^{[2]}=0.16, \quad p_{10}^{[2]}=0.16, \quad p_{11}^{[2]}=0.64,$$

and the reception events for the 1st and the 2nd inputs are independent. This is equivalent to the setting of independent sub-channels with the m-th-input-to-d_i sub-channel having success probability 0.875, 1, 0.8, and 0.8 for $(m,i)=(1,1)$, $(1,2)$, $(2,1)$, and $(2,2)$, respectively. We plot in Fig. 1 the inner and outer bounds in Lemmas 1 and 2 for this example. As can be seen, there is a non-zero gap. Further, for these particular $p_{a_1a_2}^{[m]}$ values there is no LNC solution similar to those in [5] that can achieve $(R_1,R_2)=(0.875,0.96)$, a point within the outer bound.

In this work, we will characterize the LNC capacity region for *any general*[2] *M-input MIMO broadcast PECs* without relying on the (existing) outer bound arguments in Lemma 2.

[1]In general, d_i always knows $f_t(\cdot,\cdot,\cdot)$, the overall communication scheme that is agreed upon before transmission. But d_i may not know $f_t(\cdot,\cdot,[\mathsf{rx}_1,\mathsf{rx}_2]_1^{t-1})$ since it depends on the (random) channel realization. However, in addition to the allotted n time slots, s can simply use a few extra time slots to "broadcast" the binary channel status $[\mathsf{rx}_1,\mathsf{rx}_2]_1^n$ to both destinations so that a more powerful decoder in (2) can be used. The overhead of using extra time slots to convey the *binary* receptions status $[\mathsf{rx}_1,\mathsf{rx}_2]_1^n$ to $\{d_1,d_2\}$ diminishes to zero when a sufficiently large $\mathsf{GF}(q)$ is used. As a result, we focus exclusively on decoders of the form of (2). Also see [1] and our discussion of the practical LNC implementation.

[2]It remains an open problem whether Lemma 2 is true for the cases in which the sub-channels are dependent since the results in [3] no longer hold. On the other hand, our results hold for dependent sub-channels as well.

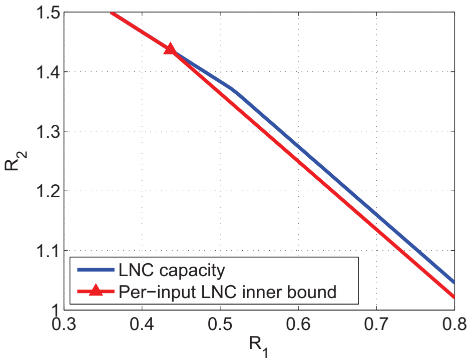

Fig. 1. The LNC capacity computed by Proposition 1 versus the per-input LNC inner bound in Lemma 1. The LNC capacity also coincides with the capacity outer bound in Lemma 2.

IV. Main Results

Consider a finite index set FTs of 18 elements:

$$\mathsf{FTs} \triangleq \{0,1,2,3,7,9,11,15,18,19,23,27,31,47,63,87,95,127\}. \quad (6)$$

It will be clear in Section V why we consider such an index set. The main results of this work can now be stated as follows.

Proposition 1: A rate vector (R_1, R_2) is in the LNC capacity region of a M-input 2-receiver MIMO broadcast PEC if and only if there exist $18M$ non-negative variables $x_{\mathbf{b}}^{[m]}$ for all $\mathbf{b} \in \mathsf{FTs}$ and $m = 1, \cdots, M$, and 7 non-negative variables y_1 to y_7 such that jointly they satisfy 4 groups of linear conditions:

• Group 1, termed the *time-sharing conditions*, has M equalities:

$$\forall m = 1, \cdots, M, \quad \left(\sum_{\forall \mathbf{b} \in \mathsf{FTs}} x_{\mathbf{b}}^{[m]} \right) = 1. \quad (7)$$

• Group 2, termed the *rank-conversion conditions*, has 7 equalities:

$$y_1 = \sum_{m=1}^{M} \left(\sum_{\substack{\forall \mathbf{b} \in \{0,1,2,3,7,9,11,15, \\ 18,19,23,27,31,47,63\}}} x_{\mathbf{b}}^{[m]} \right) \left(p_{10}^{[m]} + p_{11}^{[m]} \right) \quad (8)$$

$$y_2 = \sum_{m=1}^{M} \left(\sum_{\substack{\forall \mathbf{b} \in \{0,1,2,3,7,9,11,15, \\ 18,19,23,27,31,87,95\}}} x_{\mathbf{b}}^{[m]} \right) \left(p_{01}^{[m]} + p_{11}^{[m]} \right) \quad (9)$$

$$y_3 = R_1 + \sum_{m=1}^{M} \left(\sum_{\substack{\forall \mathbf{b} \in \{0,1,2,3,7, \\ 9,11,15,47\}}} x_{\mathbf{b}}^{[m]} \right) \left(p_{10}^{[m]} + p_{11}^{[m]} \right) \quad (10)$$

$$y_4 = R_2 + \sum_{m=1}^{M} \left(\sum_{\substack{\forall \mathbf{b} \in \{0,1,2,3,7, \\ 18,19,23,87\}}} x_{\mathbf{b}}^{[m]} \right) \left(p_{01}^{[m]} + p_{11}^{[m]} \right) \quad (11)$$

$$y_5 = \sum_{m=1}^{M} \left(\sum_{\substack{\forall \mathbf{b} \in \{0,1,2,3,9, \\ 11,18,19,27\}}} x_{\mathbf{b}}^{[m]} \right) \left(p_{10}^{[m]} + p_{01}^{[m]} + p_{11}^{[m]} \right) \quad (12)$$

$$y_6 = R_1 + \sum_{m=1}^{M} \left(x_0^{[m]} + x_1^{[m]} + x_9^{[m]} \right) \left(p_{10}^{[m]} + p_{01}^{[m]} + p_{11}^{[m]} \right) \quad (13)$$

$$y_7 = R_2 + \sum_{m=1}^{M} \left(x_0^{[m]} + x_2^{[m]} + x_{18}^{[m]} \right) \left(p_{10}^{[m]} + p_{01}^{[m]} + p_{11}^{[m]} \right). \quad (14)$$

• Group 3, termed the *rank-comparison conditions*, has 8 inequalities:

$$y_3 \leq y_6, \quad y_4 \leq y_7 \quad (15)$$

$$y_6 \leq (R_1 + R_2), \quad y_7 \leq (R_1 + R_2) \quad (16)$$

$$y_1 + y_2 - y_5 \geq 0 \quad (17)$$

$$y_5 + y_3 - y_6 \geq y_1 \quad (18)$$

$$y_5 + y_4 - y_7 \geq y_2 \quad (19)$$

$$y_6 + y_7 - (R_1 + R_1) \geq y_5. \quad (20)$$

• Group 4, termed the *decodability conditions*, has 2 equalities:

$$y_3 = y_1 \text{ and } y_4 = y_2. \quad (21)$$

Corollary 1: The capacity region of an M-input 2-receiver MIMO broadcast PEC depends only on the per-input marginal probabilities $p_{a_1 a_2}^{[m]}$. Whether the reception events of two inputs m_1 and m_2 are independent or not does not affect the capacity.

Proof: This is a direct result of Proposition 1. ■

V. Proof of Proposition 1

We first provide detailed sketches of the proof of Proposition 1 for the case of $M = 1$. That is, for each time slot, s sends a symbol $W(t) = \overline{\mathbf{X}} \cdot \mathbf{c}_t^{\mathrm{T}}$ where $\mathbf{c}_t$ is an $n(R_1 + R_2)$-dimensional (row) coding vector and $\mathbf{c}_t^{\mathrm{T}}$ is the transpose of $\mathbf{c}_t$. We will comment on how to generalize the proof for arbitrary M values in Section VI.

A. Basic Definitions

For each d_i, we define the *knowledge space* $S_i(t)$ in the end of time t by

$$S_i(t) \triangleq \mathsf{span}\{\mathbf{c}_\tau : \forall \tau \leq t \text{ s.t. } i \in \mathsf{rx}_i \text{ at time } \tau\}. \quad (22)$$

That is, $S_i(t)$ is the linear span of the vectors of those packets that have successfully arrived at d_i. One can easily see that in the end of time t, d_i is able to compute the value of $\overline{\mathbf{X}} \cdot \mathbf{v}^{\mathrm{T}}$ for all $\mathbf{v} \in S_i(t)$ by linearly combining $Z_i(\tau)$ for all $\tau \leq t$.

For $j = 1$ to $n(R_1 + R_2)$, let δ_j denote an $n(R_1 + R_2)$-dimensional elementary delta (row) vector with its j-th coordinate being one and all the other coordinates being zero. Define $\Omega \triangleq \mathsf{span}\{\delta_j : j = 1, \cdots, n(R_1 + R_2)\}$ as the *overall message space* and define $\Omega_1 \triangleq \mathsf{span}\{\delta_j : j = 1, \cdots, nR_1\}$ and $\Omega_2 \triangleq \mathsf{span}\{\delta_j : j = (nR_1 + 1), \cdots, n(R_1 + R_2)\}$ as the *individual message spaces* for d_1 and d_2, respectively. Both $S_i(t)$ and Ω_i are linear subspaces of Ω for $i = 1, 2$.

For any two linear subspaces $A, B \subseteq \Omega$, define $A \oplus B \triangleq \mathsf{span}\{\mathbf{v} : \forall \mathbf{v} \in A \cup B\}$ as the *linear sum space* of A and

B. From the discussion in the beginning of this subsection, d_i can decode the desired $X_{i,1}, \cdots, X_{i,nR_i}$ symbols if and only if in the end of time n we have $\Omega_i \subseteq S_i(n)$, or equivalently

$$(S_i(n) \oplus \Omega_i) = S_i(n). \tag{23}$$

B. Breakdown The Design Choices

In the beginning of time t, there are $q^{n(R_1+R_2)}$ different ways of designing the coding vector $\mathbf{c}_t \in \Omega$. To simplify the design choices, we consider the following 7 linear subspaces:

$$A_1 \triangleq S_1; \qquad A_2 \triangleq S_2; \tag{24}$$

$$A_3 \triangleq S_1 \oplus \Omega_1; \quad A_4 \triangleq S_2 \oplus \Omega_2; \quad A_5 \triangleq S_1 \oplus S_2; \tag{25}$$

$$A_6 \triangleq S_1 \oplus S_2 \oplus \Omega_1; \qquad A_7 \triangleq S_1 \oplus S_2 \oplus \Omega_2, \tag{26}$$

for which we use S_1 and S_2 as shorthand for $S_1(t-1)$ and $S_2(t-1)$, the knowledge spaces in the end of time $t-1$. In the subsequent discussion, we often drop the input argument "(t)" when the time instant of interest is clear in the context.

We can now partition the overall message space Ω into $2^7 = 128$ disjoint subsets depending on whether $\mathbf{c}_t$ is in A_k or not, for $k = 1, \cdots, 7$. Each subset is termed a *coding type* and can be indexed by a 7-bit string $\mathbf{b} = b_1 b_2 \cdots b_7$ where each b_k indicates whether $\mathbf{c}_t \in A_k$ or not. For example, type-0010111 contains the coding vectors that are in $A_3 \cap A_5 \cap A_6 \cap A_7$, but not in any of A_1, A_2, and A_4, which is denoted by

$$\begin{aligned} &\mathsf{TYPE}_{0010111} \\ &\triangleq (A_3 \cap A_5 \cap A_6 \cap A_7) \backslash (A_1 \cup A_2 \cup A_4) && (27) \\ &= (A_3 \cap A_5) \backslash (A_1 \cup A_4). && (28) \end{aligned}$$

Eq. (28) follows from the fact that by (24) to (26) we have $A_5 \subset (A_6 \cap A_7)$ and $A_2 \subset A_4$. We can also view $\mathbf{b}$ as a base-2 expression with b_1 (resp. b_7) being the most (resp. least) significant bit[3], e.g., type-0010111 is type-23. Note that some of the 128 coding types are always empty, which are termed the *infeasible types*. For example, type-1000000 is infeasible since there cannot be any $\mathbf{v} \in \Omega$ that is in $A_1 = S_1$ but not in $A_3 = S_1 \oplus \Omega_1 \supset A_1$. Overall, there are only 18 *Feasible Types* (FTs) and the list of them is the FTs defined in (6).

This new framework allows us to focus on the "types" of the coding choices without worrying about designing the exact values of $\mathbf{c}_t \in \Omega$. Specifically, we will focus on the following design problem: From which one of the 18 FTs should we choose $\mathbf{c}_t$ in order to maximize the throughput? We will also analyze the performance of any given scheme by quantifying how frequently a coding vector $\mathbf{c}_t$ of type-$\mathbf{b}$ is sent.

C. The "Only If" Analysis of Proposition 1

Fix any given linear network code such that d_i can decode all $X_{i,1}$ to X_{i,nR_i} in the end of time n for all $i = 1, 2$. Since the 18 disjoint FTs fully cover[4] Ω, for each time t we can always label the coding choice $\mathbf{c}_t$ of the given LNC scheme as one of the 18 FTs. Define $x^{[1]}_{\mathbf{b}} \triangleq \frac{1}{n}\mathsf{E}\left\{\sum_{t=1}^{n} 1_{\{\mathbf{c}_t \in \mathsf{TYPE}_{\mathbf{b}}\}}\right\}$ as the normalized expected number of $\mathbf{c}_t$ of type $\mathbf{b}$. Since the total number of time slots is n, this *time-sharing* argument proves that (7) holds for the case of $M = 1$.

Consider the linear spaces A_k, $k = 1$ to 7, in the beginning of time 1 and in the end of time n, and denote them by $A_k(0)$ and $A_k(n)$, respectively. Consider $A_6 = S_1 \oplus S_2 \oplus \Omega_1$ for example. By (22) we have $\mathrm{Rank}(A_6(0)) = \mathrm{Rank}(\Omega_1) = nR_1$. We then note that when source s sends a $\mathbf{c}_t \in \mathsf{TYPE}_{\mathbf{b}}$ for some $\mathbf{b}$ with b_6 being 0, then that $\mathbf{c}_t$ is not in $A_6 = S_1 \oplus S_2 \oplus \Omega_1$. *Therefore, whenever one of d_1 and d_2 receives $W(t) = \overline{\mathbf{X}} \cdot \mathbf{c}_t^{\mathrm{T}}$ successfully, the rank of A_6 will increase by one.* We thus have

$$\mathrm{Rank}(A_6(0)) + \sum_{\forall \mathbf{b} \text{ w. } b_6=0} \left(\sum_{t=1}^{n} 1_{\left\{ \substack{\mathbf{c}_t \in \mathsf{TYPE}_{\mathbf{b}}, \text{ and} \\ \text{one of } \{d_1, d_2\} \text{ receives it}} \right\}} \right) = \mathrm{Rank}(A_6(n)) \tag{29}$$

Define $y_k \triangleq \frac{1}{n}\mathsf{E}\{\mathrm{Rank}(A_k(n))\}$ as the normalized rank of $A_k(n)$. Taking the normalized expectation of (29), counting only the FTs, and by Wald's lemma, we have proven that (13) must hold for $M = 1$. By similar *rank-conversion* arguments, we can also prove (8) to (14) for $M = 1$.

For the following, we will derive the *rank comparison* inequalities in Group 3. By (24) to (26), in the end of time n we must have

$$A_3 \subset A_6, \quad A_4 \subset A_7, \quad A_6 \subset \Omega, \text{ and } A_7 \subset \Omega. \tag{30}$$

Considering the normalized expected ranks of the above inequalities in the end of time n, we have proven (15) and (16). Before continuing, we present the following lemma.

Lemma 3: For any two linear spaces B_1 and B_2, we have $\mathrm{Rank}(B_1 \oplus B_2) + \mathrm{Rank}(B_1 \cap B_2) = \mathrm{Rank}(B_1) + \mathrm{Rank}(B_2)$.

We then consider the following inequality:

$$\begin{aligned} &\mathrm{Rank}(S_1 \oplus S_2) + \mathrm{Rank}(S_1 \oplus \Omega_1) - \mathrm{Rank}(S_1 \oplus S_2 \oplus \Omega_1) \\ &= \mathrm{Rank}((S_1 \oplus S_2) \cap (S_1 \oplus \Omega_1)) && (31) \\ &\geq \mathrm{Rank}(S_1) && (32) \end{aligned}$$

where (31) follows from Lemma 3, and (32) follows from simple set operations. By taking the normalized expectation on (32) in the end of time n, we have proven (18). Similarly, we can derive the following inequalities:

$$\mathrm{Rank}(S_1) + \mathrm{Rank}(S_2) - \mathrm{Rank}(S_1 \oplus S_2) \geq 0 \tag{33}$$

$$\begin{aligned} &\mathrm{Rank}(S_1 \oplus S_2) + \mathrm{Rank}(S_2 \oplus \Omega_2) \\ &\quad - \mathrm{Rank}(S_1 \oplus S_2 \oplus \Omega_2) \geq \mathrm{Rank}(S_2) && (34) \end{aligned}$$

$$\begin{aligned} &\mathrm{Rank}(S_1 \oplus S_2 \oplus \Omega_1) + \mathrm{Rank}(S_1 \oplus S_2 \oplus \Omega_2) \\ &\quad - \mathrm{Rank}(\Omega) \geq \mathrm{Rank}(S_1 \oplus S_2), && (35) \end{aligned}$$

and use them to prove (17), (19), and (20), respectively.

Finally, since we focus on an LNC scheme that each d_i can decode its desired information symbols, we can prove (21) by the *decodability* condition (23) and by (24)–(26).

D. The "If" Analysis of Proposition 1

A critical difference between Lemma 2 and the outer bounding part in Proposition 1 is that the latter is a constructive approach while the former is implicitly a cut condition. As will

[3]We append zeros in the prefix to make the length of $\mathbf{b}$ always 7. For example, the base-2 representation of $\mathbf{b} = 6$ is 0000110.

[4]The actual set of vectors in a type, e.g., (27), evolves over time since the A_k definitions in (24) to (26) depend on the knowledge spaces S_1 and S_2 in the end of time $t-1$. However, the 18 FTs always cover Ω for any t value.

be demonstrated, the optimizing $\{x^*_{\mathbf{b}} : \forall \mathbf{b} \in \mathsf{FTs}\}$ and y^*_1 to y^*_7 in Proposition 1 can be directly translated to a capacity-achieving scheme, while it is not clear how the optimizing $R_i^{[m,k]}$ in Lemma 2 would guide the LNC design.

More explicitly, a capacity achieving LNC solution can now be designed as follows. In the beginning of time t, source s first uses the previous reception status $[\mathsf{rx}_1, \mathsf{rx}_2]_1^{t-1}$ to determine the subspaces A_1 to A_7. Then s simply needs to choose the $\mathbf{c}_t$ from one of the 18 FTs. *As long as we can make the long-term relative frequency of sending $\mathbf{c}_t$ of type-$\mathbf{b}$ equal $x^*_{\mathbf{b}}$ for all $\mathbf{b} \in \mathsf{FTs}$, then the aforementioned (outer bound) analysis guarantees that the resulting LNC scheme is throughput optimal.*

A detailed sketch of the proof is as follows. For any (R_1, R_2) in the interior of the capacity, let $\{x^R_{\mathbf{b}}\}$ (and the companying y^R_1 to y^R_7) denote the variables that satisfy the conditions in Proposition 1 and we can, without loss of generality, assume that when plugging in $\{x^R_{\mathbf{b}}\}$, the rank comparison inequalities in (15)–(20) are strict inequality, or equivalently, the set/rank inequalities (30)–(35) are strict inequalities. The central question that decides whether we can attain the target frequency $\{x^R_{\mathbf{b}}\}$ is: *Suppose we are able to achieve the target relative frequency $\{x^R_{\mathbf{b}}\}$ up to time $(t-1)$ (which leads to strict inequalities (15)–(20) at time $t-1$). Can we freely choose (from the 18 FTs) which coding type to use for time t?* If the answer is yes, then we can attain the target frequency $x^R_{\mathbf{b}}$ and maintain strict inequalities (15)–(20) iteratively throughout the entire n time slots.

To answer this question, consider type-23, as defined in (28). For s to be able to choose a $\mathbf{c}_t$ from TYPE_{23}, we must have $\mathsf{TYPE}_{23} \neq \emptyset$. Note that whether $\mathsf{TYPE}_{23} \neq \emptyset$ holds can be decided by comparing the sizes of $(A_3 \cap A_5)$ and $((A_3 \cap A_5) \cap (A_1 \cup A_4))$. Also note that $|A_3 \cap A_5| = q^{\mathrm{Rank}(A_3 \cap A_5)}$ and

$$\begin{aligned} &|(A_3 \cap A_5) \cap (A_1 \cup A_4)| \\ &\leq |A_3 \cap A_5 \cap A_1| + |A_3 \cap A_5 \cap A_4)| \\ &= q^{\mathrm{Rank}(A_3 \cap A_5 \cap A_1)} + q^{\mathrm{Rank}(A_3 \cap A_5 \cap A_4)}. \end{aligned}$$

Assuming $q \geq 3$, we thus have $\mathsf{TYPE}_{23} \neq \emptyset$ if and only if

$$\begin{aligned} &\mathrm{Rank}(A_3 \cap A_5) \\ &\quad > \max(\mathrm{Rank}(A_3 \cap A_5 \cap A_1), \mathrm{Rank}(A_3 \cap A_5 \cap A_4)). \end{aligned}$$

By (24)–(26) and Lemma 3, we can prove that

$$\begin{aligned} &\mathrm{Rank}(A_3 \cap A_5) - \mathrm{Rank}(A_3 \cap A_5 \cap A_1) \\ &= \mathrm{Rank}(S_1 \oplus \Omega_1) + \mathrm{Rank}(S_1 \oplus S_2) \\ &\quad - \mathrm{Rank}(S_1 \oplus S_2 \oplus \Omega_1) - \mathrm{Rank}(S_1) \end{aligned}$$

and

$$\begin{aligned} &\mathrm{Rank}(A_3 \cap A_5) - \mathrm{Rank}(A_3 \cap A_5 \cap A_4) \\ &= \mathrm{Rank}(S_1 \oplus S_2 \oplus \Omega_2) - \mathrm{Rank}(S_2 \oplus \Omega_2) \\ &= \mathrm{Rank}(A_7) - \mathrm{Rank}(A_4). \end{aligned}$$

The above analysis shows that when both $A_4 \subset A_7$ and (32) are strict inequalities, $\mathsf{TYPE}_{23} \neq \emptyset$ and choosing a $\mathbf{c}_t$ from TYPE_{23} is possible. By similar arguments, it can be proven that (i) Each FT is associated with a subset of inequalities of (30)–(35); and (ii) A FT is non-empty if and only if the inequalities in the corresponding subset are all strict. Therefore, when all 8 inequalities in (30)–(35) are strict, all 18 FTs are non-empty. Source s can thus freely select any FT with the goal of maintaining the target relative frequency $\{x^R_{\mathbf{b}}\}$. The central question is thus answered affirmatively.

VI. Further Discussion

To generalize the proof for arbitrary M, we simply need to consider 18^M different *joint coding types* since each coding vector $\mathbf{c}_t^{[m]}$ for the m-th input may belong to one of the 18 FTs. We can then follow similar analysis as in Section V. The final step is to notice from the resulting LP conditions that the dependence between any two input sub-channels has no effect on the overall capacity. This observation allows us to rewrite the conditions with the *marginal coding types* and reduce the number of x variables from 18^M to $18M$ as in Proposition 1. The proof for general M is thus complete.

We have used an LP solver to compute the LNC capacity region for various parameter values $p^{[m]}_{a_1 a_2}$ and in all our experiments the LNC capacity coincides with the outer bound in Lemma 2. We are currently proving the following conjecture.

Conjecture 1: The capacity outer bound in Lemma 2 holds even when the sub-channels are dependent, and it coincides with the LNC capacity for arbitrary M values.

Note that Proposition 1 has established the best LNC throughput. Even if Conjecture 1 is not true, any (R_1, R_2) outside the rate region of Proposition 1 can only be achieved by non-linear codes, which are less favorable in practice.

VII. Conclusion

This work has characterized the full LNC capacity of the 2-receiver MIMO broadcast PECs with channel output feedback. One future direction is to generalize the results for an arbitrary number of receivers. This work was supported in parts by NSF grants CCF-0845968 and CNS-0905331.

References

[1] P. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proc. 41st Annual Allerton Conf. on Comm., Contr., and Computing*. Monticello, IL, October 2003.

[2] T. M. Cover, "Comments on broadcast channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2524–2530, Oct. 1998.

[3] A. Dana and B. Hassibi, "The capacity region of multiple input erasure broadcast channels," in *Proc. IEEE Int'l Symp. Inform. Theory*. Adelaide, Australia, September 2005.

[4] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Multiuser broadcast erasure channel with feedback — capacity and algorithms," in *Proc. NetCoop*, 2010.

[5] L. Georgiadis and L. Tassiulas, "Broadcast erasure channel with feedback — capacity and algorithms," in *Proc. 5th Workshop on Network Coding, Theory, & Applications (NetCod)*. Lausanne, Switzerland, June 2009, pp. 54–61.

[6] D. Koutsonikolas, C.-C. Wang, Y. Hu, and N. Shroff, "FEC-based AP downlink transmission schemes for multiple flows: Combining the reliability and throughput enhancement of intra- and inter-flow coding," *Elsevier Performance Evaluation (PEVA)*, vol. 68, no. 11, November 2011.

[7] S.-Y. Li, R. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 371–381, February 2003.

[8] L. Ozarow and S. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 30, no. 4, July 1984.

[9] C.-C. Wang, "Capacity of 1-to-K broadcast packet erasure channels with channel output feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 957–988, February 2012.