On The Capacity Of 2-User 1-Hop Relay Erasure Networks — The Union of Feedback, Scheduling, Opportunistic Routing, and Network Coding

Wei-Cheng Kuo, Chih-Chun Wang; {wkuo, chhw}@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 2-user 1-hop relay networks, for which the sources, destinations, and the common relay are interconnected by broadcast packet erasure channels. In contrast with the existing results, this paper allows (i) transmission from a source being heard directly by its 2-hop-away destination, the so-called opportunistic routing scenario, (ii) instant channel status feedback among all network nodes, and (iii) per-slot scheduling decisions that are functions of the traffic loads and the past channel status. A new pair of inner and outer bounds is provided, and a condition is identified for the scenario in which the bounds coincide. Numerical experiments show that for commonly encountered scenarios, the gap between the inner and the outer bounds is less than 0.2%, which demonstrates the effectiveness of the proposed bounding techniques.

I. PROBLEM FORMULATION

For any positive integer \( M \), define \( [M] \triangleq \{1, \ldots, M\} \). We consider a 1-hop relay network with 2 source-destination pairs \((s_1, d_1)\) and \((s_2, d_2)\) and a common relay \( r \) that are interconnected by packet erasure channels (PECs). See Fig. 1(a) for illustration. Assume slotless transmission. Within an overall time budget of \( n \) time slots, \( s_i \) would like to convey \( nR_i \) packets \( \mathbf{W}_i \triangleq (W_{i,1}, \ldots, W_{i,nR_i}) \), \( j \in [nR_i] \), to \( d_i \) for all \( i \in \{1,2\} \). For each \( i \in \{1,2\} \), \( j \in [nR_i] \), the information packet \( W_{i,j} \) is chosen independently and uniformly randomly from \( GF(q) \).

For any time \( t \), consider a random 8-dimensional channel status vector for the entire network:

\[
\mathbf{Z}(t) = (Z_{s_1 \rightarrow d_1}(t), Z_{s_1 \rightarrow d_2}(t), Z_{s_2 \rightarrow r}(t), Z_{s_2 \rightarrow d_1}(t), Z_{s_2 \rightarrow r}(t), Z_{r \rightarrow d_1}(t), Z_{r \rightarrow d_2}(t)) \in \{0,1\}^8,
\]

where “0” and “1” represent erasure and successful reception, respectively. That is, when \( s_1 \) transmits a packet \( X_{s_1}(t) \in GF(q) \) in time slot \( t \), relay \( r \) receives \( Y_{s_1 \rightarrow r}(t) = X_{s_1}(t) \) if \( Z_{s_1 \rightarrow r}(t) = 1 \) and receives \( Y_{s_1 \rightarrow r}(t) = \ast \) if \( Z_{s_1 \rightarrow r}(t) = 0 \). For simplicity, we use \( Y_{s_1 \rightarrow r}(t) = X_{s_1}(t) \circ Z_{s_1 \rightarrow r}(t) \) as shorthand. To model interference, we further assume that only one node can be scheduled in each time slot. We use \( \sigma(t) \in \{s_1, s_2, r\} \) to denote the scheduling decision at time \( t \). For convenience, when \( s_1 \) is not scheduled at time \( t \), we simply set \( Y_{s_1 \rightarrow r}(t) = \ast \). As a result, the scheduling decision can be incorporated into the following expression of \( Y_{s_1 \rightarrow r}(t) \):

\[
Y_{s_1 \rightarrow r}(t) = X_{s_1}(t) \circ Z_{s_1 \rightarrow r}(t) \circ 1_{\{\sigma(t)=s_1\}}.
\]

Similar notation is used for all other received signals. For example, \( Y_{r \rightarrow d_2}(t) = X_r(t) \circ Z_{r \rightarrow d_2}(t) \circ 1_{\{\sigma(t)=r\}} \) is what \( d_2 \) receives from \( r \) in time \( t \), where \( X_r(t) \) is the packet sent by \( r \) in time \( t \).

We assume memoryless, stationary, erasure networks. Namely, we allow arbitrary joint distribution for the coordinates of \( Z(t) \) but assume that \( Z(t) \) is independently and identically distributed over the time axis \( t \). We also assume \( Z(t) \) is independent of the information messages \( \mathbf{W}_1 \) and \( \mathbf{W}_2 \).

The following notation turns out to be useful. We use brackets \([\bullet]\) to denote the collection from time 1 to \( t \). For example, \([\sigma, \mathbf{Z}, Y_{s_1 \rightarrow d_1}]_t \triangleq \{\sigma(t), \mathbf{Z}(t), Y_{s_1 \rightarrow d_1}(t) : \forall \tau \in [1, t]\} \). For any \( S \subseteq \{s_1, s_2, r\} \), \( T \subseteq \{r, d_1, d_2\} \), we define

\[
Y_{S \rightarrow T}(t) \triangleq \{Y_{s \rightarrow d}(t) : \forall s \in S, \forall d \in T\}.
\]

For example, \( Y_{\{s_1, r\} \rightarrow \{d_1, d_2\}}(t) \) is the collection of \( Y_{s_1 \rightarrow d_1}(t), Y_{s_1 \rightarrow d_2}(t), Y_{r \rightarrow d_1}(t), \) and \( Y_{r \rightarrow d_2}(t) \).

Given the traffic load \((R_1, R_2)\), a joint scheduling and network coding (NC) scheme is defined by \( n \) scheduling decision functions

\[
\forall t \in [n], \sigma(t) = f_{\sigma,t}(\mathbf{Z}_t^{-1}),
\]

and \( 3n \) encoding functions at \( s_1, s_2, \) and \( r \), respectively: For all \( t \in [n] \)

\[
X_{s_i}(t) = f_{\sigma,t}(\sigma(t), \mathbf{W}_i, \mathbf{Z}_t^{-1}), \forall i \in \{1,2\},
\]

\[
X_r(t) = f_{s,t}(\sigma(t), [\mathbf{Y}_{\{s_1, s_2\} \rightarrow r}, \mathbf{Z}_t^{-1}]),
\]

and \( 2 \) decoding functions at \( d_1 \) and \( d_2 \), respectively:

\[
\mathbf{W}_i = f_{d,t}(\{\sigma, \mathbf{Y}_{\{s_1, s_2, r\} \rightarrow d_i}, \mathbf{Z}_t^n\}), \forall i \in \{1,2\}.
\]

Namely, the scheduling decision at time \( t \) is based on the channel status of the entire network in time 1 to \( t-1 \). Encoding at \( s_i \) depends on the scheduling decision, the information messages, and past channel status. Encoding at \( r \) depends on the scheduling decision, what \( r \) received in the past, and past channel status. In the end, \( d_i \) decodes \( \mathbf{W}_i \) based on the past scheduling decisions, what \( d_i \) has received, and the past channel status of the entire network.

This setting models the scenario in which there is a dedicated low-rate control channel that can broadcast the scheduling decision \( \sigma(t) \) and the previous network status \( \mathbf{Z}(t-1) \) causally to all network nodes. The total amount of control information is less than \( (8 + \log_2(3)) \) bits per time slot, which is much smaller than the actual content of each packet \( \approx 10^4 \)
bits and thus can be easily implemented by buffering and piggybacking on the data packets.

**Definition 1:** Fix the distribution of \( Z(t) \). A rate vector \((R_1, R_2)\) is achievable if for any \( \epsilon > 0 \), there exists a joint scheduling and NC scheme with sufficiently large \( n \) and \( \text{GF}(q) \) such that

\[
\max_{i \in \{1, 2\}} \text{Prob}(W_i \neq W) < \epsilon.
\]

The capacity region is defined as the closure of all achievable rate vectors \((R_1, R_2)\).

**Remark 1:** In (1), the scheduling decision \( \sigma(t) \) does not depend on the information messages \( W_i \), which means that we prohibit the use of timing channels [1], [7]. Even when we allow the usage of timing channels, we conjecture that the overall capacity improvement is negligible. A heuristic argument is that each successful packet transmission gives \( \log_q(q) \) bits of information while the timing information (to transmit or not) gives roughly 1 bit of information. When focusing on sufficiently large \( \text{GF}(q) \), additional gain of timing information is thus likely to be absorbed in our timing-information-free capacity characterization.

**Remark 2:** In our setting, \( r \) is the only node that can mix packets from two different data sessions. Further relaxation such that \( s_1 \) and \( s_2 \) can hear each other and perform coding accordingly is beyond the scope of this work.

## II. Related Works

The capacity for the setting of multiple unicast sessions, termed the intersession network coding problem (INC), remains largely unsolved. Recently, [10] models the practical INC protocol [4] as a 1-hop relay network interconnected by PECs and studies the corresponding capacity region.

There are two major differences between the setting of this work and in [10]. First, a deterministic sequential scheduling policy is used in [10], which schedules nodes \( s_1, s_2, \) and \( r \) in strict order. Namely, once we finish transmission of \( s_2 \), the subsequent time slots can only be used for transmitting packets from \( r \). For comparison, our setting allows dynamically choosing the schedule \( \sigma(t) \) for each time slot \( t \). Second, in [10] no feedback is allowed when \( s_1 \) and \( s_2 \) transmit. More specifically, suppose jointly \( s_1 \) and \( s_2 \) take \( t_{s_1} + t_{s_2} \) time slots to finish transmission. Then only in the beginning of time \( (t_{s_1} + t_{s_2} + 1) \) are we allowed to send the channel status \( Z_{t_{s_1}+t_{s_2}} \) to \( r \). No further feedback is allowed until time \( n \), the end of overall transmission. For comparison, our setting allows constantly broadcasting network-wide channel status \( Z_{t_{s_1}+t_{s_2}} \) to \( s_1, s_2, \) and \( r \), as discussed in Section I. This setting thus includes the Automatic Repeat reQuest (ARQ) mechanism as a special case [3]. Broadcasting \( Z_{t_{s_1}+t_{s_2}} \) also eliminates the need of estimating/learning the reception status of the neighbors.

Our capacity region thus contains all achievable rates of the COPE protocol [4]. In Table I we use S1 to S3 to highlight the differences of the settings, and use R1 and R2 to highlight the main results.

Since we allow the schedule \( \sigma(t) \) to depend on the past reception status \( Z_{t_{s_1}+t_{s_2}} \), we also include any store-&-forward-based scheduling policies as special cases, such as the backpressure and the maximal weighted matching schemes (see [6] for references). Our results thus quantify the best achievable rates with jointly designed scheduling and coding policies.

**Remark:** From the relay’s perspective, the past channel output \( Y_{r \rightarrow \{d_1, d_2\}} \) is a function of \([X_r, \sigma, Z_{t_{s_1}+t_{s_2}}]^{t-1}\) that can be computed from \([Y_{\{s_1, s_2\} \rightarrow r}]^{t-2}\) and \([Z_{t_{s_1}+t_{s_2}}]^{t-1}\). As a result, our setting, particularly (3), also includes the setting of channel output feedback [3], [9] as a special example.

In Section III, we provide the outer/inner bounds for a 2-user 1-hop relay erasure network, and identify the condition for which the bounds coincide. Section IV provides the results of numerical experiments. Our results show that for scenarios commonly encountered in practice, the gap between the outer and inner bounds is less than .2%. The bounding techniques in this work thus effectively determine the joint scheduling-coding capacity. Section V concludes this paper.

## III. Main Results

All our results assume a fixed distribution of the channel status vector \( Z(t) \), which can be described by specifying the joint probability mass function for the \( 2^k \) possible outcomes.\(^1\)

In this work, we use \( p_{s \rightarrow s_1, T} \) to denote the probability that a packet sent from \( s \) is received by all nodes in \( T \) but not by any node in \( T_2 \) (may or may not by any node outside \( T_1 \cup T_2 \)). For example, \( p_{s \rightarrow A, T} \) denotes the probability that a packet sent from \( s_1 \) is received by \( d_1 \) and \( r \) but not by \( d_2 \). Additionally, we use \( p_{s \rightarrow T} \) to denote the probability that a packet sent by \( s \) is received by at least one node in \( T \). For example, \( p_{r \rightarrow \{d_1, d_2\}} \) is the probability that a packet sent by \( r \) is received by at least one node in \( \{d_1, d_2\} \). By definition, \( p_{s \rightarrow T} = p_{s \rightarrow r} \).

\(^1\)By allowing the coordinates of \( Z(t) \) to be correlated, our setting also models the scenario in which destinations \( d_1 \) and \( d_2 \) are situated in the same physical node and thus have perfectly correlated channel success events.
A. Capacity Outer Bound

Proposition 1: For any achievable \((R_1, R_2)\), there exist three non-negative scalars \(t_{s_1}, t_{s_2}, t_r\) satisfying
\[
t_{s_1} + t_{s_2} + t_r \leq 1
\]  
\(\forall i \in \{1, 2\}, \quad R_i \leq t_s p_{s_i;\{d_i, r\}}\)  
\[
\frac{(R_1 - t_{s_1} p_{s_1;\{d_1, d_2\}})^+}{p_{r;\{d_1, d_2\}}} + \frac{(R_2 - t_{s_2} p_{s_2;\{d_1, d_2\}})^+}{p_{r;\{d_1, d_2\}}} \leq t_r,  
\]
\[
\frac{(R_1 - t_{s_1} p_{s_1;\{d_1, d_2\}})^+}{p_{r;\{d_1, d_2\}}} + \frac{(R_2 - t_{s_2} p_{s_2;\{d_1, d_2\}})^+}{p_{r;\{d_1, d_2\}}} \leq t_r,  
\]
where \((\cdot)^+ \triangleq \max(0, \cdot)\) is the projection to non-negative reals.

Proof: For any joint scheduling and NC scheme, we choose \(t_{s_i}\) (resp. \(t_r\)) as the normalized expected number of time slots for which \(s_i\) (resp. \(r\)) is scheduled. As a result, (5) must hold. Without loss of generality, in the subsequent discussion we also assume that both the \(n\) and GF\((q)\) used by the joint scheduling and NC scheme are sufficiently large.

The intuition behind (6) is the following observation. Only for the time instants with \(\sigma(t) = s_i\) can \(s_i\) convey information directly to \(d_i\) or indirectly through \(r\). Therefore, the achievable rate \(R_i\) is upper bounded by \(t_{s_i} p_{s_i;\{d_i, r\}}\), the expected number of distinct packets received by either \(d_i\) or \(r\).

We can also prove (6) formally.
\[
I(W_1; W_1) \leq I(W_1; Y_{s_1 \to \{d_1, r\}}, Y_{s_2 \to \{d_2, r\}}, Y_r \to \{d_1, d_2\}, Z, \sigma)^n  
\]
\[
= \sum_{i=1}^n I(W_1; Y_{s_i \to \{d_i, r\}}(t), Z(t-1), \sigma(t)|B_t)  
\]
\[
= \sum_{i=1}^n I(W_1; Y_{s_i \to \{d_i, r\}}(t), Z(t-1), \sigma(t)|B_t)  
\]
\[
+ I(W_1; Y_{s_2 \to \{d_2, r\}}(t)) Y_{s_1 \to \{d_1, r\}}(t)|B_t)  
\]
where \(B_t \triangleq (Y_{s_1 \to \{d_1, r\}}, \sigma(t)|B_t)^n\); consider the summation in (10). Since the first term is non-zero only when \(\sigma(t) = s_1\), it is upper bounded by
\[
I(W_1; Y_{s_1 \to \{d_1, r\}}(t), Z(t-1), \sigma(t)|B_t) \leq \sum_{b_t} \text{Prob}(B_t = b_t) \cdot \text{Prob}(\sigma(t) = s_1|B_t = b_t) \cdot p_{s_1;\{d_1, r\}}  
\]
\[
= \text{Prob}(\sigma(t) = s_1) \cdot p_{s_1;\{d_1, r\}}  
\]
Hence the summation of the first term in (10) is upper bounded by
\[
\sum_{i=1}^n \text{Prob}(\sigma(t) = s_1) \cdot p_{s_1;\{d_1, r\}} = nt_s p_{s_1;\{d_1, r\}}  
\]
By the independence between \(W_1\) and \(Y_{s_2 \to \{d_2, r\}}(t)\) conditioning on \(Y_{s_1 \to \{d_1, d_2\}}(t)\) and \(B_t\), the second term of the summation in (10) is zero. On the other hand, Fano’s inequality gives us
\[
I(W_1; \hat{W}_1) \geq nR_1(1 - \epsilon) - H(\epsilon)  
\]
Combining (12) and (13), we have
\[
R_1(1 - \epsilon) - \frac{H(\epsilon)}{n} \leq t_s p_{s_1;\{d_1, r\}}  
\]
Let \(\epsilon \to 0\) and therefore (14) implies (6) for \(i = 1\). With symmetric derivation, we can derive (6) for \(i = 2\).

We prove (7) by similar techniques as used in [2], [8]. That is, we first add an auxiliary pipe that sends all information available at \(d_2\) directly to \(d_1\). Then the decoding function (see (4)) at \(d_1\) becomes
\[
\hat{W}_1 = f_{d_1}([\sigma, Y_{s_1, s_2, r} \to \{d_1, d_2\}, Z])  
\]
For any \(t \in [n]\), define
\[
U(t) \triangleq (W_2, [Y_{s_1, s_2} \to \{d_1, d_2\}], Y_r \to \{d_1, d_2\}, \sigma(t), \sigma(t-1), |Z|^{t-2})  
\]
By taking logarithm with base \(q\) and by Fano’s inequality, for some \(\epsilon_1 > 0\) that goes to \(0\) as \(\epsilon \to 0\), we must have
\[
n R_1 = H(W_1; W_2) \leq I(W_1; \hat{W}_1) W_2 + n \epsilon_1  
\]
\[
\leq I(W_1; \sigma, Y_{s_1, s_2, r} \to \{d_1, d_2\}, Z)|W_2| + n \epsilon_1  
\]
\[
= n \epsilon_1 + \sum_{t=1}^n I(W_1; Y_{s_1, s_2, r} \to \{d_1, d_2\}(t), \sigma(t), Z(t-1)|A_t)  
\]
\[
\leq n \epsilon_1 + \sum_{t=1}^n I(W_1; Y_{s_1 \to \{d_1, d_2\}(t)|A_t)  
\]
\[
+ I(W_1; Y_{s_2 \to \{d_2, r\}}|Y_{s_1 \to \{d_1, d_2\}(t)}, A_t)  
\]
\[
+ I(W_1; Y_{r \to \{d_1, d_2\}}|Y_{s_1 \to \{d_1, d_2\}(t)}, A_t)  
\]
\[
\leq \sum_{t=1}^n I(Y_{s_1 \to \{d_1, d_2\}(t)}, Y_{s_2 \to \{d_2, r\}}(t)|A_t)  
\]
\[
+ I(Y_{r \to \{d_1, d_2\}}|Y_{s_1 \to \{d_1, d_2\}(t)}, A_t)  
\]
\[
+ I(Y_{r \to \{d_1, d_2\}}|Y_{s_1 \to \{d_1, d_2\}(t)}, A_t)  
\]
where \(A_t \triangleq (W_2, [Y_{s_1, s_2, r} \to \{d_1, d_2\}, \sigma(t-1), |Z|^{-2}); (16)\) follows from (15); and (17) follows from the chain rule.²

Consider the summation in (17). Since the first term is non-zero only when \(\sigma(t) = s_1\), it is upper bounded by
\[
\sum_{t=1}^n \text{Prob}(\sigma(t) = s_1) \cdot p_{s_1;\{d_1, d_2\}} = nt_s p_{s_1;\{d_1, d_2\}}  
\]
By the independence between \(W_1\) and \(Y_{s_2 \to \{d_2, r\}}(t)\) conditioning on \(Y_{s_1 \to \{d_1, d_2\}}(t)\) and \(A_t\), the second term of the summation is zero. Since mutual information is non-negative, we now have
\[
(nR_1 - nt_s p_{s_1;\{d_1, d_2\}} - n \epsilon_1)^+  
\]
\[
\leq \sum_{t=1}^n I(W_1; Y_{r \to \{d_1, d_2\}}(t), \sigma(t), Z(t-1)|U(t))  
\]
\[
\leq \sum_{t=1}^n I(X_r(t); Y_{r \to \{d_1, d_2\}}(t), \sigma(t), Z(t-1)|U(t))  
\]
\[
\leq \sum_{t=1}^n (\log_q(3) + \log_q(2)^8)  
\]
\[
+ I(X_r(t); Y_{r \to \{d_1, d_2\}}(t)|U(t), \sigma(t), Z(t-1))  
\]
where (18) follows from the fact that conditioning on \(U(t)\), \(W_1 \rightarrow X_r(t) \rightarrow (Y_{r \to \{d_1, d_2\}}(t), \sigma(t), Z(t-1))\) form a²

²For simplicity, in (9) and (17), we ignore the last term involving \(Z(n)\) since it is a negligible term when a sufficiently large \(n\) is used.
Markov chain; and (19) follows from the fact that the cardinality of \( \sigma(t) \) and \( Z(t-1) \) are 3 and \( 2^8 \), respectively.

By Fano’s inequality, for some \( \epsilon_2 > 0 \) that goes to 0 as \( \epsilon \to 0 \), with similar steps as in (17), we can also prove that

\[
nR_2 = H(W_2) \\
\leq nI(W_2; [Y_{s_1}, s_2, r] \rightarrow d_2, Z, \sigma_1^n) + n\epsilon_2
\]

\[
= n\epsilon_2 + \sum_{t=1}^{n} I(W_2; Y_{s_1, s_2} \rightarrow d_2(t), Z(t-1), \sigma(t)|D_t)
\]

\[
= n\epsilon_2 + \sum_{t=1}^{n} I(W_2; Y_{s_1, s_2} \rightarrow d_2(t), Z(t-1), \sigma(t)|W_2, Y_{s_1, s_2} \rightarrow d_2(t), D_t)
\]

\[
= n\epsilon_2 + \sum_{t=1}^{n} I(W_2; Y_{s_2} \rightarrow d_2(t), Z(t-1), \sigma(t)|W_2, Y_{s_1, s_2} \rightarrow d_2(t), D_t)
\]

(20)

where \( D_t = (Y_{s_1, s_2} \rightarrow d_2(t), \sigma]\); and (20) follows from the independence between \( W_2 \) and \( Y_{s_1, s_2} \) conditioning on \( D_t \). Notice that the summation of the first term in (20) can be upper bounded by

\[
\sum_{t=1}^{n} \text{Prob}(\sigma(t) = s_1) \cdot P_{s_1} = nt_{s_1}p_{s_1} = d_2
\]

(21)

Hence

\[
(20) \leq n\epsilon_2 + nt_{s_2}p_{s_2}d_2
\]

\[
+ \sum_{t=1}^{n} (H(Y_{r} \rightarrow d_2(t), Z(t-1), \sigma(t)|Y_{s_1, s_2} \rightarrow d_2(t), D_t)
\]

\[
- H(Y_{r} \rightarrow d_2(t), Z(t-1), \sigma(t)|W_2, Y_{s_1, s_2} \rightarrow d_2(t), D_t)
\]

\[
\leq n\epsilon_2 + nt_{s_2}p_{s_2}d_2
\]

\[
+ \sum_{t=1}^{n} (H(Y_{r} \rightarrow d_2(t), Z(t-1), \sigma(t))
\]

\[
- H(Y_{r} \rightarrow d_2(t), Z(t-1), \sigma(t)|U(t))
\]

\[
= n\epsilon_2 + nt_{s_2}p_{s_2}d_2 + \sum_{t=1}^{n} I(U(t); Y_{r} \rightarrow d_2(t), Z(t-1), \sigma(t))
\]

(23)

where (23) follows from the property of entropy; since mutual information is non-negative, we now have

\[
(nR_2 - nt_{s_2}p_{s_2}d_2)^+
\]

\[
\leq n\epsilon_2 + \sum_{t=1}^{n} I(U(t); Y_{r} \rightarrow d_2(t), Z(t-1), \sigma(t))
\]

(24)

\[
\leq n\epsilon_2 + \sum_{t=1}^{n} \log_q(3) + \log_q(2^8)
\]

\[
+ I(U(t); Y_{r} \rightarrow d_2(t)|\sigma(t), Z(t-1))
\]

(25)

where (26) follows from the fact that the cardinality of \( \sigma(t) \) and \( Z(t-1) \) are 3 and \( 2^8 \), respectively.

Now we can define the time sharing random variable \( Q_1 \), where \( Q_1 = t \in [n] \) with probability \( \frac{1}{n} \), and define \( Q_1 = (Q_1, Z(\lambda)) \). By using \( q \to \infty \) to absorb the \( \log_q(3) + \log_q(2^8) \) term, we can rewrite (19) and (26) as

\[
(R_1 - t_{s_1}p_{s_1};(d_1, d_2))^+
\]

\[
\leq \epsilon_1 + \sum_{t_{q_1}=1}^{n} I(X(t_{q_1}); Y_{r} \rightarrow d_2(t_{q_1})|Q(t_{q_1})
\]

\[
U(t_{q_1}), Z(t_{q_1} - 1), \lambda = q_1, \sigma(t_{q_1})
\]

\[
= I(X(t_{q_1}); Y_{r} \rightarrow (d_1, d_2)(Q(t_{q_1}))|U(t_{q_1}), Q(t_{q_1})) + \epsilon_1
\]

(27)

\[
(R_2 - t_{s_2}p_{s_2};d_2)^+
\]

\[
\leq \epsilon_2 + \sum_{t_{q_1}=1}^{n} I(U(t_{q_1}); Y_{r} \rightarrow d_2(t_{q_1}))
\]

\[
Z(t_{q_1} - 1), \lambda = q_1, \sigma(t_{q_1})
\]

\[
= I(U(t_{q_1}); Y_{r} \rightarrow (d_1, d_2)(Q(t_{q_1}))|U(t_{q_1}), Q(t_{q_1})) + \epsilon_2
\]

(28)

Notice that only if \( \sigma(t_{q_1}) = r \), the mutual information in (27) and (28) not equals to zero (i.e. \( Y_{r} \rightarrow d_1(Q(t_{q_1})) \) and \( Y_{r} \rightarrow (d_1, d_2)(Q(t_{q_1})) \) is non-empty). And since \( Q_1 \) is uniformly distributed over \([n] \), \( \text{Prob}(\sigma(t_{q_1}) = r) = t_r \). Define \( X_r \triangleq X_r(Q(t_{q_1})), Y_{r} \rightarrow d_1 \triangleq Y_{r} \rightarrow (d_1, d_2)(Q(t_{q_1})), Y_{r} \rightarrow d_2 \triangleq Y_{r} \rightarrow (d_1, d_2)(Q(t_{q_1}), U \triangleq U(t_{q_1})) \), then we have

\[
(R_1 - t_{s_1}p_{s_1};(d_1, d_2))^+ \leq t_r I(X_r; Y_{r} \rightarrow (d_1, d_2)|U, Q) + \epsilon_1
\]

(29)

\[
(R_2 - t_{s_2}p_{s_2};d_2)^+ \leq t_r I(U; Y_{r} \rightarrow d_2|Q) + \epsilon_2
\]

(30)

(29) and (30) with \( \epsilon \to 0 \) imply that for any achievable \( (R_1, R_2) \), its derived vector \( (R_1 - t_{s_1}p_{s_1};(d_1, d_2))^+, (R_2 - t_{s_2}p_{s_2};d_2)^+ \) must lie in the convex hull of

\[
C \triangleq \{(r_1, r_2) : r_1 \leq t_r I(X_r; Y_{r} \rightarrow (d_1, d_2)|U),
\]

\[
r_2 \leq t_r I(U; Y_{r} \rightarrow d_2), \text{ for all auxiliary RVs } U\}
\]

(31)

where \( U \mapsto X_r \mapsto Y_{r} \rightarrow (d_1, d_2) \) form a Markov chain and the conditional distribution \( P_{Y_{r} \rightarrow (d_1, d_2)|X_r} \) follows the same erasure distribution as the PEC faced by \( r \) (conditioning on node \( r \) is scheduled). Equivalently, we can define the capacity function

\[
\forall \lambda > 0,
\]

\[
C(\lambda) = \max_{p(U)p(x_r|U)} \{ t_r I(U; Y_{r} \rightarrow d_2) + \lambda t_r I(X_r; Y_{r} \rightarrow (d_1, d_2)|U) \}
\]

(32)

We notice that (31) has the same form as the feedback-free erasure channel capacity with success probabilities \( p_r(d_1, d_2) \) and \( p_r(d_2) \) and overall time budget \( nt_r \) slots. We can thus have following lemmas

**Lemma 1:** If \( \{(R_1 - t_{s_1}p_{s_1};(d_1, d_2))^+, (R_2 - t_{s_2}p_{s_2};d_2)^+\} \) is in \( C \), then

\[
\frac{(R_1 - t_{s_1}p_{s_1};(d_1, d_2))^+ + (R_2 - t_{s_2}p_{s_2};d_2)^+}{p_r(d_1, d_2)} \leq t_r
\]
Proof: Notice that $U \rightarrow X_i \rightarrow Y_{r \rightarrow \{d_1, d_2\}} \rightarrow Y_{r \rightarrow d_2}$ forms a Markov chain. Then follows from [11], for the distribution $p(u)p(x|u)$ which maximizes the capacity function $C(\lambda)$ for some $\lambda > 0$, we can assume $\|d\| = \min\{\|X_i\|, \|Y_{r \rightarrow d_1} \cup Y_{r \rightarrow d_2}\|, \|Y_{r \rightarrow d_2}\|\} = \|X_i\|$. Let

$$E = \begin{cases} 0 & \text{if } u = x_r; \\ 1 & \text{if } u \neq x_r, \end{cases}$$

with $\text{Prob}(E = 1) = \beta$ and $\text{Prob}(E = 0) = 1 - \beta$, then we have

$$\frac{(R_1 - t_{s_1}p_{s_1;\{d_1, d_2\}}^+) + (R_2 - t_{s_2}p_{s_2;\{d_1, d_2\}}^+)}{p_{r;\{d_1, d_2\}}} \leq t_r.$$

Lemma 2: If $((R_1 - t_{s_1}p_{s_1;\{d_1, d_2\}}^+), (R_2 - t_{s_2}p_{s_2;\{d_1, d_2\}}^+))$ is in the convex hull of $C$, then

$$\frac{(R_1 - t_{s_1}p_{s_1;\{d_1, d_2\}}^+)}{p_{r;\{d_1, d_2\}}} + \frac{(R_2 - t_{s_2}p_{s_2;\{d_1, d_2\}}^+)}{p_{r;\{d_2\}}} \leq t_r.$$

Proof: Since $((R_1 - t_{s_1}p_{s_1;\{d_1, d_2\}}^+), (R_2 - t_{s_2}p_{s_2;\{d_1, d_2\}}^+))$ is in the convex hull of $C$, there exists finite number of pairs $(r_{1,1}, r_{1,2}), (r_{2,1}, r_{2,2}), ..., (r_{N,1}, r_{N,2})$ in $C$. And $N$ nonnegative real numbers $c_1, c_2, ..., c_N$ satisfying $\sum_{i=1}^{N} c_i = 1$ such that $((R_1 - t_{s_1}p_{s_1;\{d_1, d_2\}}^+), (R_2 - t_{s_2}p_{s_2;\{d_1, d_2\}}^+)) = c_1(r_{1,1}, r_{1,2}) + c_2(r_{2,1}, r_{2,2}) + ... + c_N(r_{N,1}, r_{N,2}).$

Follows from Lemma 1 and the fact that $\sum_{i=1}^{N} c_i = 1$, we get

$$\frac{(R_1 - t_{s_1}p_{s_1;\{d_1, d_2\}}^+)}{p_{r;\{d_1, d_2\}}} + \frac{(R_2 - t_{s_2}p_{s_2;\{d_1, d_2\}}^+)}{p_{r;\{d_2\}}} \leq t_r.$$

Finally, the following lemma gives us the capacity outer bound

Lemma 3: If $((R_1 - t_{s_1}p_{s_1;\{d_1, d_2\}}^+), (R_2 - t_{s_2}p_{s_2;\{d_1, d_2\}}^+))$ is in the closure of the convex hull of $C$, then

$$\frac{(R_1 - t_{s_1}p_{s_1;\{d_1, d_2\}}^+)}{p_{r;\{d_1, d_2\}}} + \frac{(R_2 - t_{s_2}p_{s_2;\{d_1, d_2\}}^+)}{p_{r;\{d_2\}}} \leq t_r.$$

Proof: Since $((R_1 - t_{s_1}p_{s_1;\{d_1, d_2\}}^+), (R_2 - t_{s_2}p_{s_2;\{d_1, d_2\}}^+))$ is in the closure of the convex hull of $C$, there exists a sequence of pairs $(r_{i,1}, r_{i,2})$ in the convex hull of $C$, $i = 1, 2, 3, ..., \lim_{i \to \infty} (r_{i,1}, r_{i,2}) = (R_1 - t_{s_1}p_{s_1;\{d_1, d_2\}}^+), (R_2 - t_{s_2}p_{s_2;\{d_1, d_2\}}^+)$. From Lemma 2, for any $i = 1, 2, 3, ...$

$$\frac{r_{i,1}}{p_{r;\{d_1, d_2\}}} + \frac{r_{i,2}}{p_{r;\{d_2\}}} \leq t_r.$$

Taking the limit as $i \to \infty$ on both sides, we get

$$\frac{(R_1 - t_{s_1}p_{s_1;\{d_1, d_2\}}^+)}{p_{r;\{d_1, d_2\}}} + \frac{(R_2 - t_{s_2}p_{s_2;\{d_1, d_2\}}^+)}{p_{r;\{d_2\}}} \leq t_r.$$
In the beginning of time $t$, arbitrarily choose one $v \in Q^{[1]}_{r,d_1}$. $s_i$ then sends a coded symbol $vW^T_1$, where $W_1$ is the column vector of all $s_1$ information packets.

In the end of time $t$, let $R_t$ denote the set of destinations successfully receiving $vW^T_1$. We have two cases depending on $R_t$: Case 1: $d_1 \in R_t$; remove $v$ from $Q^{[1]}_i$ and add it to $Q^{[1]}_i$. We use $Q^{[1]}_i \leftarrow Q^{[1]}_i$ as shorthand; Case 2: $d_1 \notin R_t$ and $R_t \neq \emptyset$: do $Q^{[1]}_i \leftarrow Q^{[1]}_i$, $R_t \leftarrow R_t \setminus \{d_1, d_2, r\}$.

4: end while

**Phase 1.2:** ($s_1$ sends packets that benefit $\{r, d_1\}$ while promoting overhearing at $d_2$.)

1: while $|Q^{[1]}_{r,d_1,d_2}^{[1]}| > 0$ do

2: In the beginning of time $t$, arbitrarily choose one $v_2 \in Q^{[1]}_{r,d_2}$. and one $v_r \in Q^{[1]}_{r,d_2}$. Set $v = c_2v_2 + c_rv_r$ where $c_2$ and $c_r$ are chosen independently and uniformly randomly from $GF(q)$. $s_1$ then sends $vW^T_1$.

3: In the end of time $t$, there are five possible actions: Action 1: $Q^{[1]}_{r,d_2} \leftarrow Q^{[1]}_{r,d_2}$; Action 2: $Q^{[1]}_{r,d_2} \leftarrow Q^{[1]}_{r,d_2}$; Action 3: $Q^{[1]}_{r,d_2} \leftarrow Q^{[1]}_{r,d_2}$; Action 4: Remove $v_2$ from $Q^{[1]}_{r,d_2}$ and add $v$ to $Q^{[1]}_i$. We use $Q^{[1]}_{r,d_2} \leftarrow Q^{[1]}_{r,d_2}$ as shorthand; Action 5: $Q^{[1]}_{r,d_2} \leftarrow Q^{[1]}_{r,d_2}$.

4: Let $R_t$ denote the receiving set. We then have six cases depending on $R_t$: Case 1: $\{d_1, d_2\} \subseteq R_t$; do Actions 1 and 4; Case 2: $R_t = \{d_1, r\}$: if $v_r \neq 0$, do Actions 2 and 5, else do Action 3; Case 3: $R_t = \{d_1\}$: do Action 4; Case 4: $R_t = \{d_2, r\}$: do Actions 1 and 2; Case 5: $R_t = \{d_2\}$: do Action 1; Case 6: $R_t = \{r\}$, do Action 2.

5: end while

**Phase 1.3:** ($s_1$ maximizes overhearing by retransmitting packets heard by $r$ but not by $d_1$.)

1: while $s_1$ has not used up its allocated time budget $nt_{s_1}$ time slots do

2: In the beginning of time $t$, arbitrarily choose one $v_r \in Q^{[1]}_{r,d_2}$ and one $v_{d_2} \in Q^{[1]}_{r,d_2}$. If $v_r \neq 0$, $s_1$ sends $v_r W_1$; else $s_1$ sends $v_{d_2} W^T_1$.

3: In the time of time $t$, there are two cases depending on $R_t$: Case 1: $d_1 \in R_t$: if $v_r \neq 0$, do $Q^{[1]}_{r,d_2,d_2} \leftarrow Q^{[1]}_{d_1}$. else do $Q^{[1]}_{r,d_2,d_2} \leftarrow Q^{[1]}_{d_1}$; Case 2: $d_1 \notin R_t$ but $d_2 \in R_t$: do $Q^{[1]}_{r,d_2,d_2} \leftarrow Q^{[1]}_{d_1,d_2}$.

4: end while

**Phases 2.1 to 2.3** are symmetric versions of Phases 1.1 to 1.3 with reversed roles of ($s_1, d_1$) and ($s_2, d_2$).

**Phase 3.1:** ($r$ sends $s_1$-packets that are not heard by $d_2$.)

1: while $|Q^{[1]}_{r,d_1}^{[1]}| > 0$ do

2: In the beginning of time $t$, arbitrarily choose one $v \in Q^{[1]}_{r,d_1,d_2}$ and $r$ sends $vW^T_1$.

3: In the end of time $t$, there are two cases depending on $R_t$: Case 1: $d_1 \in R_t$: do $Q^{[1]}_{r,d_1,d_2} \leftarrow Q^{[1]}_{d_1}$; Case 2: $R_t = \{d_2\}$: do $Q^{[1]}_{r,d_1,d_2} \leftarrow Q^{[1]}_{d_2}$.

4: end while

**Phase 3.2** is symmetric to Phase 3.1 with reversed roles of $d_1$ and $d_2$.

**Phase 3.3:** ($r$ mixes observed $s_1$- and $s_2$-packets.)

1: while $\max \{ |Q^{[1]}_{r,d_1}^{[1]}|, |Q^{[1]}_{r,d_2}^{[1]}| \} > 0$ do

2: In the beginning of time $t$, arbitrarily choose one $v_1 \in Q^{[1]}_{r,d_1}$. and one $v_2 \in Q^{[1]}_{r,d_2}$. $r$ sends $v_1 W^T_1 + v_2 W^T_2$.

3: In the end of time $t$, for $i = 1, 2$ and $j \neq i$, if $d_i \in R_t$, then do $Q^{[1]}_{r,d_i} \leftarrow Q^{[1]}_{d_i}$.

4: end while

The intuition of the proposed scheme is as follows. Claim(i) For any $i \neq j$, queues $Q^{[1]}_i$, $Q^{[1]}_{r,d_1,d_2}$, $Q^{[1]}_{r,d_2}$, $Q^{[1]}_i$, and $Q^{[1]}_0$ store the $s_i$-packets (vectors) that are known to $d_i$, known to $r$ but not to $\{d_1, d_2\}$, known to $d_j$ but not to $\{d_1, r\}$, known to $\{d_j, r\}$ but not known to $d_i$, and known to no nodes, respectively. Since each vector removal from a queue is company by a vector addition to another queue, we have claim(ii) for any time $t$, the above 5 queues contain totally $nR_t$ vectors. By the random linear network coding (RLNC) arguments as used in [9], we can prove that claim(iii) with sufficiently large $GF(q)$, these $nR_t$ vectors remain linearly independent and form the basis of the information space of $s_i$.

The main coding parts are in Phases 1.2, 2.2, and 3.3. For example, the goal of Phase 1.2 is two-fold. Firstly, $s_1$ would like to send packets that are beneficial to both $\{d_1, r\}$ so that if $d_1$ hears it, $d_1$ can obtain new information; or if $r$ hears it, $r$ can relay it to $d_1$ in the subsequent phase. One such candidate is those packets in $Q^{[1]}_{r,d_1,d_2}$, packets that have not been heard by any of $\{d_1, r\}$. Secondly, $s_1$ would also like to maximize $Q^{[1]}_{r,d_1,d_2}$ since this queue contains those $s_1$-packets at $r$ that are overheard by $d_2$ and thus can later be mixed at $r$ with those $s_2$-packets overheard by $d_1$ (as described in Line 2 of Phase 3.3). To promote overhearing, a good candidate is to send a packet in $Q^{[1]}_{r,d_1,d_2}$ so that if $d_2$ receives it, that packet is now heard by both $\{d_2, r\}$. We then use RLNC to achieve these two goals simultaneously and send a linear sum of the packets from $Q^{[1]}_{r,d_1,d_2}$ and $Q^{[1]}_{r,d_2}$.

In addition of (i) to (iii), proving the correctness of our construction also consists of proving the following statements for sufficiently large $n$: claim(iv) We can finish Phases 1.1 and 2.2 of $s_i$ within the allocated $nt_{s_i}$ time slots for $i = 1, 2$; claim(v) We can finish Phases 3.1, 3.2 and 3.3 of $r$ within the allocated $nt_r$ time slots; claim(vi) After Phase 3.3, we have $Q^{[1]}_{r,d_i} = nR_t$ for $i = 1, 2$. Then by (i) and (iii), all basis vectors in $Q^{[1]}_{d_i}$ are known to $d_i$. The information at $s_i$ can thus be successfully decoded at $d_i$.

**Proof of claim(i):** For $i \in \{1, 2\}$, $j = 3 - i$, we can see that every vector in $Q^{[1]}_{d_i}$ is indeed received by $d_i$. In Phase 1.1, we also notice that if the vector is not yet received by any of $\{d_1, d_2, r\}$, then it is in $Q^{[1]}_0$. For $Q^{[1]}_{r,d_1,d_2}$ in Phase 1.1, if the vector is in $Q^{[1]}_{r,d_1,d_2}$, then it is indeed received by $r$ only. Notice that in Phase 1.2, every time we move out the vector, $v_r$ from $Q^{[1]}_{r,d_1,d_2}$, it is indeed received by $d_1$ or $v_r$ is received by $d_1$ and...
r receives \(v_2\) (hence we do not need \(v_i\), anymore). Similar argument can be applied on Phases i.3 and 3.i. Hence \(Q^{[i]}_{r_1d_1,d_2}\) contains the vectors which are received by \(r\) only. Similar argument can be applied on \(Q^{[i]}_{r_1d_1,d_2}\). For \(Q^{[i]}_{r_1d_1,d_2}\), follows from the actions in the scheme, we notice that every time we put one vector into \(Q^{[i]}_{d_1r_1}\), both \(r\) and \(d_j\) knows the vector.

Proof of claim(iv): Since we assume \(n\) is large, in the following discussion, we always implicitly use Large Number Law. For \(i \in \{1,2\}\), \(j = 3 - i\), let \(N_{i,1}\) and \(N_{i,2}\) denotes the number of required time slots to finish Phase i.1 and Phase i.2, respectively. In Phase i.1, notice that if \(R_i \neq \emptyset\), then we move out one vector from \(Q^{[i]}_{0}\). Since initially, there are \(nR_i\) vectors in \(Q^{[i]}_{0}\), hence

\[
N_{i,1} = \frac{nR_i}{p_{s_1};(d_1,d_2,r)}
\]

In Phase i.2, notice that in Case 1, 2, 3, 4, and 6, we move out one vector from \(Q^{[i]}_{d_1d_2}\) and the probability that any of Case 1, 2, 3, 4, and 6 happens is \(p_{s_1};(d_1,r)\). Since after Phases i.1, there are \(\frac{nR_i}{p_{s_1};(d_1,d_2,r)}\) vectors in \(Q^{[i]}_{d_1d_2}\). Hence

\[
N_{i,2} = \frac{nR_i}{p_{s_1};(d_1,d_2,r)} p_{s_1};(d_1,d_2,r) \cdot p_{s_1};(d_1,r)
\]

It remains to show that \(N_{i,1} + N_{i,2} \leq nt_{s_1}\)

\[
N_{i,1} + N_{i,2} = \frac{nR_i}{p_{s_1};(d_1,d_2,r)} + \frac{nR_i}{p_{s_1};(d_1,d_2,r)} p_{s_1};(d_1,d_2,r) \cdot p_{s_1};(d_1,r) < nt_{s_1}
\]

where (39) follows from the fact that \(p_{s_1};(d_1,r) + p_{s_1};(d_1,d_2,r) = p_{s_1};(d_1,d_2,r)\) and (34).

Proof of claim(v): Let \(N_{3,1}\), \(N_{3,2}\), and \(N_{3,3}\) be the number of required time slots to finish Phase 3.1, 3.2, and 3.3, respectively. To show \(N_{3,1} + N_{3,2} + N_{3,3} \leq nt_{s_1}\), we first need to quantify (i) \(Q^{[1]}_{r_1d_1d_2}\) at the beginning of Phase 3.1, (ii) \(Q^{[2]}_{r_1d_1d_2}\) at the beginning of Phase 3.2, and (iii) \(Q^{[3]}_{r_1d_1d_2}\) at the beginning of Phase 3.3 and (iv) \(Q^{[4]}_{r_1d_1d_2}\) at the beginning of Phase 3.3.

We first discuss the first quantity. Notice that after Phase 1.1,

\[
Q^{[1]}_{r_1d_1d_2} = N_{1,1}p_{s_1};r_1d_2 = \frac{nR_1}{p_{s_1};(d_1,d_2,r)} p_{s_1};r_1d_2
\]

Also notice that in Phase 1.2, if any of Case 1, 2, 4, and 5 happens, then we move out one vector from \(Q^{[1]}_{r_1d_1d_2}\). The probability that any of Case 1, 2, 4, and 5 happens is \(p_{s_1};(d_1,d_2) - p_{s_1};(d_1,d_2,r)\). Hence after Phase 1.2, we move out

\[
N_{1,2}(p_{s_1};(d_1,d_2) - p_{s_1};(d_1,d_2,r)) = \frac{nR_1}{p_{s_1};(d_1,d_2,r)} p_{s_1};(d_1,d_2,r) - p_{s_1};(d_1,d_2,r)
\]

The remaining time slot for Phase 1.3 is

\[
nt_{s_1} - N_{1,1} - N_{1,2} = nt_{s_1} - \frac{nR_1}{p_{s_1};(d_1,r)}
\]

In Phase 1.3, if any of \(d_1\) and \(d_2\) in \(R_i\), then we move out one vector from \(Q^{[1]}_{r_1d_1d_2}\). Hence after Phase 1.3, we move out

\[
(nR_1 - nt_{s_1})p_{s_1};(d_1,d_2) = (\frac{nR_1}{p_{s_1};(d_1,d_2,r)} p_{s_1};(d_1,d_2) - p_{s_1};(d_1,d_2)r)
\]

With (40), (41), and (43), the first quantity is

\[
\frac{nR_1}{p_{s_1};(d_1,d_2,r)} p_{s_1};(d_1,d_2) - p_{s_1};(d_1,d_2)r
\]

Similarly, we have the second quantity, \(Q^{[2]}_{r_1d_1d_2}\) at the beginning of Phase 3.2, is

\[
(nR_2 - nt_{s_2})p_{s_2};(d_1,d_2) + \frac{nR_2}{p_{s_2};(d_1,d_2,r)} p_{s_2};(d_1,d_2) - p_{s_2};(d_1,d_2)r
\]

Hence

\[
N_{3,2} = \frac{nR_2 - nt_{s_2}}{p_{r};(d_1,d_2)}
\]

We then discuss the third and forth quantity, \(Q^{[3]}_{r_1d_1d_2}\) and \(Q^{[4]}_{r_1d_1d_2}\) at the beginning of Phase 3.3. Notice that after Phase 3.1 and 3.2, \(Q^{[3]}_{r_1d_1d_2}\) and \(Q^{[4]}_{r_1d_1d_2}\) are empty sets. Hence by claim(ii), \(Q^{[1]}_{r_1d_1d_2} = nR_1 - Q^{[1]}_{d_1d_2}\) and \(Q^{[2]}_{r_1d_1d_2} = nR_2 - Q^{[2]}_{d_1d_2}\). Thus it is enough to quantify \(Q^{[3]}_{d_1d_2}\) and \(Q^{[4]}_{d_1d_2}\) at the beginning of Phases 3.3.

For \(i \in \{1,2\}\) and \(j = 3 - i\), since during Phase \(i\), we do not repeat sending the vectors which are already received by \(d_i\), hence after Phase \(i\), there are \(nt_{s_1}p_{s,i};d_i\) vectors in \(Q^{[i]}_{d_1}\). In Phase 3.1, \(d_i\) can further receive \(N_{3,i}p_{r};d_i\). Hence at the beginning of Phase 3.3,

\[
Q^{[i]}_{d_1} = \min\{nt_{s_1}p_{s,i};d_i + N_{3,i}p_{r};d_i, nR_i\}
\]

\[
Q^{[i]}_{d_1} = (nR_i - nt_{s_1}p_{s,i};d_i - N_{3,i}p_{r};d_i)^+
\]
Notice that in Phase 3.3, if \( d_i \in R_t \), then do \( Q[i]_{d_i,r} \Rightarrow Q[i]_{d_i} \). Hence we have
\[
N_{3.3} = \max \left\{ \frac{Q[i]_{d_i,r}}{p_r:d_i}, \frac{Q[i]_{d_i,r}}{p_r:d_2} \right\}
\]
(50)

Now we have the required 4 quantities and \( N_{3.1}, N_{3.2}, \) and \( N_{3.3} \). It remains to show that \( N_{3.1} + N_{3.2} + N_{3.3} \leq n_t \).

Suppose \( \frac{Q[i]_{d_i,r}}{p_r:d_1} \geq \frac{Q[i]_{d_i,r}}{p_r:d_2} \), we have
\[
N_{3.1} + N_{3.2} + N_{3.3} = N_{3.1} + N_{3.2} + \frac{(nR_1 - nt_{s_1}p_{s_1:d_1} - N_{3.1}p_r:d_1)^+}{p_r:d_1} + \frac{\tilde{R}_2 - t_{s_2}p_{s_2:d_2}}{p_r:d_2} < n_t
\]
(51)

Similarly, suppose \( \frac{Q[i]_{d_i,r}}{p_r:d_1} < \frac{Q[i]_{d_i,r}}{p_r:d_2} \), we have
\[
N_{3.1} + N_{3.2} + N_{3.3} = N_{3.1} + N_{3.2} + \frac{(nR_2 - nt_{s_2}p_{s_2:d_2} - N_{3.2}p_r:d_2)^+}{p_r:d_2} + \frac{\tilde{R}_1 - t_{s_1}p_{s_1:d_1}}{p_r:d_1} < n_t
\]
(52)

**Proof of claim (vi):** Notice that after Phase 3.3, all the queues except \( Q[i]_{d_1} \) and \( Q[i]_{d_2} \) are empty sets. Follows from claim (ii), there are \( nR_1 \) vectors in \( Q[i]_{d_1} \) and \( nR_2 \) vectors in \( Q[i]_{d_2} \).

**IV. Numerical Results**

In the numerical experiments, we first place the relay \( r \) in the center of a unit disk. Then we place the source \( s_i \) and destination \( d_i \) uniformly randomly in the disk. To model a practical scenario, we impose a condition that \( d_i \) must be in the 90-degree pie area opposite of \( s_i \). See Fig. 1(b) for illustration. Repeat the node placement for \( i = 1, 2 \). See Fig. 1(c) for one realization of our random node placement.

After node placement, we use the distance \( D \) between each pair of nodes and the Rayleigh model to decide the packet overheading probability:
\[
\text{Prob(success)} = \int_T^{\infty} 2 \frac{x}{\gamma} e^{-x^2/\gamma} dx \quad \text{where} \quad \gamma \triangleq 1 \left( \frac{4x^2}{\gamma^2} \right)^2
\]
\( \alpha \) is the path loss factor, and \( T^* \) is the decodable SNR threshold. To fit the experience in practice, we choose \( \alpha = 2.5 \) and \( T^* = 0.006 \) such that the probability that a 1-hop neighbor (resp. a 2-hop neighbor) hears a transmission is around 0.7–0.8 (resp. 0.2–0.3). We assume the success events between different node pairs are independent.

Once the success probabilities are determined, we can maximize \( (R_1 + R_2) \) subject to the outer and inner bounds in Propositions 1 and 2. To enforce fairness, we also impose an additional constraint \( R_i = \beta \min (p_{s_i:d,r}, p_{s_i:d_2} + p_{r:d}) \) for \( i = 1, 2 \) with a common \( \beta \), which requires \( R_i \) being proportional to the min-cut value from \( s_i \) to \( d_i \) assuming no other sessions are transmitting and \( s_i \) and \( r \) are scheduled with the same frequency. We repeat the above process for 1000 times and then take the average of the maximized sum-rate.

Our capacity characterization can be easily modified for schemes with different capabilities. More explicitly, we use “OpR” to stand for opportunistic routing. “CSF” stands for instant channel status feedback. “Sch” allows dynamically scheduling different nodes, in contrast with a fixed schedule for which each node is scheduled with frequency 1/3. Table II summarizes the sum-rate capacity for different NC schemes. For schemes equipped with OpR, we list both the inner and outer bounds. For schemes without OpR, our results describe the full capacity. As can be seen, the largest throughput gain follows from allowing OpR. Note that CSF strictly increases the capacity for all cases. Moreover, with CSF the gap between the inner and outer bounds is much smaller (less than 0.05) when OpR is used. When compared to the “NC only” scheme, the joint use of OpR, CSF, and Sch enhances capacity by 45%. This thus illustrates the importance of jointly designing OpR, CSF, and Sch together with the NC solution.

**V. Conclusion and Future Work**

This paper has provided a pair of inner and outer bounds for the capacity of 2-user 1-hop relay erasure network that allows opportunistic routing, instant channel status feedback, and per-slot scheduling. The conditions for which the outer and inner bounds coincide have been identified. In our numerical experiments, the gap between the inner and outer bounds is within 0.2%. In the future, we will generalize the results for \( K \)-user 1-hop relay erasure networks with arbitrary \( K \) values.

This work was supported in parts by NSF grants CCF-0845968 and CNS-0905331.

**REFERENCES**


