A Simple 1-Hop Broadcast Scheme That Achieves The Optimal Deadline-Constrained Throughput For Two Independent Sessions With Over-Provisioned Bandwidth.

1 The Setting

In this work, we are interested in the achievable throughput of network coding schemes for the setting of two unicast sessions under the sequential deadline constraints of stored-video streaming.

We consider the downlink of a single cell in which the base station (BS) broadcasts two video files to 2 users, d_1 and d_2, respectively. Each video file contains N packets, and we use $\{X_{1,n}\}_{n=1}^{N}$, $\{X_{2,n}\}_{n=1}^{N}$ to represent them. We use session 1 and session 2 to denote (the transmission of) the data packets for d_1 and d_2, respectively.

We define the time when the BS begins transmitting the first packet as the time origin, and assume that all packets are available at the source (the video-file server) at the time origin. We also assume slotted transmission. Each packet $X_{j,n}$ $(j = 1, 2)$ has a deadline $\tau_{j,n}$ such that after time slot $\tau_{j,n}$ the packet $X_{j,n}$ is no longer useful for user j. For ease of exposition, we assume that the deadlines of $X_{j,n}$ are the same for both sessions and they have the following form:

$$\forall j \in \{1, 2\},\ \tau_{j,n} = \delta \cdot n \text{ , where } \delta \text{ is a fixed positive integer.}$$

We consider random and unreliable wireless channels. Both users can overhear the transmission to the other user. That is, a transmitted packet may be received by both users, by only one user, or by neither users. Consider a packet is transmitted in the t-th time slot. For $j = 1, 2$, we use $C_j(t) = 1$ to denote the event that user j can receive the packet successfully, and $C_j(t) = 0$, otherwise. The successful probability for each channel is p. In this manuscript, we consider the models in which channels are independently and identically distributed (i.i.d.) and both channels $C_1(t)$ and $C_2(t)$ are independent with each other. We also assume that in the end of each time slot, the BS has perfect feedback from both destinations regarding whether the transmitted packet has been successfully received by each user. Such information will be used to decide what to transmit in the next time slot.

If coding is not allowed, the source can only transmit uncoded packets. Suppose packet $X_{1,n}$ is transmitted at time t, and user 1 does not receive it. After receiving the feedback at the end of time t, the BS may decide to retransmit the same packet $X_{1,n}$; or may decide to move to the next packet $X_{1,n+1}$ to enhance the chance that packet $X_{1,n+1}$ can be received before its deadline; or it may decide to send the packet $X_{2,n'}$ for the other session instead. If coding across different packets is allowed, then in one slot, the BS can encode a set of unexpired packets together and broadcast it to all users. Here we allow coding operation across different sessions. When coding is used, we require an information packet to be "decoded" before the corresponding deadline.

Our goal is to design a coding/scheduling policy that maximizes the number of successful (unexpired) packet receptions. Let $D_j(n) = 1$ if user j can successfully decode/recover the n-th information packet for session j before its deadline $\tau_{j,n} = \delta n$; and $D_j(n) = 0$, otherwise. We define the total number of successes N_{success} by $N_{\text{success}} \triangleq \sum_{n=1}^{N} \sum_{j=1}^{2} D_j(n)$. Our goal is to maximize the normalized expected throughput $\frac{\mathsf{E}\{N_{\text{success}}\}}{2N}$.

2 The Scheme

Next we present a simple network coding scheme for the above setting.

To begin with, we will introduce some definitions. The source s keeps two registers n_1 and n_2. One can view the purpose of n_i as to keep track of the next uncoded packet to be sent for session i. Since both n_1 and n_2 evolve over time, we sometimes use $n_i(t)$ to denote the value of n_i in the end of time t. The source s also keeps two lists of packets: $\tilde{v}_X^{[2]}$ and $\tilde{v}_Y^{[1]}$. $\tilde{v}_X^{[2]}$ is a list that contains all unexpired packets of session 1 that have been received by user 2 but not by user 1. Symmetrically, $\tilde{v}_Y^{[1]}$ is a list that contains all unexpired packets for session 2 that have been received by user 1 but not by user 2. We use $\tilde{s}_1$ and $\tilde{s}_2$ to denote the sets of packets to be transmitted or sessions 1 and 2, respectively.

1: Set $n_1 \leftarrow 1$, $n_2 \leftarrow 1$, $\tilde{v}_X^{[2]} \leftarrow \emptyset$, $\tilde{v}_Y^{[1]} \leftarrow \emptyset$, $\tilde{s}_1 \leftarrow \{X_{1,1}, \cdots, X_{1,N}\}$, and $\tilde{s}_2 \leftarrow \{X_{2,1}, \cdots, X_{2,N}\}$.
2: **for** $t = 1$ to δN **do**
3: In the beginning of the t-th time slot, do the following:
4: **if** $n_2 \leq N$ **then**

if both $\tilde{v}_X^{[2]}$ and $\tilde{v}_Y^{[1]}$ are non-empty **then**
Choose the oldest packet X_{1,j_1^*} from $\tilde{v}_X^{[2]}$ and the oldest packet X_{2,j_2^*} from $\tilde{v}_Y^{[1]}$. Broadcast the linear sum $[X_{1,j_1^*} + X_{2,j_2^*}]$.
else if $n_1 = n_2$ **then**
Send uncoded packet X_{1,n_1} directly.
else if $n_1 > n_2$ **then**
Send uncoded packet X_{2,n_2} directly.
end if
else
Choose the oldest packet in $\tilde{v}_X^{[2]} \cup \tilde{v}_Y^{[1]}$ and send that packet uncodedly.
end if
In the end of the t-th time slot,
if an uncoded packet X_{1,n_1} was sent and received by at least one user **then**
$n_1 \leftarrow n_1 + 1$.
Remove X_{1,n_1} from $\tilde{s}_1$. If X_{1,n_1} was received only by d_2, then add X_{1,n_1} to $\tilde{v}_X^{[2]}$.
else if an uncoded packet X_{2,n_2} was sent and received by at least one user **then**
$n_2 \leftarrow n_2 + 1$.
Remove X_{2,n_2} from $\tilde{s}_2$. If X_{2,n_2} was received only by d_1, then add X_{2,n_2} to $\tilde{v}_X^{[2]}$.
else
Depending on whether the coded transmission $[X_{1,j_1^*} + X_{2,j_2^*}]$ was received by d_1, remove X_{1,j_1^*} from $\tilde{v}_X^{[2]}$.
Depending on whether the coded transmission $[X_{1,j_1^*} + X_{2,j_2^*}]$ was received by d_2, remove X_{2,j_2^*} from $\tilde{v}_Y^{[1]}$.
end if
Remove all expired packets (those with index $\leq \frac{t}{\delta}$) from $\tilde{v}_X^{[2]}$, $\tilde{v}_Y^{[1]}$, $\tilde{s}_1$, and $\tilde{s}_2$.
end for

By noting that the above scheme actually schedules the non-coded transmission in a round-robin fashion, we have the following self-explanatory lemma.

Lemma 1. *For any time slot t, we have $0 \leq n_1(t) - n_2(t) \leq 1$.*

3 Performance Analysis

The performance of the proposed simple scheme is characterized as follows.

Proposition 2. *The proposed scheme satisfies the following:*

$$\lim_{N\to\infty} \frac{\mathsf{E}\{N_{success}\}}{2N} = 1 \tag{1}$$

for all p values satisfying $p \geq p^\star$, where $p^\star \triangleq \frac{2\delta+1-\sqrt{4\delta^2-8\delta+1}}{2\delta}$.

Proposition 3. *For any $p < p^\star$, any scheme, coded or non-coded, must have*

$$\lim_{N\to\infty} \frac{\mathsf{E}\{N_{success}\}}{2N} = 1-\epsilon \tag{2}$$

for some ϵ that depends only on p.

The above two propositions show that the proposed scheme is asymptotically throughput optimal in an over-provisioned environment ($p > p^\star$). We first prove Proposition 3.

Proof. Let us temporarily relax the sequential deadline constraint and set

$$\forall i \in \{1,2\}, n \in \{1,\cdots,N\}\tau_{i,n} = \delta \cdot N. \tag{3}$$

Then the question becomes how many packets out of the overall $2N$ packets (N for d_1 and N for d_2) can be sent to their desired destinations within $\delta \cdot N$ time slots. Let $\tilde{N}_1$ and $\tilde{N}_2$ denote the number of packets that are successfully received by each user, respectively. In [1], it was proven that $\tilde{N}_1$ and $\tilde{N}_2$ and must satisfy

$$\frac{\tilde{N}_1}{p} + \frac{\tilde{N}_2}{1-(1-p)^2} \leq \delta N \tag{4}$$

$$\frac{\tilde{N}_1}{1-(1-p)^2} + \frac{\tilde{N}_2}{p} \leq \delta N. \tag{5}$$

With the objective function being $\max\left(\tilde{N}_1 + \tilde{N}_2\right)$, we can solve the above linear-programming problem and prove that

$$\lim_{N\to\infty} \frac{\mathsf{E}\{\tilde{N}_1+\tilde{N}_2\}}{2N} \leq \frac{(2p-p^2)\delta}{3-p}. \tag{6}$$

The remaining step is to observe that for all $p < p^\star$, the left-hand side of (6) is strictly bounded away from 1. The proof is complete. □

Before proving Proposition 2, we present the following key lemma, which is critical to our proof.

Lemma 4. *For any $\epsilon > 0$, there exists a $B > 0$ such that for all fixed t_1 and t_2 satisfying $(t_2 - t_1) > B$,*

$$\mathsf{E}\Big\{n_2(t_2) - n_2(t_1)\,|t_2 < \delta n_2(t_1)\,\Big\} \leq (t_2 - t_1)\left(\frac{2p - p^2}{3 - p}\right)(1 + \epsilon).$$

Proof. The following discussion is conditioned on the event that in the end of time t_1, we have $\mathcal{A}_{t_1} \triangleq \{t_2 < \delta n_2(t_1)\}$. Define $\Delta n = \left\lfloor (t_2 - t_1)\frac{2p-p^2}{3-p}\right\rfloor$. From the beginning of time t_1+1, let us temporarily suspend the "expiration mechanism" and use our proposed scheme to transmit packets while allowing the supposedly expired packets to remain in the system. We first examine how long it takes before the register $n_2(t)$ evolves from its current value $n_2(t_1)$ to a different value $n_1(t_1) + \Delta n$. More specifically, we use t_3 to denote the (random) time slot for which in the end of time t_3, $n_2(t)$ changes to $n_1(t_1) + \Delta n$ for the first time.

We define $\mathsf{UT}_1[n_1(t_1), n_1(t_1)+\Delta n-1]$ (which stands for "Uncoded Transmission") as the number of time slots in $[t_1+1, t_3]$ when the proposed scheme schedules an *uncoded* packet transmission for Session 1. Note that by our definitions, all those uncoded transmissions must be used to transmit $X_{1,n}$ for some $n \in [n_1(t_1), n_1(t_1) + \Delta n - 1]$. This is why we append $[n_1(t_1), n_1(t_1) + \Delta n - 1]$ to UT_1. Similarly, we also define $\mathsf{UT}_2[n_2(t_1), n_1(t_1) + \Delta n - 1]$ as the number of time slots in $[t_1 + 1, t_3]$ when the proposed scheme schedules an uncoded packet transmission for Session 2 packets $X_{2,n}$ with the indices being $n \in [n_2(t_1), n_1(t_1)+\Delta n-1]$. We use UT_1 and UT_2 as shorthand in the subsequent discussion.

Define

$$H_{1,n} = |\{t > t_1 : \text{in the beginning of time } t\text{, the scheme schedules an uncoded transmission of } X_{1,n}\}|\,. \quad (7)$$

Since we stop an uncoded transmission if any one of the destinations successfully receives it, we have

$$\mathsf{E}\{H_{1,n}|\mathcal{A}_{t_1}\} = \frac{1}{2p - p^2} \quad (8)$$

for all $n \geq n_1(t_1)$. As a result, the total number of time slots to transmit the uncoded session-1 packets is

$$\mathsf{UT}_1 = \sum_{i=n_1(t_1)}^{n_1(t_1)+\Delta n-1} H_{1,i}.$$

Similarly, the total number of time slots to transmit the uncoded session 2 packets in time $[t_1+1, t_3]$ is at least

$$\mathsf{UT}_2 = \sum_{i=n_2(t_1)}^{n_1(t_1)+\Delta n-1} H_{2,i} \geq \sum_{i=n_1(t_1)}^{n_1(t_1)+\Delta n-1} H_{2,i}.$$

Since each $H_{1,i}$ and $H_{2,j}$ are i.i.d. (conditional) geometric distribution with expectation (8), for any $\epsilon_1, \delta_1 > 0$, we can choose a sufficiently large B_1 such that if $\Delta n > B_1$, then

$$\begin{aligned}
&\mathsf{P}\left(\mathsf{UT}_1 + \mathsf{UT}_2 > (1-\epsilon_1)\frac{2\Delta n}{2p-p^2} \middle| \mathcal{A}_{t_1}\right) \\
&\geq \mathsf{P}\left(\sum_{i=n_1(t_1)}^{n_1(t_1)+\Delta n-1} (H_{1,i} + H_{2,i}) > (1-\epsilon_1)\frac{2\Delta n}{2p-p^2} \middle| \mathcal{A}_{t_1}\right) \\
&= \mathsf{P}\left(\sum_{i=1}^{2\Delta n} H_i > (1-\epsilon_1)\frac{2\Delta n}{2p-p^2}\right) > 1-\delta_1, \qquad (9)
\end{aligned}$$

where $\{H_i\}$ are i.i.d. geometric random variables with expectation $\frac{1}{2p-p^2}$ and (9) follows from the weak law of large numbers.

Let $O_{1,n}$ denote a Bernoulli random variable that is 1 if when sending $X_{1,n}$ uncodedly, it was d_2 that received $X_{1,n}$ first. Symmetrically, we define the Bernoulli random variable $O_{2,n}$ such that $O_{2,i}$ is 1 if when sending $X_{2,i}$ uncodedly, it was d_1 that received $X_{2,i}$ first.

We now define $\mathsf{CT}_{1,n}$ as follows:

$$\begin{aligned}
\mathsf{CT}_{1,n} \triangleq \big|\{t > t_1 : &\text{ in time } t, \text{ packet } X_{1,n} \text{ is mixed (coded) with} \\
&\text{some other } X_{2,n'} \text{ packets.}\}\big|, \qquad (10)
\end{aligned}$$

where $\mathsf{CT}_{1,n}$ stands for the coded transmission for packet $X_{1,n}$. Define TCT as the total number of coded transmission in time $[t_1+1, t_3]$. We then notice the following facts: (i) In the beginning of time t_3, the scheme must transmit an uncoded packet $X_{2,n_1(t_1)+\Delta n-1}$ and it is received by one of the destinations (that is why $n_2(t)$ changes to $n_1(t_1)+\Delta n$). (ii) Therefore, in the end of time t_3-1, there must have $\min\left(\tilde{v}_X^{[2]},\tilde{v}_Y^{[1]}\right)=0$. That are no packets to be coded in the end of time t_3-1. (iii) Therefore, in the end of time t_3-1, either (a) there is no $\{X_{1,n}:n\in[n_1(t_1),n_1(t_1)+\Delta n-1]\}$ in $\tilde{v}_X^{[2]}$, or (b) there is no $\{X_{2,n}:n\in[n_2(t_1),n_1(t_1)+\Delta n-2]\}$ in $\tilde{v}_Y^{[1]}$. From the above three facts, we have

$$\mathsf{TCT}=\min\left(\sum_{i=1}^{n_1(t_1)+\Delta n-1}\mathsf{CT}_{1,i},\ \sum_{i=1}^{n_1(t_1)+\Delta n-2}\mathsf{CT}_{2,i}\right). \tag{11}$$

For the following, we will prove that for any $\epsilon_2,\delta_2>0$, we can choose a sufficiently large B_2 such that if $\Delta n>B_2$, we have

$$\mathsf{P}\left(\mathsf{TCT}>\Delta n\left(\frac{1-p}{2p-p^2}\right)(1-\epsilon_2)\middle|\mathcal{A}_{t_1}\right)>1-\delta_2. \tag{12}$$

To that end, we use the following union-bound arguments and focus on the sub-series of the summations:

$$\begin{aligned}
&\mathsf{P}\left(\mathsf{TCT}>\Delta n\left(\frac{1-p}{2p-p^2}\right)(1-\epsilon_2)\middle|\mathcal{A}_{t_1}\right)\\
&=\mathsf{P}\left(\text{Eq. (11)}>\Delta n\left(\frac{1-p}{2p-p^2}\right)(1-\epsilon_2)\middle|\mathcal{A}_{t_1}\right)\\
&\geq 1-\mathsf{P}\left(\sum_{i=n_1(t_1)}^{n_1(t_1)+\Delta n-1}\mathsf{CT}_{1,i}\leq\Delta n\left(\frac{1-p}{2p-p^2}\right)(1-\epsilon_2)\middle|\mathcal{A}_{t_1}\right)\\
&\quad-\mathsf{P}\left(\sum_{i=n_2(t_1)}^{n_1(t_1)+\Delta n-2}\mathsf{CT}_{2,i}\leq\Delta n\left(\frac{1-p}{2p-p^2}\right)(1-\epsilon_2)\middle|\mathcal{A}_{t_1}\right).
\end{aligned} \tag{13}$$

Note that for any $i\geq n_1(t_1)$, $\mathsf{CT}_{1,i}=0$ if $O_{1,i}=0$, and conditioning on $O_{1,i}=1$, the random variable $\mathsf{CT}_{1,i}$ is geometrically distributed with success probability p. Moreover, $\mathsf{CT}_{1,i}$ is i.i.d. with epectation $\left(\frac{1-p}{2-p}\cdot\frac{1}{p}\right)$for any $i\geq n_1(t_1)$ (recall that we have temporarily suspended "expiration"). The weak law of large numbers thus implies that for any $\delta_3 > 0$, there exists a B_3 such that if $\Delta n > B_3$, we have

$$\mathsf{P}\left(\sum_{i=n_1(t_1)}^{n_1(t_1)+\Delta n-1} \mathsf{CT}_{1,i} \leq \Delta n \left(\frac{1-p}{2-p} \cdot \frac{1}{p} \right) (1-\epsilon_2) \middle| \mathcal{A}_{t_1} \right) \leq \delta_3. \quad (14)$$

Similarly by the weak law of large numbers, we also have for any $\delta_4 > 0$, there exists a B_4 such that if $\Delta n > B_4$, we have

$$\mathsf{P}\left(\sum_{i=n_2(t_1)}^{n_1(t_1)+\Delta n-2} \mathsf{CT}_{2,i} \leq (\Delta n - 1) \left(\frac{1-p}{2-p} \cdot \frac{1}{p} \right) (1-\epsilon_2) \middle| \mathcal{A}_{t_1} \right) \leq \delta_4. \quad (15)$$

Jointly (14) and (15) imply that (13) can be made arbitrarily close to one by choosing a sufficiently large B_3 and B_4 and setting $B_2 = \max(B_3, B_4)$. Eq. (12) is thus proven.

Since for any time slot in $[t_1+1, t_3]$ we either send an uncoded or a coded transmission, we must have $t_3 - t_1 = \mathsf{UT}_1 + \mathsf{UT}_2 + \mathsf{TCT}$. By (9) and (12), we have thus proven that for any $\epsilon_5, \delta_5 > 0$, there exists a $B_5 > 0$ such that if $\delta n > B_5$, we have

$$\mathsf{P}\left((t_3 - t_1) > \Delta n \left(\frac{2}{2p-p^2} + \frac{1-p}{2p-p^2} \right) (1-\epsilon_5) \middle| \mathcal{A}_{t_1} \right) > 1 - \delta_5. \quad (16)$$

By the definition of $\Delta n = \left\lfloor (t_2 - t_1) \frac{2p-p^2}{3-p} \right\rfloor$, we thus also have

$$\mathsf{P}\left((t_3 - t_1) > (t_2 - t_1)(1-\epsilon_5) | \mathcal{A}_{t_1} \right) > 1 - \delta_5. \quad (17)$$

Namely, with close to one probability, the random time t_3, in the end of which $n_2(t)$ changes to $n_1(t_1) + \Delta n$ for the first time, is no less than $t_1 + (t_2 - t_1)(1-\epsilon_5)$. Therefore, in the end of time $t_1 + (t_2 - t_1)(1-\epsilon_5)$, $n_2(t)$ must be no larger than $n_1(t_1) + \Delta n$ with close-to-one probability since we have not reached t_3 yet. (17) thus implies

$$\mathsf{P}\left(n_2(t_1 + (t_2 - t_1)(1-\epsilon_5)) \leq n_1(t_1) + \Delta n | \mathcal{A}_{t_1} \right) > 1 - \delta_5. \quad (18)$$

We then notice the following two facts: (i) the difference between t_2 and $(t_1 + (t_2 - t_1)(1-\epsilon_5))$ is $(t_2 - t_1)\epsilon_5$; and (ii) for any $t'_1 < t'_2$ the difference $n_2(t'_2) -$

$n_2(t_1')$ is no larger than $t_2' - t_1'$ since the register $n_2(t)$ at most increments by one in every time slots. As a result, (18) implies

$$\mathsf{P}\left(n_2(t_2) - n_1(t_1) \leq \Delta n + (t_2 - t_1)\epsilon_5 \middle| \mathcal{A}_{t_1} \right) > 1 - \delta_5. \tag{19}$$

We can then reuse the above fact (ii) to upper bound the expectation of $n_2(t_2) - n_1(t_1)$:

$$\mathsf{E}\left\{ n_2(t_2) - n_1(t_1) | \mathcal{A}_{t_1} \right\} \leq \left(\Delta n + (t_2 - t_1)\epsilon_5 \right)(1 - \delta_5) + \delta_5(t_2 - t_1). \tag{20}$$

By Lemma 1, the difference between $n_1(t_1)$ and $n_2(t_1)$ is no larger than 1. Also by noticing that Δn is linearly proportional to $(t_2 - t_1)$ while all other terms are sub-linear (with either a ϵ or a δ coefficient), (20) thus implies that for any $\epsilon > 0$, there exists a sufficiently large B such that if $t_2 - t_1 > B$, then

$$\mathsf{E}\left\{ n_2(t_2) - n_1(t_1) | \mathcal{A}_{t_1} \right\} \leq \Delta n (1 + \epsilon). \tag{21}$$

In the above analysis, we have not considered the impact of when allowing expiration. In the following, we will include expiration back to our analysis. To that end, we first notice that we can still define $H_{1,n}$, $H_{2,n}$, $\mathsf{CT}_{1,n}$, $\mathsf{CT}_{2,n}$ as in (7) and (10), respectively. Note that now these four random variables are no longer independently distributed as the results of one, say $H_{1,n}$, may affect the other, say $\mathsf{CT}_{2,n'}$, due to expiration. Define a set of *shadow random variables* $\tilde{H}_{1,n}$, $\tilde{H}_{2,n}$, $\tilde{\mathsf{CT}}_{1,n}$, $\tilde{\mathsf{CT}}_{2,n}$ that characterize the behaviors when there is no expriation involved. More specifically, we choose $\tilde{H}_{1,n} = H_{1,n}$ if $H_{1,n}$ stops "growing" due to the $X_{1,n}$ packet being received by one of the two destinations. If $H_{1,n}$ stops growing due to the expiration of $X_{1,n}$, then we let $\tilde{H}_{1,n}$ continue to grow as an independent geometric random variable with success probability $(2p - p^2)$. In this way, $\tilde{H}_{1,n}$ mimics the behavior of a system with no expiration and $\tilde{H}_{1,n}$ is independent from all other random variables. Similarly, we choose $\tilde{\mathsf{CT}}_{1,n} = \mathsf{CT}_{1,n}$ if $\mathsf{CT}_{1,n}$ stops growing due to the mixed coded transmission involving $X_{1,n}$ being received d_1. If $\mathsf{CT}_{1,n}$ stops growing due to the expiration of $X_{1,n}$, then we let $\tilde{\mathsf{CT}}_{1,n}$ continue to grow as an independent geometric random variable with success probability p. In this way, $\tilde{\mathsf{CT}}_{1,n}$ mimics the behavior of a system with no expiration and $\tilde{\mathsf{CT}}_{1,n}$ is independent from all other random variables.

Then we need to prove the following version of (17): For any $\epsilon_5, \delta_5 > 0$, there exists a sufficiently large B_5 such that for any $t_2 - t_1 > B_5$, we have

$$\begin{aligned}\delta_5 &\geq \mathsf{P}\left(\mathsf{UT}_1 + \mathsf{UT}_2 + \mathsf{TCT} \leq (t_2 - t_1)(1-\epsilon_5) | \mathcal{A}_{t_1}\right) \\ &= \mathsf{P}\Bigg(\sum_{i=n_1(t_1)}^{n_1(t_1)+\Delta n-1} H_{1,i} + \sum_{j=n_2(t_1)}^{n_1(t_1)+\Delta n-1} H_{2,j} + \\ &\quad \min\left(\sum_{k=1}^{n_1(t_1)+\Delta n-1} \mathsf{CT}_{1,k}, \sum_{l=1}^{n_1(t_1)+\Delta n-2} \mathsf{CT}_{2,l} \right) \leq (t_2-t_1)(1-\epsilon_5) \Bigg| \mathcal{A}_{t_1} \Bigg)\end{aligned} \tag{22}$$

Note that conditioning on $\mathcal{A}_{t_1} = \{t_2 < \delta n_2(t_1)\}$, during time $[t_1, t_1 + (t_2 - t_1)(1-\epsilon_5)]$, no packets with indices $\geq n_2(t_1)$ will expire. Therefore, conditioning on $\mathcal{A}_{t_1}$ any realization of $H_{1,i}$, $H_{2,j}$, and $\mathsf{CT}_{1,k}$, and $\mathsf{CT}_{2,l}$ in (22) must not result in any expiration for packets with indices $\geq n_2(t_1)$. As a result, we have

$$\begin{aligned}&\mathsf{P}\Bigg(\sum_{i=n_1(t_1)}^{n_1(t_1)+\Delta n-1} H_{1,i} + \sum_{j=n_2(t_1)}^{n_1(t_1)+\Delta n-1} H_{2,j} + \\ &\quad \min\left(\sum_{k=1}^{n_1(t_1)+\Delta n-1} \mathsf{CT}_{1,k}, \sum_{l=1}^{n_1(t_1)+\Delta n-2} \mathsf{CT}_{2,l} \right) \leq (t_2-t_1)(1-\epsilon_5) \Bigg| \mathcal{A}_{t_1} \Bigg) \\ &\leq \mathsf{P}\Bigg(\sum_{i=n_1(t_1)}^{n_1(t_1)+\Delta n-1} \tilde{H}_{1,i} + \sum_{j=n_2(t_1)}^{n_1(t_1)+\Delta n-1} \tilde{H}_{2,j} + \\ &\quad \min\left(\sum_{k=n_1(t_1)}^{n_1(t_1)+\Delta n-1} \tilde{\mathsf{CT}}_{1,k}, \sum_{l=n_2(t_1)}^{n_1(t_1)+\Delta n-2} \tilde{\mathsf{CT}}_{2,l} \right) \leq (t_2-t_1)(1-\epsilon_5) \Bigg| \mathcal{A}_{t_1} \Bigg)\end{aligned} \tag{23}$$

since for those realizations, the shadow random variables and the actual random variables for packets with indices $\geq n_2(t_1)$ have the same probability. Since (17) holds for the case without expiration, (23) can thus be made smaller than δ_5 with sufficiently large B_5. (22) is thus proven. We can then follow the same analysis as in (17) to (21). The proof of Lemma 4 is complete. □

For the following, we prove Proposition 2.

Proof. We define $q_2(t) = n_2(t) - \frac{t}{\delta}$. By our definition of $p^\star$, for any $\epsilon_1 > 0$ we can choose a $p < p^\star$ such that $\frac{1}{\delta}(1-\epsilon_1) < \frac{2p-p^2}{3-p} < \frac{1}{\delta}$. We fix that one such p throughout this proof. By definition, we can now rewrite Lemma 4 by letting $t_2 = t_1 + B_0$. More specifically, for any $\epsilon > 0$, there exists a $B > 0$ such that for any $B_0 > B$

$$\begin{aligned}
&\mathsf{E}\Big\{q_2(t_1+B_0) - q_2(t_1)\Big| q_2(t_1) > \frac{B_0}{\delta}\Big\} \\
&= \mathsf{E}\Big\{n_2(t_1+B_0) - n_2(t_1)\Big| q_2(t_1) > \frac{B_0}{\delta}\Big\} - \frac{B_0}{\delta} \\
&\leq \frac{B_0(2p-p^2)}{3-p}(1+\epsilon) - \frac{B_0}{\delta} < 0,
\end{aligned}$$

where the negativeness is established by choosing a sufficiently small $\epsilon > 0$. As a result, $q_2(t)$ has a negative drift. Since $q_2(t)$ has a negative drift, it implies that for any $\epsilon_1, \epsilon' > 0$, there exists a $t_0 > 0$ such that $\mathsf{P}\left(q_2(t) < \epsilon' t\right) > 1 - \epsilon_1$, for all $t > t_0$. We can also define $q_1(t) = n_1(t) - \frac{t}{\delta}$. By Lemma 1, the difference between $q_1(t)$ and $q_2(t)$ is no larger than 1. As a result, $q_1(t)$ also has a negative drift.

Using the negative drift of $q_1(t)$, we have for any $t > t_0$,

$$\begin{aligned}
\mathsf{E}\{n_1(t)\} &= \mathsf{E}\left\{\frac{t}{\delta} + q_1(t)\right\} && (24) \\
&= \mathsf{E}\left\{\frac{t}{\delta} + q_1(t)\Big| q_1(t) < \epsilon' t\right\} \mathsf{P}(q_1(t) < \epsilon' t) \\
&\quad + \mathsf{E}\{n_1(t) | q_1(t) \geq \epsilon' t\} \mathsf{P}(q_1(t) \geq \epsilon' t) \\
&\leq \Big(\frac{t}{\delta} + \epsilon' t\Big)(1-\epsilon_1) + \Big(\frac{t}{2} + 1\Big)\epsilon_1, && (25)
\end{aligned}$$

where (25) is because, when $q_1(t) < \epsilon' t$, $n_1(t) < \frac{t}{\delta} + \epsilon' t$; when $q_1(t) \geq \epsilon' t$, $n_1(t)$ is bounded by $\frac{t}{2} + 1$, since in t time slots, at most $\frac{t}{2} + 1$ packets in $\{X_{1,n}\}_{n=1}^N$ have been transmitted uncodedly. Namely, when we let $t \to \infty$, the expectation $\mathsf{E}\{n_1(t)\}$ is upper bounded by $\frac{t}{\delta}(1+\epsilon)$ for any arbitrary $\epsilon > 0$. By the same approach, we can also derive that for any $\epsilon > 0$,

$$\mathsf{E}\{n_2(t)\} \leq \frac{t}{\delta}(1+\epsilon) + o(t), \qquad (26)$$

where $o(t)$ is a sublinear term.

We define $T_1(t)$ as the number of time slots when the BS transmits an uncoded packet for session 1 up to time t; $T_2(t)$ as the number of time slots when the BS transmits an uncoded packet for session 2 up to time t. Since the BS will transmit every uncoded packet until it has been received by at least one user, we have

$$\mathsf{E}\{T_1(t)\} \leq \mathsf{E}\{n_1(t)\}\frac{1}{2p-p^2}, \tag{27}$$

$$\mathsf{E}\{T_2(t)\} \leq \mathsf{E}\{n_2(t)\}\frac{1}{2p-p^2}, \tag{28}$$

where the inequality is because some uncoded packets are expired before they can be received by any user, so the expected transmission time for each packet is shortened.

Note that when we transmit an uncoded packet, the expected "reward" is p since only one destination can benefit from this transmission. When we transmit a coded packet, the expected reward is $2p$ since both destinations can benefit. As a result, for sufficiently large t, the expected total rewards is lower bounded by

$$\begin{aligned}
&\mathsf{E}\{N_{\text{success}}\} && (29)\\
&= p\mathsf{E}\{T_1(t)\} + p\mathsf{E}\{T_2(t)\} + 2p\mathsf{E}\{t - T_1(t) - T_2(t)\} && (30)\\
&= 2pt - p\mathsf{E}\{T_1(t)\} - p\mathsf{E}\{T_2(t)\} && (31)\\
&\geq 2pt - p\frac{t}{\delta}\frac{2}{2p-p^2} + o(t) &&\\
&= \frac{2t}{\delta}\left(p\delta - \frac{1}{2-p}\right) + o(t). && (32)
\end{aligned}$$

The above result implies that at time $t = \delta N$ (the end of transmission),

$$\lim_{N\to\infty}\frac{\mathsf{E}\{N_{\text{success}}\}}{2N} \geq \frac{2\delta N}{\delta}\frac{1}{2N}\left(p\delta - \frac{1}{2-p}\right) = \left(p\delta - \frac{1}{2-p}\right). \tag{33}$$

Recall that we can choose p such that $\frac{2p-p^2}{3-p}$ as close to $\frac{1}{\delta}$ as possible, which implies that (33) can be made arbitrarily close to 1.

Now we have shown that when p approaches $p^\star$ from the left, we have $\lim_{N\to\infty}\frac{\mathsf{E}\{N_{\text{success}}\}}{2N} = 1$. Next we are going to show that when $1 \geq p > p^\star$, we also have $\lim_{N\to\infty}\frac{\mathsf{E}\{N_{\text{success}}\}}{2N} = 1$.

For any $p > p^\star$, we slighly modify the definition and let $q'_j(t) = n_j(t) - \frac{t}{\delta'}$, $j = 1, 2$, for some new $\delta' < \delta$ satisfying $\frac{1}{\delta'}(1 - \epsilon_1) < \frac{2p-p^2}{3-p} < \frac{1}{\delta'}$ for some arbitrarily small ϵ_1. Note that Lemma 4 implies that $q_j(t)$ has negative drift. By the same arguments, Lemma 4 also implies that $q'_j(t)$ has negative drifts as well. We can then follow the same analysis as from (24) to (33). Namely, when $p > p^\star$, the proposed scheme can achieve the capacity $\lim_{N\to\infty} \frac{\mathsf{E}\{N_{\text{success}}\}}{2N} = 1$ within $\delta' N < \delta N$ time slots. The proof is complete.

□

References

[1] 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.