

Capacity and Stable Throughput Regions for Broadcast Erasure Channel with Feedback -An Unusual Union-

Yalin Evren Sagduyu *Member, IEEE*, Leonidas Georgiadis *Senior Member, IEEE*, Leandros Tassioulas *Fellow, IEEE*, Anthony Ephremides *Life Fellow, IEEE*

Abstract

We consider a source node broadcasting to two receivers over a general erasure channel with receiver feedback. We characterize the capacity region of the channel and construct algorithms based on linear network coding (either randomized or depending on channel dynamics) that achieve this capacity. We then consider stochastic arrivals at the source for the two destinations and characterize the stable throughput region achieved by adapting the same algorithms that achieve capacity. Next, we modify these algorithms to improve their delay performance and characterize their stable throughput regions. Although the capacity and stability regions obtained by the algorithms are not always identical (because of the extra overhead needed for the algorithms to handle stochastic traffic), they are within a few bits of each other and have similar forms. This example exhibits an unusual relationship between capacity and stability regions and extends similar prior studies for multiple access channels.

Index Terms

Network coding, broadcast erasure channel, capacity region, stability region, feedback, overhead.

I. INTRODUCTION

The quest for understanding the relationship between information-theoretic capacity and stable throughput regions for networks in general, and for wireless networks in particular, has received considerable attention in recent years [1]–[3] and a small body of knowledge has been developed primarily for multi-access channels. These two regions, measured in terms of bits/sec, may be identical or not and we do not yet know general conditions under which they coincide.

In this paper we consider a two-receiver broadcast channel with feedback and essentially address the same issue. The erasure channel model applies broadly to packet transmission and packet dropping mechanisms in the Internet and in general communication protocols. The usual extension to wireless communications is based on the broadcast erasure channel with multiple receiving ends. However, the capacity problem in general broadcast channels [4] is unresolved yet and the capacity region is only known for special cases including degraded broadcast channels [5]. For a single multicast session the capacity of wireless erasure channels session has been determined in [6] and a capacity achieving algorithm for the same model has been formulated in [7] using random linear coding. Different throughput and delay benefits of using feedback for network coding are presented in [8]–[15] with the typical assumption of single multicast session or backlogged packet traffic. The capacity region for multiple unicast sessions has been considered in [16] for backlogged packet traffic with a different perspective of handling feedback overhead.

Here, our focus is on deriving the capacity region (with backlogged packet traffic) and stable throughput region (with stochastic packet traffic) for multiple unicast sessions (as well as combination of unicast and multicast sessions) over the broadcast erasure channel with feedback. It turns out that the capacity region can be easily characterized by bounding it through the degraded version of the channel and it is remarkable that two simple algorithms that are based on linear network coding over blocks of packets actually achieve the upper bound to the capacity region.

The calculation of the capacity region is made under the usual assumption of unlimited reservoir of traffic at the source and under scenaria of both common as well as individual traffic streams for the two destinations. In addition the channel model has an arbitrary erasure structure that allows for independent or correlated erasures to the two receivers.

This paper was presented in part at the 44th Annual Allerton Conference on Communication, Control, and Computing, 2006, and at the 5th Workshop on Network Coding, Theory and Applications, 2009.

Y. E. Sagduyu is with the Institute for Systems Research, University of Maryland, College Park, MD 20742, and with the Intelligent Automation Inc., Rockville, MD 20855, USA. Email: yalinsagduyu@ieee.org

L. Georgiadis is with the Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, 54 124, Greece and with CERTH-ITI. Email: leonid@auth.gr

L. Tassioulas is with the Computer Engineering and Telecommunications Department, University of Thessaly, Volos, 38 221. Greece and with CERTH-ITI. Email:leandros@uth.gr

A. Ephremides is with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742. Email: etony@umd.edu
Copyright (c) 2011 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

We then turn our attention to the case, where traffic for the two receivers is generated randomly through stochastic arrival processes at the source. The goal is to maximize the achievable rates for both receivers while stabilizing the packet queue at the source. We consider again both common as well as separate traffic streams for the two destinations. We consider a slightly adapted form of the capacity-achieving algorithms and characterize the stable throughput region they achieve. Finally we modify these algorithms to enhance their delay performance (for a broad class of erasure probabilities and arrival rates) by switching from coding over packet blocks to dynamic coding at the individual packet level. The latter case reduces the average delay by eliminating the need for packet accumulation. Then, we characterize the resulting stable throughput regions that continue to be closely related to the ones obtained before. The stability regions bear remarkable similarity to the capacity region, although they need not coincide with it. This is mainly because of the extra overhead needed for the algorithms to adapt to random packet traffic.

The results of this study show a number of interesting facts. First, random linear coding and simple transmission management schemes achieve capacity. Second, essentially the same algorithms, when messages are generated randomly at the source, achieve a stable throughput region that is almost identical to the capacity region. Finally, a further modification of the algorithms through a queue management scheme that improves the delay performance, retains the shape and form of the stable throughput region. As a consequence, the results confirm the similarity between the two regions that have been observed in multi-access channels and provide additional evidence that there is a deeper relationship between the two regions that are both measuring rates but are defined in terms of different sets of assumptions and traffic scenarios.

The rest of the paper is organized as follows. Section II introduces the system model for broadcast erasure channels with feedback. We characterize the upper bound to the capacity region in section III and present the capacity achieving algorithms with linear network coding in Section IV. This is followed by the extension of the model to stochastic traffic in Section V, where we derive the stability region along with queue stabilizing algorithms with different levels of complexity, overhead requirement, and delay expectation. Finally, we draw conclusions in Section VI.

II. SYSTEM MODEL

We consider a slotted system where packets of length L bits are transmitted within each slot. We assume that the unit of time is the time needed to transmit one bit. Time slot $[(l-1)L, lL)$, $l = 1, 2, \dots$, is referred to as ‘‘slot l ’’.

The system consists of a single transmitter and two receivers. A packet transmitted in slot l is broadcasted to (may be heard by) both receivers [17]. At the end of slot l , a receiver may either receive the packet correctly, or the packet may be lost (dropped) for this receiver, in which case we say that an erasure occurred - we denote this event by the symbol E which is distinct from any of the regular packets. Define

$$Z_{i,l} = \begin{cases} 1 & \text{if an erasure occurs at receiver } i \text{ in the } l\text{th slot} \\ 0 & \text{otherwise} \end{cases}.$$

The random pair $(Z_{1,l}, Z_{2,l})$, $l = 1, 2, \dots$, has an arbitrary distribution, and the sequence $\{(Z_{1,l}, Z_{2,l})\}_{l=1}^{\infty}$ consists of independent identically distributed pairs. Denote by (Z_1, Z_2) a generic pair of $\{0, 1\}$ random variables having the distribution of $(Z_{1,l}, Z_{2,l})$ and define

$$\Pr(Z_1 = 1) = \varepsilon_1, \Pr(Z_2 = 1) = \varepsilon_2, \Pr(Z_1 Z_2 = 1) = \varepsilon_{12}.$$

In the sequel, to avoid trivial cases we assume that $\varepsilon_i < 1$, $i = 1, 2$.

According to the definitions above, if packet X_l is transmitted in time slot l then the output received by receiver i is,

$$Y_{i,l} = Z_{i,l}E + (1 - Z_{i,l})X_l. \quad (1)$$

At the end of slot l receiver $i = 1, 2$, provides feedback to the transmitter as to whether the packet has been received or erased, which according to (1) is equivalent to having each receiver inform the transmitter about the value of $Y_{i,l}$.

In the next section we provide an upper bound to the information theoretic capacity of this channel.

III. UPPER BOUND TO CHANNEL CAPACITY

The definitions and notation below are adapted from [17] and [18]. A generic sequence X_1, \dots, X_n is denoted by X^n . We denote by \mathcal{X} the set of transmitted messages, $\mathcal{X} = \mathcal{F}_q$, where \mathcal{F}_q is the finite field of size $q = 2^L$. We also denote the set of received symbols, common to both receivers, by $\mathcal{Y} = \mathcal{X} \cup \{E\}$.

Definition I: Let \mathcal{W}_i , $i = 1, 2$, be message index sets of size $\lceil 2^{nR_i} \rceil$ and let $\mathbf{W} \doteq (W_1, W_2) \in \mathcal{W}_1 \times \mathcal{W}_2 \doteq \mathcal{W}$. An $(\lceil 2^{nR_1} \rceil, \lceil 2^{nR_2} \rceil, n)$ code for the broadcast erasure channel with feedback consists of

- an encoder that at slot l transmits message X_l which is a function of $\mathbf{W}, X^{l-1}, Y_1^{l-1}, Y_2^{l-1}$ - when $l = 1$, X_1 is a function of \mathbf{W} only.
- two decoders, one for each of the receivers, represented by the functions,

$$g_i : \mathcal{Y}^n \rightarrow \mathcal{W}_i, \quad i = 1, 2.$$

Definition II: For a given channel code $(\lceil 2^{nR_1} \rceil, \lceil 2^{nR_2} \rceil, n)$, the conditional probability of erroneous decoding is defined as

$$\lambda_n(\mathbf{W}) = \Pr \left(\bigcup_{i=1}^2 \{g_i(Y_i^n) \neq W_i\} \mid \mathbf{W} \right), \mathbf{W} \in \mathcal{W}.$$

Definition III: A vector rate $\mathbf{R} = (R_1, R_2)$ is achievable if there exists a sequence of channel codes $(\lceil 2^{nR_1} \rceil, \lceil 2^{nR_2} \rceil, n)$, $n = 1, 2, \dots$, such that $\max_{\mathbf{W} \in \mathcal{W}} \lambda_n(\mathbf{W}) \rightarrow 0$ for $n \rightarrow \infty$.

Definition IV: The capacity region of the broadcast erasure channel with feedback is defined as the closure of the set of achievable rate vectors.

In case there is multicast traffic (intended to both receivers), then an additional message index W_0 needs to be transmitted, and this index must be received correctly by both receivers. The definitions above can be extended in a straightforward fashion to cover this case, for details see [17, Chapter 15.6.1].

The approach to the development of the outer bound for the capacity region of this channel is analogous to the methodology developed in [19] for the Gaussian broadcast channel with feedback. Specifically the approach consists of the following.

- Construct modified broadcast erasure channels with feedback, whose capacity region is superset of the capacity region of the system under consideration.
- Show that the modified systems are physically degraded (see [17, Chapter 15.6.2] for the definition). It follows, according to the result in [18], that the capacity region of the modified systems cannot be improved by feedback. Since the capacity region of the broadcast erasure channel without feedback is known [20], the capacity region of the modified systems with feedback is also known.
- Take the intersection of the capacity regions of the modified systems, which according to the above is an upper bound to the capacity region of the original system.

To apply the methodology described above, denote by C the channel under consideration and define a modified channel \widehat{C}_1 with $\widehat{Z}_{1,l} = Z_{1,l}$ and $\widehat{Z}_{2,l} = Z_{1,l}Z_{2,l}$. For channel \widehat{C}_1 an erasure to receiver 2 occurs only when erasures to both receivers in channel C occur. The following hold.

Lemma 1. *If \mathcal{C}_f and $\widehat{\mathcal{C}}_{1,f}$ are the feedback capacity regions of channels C and \widehat{C}_1 respectively, then*

$$\mathcal{C}_f \subseteq \widehat{\mathcal{C}}_{1,f}.$$

Proof: Channel \widehat{C}_1 can be viewed as an improvement of channel C , where receiver 1 informs receiver 2 about its feedback at each time slot. Hence any code designed for channel C can also be applied to \widehat{C}_1 with the same error probability ■

By direct calculation of the channel transition probabilities required for the definition of degraded channel [17, p. 564] it can be seen that,

Lemma 2. *Channel \widehat{C}_1 is physically degraded.*

Based on Lemma 1 and Lemma 2, Theorem 3 characterizes the outer bound to the feedback capacity region of broadcast erasure channels.

Theorem 3. *Let*

$$\begin{aligned} \widehat{\mathcal{C}}_1 &= \left\{ (r_1, r_2) \geq 0 : \frac{r_1}{1 - \varepsilon_1} + \frac{r_2}{1 - \varepsilon_{12}} \leq L \right\}, \\ \widehat{\mathcal{C}}_2 &= \left\{ (r_1, r_2) \geq 0 : \frac{r_1}{1 - \varepsilon_{12}} + \frac{r_2}{1 - \varepsilon_2} \leq L \right\}. \end{aligned}$$

Then

$$\mathcal{C}_f \subseteq \widehat{\mathcal{C}} \doteq \widehat{\mathcal{C}}_1 \cap \widehat{\mathcal{C}}_2 = \left\{ (r_1, r_2) \geq 0 : \max \left(\frac{r_1}{1 - \varepsilon_1} + \frac{r_2}{1 - \varepsilon_{12}}, \frac{r_1}{1 - \varepsilon_{12}} + \frac{r_2}{1 - \varepsilon_2} \right) \leq L \right\}.$$

Proof: The erasure probabilities for channel \widehat{C}_1 are, $\widehat{\varepsilon}_1 = \varepsilon_1$ and $\widehat{\varepsilon}_2 = \varepsilon_{12}$. Since channel \widehat{C}_1 is physically degraded, its capacity region is equal to its capacity region without feedback, i.e., $\widehat{\mathcal{C}}_{1,f} = \widehat{\mathcal{C}}_1$ [20]. Reversing the roles of receivers 1 and 2 we obtain $\widehat{\mathcal{C}}_{2,f} = \widehat{\mathcal{C}}_2$. The theorem follows from Lemma 1. ■

Including Multicast Traffic (Common Information)

Consider the situation where multicast traffic with rate r_{12} needs to be sent over the broadcast erasure channel. By combining the results in [17, Theorem 15.6.4] with [21] and [20], the capacity region of the physically degraded broadcast channel \widehat{C}_1 without feedback can then be computed as

$$\widehat{\mathcal{C}}_1^m = \left\{ (r_1, r_2, r_{12}) \geq 0 : \frac{r_1}{1 - \varepsilon_1} + \frac{r_2}{1 - \varepsilon_{12}} + \frac{r_{12}}{1 - \varepsilon_1} \leq L \right\}.$$

Extending the proof in [18] it can be shown that the capacity region of \widehat{C}_1 with feedback when multicast traffic exists, is again \widehat{C}_1^m . Repeating the arguments used in the previous sections, Theorem 4 characterizes the outer bound to the feedback capacity region of broadcast erasure channels with multicast traffic.

Theorem 4. *Let C_f^m be the capacity region of the broadcast channel with feedback when multicast traffic exists. Then,*

$$C_f^m \subseteq \widehat{C}^m = \left\{ (r_1, r_2, r_{12}) \geq 0 : \max \left(\frac{r_1}{1-\varepsilon_1} + \frac{r_2}{1-\varepsilon_{12}} + \frac{r_{12}}{1-\varepsilon_1}, \frac{r_1}{1-\varepsilon_{12}} + \frac{r_2}{1-\varepsilon_2} + \frac{r_{12}}{1-\varepsilon_2} \right) \leq L \right\}.$$

IV. CAPACITY ACHIEVING ALGORITHMS

In this section we present two algorithms whose rate region is very close to the upper bound to system capacity described in Theorem 3. The first algorithm is based on forming random linear combinations of sets of packets which are determined by the received feedback. The second algorithm is based on performing XOR operations on pairs of packets, depending on the received feedback. We compare the two algorithms at the end of this section. Note that this setup has a natural connection to index coding [22], where the receivers have some prior information and the source transmits linearly coded packets until all receivers can decode their intended packets. In our case, there is no *a priori* side information available but the source adjusts transmissions through network coding while the side information builds up at the receivers dynamically depending on random channel erasure events.

A. Algorithm Based on Random Linear Coding

In this section we present a coding algorithm whose rate region is very close to region \widehat{C} for reasonably large packet sizes. We assume that a fixed number of packets for each receiver must be transmitted. For ease of presentation, instead of considering a fixed block code length n (i.e., fixed number of transmission slots within which all packets must be transmitted or an error occurs) we will consider that packets are transmitted until both receivers decode correctly the packets destined to them and will analyze the number of slots required to do so.

The algorithm operates in two phases. Phase 1 is subdivided in Phase 1.1 and Phase 1.2. In Phases 1.1 and 1.2, the transmitter transmits packets destined to receivers 1 and 2 respectively, possibly retransmitting a packet until *at least one* of the receivers receives the packet. Phase 2 consists of transmitting coded packets, where each packet is a random linear combination [23] of packets destined to either one of the receivers. The latter packets are determined by the feedback received during Phase 1.

Algorithm I

The packets destined to receiver $i = 1, 2$ are placed in queue $Q_{0,i}^t$ at the transmitter. The number of packets in queue $Q_{0,i}^t$ is denoted by k_i . The numbers k_1 and k_2 are known to the receivers.

The following queues, initially empty, are maintained. The content of these queues are determined by the operation of the algorithm.

- **Transmitter:** Two packet queues Q_1^t and Q_2^t . Queue Q_1^t holds transmitted packets destined to receiver 1, erased by receiver 1 and seen by receiver 2. Queue Q_2^t is similarly defined with the roles of receivers reversed.
- **Receiver i :** A packet queue Q^i . Queue Q^i contains all packets received by receiver i , irrespective of their destination. Also, a queue R_i where the packets destined to receiver i and received during Phase 1.1 are placed, after processing performed at the end of Phase 1 as will be described below.

A single control bit, b_1 , from each packet is reserved to convey control information to the receivers. When $b_1 = 1$, the bit indicates the end of Phase 1 and the information part of the packet conveys to the receivers the feedback information during Phase 1, as will be described below. Otherwise, $b_1 = 0$.

1) Phase 1.1:

- **Transmitter:** Transmit k_1 receiver 1 packets from queue $Q_{0,1}^t$, as follows.
 - a) Set $b_1 = 0$.
 - b) A packet is (re)transmitted until it is received by *at least one* of the receivers, in which case the packet is removed from $Q_{0,1}^t$.
 - c) If a transmitted packet is received by receiver 2 and erased at receiver 1, it is placed in queue Q_1^t . Hence Q_1^t contains all packets destined to receiver 1 and seen only by receiver 2.
 - d) Keep a log of the feedback received by both receivers. For each transmission, the log consists of two bits, where bit $i = 1, 2$ indicates whether the transmitted packet has been received by receiver i (bit set to 1) or has been erased at this receiver (bit set to 0).
- **Receiver i :** Place every received packet in queue Q^i . Send the appropriate feedback (ACK if the packet is received, or NACK if the packet is erased).

2) **Phase 1.2:** Same rules as Phase 1.1, with the roles of receiver 1 and 2 reversed.

3) When both queues $Q_{0,1}^t, Q_{0,2}^t$ become empty, Phase 1 is complete. The transmitter creates packets where the feedback information kept on the log is placed. For these packets, $b_1 = 1$. The packets are transmitted until they are received by

both receivers. Since the receivers know k_1 and k_2 , by observing the feedback log they can obtain information about the packets stored in their corresponding queues Q^i , $i = 1, 2$. Specifically, receiver i can decide where Phase 1.1 and Phase 1.2 ends, whether a received packet is destined to itself and the order of this packet in queue $Q_{0,i}^t$, whether a transmitted packet destined to itself has been erased at this receiver and received by the other receiver, and whether a received packet is destined to and has been erased at the other receiver. Hence receiver i can reconstruct queue Q_j^t , $j \neq i$, and can place all received packets destined to itself in queue R_i . Receiver i can also infer the order in queue $Q_{0,i}^t$ of each packet in queue Q_i^t . There may be more efficient ways to convey the log to the receivers, but the method presented here suffices for our purposes. Note that at this point, receiver i needs to recover only the packets in queue Q_i^t .

4) **Phase 2:** Let M_i be the number of packets in queue R_i at the end of Phase 1 (end of Phase 1.2). Let K_i^r be the number of packets in queue Q_i^t , $i = 1, 2$, at the beginning of Phase 2. Note that since receiver i can reconstruct Q_j^t , $j \neq i$, it knows K_j^r . Also, since $k_i = M_i + K_i^r$, receiver i knows K_i^r .

- **Transmitter:** The following actions are taken

- a) Set $b_1 = 0$.

- b) Transmit random linear combinations of the $L - 1$ information bits of the $K_1^r + K_2^r$ packets in queues Q_i^t , $i = 1, 2$ -bit b_1 is not encoded. That is, the $L - 1$ information bits of each packet are considered to form an element of $\mathcal{F}_{2^{L-1}}$; for each such elements p_i , a uniformly random coding coefficient a_i from $\mathcal{F}_{2^{L-1}}$ is selected and then the linear combination $p = \sum_{i=1}^{K_1^r+K_2^r} a_i p_i$ is formed. The transmitted packet consists of the concatenation (b_1, p) . The coding coefficients are generated by a random number generation algorithm known *a priori* to both receivers, with the same seed, and hence these coefficients do not need to be transmitted. Transmissions are taking place until a *single* feedback (see actions of the receivers below) declaring successful decoding is received by both receivers - in case $K_i^r = 0$ no feedback from receiver i is required since the receiver has already received the packets destined to it.

- **Receiver 1:** Since the receiver knows the content of the K_2^r packets in queue Q_2^t , only the K_1^r packets in Q_1^t destined to receiver 1 need to be recovered from the observed linear combinations of the $K_1^r + K_2^r$ packets. The receiver observes the received packets until it has enough information (dimensions) to recover the K_1^r packets through the associated system of linear equations. Feedback is sent to the transmitter *only* when this recovering is successful. At this point the receiver knows the contents of all packets in queue $Q_{0,1}^t$ as well as their order in this queue (obtained earlier though the received feedback log) and hence decoding is successful.

- **Receiver 2:** The corresponding actions as receiver 1.

1) *Analysis of Algorithm 1:* We assume k_1, k_2 large so that we can invoke the law of large numbers to replace averages with sample values. We skip technical details involved in this replacement since, while based on known techniques, they are tedious and distract from the main results.

Consider first Phases 1.1 and 1.2. The number, N_i , of transmissions needed to complete Phase 1. i of the algorithm, i.e., for receiver i packets to be correctly received by at least one of the receivers, is $N_i = \frac{k_i}{1-\varepsilon_{12}}$, $i = 1, 2$. The number, M_1 , of receiver 1 packets transmitted during Phase 1 that were delivered to receiver 1, is $M_1 = k_1 \frac{1-\varepsilon_1}{1-\varepsilon_{12}}$. Similarly, for M_2 we have $M_2 = k_2 \frac{1-\varepsilon_2}{1-\varepsilon_{12}}$, and hence for $i = 1, 2$, $K_i^r = k_i - M_i = k_i \frac{\varepsilon_i - \varepsilon_{12}}{1-\varepsilon_{12}}$.

Regarding the log-containing packets that need to be transmitted at the end of Phase 1, note that the number of bits contained in the log is $2(N_1 + N_2)$ and hence the number of packets that need to be transmitted is

$$M_{12} = \left\lceil \frac{2(k_1 + k_2)}{(L-1)(1-\varepsilon_{12})} \right\rceil.$$

If the transmitter broadcasts these packets by transmitting random linear combinations of them, then for sufficiently large k_1, k_2 , the number of transmissions needed until both receives are able to decode the packets is $N_{12} = M_{12}/(1-\varepsilon_0)$, where $\varepsilon_0 = \max\{\varepsilon_1, \varepsilon_2\}$.

Consider now Phase 2. Since receiver 1 knows K_2^r out of the $K_1^r + K_2^r$ packets involved in each of the random linear combinations received, the receiver will be able to recover the K_1^r packets that it needs for successful decoding, after receiving correctly $G_1^r \leq \frac{2^{L-1}}{2^{L-1}-1} K_1^r$ such linear combinations - the coefficient $\frac{2^{L-1}}{2^{L-1}-1}$ accounts for the possibility of receiving "non-innovative" packets [12]. The number of transmissions needed in order to correctly receive these G_1^r packets is $G_1^r/(1-\varepsilon_1)$. A similar argument holds for receiver 2. Hence the number of transmissions needed so that both receivers recover their undelivered packets during Phase 2 is,

$$\begin{aligned} N_3 &= \max \left\{ \frac{G_1^r}{1-\varepsilon_1}, \frac{G_2^r}{1-\varepsilon_2} \right\} \\ &\leq \frac{2^{L-1}}{2^{L-1}-1} \max \left\{ \frac{k_1(\varepsilon_1 - \varepsilon_{12})}{(1-\varepsilon_1)(1-\varepsilon_{12})}, \frac{k_2(\varepsilon_2 - \varepsilon_{12})}{(1-\varepsilon_2)(1-\varepsilon_{12})} \right\}. \end{aligned}$$

From the above we conclude that the number of slots needed for successful reception of the k_1, k_2 packets by the respective

receivers is,

$$\begin{aligned}
T &= N_1 + N_2 + N_{12} + N_3 \\
&\leq \frac{k_1}{1-\varepsilon_{12}} + \frac{k_2}{1-\varepsilon_{12}} + \frac{1}{1-\varepsilon_0} \left\lceil \frac{2(k_1+k_2)}{(L-1)(1-\varepsilon_{12})} \right\rceil + \frac{2^{L-1}}{2^{L-1}-1} \max \left\{ \frac{k_1(\varepsilon_1-\varepsilon_{12})}{(1-\varepsilon_1)(1-\varepsilon_{12})}, \frac{k_2(\varepsilon_2-\varepsilon_{12})}{(1-\varepsilon_2)(1-\varepsilon_{12})} \right\} \\
&\leq \frac{2^{L-1}}{2^{L-1}-1} \left(\max \left\{ \frac{k_1}{1-\varepsilon_1} + \frac{k_2}{1-\varepsilon_{12}}, \frac{k_1}{1-\varepsilon_{12}} + \frac{k_2}{1-\varepsilon_2} \right\} \right) + \frac{2}{(L-1)(1-\varepsilon_0)} \frac{k_1+k_2}{1-\varepsilon_{12}} + \frac{2}{(L-1)(1-\varepsilon_0)} \quad (2)
\end{aligned}$$

$$< \frac{2^{L-1}}{2^{L-1}-1} \left(1 + \frac{2}{(L-1)(1-\varepsilon_0)} \right) \left(\max \left\{ \frac{k_1}{1-\varepsilon_1} + \frac{k_2}{1-\varepsilon_{12}}, \frac{k_1}{1-\varepsilon_{12}} + \frac{k_2}{1-\varepsilon_2} \right\} \right) + \frac{2}{(L-1)(1-\varepsilon_0)}. \quad (3)$$

Since the receiver i rate (in information bits per packet transmission) is $r_i = k_i(L-1)/T$, we conclude from the last inequality that

$$\left(1 + \frac{2}{(L-1)(1-\varepsilon_0)} \right) \max \left\{ \frac{r_1}{1-\varepsilon_1} + \frac{r_2}{1-\varepsilon_{12}}, \frac{r_1}{1-\varepsilon_{12}} + \frac{r_2}{1-\varepsilon_2} \right\} + \frac{2}{(L-1)(1-\varepsilon_0)T} \geq \frac{2^{L-1}-1}{2^{L-1}} (L-1),$$

and since T can be made arbitrarily large by picking a large value of $k_1 + k_2$, the region of achievable rates is at least,

$$\mathcal{R}^I = \left\{ (r_1, r_2) \geq 0 : M(r_1, r_2) \leq \frac{2^{L-1}-1}{2^{L-1}} \frac{L-1}{1 + \frac{2}{(L-1)(1-\varepsilon_0)}} \right\},$$

where

$$M(r_1, r_2) = \max \left\{ \frac{r_1}{1-\varepsilon_1} + \frac{r_2}{1-\varepsilon_{12}}, \frac{r_1}{1-\varepsilon_{12}} + \frac{r_2}{1-\varepsilon_2} \right\}.$$

For moderately large L , region \mathcal{R}^I is a close approximation of the region defined by Theorem 3. In fact, in units of information bits per transmitted bit (i.e., scaling by L) the region \mathcal{R}^I becomes arbitrarily close to the bound given by Theorem 3.

2) *Including Multicast Traffic:* Algorithm I can be modified to handle multicast traffic. We present the modification below.

Algorithm I.1:

Let k_1, k_2, k_{12} be the number of packets to be transmitted from each of the three sessions (receiver 1, receiver 2 and multicast). These numbers are known to the receivers.

- Employ Phases 1.1, 1.2 of Algorithm I
- In Phase 2 employ random linear coding of packets K_1^T, K_2^T, k_{12} , i.e., include the multicast session packets in the process.

An analysis similar to the analysis of Algorithm I shows that the rate region Algorithm I.1 is at least

$$\mathcal{R}^{I.1} = \left\{ (r_1, r_2, r_{12}) \geq 0 : M_1(r_1, r_2, r_{12}) \leq \frac{2^{L-1}-1}{2^{L-1}} \frac{L-1}{1 + \frac{2}{(L-1)(1-\varepsilon_0)}} \right\}, \quad (4)$$

where

$$M_1(r_1, r_2, r_{12}) = \max \left\{ \frac{r_1}{1-\varepsilon_1} + \frac{r_2}{1-\varepsilon_{12}} + \frac{r_{12}}{1-\varepsilon_1}, \frac{r_1}{1-\varepsilon_{12}} + \frac{r_2}{1-\varepsilon_2} + \frac{r_{12}}{1-\varepsilon_2} \right\}.$$

Again, we see from (4) that the rate of Algorithm I.1 is very close to the upper bound on system capacity described in Theorem 4, for reasonably large packet sizes.

B. Algorithm Using Only XOR Operations

In this section we provide an algorithm that codes using only XOR operations between packets and also approximates closely the system capacity for reasonably large L .

Algorithm II

Two integers k_1 and k_2 are selected. These integers denote respectively the number of packets from each receiver that must be transmitted. These packets are placed in queues $Q_{0,1}^t$ and $Q_{0,2}^t$ respectively.

The following buffers, initially empty, are maintained. The contents of these buffers are determined by the operation of the algorithm.

- **Transmitter:** Two single-packet buffers B_1^t, B_2^t . Initially these buffers are empty. At the beginning of slot l , buffer B_1^t is either empty or contains a receiver 1 packet that has been transmitted at some prior time, erased at receiver 1 and received by receiver 2. Buffer B_2^t is similarly defined with the roles of 1 and 2 reversed. We denote by $B_{i,l}^t$ the contents of buffer B_i^t at the beginning of slot l .
- **Receiver 1:** A single-packet buffer B_2^1 . Initially B_2^1 is empty. We denote by $B_{2,l}^1$ the contents of buffer B_2^1 at the beginning of slot l .
- **Receiver 2:** A single-packet B_1^2 . Initially B_1^2 is empty. We denote by $B_{1,l}^2$ the contents of buffer B_1^2 at the beginning of slot l .

Two bits b_1, b_2 are reserved for control purposes. These bits are set as follows.

- b_1 indicates whether the transmitted packet is the result of an XOR operation, i.e., $b_1 = 0$: non-XORed packet, $b_1 = 1$: XORed packet
- b_2 indicates the receiver to which the packet is destined (if it is not the result of an XOR operation), i.e. $b_2 = 0$: receiver 1 packet, $b_2 = 1$: receiver 2 packet

The algorithm consists of two phases as follows. Let $R_i(l)$ be the number of receiver i packets at the beginning of slot l , that have not been received yet by receiver i .

Phase 1: This phase lasts as long as $R_i(l) > 0$ for both $i = 1, 2$.

- **Transmitter:** Whenever a packet is transmitted in slot l and an erasure occurs to both receivers, i.e., $Z_{1,l} = 0, Z_{2,l} = 0$, the packet is retransmitted. If the packet is destined to receiver 1 and is received by this receiver, i.e., $Z_{1,l} = 1, Z_{2,l} = 0$ or 1, the packet is removed from queue $Q_{0,1}^t$. Similar actions are taken if the packet is destined to receiver 2. The rest of the cases are described below.
 - 1) If $B_{1,l}^t = \emptyset, B_{2,l}^t = \emptyset$ then transmit a receiver 1 packet in slot l , setting $b_1 = 0, b_2 = 0$; At the end of slot l ,
 - a) If $Z_{1,l} = 1, Z_{2,l} = 0$ then remove the transmitted packet from queue $Q_{0,1}^t$ and place it in buffer B_1^t .
 - 2) If $B_{1,l}^t \neq \emptyset, B_{2,l}^t = \emptyset$ then transmit a receiver 2 packet in slot l , setting $b_1 = 0, b_2 = 1$; At the end of slot l ,
 - a) If $Z_{1,l} = 0, Z_{2,l} = 1$ then remove the transmitted packet from queue $Q_{0,2}^t$ and place it in buffer B_2^t .
 - 3) If $B_{1,l}^t = \emptyset, B_{2,l}^t \neq \emptyset$ then transmit a receiver 1 packet in slot l , setting $b_1 = 0, b_2 = 0$; At the end of slot l ,
 - a) If $Z_{1,l} = 1, Z_{2,l} = 0$ then remove the transmitted packet from queue $Q_{0,1}^t$ and place it in buffer B_1^t .
 - 4) If $B_{1,l}^t \neq \emptyset, B_{2,l}^t \neq \emptyset$ then perform a bitwise XOR of the *information part* (i.e., $L - 2$ bits) of the packets in $B_{1,l}^t, B_{2,l}^t$, and transmit the result in slot l , setting $b_1 = 1$; At the end of slot l ,
 - a) If $Z_{1,l} = 1, Z_{2,l} = 0$ then remove the packet in buffer B_1^t . Else,
 - b) If $Z_{1,l} = 0, Z_{2,l} = 1$ then remove the packet in buffer B_2^t . Else,
 - c) If $Z_{1,l} = 0, Z_{2,l} = 0$ then remove the packet from both buffers B_1^t, B_2^t .
- **Receiver 1:** At the end of each slot the receiver sends the appropriate (ack, nack) feedback to the transmitter.
 - 1) If at the end of slot l receiver 1 receives a packet destined to receiver 2 ($b_1 = 0, b_2 = 1$), it places this packet in B_2^1 , replacing any other packet that may exist.
 - 2) If at the end of slot l receiver 1 receives an XORed packet ($b_2 = 1$) then it XORs the information part of this packet with the information part of the packet stored in buffer B_2^1 - the result is a receiver 1 packet.
 - 3) If at the end of slot l receiver 1 receives a packet destined to itself ($b_1 = 0, b_2 = 0$) it accepts the packet.
- **Receiver 2:** Similar actions, with the roles of 1 and 2 reversed.

Phase 2: At this phase, packets destined to only one of the receivers, say receiver 1, remain to be transmitted. In this case, the transmitter sets $b_1 = 0, b_2 = 0$, and retransmits the packets until it is ensured, through the received feedback that all packets have been received correctly by receiver 1. The receivers act as in Phase 1.

Notes:

- When $B_{1,l}^t = \emptyset, B_{2,l}^t = \emptyset$, the transmitter may transmit receiver 2 packets, or randomly packets destined to either of the receiver. This does not alter the throughput region of the algorithm.
- With this algorithm, the receivers do not need to know the number of packets k_1, k_2 .

In the next section we discuss the correctness and performance analysis of this algorithm.

1) *Analysis of Algorithm II:* We must show first that the operation of the algorithm is correct. From the description of the algorithm it follows that at any time during Phase 1, B_1^t and B_2^t contain at most one packet. Note that according to the algorithm, (the information part of) an XORed packet q sent at time slot l is of the form $q = B_{1,l}^t \oplus B_{2,l}^t$. When a receiver, say receiver 1, receives q in slot l , it performs the operation $\tilde{q}_1 = q \oplus B_{2,l}^1$. In order for receiver 1 to recover $B_{1,l}^t$, it must hold that $\tilde{q}_1 = B_{1,l}^t$, i.e., it must be ensured that whenever an XORed packet is received by receiver 1, $B_{2,l}^1 = B_{2,l}^t$. This is not true always since it may happen that if the transmitted packet is not an XORed packet then $B_{2,l}^1 \neq B_{2,l}^t$ - this will be the case if the transmitted packet is destined to transmitter 2 and is received by both receivers. However, from the description of the algorithm, it follows that whenever $B_{1,l}^t \neq \emptyset$, then $B_{2,l}^1 = B_{1,l}^t$ and similarly, whenever $B_{2,l}^t \neq \emptyset$, then $B_{2,l}^1 = B_{2,l}^t$. Since the transmitter transmits an XORed packet during Phase 1 only if $B_{1,l}^t \neq \emptyset$ and $B_{2,l}^t \neq \emptyset$, it follows that

Lemma 5. *During Phase 1, whenever receiver i receives an XORed packet in slot l , it can correctly recover its own packet by XORing the received packet with packet in $B_{i,l}^t$.*

We now proceed with the analysis of Algorithm II. As with Algorithm I, we assume that k_1 and k_2 are large so that we can invoke the law of large numbers to replace averages with sample values. During Phase 1 the operation of the algorithm is described as a Markov Chain with 4 states. The states describe the content of buffers $|B_1^t|, |B_2^t|$, as follows.

$$\begin{aligned} A &= (|B_1^t|, |B_2^t|) = (0, 0), \quad B = (1, 0), \\ C &= (0, 1), \quad D = (1, 1). \end{aligned}$$

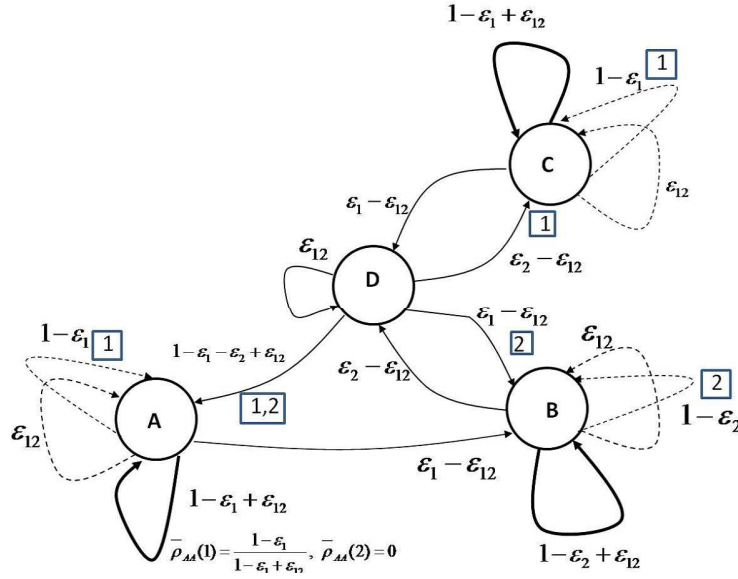


Figure 1. Markov Chain description of Phase 1 of Algorithm II.

A state transition occurs each time a transmission takes place. The rewards consist of pairs $(\rho_{ss'}(1), \rho_{ss'}(2))$ where for a transition from state s to state s' , $\rho_{ss'}(i)$ corresponds to the number of receiver i packets received *correctly* by receiver i during the corresponding transmission. For transitions $A \rightarrow A$, $B \rightarrow B$ and $C \rightarrow C$ this number is random - see the description of state transitions in the next paragraph - and we denote with $\bar{\rho}_{ss'}(i)$ its average value. The complete state diagram is described in Figure 1. A number i in a square next to a transition arrow indicates that a receiver i packet has been received correctly by receiver i under the corresponding transition, hence $\rho_{ss'}(i) = 1$. If the number i does not exist next to a transition, then the corresponding reward is zero. State transitions are indicated by solid arrows. Dotted arrows indicate sub-transitions that compose the transition indicated by the corresponding solid bold arrow - see the description in the next paragraph. The rewards of the bold arrows are averages of the costs of the corresponding dotted ones. To avoid overloading the figure, these latter rewards are indicated only for the transition $A \rightarrow A$. The formulas next to each arrow indicate transition probabilities.

We explain next how the transitions and rewards are defined when in state A . In this state, according to the algorithm receiver 1 packets are transmitted. From this state, a transition to the same state, $A \rightarrow A$, occurs if either a) the transmitted packet is erased at both receivers, an event of probability ε_{12} , or b) it is correctly received by receiver 1, an event of probability $1 - \varepsilon_1$. Hence the overall transition probability is, $1 - \varepsilon_1 + \varepsilon_{12}$. Given that this transition occurs, a reward of one receiver 1 packet is assigned only if the packet is correctly received by receiver 1, an event of conditional probability $(1 - \varepsilon_1)/(1 - \varepsilon_1 + \varepsilon_{12})$. There are no receiver 2 rewards since the transmitted packets are destined to receiver 1. Hence the average rewards for this transition are $\bar{\rho}_{AA}(1) = (1 - \varepsilon_1)/(1 - \varepsilon_1 + \varepsilon_{12})$, $\bar{\rho}_{AA}(2) = 0$. A transition $A \rightarrow B$ occurs when a transmitted receiver 1 packet is erased at receiver 1 and received correctly at receiver 2, an event of probability $\varepsilon_1 - \varepsilon_{12}$. In this case, $\bar{\rho}_{AB}(1) = \rho_{AB}(1) = \rho_{AB}(2) = \bar{\rho}_{AB}(2) = 0$ since the packet has not been received correctly by the intended destination. The rest of the transitions and rewards are similarly obtained.

Let π_s , $s \in \mathcal{S} = \{A, B, C, D\}$, be the steady state probability of the Markov Chain described above. The long term average number of successfully transmitted receiver i packets is then,

$$\bar{\rho}_i = \sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} \bar{\rho}_{ss'}(i) p_{ss'} \pi_s, \quad (5)$$

where $p_{ss'}$ denote transition probabilities. Hence, for large k_i the number of slots needed to transmit the k_i packets is $k_i/\bar{\rho}_i$ and the number of time slots needed for Phase 1 to complete is $N_1 = \min \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}$.

Let now $k_1/\bar{\rho}_1 \leq k_2/\bar{\rho}_2$. Then, Phase 1 lasts $N_1 = k_1/\bar{\rho}_1$ slots. During this phase, receiver 1 receives all its packets and the number of receiver 2 packets transmitted is $N_1\bar{\rho}_2$. Hence there are

$$\hat{K}_2 = k_2 - N_1\bar{\rho}_2 = k_2 - k_1 \frac{\bar{\rho}_2}{\bar{\rho}_1},$$

receiver 2 packets left to be transmitted during phase 2. The number of slots needed for the latter packets to be transmitted is

$$N_2 = \frac{\hat{K}_2}{1 - \varepsilon_2} = \frac{k_2}{1 - \varepsilon_2} - \frac{k_1\bar{\rho}_2}{(1 - \varepsilon_2)\bar{\rho}_1}.$$

Hence the total number of slots needed for Algorithm II to complete is

$$\begin{aligned}
T &= N_1 + N_2 \\
&= \frac{k_2}{1 - \varepsilon_2} + k_1 \left(\frac{1}{\bar{\rho}_1} - \frac{\bar{\rho}_2}{(1 - \varepsilon_2)\bar{\rho}_1} \right).
\end{aligned} \tag{6}$$

The probabilities π_s and then $\bar{\rho}_1, \bar{\rho}_2$, can be computed easily based on the transition diagram in Figure 1. It can then be seen that the following hold.

$$\frac{\bar{\rho}_1}{\bar{\rho}_2} = \frac{(\varepsilon_1 - 1)(\varepsilon_2 - \varepsilon_{12})}{(\varepsilon_2 - 1)(\varepsilon_1 - \varepsilon_{12})}, \quad \frac{1}{\bar{\rho}_1} - \frac{\bar{\rho}_2}{(1 - \varepsilon_2)\bar{\rho}_1} = \frac{1}{1 - \varepsilon_{12}}. \tag{7}$$

Replacing in (6) we conclude,

$$T = \frac{k_2}{1 - \varepsilon_2} + \frac{k_1}{1 - \varepsilon_{12}} \geq \frac{k_2}{1 - \varepsilon_{12}} + \frac{k_1}{1 - \varepsilon_1}. \tag{8}$$

Similarly, if $k_1/\bar{\rho}_1 > k_2/\bar{\rho}_2$, we conclude,

$$T = \frac{k_2}{1 - \varepsilon_{12}} + \frac{k_1}{1 - \varepsilon_1} \geq \frac{k_2}{1 - \varepsilon_2} + \frac{k_1}{1 - \varepsilon_{12}}. \tag{9}$$

From (8) and (9) we conclude that Algorithm II achieves rates (in bits per transmitted packet),

$$\mathcal{R}^{II} = \{r \geq \mathbf{0} : M(r_1, r_2) \leq L - 2\},$$

which is essentially the same region as that of Algorithm I for large L .

2) *Comparison of Algorithms I and II:* We have seen that the throughput region of both algorithms approximates the capacity region of the system very closely for reasonably large packet sizes. Both algorithms use the idea of performing linear combinations of packets in situations where both receivers are able to obtain useful information by observing the same packet. Algorithm I performs random linear combinations of packets, considering each one as an element of the field $\mathcal{F}_{2^{L-1}}$. Algorithm II on the other hand performs (deterministic) XOR operation on two packets at a time, which can be viewed as modulo 2 addition of vectors whose coordinates are in \mathcal{F}_2 , i.e., the information part of each packet is considered as a vector of $L - 2$ bits $\mathbf{p} = (b_3, \dots, b_L)$ - bits b_1 and b_2 are control bits. These differences have the following implications regarding implementation complexity and flexibility:

- Algorithm II is simpler to implement. It imposes minimal requirements on the receivers as they are only required to maintain single-packet buffers for packets that may need to be XORed. Moreover, the XOR operations are very simple to implement both at the transmitter and the receivers. Algorithm I on the other hand, requires operations in the field $\mathcal{F}_{2^{L-1}}$. This part of the complexity can be reduced by performing operations in a smaller field \mathcal{F}_{2^l} , $l < L$, with minimal effect on the throughput region (the factor $(2^{L-1} - 1)/2^{L-1}$ is replaced by $(2^l - 1)/2^l$). However the task to recover packets from the received linear combinations in Phase 2, i.e., inverting matrices, still remains.
- In Phase 2, Algorithm I requires only a single feedback from each receiver, while Algorithm II requires feedback at each slot until completion. Note that per packet feedback is needed by both algorithms during the corresponding Phase 1.
- Multicast traffic can be easily handled by Algorithm I, without increasing the overhead required for transferring control information to the receivers. While Algorithm II can be modified to handle multicast traffic, this modification renders Lemma 5 invalid, implying that now the receivers need to maintain multi-packet queues instead of single-packet buffers and the required control overhead increases. We defer the inclusion of multicast traffic in an algorithm using only XOR operations to Section V where we address the situation of handling a system where packets arrive randomly at each slot.
- Algorithm I is more amenable to generalization to a system with more than two receivers. In fact, recent work [24], [25] demonstrates that using linear coding, algorithms can be developed whose rate region is very close to the upper bound to the capacity for systems with 3 receivers and under certain statistical assumptions for systems with $N > 3$ receivers.

V. STOCHASTIC ARRIVALS

In this section we assume that packets arrive randomly to the system instead of assuming backlogged packet queues as was done above. We first show how the algorithms of the previous section can be adapted to such an environment to obtain a stability region that is close to the capacity region of the system. We note that the principles on which this adaptation is based have potential applications to other communication systems as well. Then we present additional algorithms with improved delay performance for the system under consideration.

Let $A_i(l)$, $i = 1, 2$, be the number of packet arrivals during slot l , with destination receiver i . We assume that $\{A_1(l), A_2(l)\}_{l=1}^{\infty}$ are i.i.d. with arrival rates $\lambda_i = \mathbb{E}[A_i(l)]$, $i = 1, 2$. To simplify the discussion, we present the stability results measuring rates in “packets per slot” without taking into account the overhead required to transfer control information to the receivers; subsequently, we discuss issues related to overhead. In packets per slot, the upper bound on the capacity region in Theorem 4 becomes,

$$\mathcal{R} = \left\{ (r_1, r_2) \geq \mathbf{0} : \max \left\{ \frac{r_1}{1 - \varepsilon_1} + \frac{r_2}{1 - \varepsilon_{12}}, \frac{r_1}{1 - \varepsilon_{12}} + \frac{r_2}{1 - \varepsilon_2} \right\} \leq 1 \right\},$$

and in case multicast traffic also exists,

$$\mathcal{R}^m = \left\{ (r_1, r_2, r_{12}) \geq 0 : \max \left(\frac{r_1}{1 - \varepsilon_1} + \frac{r_2}{1 - \varepsilon_{12}} + \frac{r_{12}}{1 - \varepsilon_1}, \frac{r_1}{1 - \varepsilon_{12}} + \frac{r_2}{1 - \varepsilon_2} + \frac{r_{12}}{1 - \varepsilon_2} \right) \leq 1 \right\}.$$

In the following we will show that for any arrival rates in the interior of \mathcal{R} , the proposed algorithms stabilize the system - barring the overhead bits needed to transfer control information which, as will be seen, is small.

Before proceeding we define stability for the systems under consideration. There are many definitions of stability in the literature. For our purposes we adopt the following. For vectors quantities, the notation $\mathbf{X}(l) \leq \mathbf{x}$ means coordinate-wise inequalities.

Definition (Process Stability). Let $\{\mathbf{X}(l)\}_{l=1}^{\infty} = \{(X_1(l), \dots, X_m(l))\}_{l=1}^{\infty}$ be a multi-dimensional stochastic process. The process is stable if it converges in distribution to an ‘‘honest’’ distribution, i.e., the following holds at all points of continuity of some distribution function $F(\mathbf{x})$,

$$\lim_{l \rightarrow \infty} \Pr \{\mathbf{X}(l) \leq \mathbf{x}\} = F(\mathbf{x}), \text{ and } \lim_{\min\{x_1, \dots, x_m\} \rightarrow \infty} F(\mathbf{x}) = 1.$$

The process is called substable if

$$\lim_{\min\{x_1, \dots, x_m\} \rightarrow \infty} \liminf_{l \rightarrow \infty} \Pr \{\mathbf{X}(l) \leq \mathbf{x}\} = 1.$$

It is well known that if $\{X_i(l)\}_{l=1}^{\infty}$ is substable for all $i \in \{1, \dots, m\}$, then the process is substable and that if the process is an irreducible aperiodic homogeneous Markov Chain with countable state space, then substability implies ergodicity, see, e.g., [26]. In our case $\{\mathbf{X}(l)\}_{l=1}^{\infty}$ will represent the process of queue lengths.

The adaptations of Algorithms I, I.1 and II to operate in the current environment are similar. For the sake of definiteness we concentrate on the adaptation of Algorithm II; however as will be indicated the approach can be used for the other two algorithms with some simple modifications.

A. Adaptation of Algorithm II to Stochastic Traffic

The transmitter has two queues of infinite size, $Q_{0,1}^t, Q_{0,2}^t$, where the exogenously arriving packets for receivers 1 and 2 respectively, are queued. The following algorithm is a natural adaptation of Algorithm II to the current environment.

Algorithm III

Let $K_i(T)$, $i = 1, 2$, be the number of packets queued at $Q_{0,i}^t$ at time T . The algorithm operates in *epochs*. Epoch 1 starts at time $T_1 = 0$. If $K_1(0) = K_2(0) = 0$, the epoch ends at the time the first slot ends, T_2 , and epoch 2 starts at the same time. Otherwise, Algorithm II with packets $k_i = K_i(T_1)$, $i = 1, 2$, is used for the transmission of these packets. New arrivals are queued in queues $Q_{0,1}^t, Q_{0,2}^t$, but are not transmitted until the epoch ends.

In general, after the end of epoch j at time T_{j+1} , epoch $j + 1$ starts at the same time, employing the same procedure as in epoch 1, with packets $k_i = K_i(T_{j+1})$, $i = 1, 2$.

1) *Stability analysis of Algorithm III:* Under the stated assumptions on the process of packet arrivals, the process of number of queued packets at epoch beginnings, $\{(K_1(T_j), K_2(T_j))\}_{j=1}^{\infty}$, forms an irreducible aperiodic homogeneous Markov chain with countable state space. Below we will show using Lyapunov function drift analysis that if the arrival rates are within the region \mathcal{R} (measured in packet arrivals per slot) then this Markov chain is ergodic. This will imply that the stochastic process $\{Q_1(l), Q_2(l)\}_{l=1}^{\infty}$ representing the number of packets destined to receiver 1, 2 that are in the system at slot l is stable.

We will use the following theorem from [27] (see also [28]) which we present in a form appropriate for the problem under consideration. In the stability analysis below we will use capital letters to denote random variables and small case letters to denote values of random variables.

Theorem 6. Let $\{X_n\}_{n=1}^{\infty}$ be an irreducible and aperiodic homogeneous Markov Chain with countable state space \mathcal{S} . Let $v(x)$ be a nonnegative real function defined on the state space (Lyapunov function). If there exists a finite set of states $\mathcal{A} \subseteq \mathcal{S}$ such that $v(x) \geq \epsilon > 0$, $x \in \mathcal{A}^c$,

$$\mathbb{E}[v(X_2) | X_1 = x] < \infty, \quad x \in \mathcal{A},$$

and for some $\delta > 0$,

$$\mathbb{E}[v(X_2) | X_1 = x] \leq (1 - \delta)v(x), \quad x \in \mathcal{A}^c,$$

then the Markov Chain is geometrically ergodic and $\mathbb{E}[v(\hat{X})] < \infty$, where \hat{X} has the steady-state distribution of $\{X_n\}_{n=1}^{\infty}$.

In our setup the state space \mathcal{S} consists of pairs of nonnegative integers, (k_1, k_2) . The Lyapunov function we use is given by

$$v(k_1, k_2) = \max \left\{ \frac{k_1}{1 - \varepsilon_1} + \frac{k_2}{1 - \varepsilon_{12}}, \frac{k_1}{1 - \varepsilon_{12}} + \frac{k_2}{1 - \varepsilon_2} \right\}.$$

Denote $V_2 = v((K_1(T_2), K_2(T_2)))$. To apply Theorem 6 to the Markov chain $\{(K_1(T_j), K_2(T_j))\}_{j=1}^{\infty}$, we must show that for (k_1, k_2) outside a bounded region, the expected drift of the Lyapunov function satisfies for some $\delta > 0$ and $i \geq 2$,

$$\mathbb{E}[V_2 | (K_1(T_i), K_2(T_i)) = (k_1, k_2)] \leq (1 - \delta) v(k_1, k_2). \quad (10)$$

Based on the discussion below, the rest of the conditions are easily demonstrated.

Let k_1, k_2 be the number of packets in queues $Q_{0,1}^t, Q_{0,2}^t$ respectively at the beginning of epoch 1. Let N_{k_1, k_2} be the number of slots needed to complete this epoch.

Based on the operation of Algorithm II, Theorem 7 evaluates N_{k_1, k_2} in the asymptotic regime when the sum of queue backlogs k_1 and k_2 grow to infinity. We denote, $|(k_1, k_2)| \doteq \sqrt{k_1^2 + k_2^2}$.

Theorem 7. *The following holds with probability 1 (pathwise).*

$$\lim_{|(k_1, k_2)| \rightarrow \infty} \frac{N_{k_1, k_2}}{v(k_1, k_2)} = 1.$$

The limit also holds in expectation.

Proof: The proof of this theorem is rather lengthy and is given in the Appendix. ■

Assume that there are k_1, k_2 packets in queues $Q_{0,1}^t, Q_{0,2}^t$ respectively at the beginning of epoch 1 and denote by $\tilde{A}_i(k_1, k_2)$ the number of packets destined to receiver i that arrived to the system during epoch 1. Also, to make explicit the dependence on (k_1, k_2) , denote $V_2(k_1, k_2) = v(K_1(T_2), K_2(T_2))$. According to the definitions and the operation of the algorithm, we have,

$$\begin{aligned} K_i(T_2) &= \tilde{A}_i(k_1, k_2) = \sum_{l=1}^{N_{k_1, k_2}} A_i(l), \quad i = 1, 2, \\ V_2(k_1, k_2) &= \max \left\{ \frac{\tilde{A}_1(k_1, k_2)}{1 - \varepsilon_1} + \frac{\tilde{A}_2(k_1, k_2)}{1 - \varepsilon_2}, \frac{\tilde{A}_1(k_1, k_2)}{1 - \varepsilon_{12}} + \frac{\tilde{A}_2(k_1, k_2)}{1 - \varepsilon_2} \right\}. \end{aligned} \quad (11)$$

Based on Theorem 7, Lemma 8 evaluates the Lyapunov drift in the asymptotic regime when the sum of queue backlogs k_1 and k_2 grow to infinity.

Lemma 8. *The following hold with probability 1*

$$\lim_{|(k_1, k_2)| \rightarrow \infty} \frac{V_2(k_1, k_2)}{v(k_1, k_2)} = \max \left\{ \frac{\lambda_1}{(1 - \varepsilon_1)} + \frac{\lambda_2}{(1 - \varepsilon_{12})}, \frac{\lambda_1}{(1 - \varepsilon_{12})} + \frac{\lambda_2}{(1 - \varepsilon_2)} \right\}.$$

The limit also holds in expectation.

Proof: Since $\lim_{|(k_1, k_2)| \rightarrow \infty} N_{k_1, k_2} = \infty$, taking into account (11), Theorem 7 and using the Strong Law of Large Numbers for the packet arrival process $\{A_1(l), A_2(l)\}_{l=1}^{\infty}$, we have,

$$\begin{aligned} \lim_{|(k_1, k_2)| \rightarrow \infty} \frac{\tilde{A}_i(k_1, k_2)}{v(k_1, k_2)} &= \lim_{|(k_1, k_2)| \rightarrow \infty} \frac{\tilde{A}_i(k_1, k_2)}{N_{k_1, k_2}} \frac{N_{k_1, k_2}}{v(k_1, k_2)} \\ &= \lambda_i. \end{aligned} \quad (12)$$

Also, since by Wald's equation [29, Theorem 14], $\mathbb{E}[\tilde{A}_i(k_1, k_2)] = \mathbb{E}[\sum_{l=1}^{N_{k_1, k_2}} A_i(l)] = \lambda_i \mathbb{E}[N_{k_1, k_2}]$, we have,

$$\begin{aligned} \lim_{|(k_1, k_2)| \rightarrow \infty} \frac{\mathbb{E}[\tilde{A}_i(k_1, k_2)]}{v(k_1, k_2)} &= \lambda_i \lim_{|(k_1, k_2)| \rightarrow \infty} \frac{\mathbb{E}[N_{k_1, k_2}]}{v(k_1, k_2)} \\ &= \lambda_i \text{ according Theorem 7.} \end{aligned}$$

This implies [30, Corollary p. 218] that the sequences

$$\frac{\tilde{A}_i(k_1, k_2)}{v(k_1, k_2)}, \quad i = 1, 2,$$

are uniformly integrable. Since the sum and the maximum of uniformly integrable functions are also uniformly integrable, we conclude that the sequence

$$\frac{V_2(k_1, k_2)}{v(k_1, k_2)},$$

is uniformly integrable. In addition, it holds

$$\begin{aligned} \lim_{|(k_1, k_2)| \rightarrow \infty} \frac{V_2(k_1, k_2)}{v(k_1, k_2)} &= \lim_{|(k_1, k_2)| \rightarrow \infty} \max \left\{ \frac{\tilde{A}_1(k_1, k_2)}{(1 - \varepsilon_1) v(k_1, k_2)} + \frac{\tilde{A}_2(k_1, k_2)}{(1 - \varepsilon_{12}) v(k_1, k_2)}, \frac{\tilde{A}_1(k_1, k_2)}{(1 - \varepsilon_{12}) v(k_1, k_2)} + \frac{\tilde{A}_2(k_1, k_2)}{(1 - \varepsilon_2) v(k_1, k_2)} \right\} \\ &= \max \left\{ \frac{\lambda_1}{(1 - \varepsilon_1)} + \frac{\lambda_2}{(1 - \varepsilon_{12})}, \frac{\lambda_1}{(1 - \varepsilon_{12})} + \frac{\lambda_2}{(1 - \varepsilon_2)} \right\}, \text{ by (12),} \end{aligned}$$

and the uniform integrability of this sequence, implies that we also have,

$$\lim_{|(k_1, k_2)| \rightarrow \infty} \frac{\mathbb{E}[V_2(k_1, k_2)]}{v(k_1, k_2)} = \max \left\{ \frac{\lambda_1}{(1 - \varepsilon_1)} + \frac{\lambda_2}{(1 - \varepsilon_{12})}, \frac{\lambda_1}{(1 - \varepsilon_{12})} + \frac{\lambda_2}{(1 - \varepsilon_2)} \right\}.$$

■

We now have all the ingredients to show that the system under consideration can be stabilized by Algorithm III for any arrival rates in the interior of the region \mathcal{R} .

Theorem 9. *If the arrival rates (λ_1, λ_2) are in the interior of the region \mathcal{R} , then the Markov Chain $\{(K_1(T_j), K_2(T_j))\}_{j=1}^{\infty}$ is geometrically ergodic and the process of queue lengths $\{Q_1(l), Q_2(l)\}_{l=1}^{\infty}$ is stable.*

Proof: Pick $(\lambda_1, \lambda_2) \geq \mathbf{0}$ in the interior of the region, so that for some $\epsilon > 0$,

$$\max \left\{ \frac{\lambda_1}{1 - \varepsilon_1} + \frac{\lambda_2}{1 - \varepsilon_{12}}, \frac{\lambda_1}{1 - \varepsilon_{12}} + \frac{\lambda_2}{1 - \varepsilon_2} \right\} \leq 1 - \epsilon.$$

Using Lemma 8, pick $k(\epsilon)$ large enough so that for $|(k_1, k_2)| > k(\epsilon)$ it holds

$$\begin{aligned} \frac{\mathbb{E}[V_2(k_1, k_2)]}{v(k_1, k_2)} &\leq \max \left\{ \frac{\lambda_1}{1 - \varepsilon_1} + \frac{\lambda_2}{1 - \varepsilon_{12}}, \frac{\lambda_1}{1 - \varepsilon_{12}} + \frac{\lambda_2}{1 - \varepsilon_2} \right\} + \epsilon/2 \\ &\leq 1 - \epsilon + \epsilon/2 = 1 - \epsilon/2. \end{aligned}$$

The theorem now follows from Theorem 6 by selecting

$$A = \{(k_1, k_2) : |(k_1, k_2)| \leq k(\epsilon)\}.$$

Note next that $\{Q_1(l), Q_2(l)\}_{l=1}^{\infty}$ is regenerative with respect to the renewal process representing the number of slots required for successive returns of the process $\{(K_1(T_j), K_2(T_j))\}_{j=1}^{\infty}$ to the state $(0, 0)$. Since the renewal process is nonlattice, this implies stability of $(Q_1(l), Q_2(l))$, $l = 1, 2$ [29, Theorem 20 p.120]. ■

Accounting for overhead. Algorithm III requires the transmission of two control bits. Hence, if each arriving packet is of length L_a bits, then each transmitted packet is of length $L = L_a + 2$ bits. Hence, if we measure rates in bit arrivals per slot, the region described in Theorem 9 becomes \mathcal{R}^{II} .

B. Adaptation of Algorithms I and I.1 to Stochastic Traffic

These algorithms operate also in epochs, in analogous manner to the one described in Section V-A. The main difference is that the receivers need to know now the number of accumulated packets at the beginning of each epoch, $\{(K_1(T_j), K_2(T_j))\}_{j=1}^{\infty}$. This can be accomplished by using a sub-epoch at the beginning of each epoch, T_j , to broadcast the information $K_1(T_j), K_2(T_j)$ to the receivers. This requires transmission of $O\left(\frac{\log(K_1(T_j)) + \log(K_2(T_j))}{L}\right)$ packets during the sub-epoch. Therefore, if k_1, k_2 , are the number of packets in the beginning of an epoch, the subepoch lasts $O(\log(k_1) + \log(k_2))$ slots. Since as seen from Theorem 7 an epoch lasts $O(k_1, k_2)$ slots, the addition of the sub-epoch does not alter the stability region.

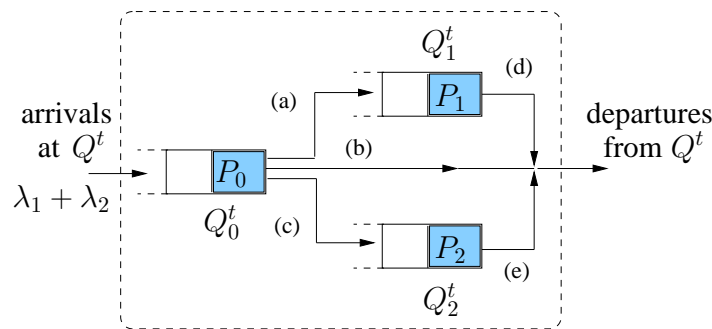
While the algorithms presented in this section are stable for any rates in the interior of \mathcal{R} , they may induce unnecessary delays. To see this, assume that one of the receivers, say receiver 1, receives all packets destined to itself during an epoch. Then all three algorithms proceed to complete transmission of the remaining packets for receiver 2. However, at this time receiver 1 packets may have arrived exogenously to the system and including these packets in the transmission may increase coding opportunities and shorten delays. In the next section we present algorithms that address this problem.

C. Queue-based Dynamic Network Coding with Improved Delay

In this section we present an algorithm which induces smaller delays relative to the algorithms presented in Sections V-A and V-B. As discussed above, the algorithms presented so far operate in epochs, preventing transmissions of new arrivals during an epoch. The algorithm to be presented below allows such transmissions while maintaining stability and as a result achieves better delays. The algorithm operates as follows.

Algorithm IV:

Queue Structure at the Transmitter



- (a) Addressed to user 1, erased at user 1, and received by user 2
- (b) Received by both users 1 and 2
- (c) Addressed to user 2, erased at user 2, and received by user 1
- (d) Received by user 1
- (e) Received by user 2

Figure 2. Packet queues Q_0^t , Q_1^t and Q_2^t at the transmitter for Algorithm V.

We define the coding mechanism at the transmitter by introducing three queues Q_0^t , Q_1^t , and Q_2^t . Based on the receiver feedback, the transmitter assigns each packet to one of these queues depending on which receiver has successfully received the transmitted packet. The queue structure at the transmitter is illustrated in Figure 2 and the queue management policy is the following:

- All packets (regardless whether they are addressed to receiver 1 or 2) arrive at queue Q_0^t and are transmitted on a first-come-first-served basis.¹
 - 1) If a packet transmitted from Q_0^t is destined to receiver i and is successfully received by this receiver, the packet leaves the system (independent of the channel outcome at the other receiver $j \neq i$ in the same time slot).
 - 2) If the packet has not been received by either of the two receivers, the packet remains in queue Q_0^t .
 - 3) A packet transmitted from Q_0^t and destined to receiver $i = 1, 2$, is placed at queue Q_i^t , if it is erased at receiver i and seen by the other receiver $j \neq i$.
- An uncoded packet transmitted from Q_i^t , $i = 1, 2$, leaves this queue (and consequently leaves the system), if it is successfully received by receiver i .

Network Coding Mechanism at the Transmitter

At any time slot l , the network coding mechanism is given as follows:

- If $Q_0^t \neq \emptyset$, a (uncoded) packet is transmitted from queue Q_0^t .
- Else, if $Q_0^t = \emptyset$ and $Q_i^t \neq \emptyset$, $i = 1, 2$, then packets from both queues Q_1^t and Q_2^t are combined by linear network coding before transmission, i.e., if queue Q_0^t is empty and P_1 and P_2 are the first in line packets in queues Q_1^t and Q_2^t , respectively, then the packet $P_1 \oplus P_2$ is transmitted. Here, $P_1 \oplus P_2$ corresponds to pairwise XOR operation of the information bits at the same position in two packets P_1 and P_2 . If this coded packet is successfully received by receiver 1 (with probability $1 - \varepsilon_1$) or by receiver 2 (with probability $1 - \varepsilon_2$), it leaves queue Q_2^t or Q_1^t respectively, since the corresponding receiver can recover the packet destined to itself by performing XOR operation. Hence, in this case it is possible to deliver one packet to each receiver in a single transmission. This cannot be realized with a plain retransmission policy that does not perform coding.
- Else if $Q_0^t = \emptyset$, $Q_i^t \neq \emptyset$ and $Q_j^t = \emptyset$, $j \neq i$, $i = 1, 2$, the first available packet in Q_i^t is transmitted uncoded.

Decoding Mechanism at the Receivers

Each receiver $i = 1, 2$ constructs a separate queue Q_j^i , $j \neq i$, that tracks the transmitter queue Q_j^t . The decoding mechanism at receiver i is given by:

- If receiver i receives an uncoded packet P addressed to itself, the packet P leaves the system.
- If receiver i receives an uncoded packet P addressed to the other receiver j but erased at receiver j , the packet P is moved to queue Q_j^i at receiver i .
- If receiver i receives a coded packet $P_1 \oplus P_2$ (combined from queues Q_1^t and Q_2^t), receiver i decodes the packet addressed to itself, P_i , through the operation $(P_1 \oplus P_2) \oplus P_j$ by using a packet P_j that is already available in Q_j^i .

Note that for this mechanism to work, if receiver i receives an uncoded packet that is destined to the other receiver j , then receiver i needs to know whether the packet has been received or not by j . This information is conveyed by overhead bits that

¹Any other work conserving queue management policy that distinguishes two separate queues $Q_{0,1}^t$ and $Q_{0,2}^t$, where newly arriving packets addressed to the two receivers are placed, would not change the stability result although the analysis may become more complex.

are placed in the packet headers. We postpone overhead issues to Section V-C2 and proceed next with the stability analysis of the algorithm.

1) *Stability Analysis of Algorithm IV*: The process $\{\mathbf{Q}^t(l)\}_{l=1}^{\infty} = \{(Q_0^t(l), Q_1^t(l), Q_2^t(l))\}_{l=1}^{\infty}$ of queue sizes at the beginning of slot l , constitutes an irreducible, aperiodic homogeneous Markov Chain with countable state space. Here, for simplicity in notation, we denote “queue size” and “queue” with the same letter. We have the following theorem showing that any arrival rates in the interior of the region \mathcal{R} can be stabilized by Algorithm IV.

Theorem 10. *If the arrival rates (λ_1, λ_2) are in the interior of the region \mathcal{R} then the process $\{\mathbf{Q}^t(l)\}_{l=1}^{\infty}$ is stable.*

Proof: As discussed in Section V, to show stability, it suffices to show that under some given initial condition, $\{\mathbf{Q}^t(l)\}_{l=1}^{\infty}$ is stable, or equivalently all three queue processes $\{Q_0^t(l)\}_{l=1}^{\infty}$, $\{Q_1^t(l)\}_{l=1}^{\infty}$, $\{Q_2^t(l)\}_{l=1}^{\infty}$ are stable. Assume that the arrival rates (λ_1, λ_2) are in the interior of the region \mathcal{R} , so that for some $\epsilon > 0$,

$$\max \left\{ \frac{\lambda_1}{1 - \epsilon_1} + \frac{\lambda_2}{1 - \epsilon_{12}}, \frac{\lambda_1}{1 - \epsilon_{12}} + \frac{\lambda_2}{1 - \epsilon_2} \right\} \leq 1 - \epsilon. \quad (13)$$

The arrival rate to queue Q_0^t is $\lambda_1 + \lambda_2$ and the departure rate is $1 - \epsilon_{12}$. Taking into account that $\epsilon_{12} \leq \min\{\epsilon_1, \epsilon_2\}$, we conclude,

$$\frac{\lambda_1 + \lambda_2}{1 - \epsilon_{12}} \leq \max \left\{ \frac{\lambda_1}{1 - \epsilon_1} + \frac{\lambda_2}{1 - \epsilon_{12}}, \frac{\lambda_1}{1 - \epsilon_{12}} + \frac{\lambda_2}{1 - \epsilon_2} \right\} \leq 1 - \epsilon.$$

Hence, $\{Q_0^t(l)\}_{l=1}^{\infty}$ is stable [31]. We may assume that we start the system from an appropriate initial condition $Q_0^t(1)$, so that the departure process from Q_0^t is a stationary process with rate $\lambda_1 + \lambda_2$. A departing packet from this queue (recall that departure occurs only if the packet is seen by at least one receiver) is placed at queue Q_i^t , if a) it is destined to receiver i (i.e., with probability $\lambda_i/(\lambda_1 + \lambda_2)$), and, b) it is erased at receiver i and received by receiver $j \neq i$ (which occurs with probability $(\epsilon_i - \epsilon_{12})/(1 - \epsilon_{12})$). Therefore, the arrival rate to queue Q_i^t , $i = 1, 2$, is a stationary process with rate

$$\lambda(Q_i^t) = (\lambda_1 + \lambda_2) \frac{\lambda_i}{\lambda_1 + \lambda_2} \frac{\epsilon_i - \epsilon_{12}}{1 - \epsilon_{12}} = \frac{\lambda_i(\epsilon_i - \epsilon_{12})}{1 - \epsilon_{12}}.$$

Packets are transmitted (either in coded or uncoded form) from queue Q_i^t only if Q_0^t is empty, which occurs with probability $1 - \frac{\sum_{m=1}^2 \lambda_m}{1 - \epsilon_{12}}$ according to Little’s law [32, p. 157]. Whenever a packet is transmitted from Q_i^t (either uncoded or coded with a packet from Q_j^t , $j \neq i$), it is successfully received and decoded by receiver i with probability $1 - \epsilon_i$ - note that using the network coding mechanism both queues Q_1^t and Q_2^t can be served simultaneously. Therefore, the service rate for queue Q_i^t is given by

$$\mu(Q_i^t) = \left(1 - \frac{\sum_{m=1}^2 \lambda_m}{1 - \epsilon_{12}} \right) (1 - \epsilon_i).$$

We then have, with $j \neq i$,

$$\begin{aligned} \lambda(Q_i^t) - \mu(Q_i^t) &= \frac{\lambda_i(\epsilon_i - \epsilon_{12})}{1 - \epsilon_{12}} - \left(1 - \frac{\sum_{m=1}^2 \lambda_m}{1 - \epsilon_{12}} \right) (1 - \epsilon_i) \\ &= \left(\frac{\lambda_i}{1 - \epsilon_i} + \frac{\lambda_j}{1 - \epsilon_{12}} - 1 \right) (1 - \epsilon_i) \\ &\leq -\epsilon(1 - \epsilon_i) \text{ according to (13)}. \end{aligned}$$

Using again Loynes’ results [31], we conclude that $\{Q_1^t(l)\}_{l=1}^{\infty}$, $\{Q_2^t(l)\}_{l=1}^{\infty}$ are stable. ■

2) *Protocol Overhead*: We specify next the packet overhead that is required for transmitting control information so that the receivers can keep track and construct the transmitter queues. The control bits should inform each receiver whether the packet transmission in a previous slot is successful at the other receiver, or not. Since it takes a random number of slots until this overhead can be delivered (because of channel erasures), receiver i needs to learn how many packets have been moved into transmitter queue Q_i^t and out of transmitter queue Q_j^t , $j \neq i$, in the meantime. This way, each receiver i can keep track of both transmitter queues Q_1^t and Q_2^t . Since the transmissions from Q_1^t and Q_2^t are interleaved with new packet transmissions from Q_0^t , packets may arrive in mixed order and it is necessary to inform receiver i of the order of the received packets destined to itself.

To construct the overhead, one option is to take advantage of packet ID fields that are reserved at the packet header under several communication protocols. If packet IDs reflect order of arrival, then reserving two packet ID fields, F_1 , F_2 , at the header of each packet suffices: Field F_i contains the ID of receiver i packet that is used in the encoding of the transmitted packet - if the packet is uncoded and is destined to user $j \neq i$ then F_i takes the value 0.

If packet IDs are not available, two types of overhead can be used as follows. Overhead type 1 informs the receivers whether the packet under transmission is

- 1) transmitted from queue Q_0^t and addressed to receiver $i = 1, 2$, or

- 2) transmitted from queue Q_i^t , $i = 1, 2$, in uncoded form, or
- 3) coded from queues Q_1^t and Q_2^t .

Overhead type 1 distinguishes five separate cases and therefore requires $\eta_1 = 3$ bits. As for overhead type 2, receiver i needs to additionally know 1) the number of packet moved out of Q_j^t , $j \neq i$, between two successful packet receptions and 2) the number of packets moved to queue Q_i^t between two successful packet receptions. The average size of these numbers is between 0 and $\frac{1}{1-\varepsilon_i}$. Hence, if the transmitter sends these numbers, the average number of overhead bits that will be required for user i is at most $-2 \log(1 - \varepsilon_i)$. Therefore, overhead type 2 requires $\eta_2 = -\sum_{i=1}^2 2 \log(1 - \varepsilon_i)$ bits. As a result, the total overhead length $\eta_1 + \eta_2$ sufficient for successful decoding is given by

$$\eta = 3 - \sum_{i=1}^2 2 \log(1 - \varepsilon_i)$$

bits per packet such that at least $L - \eta$ bits out of L bits per packet carry data on the average. In protocol implementation, variable-length overhead type 2 needs to be separated from the data bits, i.e., it is necessary to indicate when overhead type 2 ends. This separation can be realized by inserting additional control bits in the overhead. Note that the total overhead is independent of the packet length L (in bits) and its relative effect decreases with increasing L .

Incorporating this random overhead length in the stability analysis of the algorithm would introduce additional complexities. However as stated above this overhead is generally small, e.g., for $\varepsilon_i \leq 0.5$ the average number of overhead bits is not more than 7 bits. Hence, even reserving a constant number of bits, say 25, for overhead should suffice for the operation of the algorithm in practice. For 200-byte packets this represents a loss in throughput less than 1.56%.

3) *Adaptation of Algorithm IV to Include Multicast Traffic:* In addition to unicast traffic addressed to one receiver only, we can also assume that multicast packets addressed to both receivers arrive at queue Q_0^t with rate λ_{12} . Next, we adapt Algorithm IV to multicast traffic and derive the stability region.

As with Algorithm IV, all newly arrived packets (unicast or multicast) are placed in queue Q_0^t and are served on a first-come-first-served basis. If a transmitted multicast packet is received by both receivers, it leaves the system. Otherwise, if a transmitted multicast packet is erased at receiver $i = 1, 2$ but received at receiver $j \neq i$, the packet is placed in queue Q_i^t . Then, all packets in queue Q_i^t are treated the same way as in unicast traffic. The necessary packet overhead is the same as before except that overhead type 1 should also distinguish the case when a packet transmitted from Q_0^t is of multicast type.

The stability analysis in this case is similar to the stability analysis of Algorithm IV. We give a summary of the steps next. Assume that $(\lambda_1, \lambda_2, \lambda_{12})$ is in the interior of the region \mathcal{R}^m .

The total arrival rate for Q_0^t is $\lambda_{12} + \sum_{m=1}^2 \lambda_m$. A packet transmitted from Q_0^t leaves the system, if it is received by at least one receiver. Therefore, the service rate for Q_0^t is $1 - \varepsilon_{12}$. Since $(\lambda_1, \lambda_2, \lambda_{12})$ is in the interior of \mathcal{R}^m , we have

$$\lambda_{12} + \sum_{m=1}^2 \lambda_m < 1 - \varepsilon_{12}.$$

and hence queue Q_0^t is stable. Using similar arguments as in the proof of Theorem 10, we can define initial conditions so that the departure process from Q_0^t is stationary, and under these conditions we have that the arrival rate to Q_i^t is

$$\lambda(Q_i^t) = \frac{(\lambda_{12} + \lambda_i)(\varepsilon_i - \varepsilon_{12})}{1 - \varepsilon_{12}},$$

and the service rate is

$$\mu(Q_i^t) = \left(1 - \frac{\lambda_{12} + \sum_{m=1}^2 \lambda_m}{1 - \varepsilon_{12}}\right) (1 - \varepsilon_i).$$

Since $(\lambda_1, \lambda_2, \lambda_{12})$ is in the interior of \mathcal{R}^m , it can be calculated that

$$\lambda(Q_i^t) - \mu(Q_i^t) < -\epsilon(1 - \varepsilon_i),$$

for some $\epsilon > 0$ and hence both queues Q_1^t , Q_2^t are stable.

Note: Algorithm IV is designed so that it induces small delays and, as will be seen in the next section, its delay performance is in fact better than that of Algorithm III. On the other hand, the approach used in the design of Algorithm III can be applied to other systems as well and as such it seems more suitable for generalization to a system with more than two receivers.

D. Average Packet Delay

The algorithms presented above in Section V achieve in effect the same stability region with different levels of complexity, overhead, and feedback requirements. However, their delay performance is different and this difference increases as arrival rates increase. In this section we study by simulations the delay performance of these algorithms. We also study by simulations two variants of the algorithms with improved delays and/or reduced overhead. The stability analysis of the latter algorithms requires additional techniques and is not attempted here. We first describe the two variant algorithms.

Algorithm IV.b:

This algorithm improves the delays of Algorithm IV for high arrival rates, by scheduling coded packets from queues Q_1^t and Q_2^t whenever these queues are both nonempty, instead of always giving priority to packets in queue Q_0^t . Specifically, the queue structure, protocol overhead, and the decoding mechanism at the receivers are the same as in Algorithm IV. The network coding mechanism at the transmitter is changed as follows by giving priority to network-coded packet transmissions Q_1^t and Q_2^t over uncoded packet transmissions from Q_0^t :

- $Q_i^t \neq \emptyset$, $i = 1, 2$, packet pairs from queues Q_1^t and Q_2^t are XORed and transmitted.
- Else if $Q_1^t = \emptyset$ or $Q_2^t = \emptyset$, and $Q_0^t \neq \emptyset$, a packet is transmitted uncoded from queue Q_0^t .
- Else if $Q_0^t = \emptyset$, $Q_i^t \neq \emptyset$ and $Q_j^t = \emptyset$, $j \neq i$, $i = 1, 2$, the first available packet in Q_i^t is transmitted uncoded.

The overhead requirements of this algorithm are similar to those of Algorithm IV and it can be similarly adapted to operate with multicast traffic.

Algorithm V:

This algorithm is a combination of the desirable features of Algorithm III and IV.b. As with Algorithm IV.b, Algorithm V does not have coding epochs. On the other hand, as with Algorithm III, Algorithm V requires 2 bits of overhead and operates with single packet buffers at the receivers. Specifically,

- The queues and buffer structures are the same as in Algorithm III, (i.e., the same as Algorithm II).
- As with Algorithms III, two control bits b_1, b_2 are used, having the same meaning as the corresponding bits in Algorithms III. That is, b_1 indicates whether the packet is an XORed packet or not and b_2 indicates the receiver to which the packet is intended in case it is not an XORed packet.
- The receivers operate as described in Algorithms II and III.
- The transmitter operates in manner similar to Algorithm III, with the main difference that *new packet arrivals may be transmitted* as opposed to waiting for the completion of an epoch. Specifically,
 - 1) If $B_1^t = \emptyset$ and $B_2^t = \emptyset$, then packets from queues $Q_{0,i}^t$, $i = 1, 2$, are transmitted (the transmission order does not matter)
 - 2) If $B_i^t \neq \emptyset$ and $Q_{0,j}^t \neq \emptyset$, $i \neq j$, then a packet from queue $Q_{0,j}^t$ is transmitted
 - 3) If $B_i^t \neq \emptyset$ and $Q_{0,j}^t = \emptyset$, $i \neq j$, then the transmitter keeps (re)transmitting the packet in B_i^t .

The last condition (item 3 above) ensures that single packet buffers suffice for the operation. If $B_i^t \neq \emptyset$ and $Q_j^t \neq \emptyset$, Algorithm V operates as Algorithm IV.b, hence creating additional coding opportunities relative to Algorithm III (which blocks new arrivals from transmission until the end of an epoch). However, compared to Algorithm IV.b it still misses coding opportunities when $B_i^t \neq \emptyset$, $Q_{0,j}^t = \emptyset$, $i \neq j$, and $Q_{0,i}^t \neq \emptyset$: in this case Algorithm V keeps retransmitting the packet in B_i^t while packets from Q_i^t are transmitted by Algorithm IV.b, hence providing the possibility for future coding opportunities in case the transmitted packet is erased by receiver i and received by receiver $j \neq i$.

In the simulation below, we measure the delay of each packet as the length of interval (in terms of slots) from the time the packet arrives at the source until it is *successfully decoded* by the intended receiver. Figure 3 shows the average packet delay as a function of the packet arrival rates (common for both multicast sessions) under channel erasure probabilities $\varepsilon_1 = \varepsilon_2 = 0.4$ and $\varepsilon_{12} = 0.2$. We consider Poisson arrivals and average the delay results over 200 simulations, each with length of 10^4 time slots. As expected, Algorithm III has poor delay performance at high arrival rates since it defers transmission of newly arriving packets until the next epoch. On the other hand, Algorithms IV, IV.b, and V allow dynamic network coding based on the instantaneous queue contents and therefore they can reduce the average packet delay compared to Algorithm III. The delay gain grows with the arrival rates.

VI. CONCLUSION

We have analyzed the rate performance of a simple broadcast channel both in information-theoretic terms (i.e. capacity region) as well as network-theoretic terms (i.e. stable throughput region). The algorithms that achieve capacity, and those that achieve a stable throughput region that is almost identical to the former, are based on linear network coding (either random or based on dynamic queue contents) and simple queue management schemes. The algorithms proposed for stochastic packet traffic differ in terms of complexity and overhead requirement, and their average packet delay performance depends on how effectively they exploit the network coding opportunities based on the instantaneous queue contents.

Thus we observe a similarity in the relationship between the capacity and stable throughput regions that has been observed for the case of multi-access channels, adding credibility to the quest for a union between information-theoretic and networking treatments of multi-receiver systems.

As discussed in Section IV-B2, the recent work in [24], [25], extends the capacity results to systems with more than two receivers. Regarding the extension of stability results to such systems, an approach based on coding epochs similar to the one followed in the design of Algorithm III should provide algorithms with stability region close to the capacity region. However, the design of algorithms with maximal stability region and improved delay characteristics for such systems, remains an open problem.

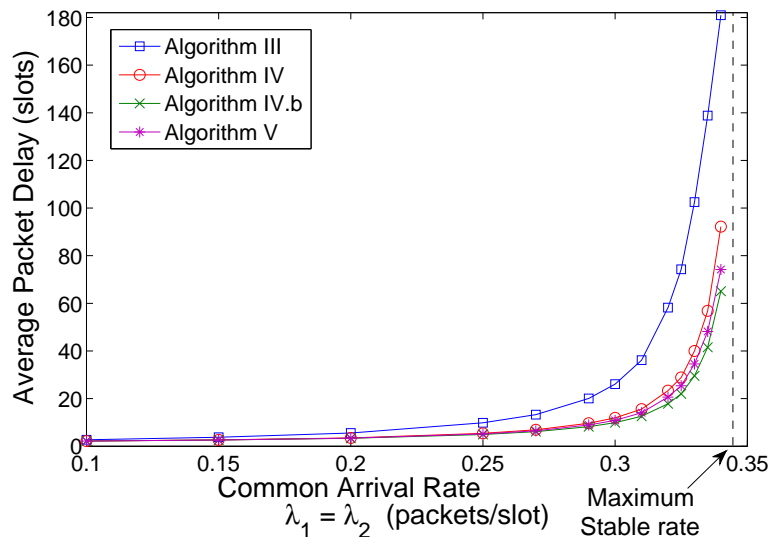


Figure 3. Average packet delay performance of Algorithms III-VI for stochastic packet traffic.

ACKNOWLEDGMENT

We wish to thank professors Gerhard Kramer, Randall Berry, and Dongning Guo for helpful discussions.

REFERENCES

- [1] R. Rao and A. Ephremides, "On the stability of interacting queues in a multiple-access system," *IEEE Transactions on Information Theory*, vol. 34, pp. 918–930, September 1988.
- [2] A. Ephremides and B. Hajek, "Information theory and communication networks: An unconsumed union," *IEEE Transactions on Information Theory*, vol. 44, pp. 2416–2434, October 1998.
- [3] J. Luo and A. Ephremides, "On the throughput, capacity and stability regions of random multiple access," *IEEE Transactions on Information Theory*, vol. 52, pp. 2593–2607, June 2006.
- [4] T. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. 18, pp. 2–14, January 1972.
- [5] P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Transactions on Information Theory*, vol. 19, pp. 197–207, March 1973.
- [6] A. F. Dana, R. Gowaika, R. Palanki, B. Hassibi, and M. Effros, "Capacity of wireless erasure networks," *IEEE Transactions on Information Theory*, vol. 52, pp. 789–804, March 2006.
- [7] D. S. Lun, M. Medard, R. Koetter, and M. Effros, "On coding for reliable communication over packet networks," *Physical Communication*, vol. 1, pp. 3–20, March 2008.
- [8] P. Larsson and N. Johansson, "Multi-user arq," in *Proc. of the IEEE Vehicular Technology Conference*, May 2007.
- [9] M. Durvy, C. Fragouli, and P. Thiran, "Towards reliable broadcasting using ACKs," in *Proc. of 2007 IEEE International Symposium on Information Theory (ISIT)*, 2007.
- [10] C. Fragouli, D. S. Lun, M. Medard, and P. Pakzad, "On feedback for network coding," in *Proc. of 2007 Conference on Information Sciences and Systems (CISS)*, March 2007.
- [11] J. K. Sundarajan, D. Shah, and M. Medard, "ARQ for network coding," in *Proc. of 2008 IEEE International Symposium on Information Theory (ISIT)*, 2008.
- [12] A. Eryilmaz, A. Ozdaglar, M. Medard, and E. Ahmed, "On the delay and throughput gains of coding in unreliable networks," *IEEE Transactions on Information Theory*, vol. 54, pp. 5511–5524, December 2000.
- [13] Y. Shi and M. Sheng, "Study of a scheduling algorithm based on simple network coding with high throughput and low delay," in *Proc. of the Conference on Advanced Information Networking and Application*, March 2008.
- [14] D. Nguyen, T. Tran, T. Nguyen, and B. Bose, "Wireless broadcast using network coding," *IEEE Transactions on Information Theory*, vol. 58, pp. 914–925, February 2009.
- [15] Y. E. Sagduyu and A. Ephremides, "On broadcast stability of queue-based dynamic network coding over erasure channels," *IEEE Transactions on Information Theory*, vol. 55, pp. 5463–5478, December 2009.
- [16] C. C. Wang, "Capacity of 1-to-k broadcast packet erasure channel with output feedback," in *Proc. of the Forty-Eighth Annual Allerton Conference on Communication, Control, and Computing*, September 2010.
- [17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 2nd ed., 2006.
- [18] A. E. Gamal, "The feedback capacity of degraded broadcast channels," *IEEE Transactions on Information Theory*, vol. 25, pp. 379–381, March 1978.
- [19] L. Ozarow and S. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback," *IEEE Transactions on Information Theory*, vol. 30, pp. 667–671, July 1984.
- [20] M. Effros, M. Medard, T. Ho, S. Ray, D. Karger, and R. Koetter, "Linear network codes: a unified framework for source channel, and network coding," in *Proc. of the DIMACS workshop on network info. theory*, 2003.
- [21] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Problemy Peredaci Informatsii*, vol. 10, no. 3, pp. 3–14, 1974.
- [22] Z. Bar-Yossef, Y. Birk, T. S. Jayram, and T. Kol, "Index coding with side information," in *Proc. of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2006.
- [23] T. Ho, M. Medard, R. Koetter, D. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol. 52, pp. 4413–4430, October 2006.
- [24] C.-C. Wang, "Capacity of 1-to-k broadcast packet erasure channels with channel output feedback," [Online]. Available: <http://arxiv.org/abs/1010.2436v1> submitted to *IEEE Transactions on Information Theory*, 2010.

- [25] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Multiuser broadcast erasure channel with feedback - capacity and algorithms," [Online]. Available: arxiv.org/abs/1009.1254 submitted to *IEEE Transactions on Information Theory*, 2010.
- [26] L. Georgiadis and W. Szpankowski, "Stability of token passing rings," *Queueing Systems*, vol. 7, pp. 7–33, 1992.
- [27] N. N. Popov, "Conditions for geometric ergodicity of countable markov chain," *Soviet Math. Dokl.*, vol. 18, pp. 676–679, 1977.
- [28] D. Issacson, "A characterization of geometric ergodicity," *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, vol. 49, pp. 267–273, 1979.
- [29] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*. Prentice Hall, 1989.
- [30] P. Billingsley, *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics, 1995.
- [31] R. Loynes, "The stability of a queue with non-interdependent inter-arrival and service times," *Proc. Camb. Philos. Soc.*, vol. 58, pp. 497–520, 2006.
- [32] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1992.

APPENDIX A PROOF OF THEOREM 7

The nature of the Lyapunov function $v(k_1, k_2)$ involved in the theorem, requires some tedious estimates to handle limits in various regions in the two-dimensional space. Also, one has to consider the evolution of the system during the two phases of Algorithm II. Moreover, more details are required to transform convergence with probability 1 to convergence in expectation. These complicate the proof. The arguments are based on the regenerative structure of the Markov Chain describing the evolution of the system in Phase 1.

Next we outline the steps of the proof. Lemma 11 presents some relations between the long-term rewards of the Markov Chain describing the evolution of the system during Phase 1. These relations are needed in the sequel. Lemma 12 provides asymptotic form of the time needed to transmit the k_i , $i = 1, 2$ packets to receiver i within Phase 1. Based on the fact that the duration of Phase 1 is the minimum of these times, Lemma 13 provides asymptotic forms of the time needed to complete Phase 1, the number of packets successfully transmitted during Phase 1 to each receiver, and the receiver with packets remaining to be transmitted during Phase 2. Based on the latter derivations, the evolution of the systems during Phase 2 can be described, and Theorem 7 is concluded by putting the above results together.

Lemma 11. *The following relations hold*

$$\frac{\bar{\rho}_1}{\bar{\rho}_2} = \frac{(1 - \varepsilon_1)(\varepsilon_2 - \varepsilon_{12})}{(1 - \varepsilon_2)(\varepsilon_1 - \varepsilon_{12})}, \quad (14)$$

$$\frac{1}{\bar{\rho}_1} - \frac{\bar{\rho}_2}{(1 - \varepsilon_2)\bar{\rho}_1} = \frac{1}{\bar{\rho}_2} - \frac{\bar{\rho}_1}{(1 - \varepsilon_1)\bar{\rho}_2} = \frac{1}{1 - \varepsilon_{12}}, \quad (15)$$

$$k_2 \left(\frac{1}{\bar{\rho}_2} - \frac{\bar{\rho}_1}{(1 - \varepsilon_1)\bar{\rho}_2} \right) + \frac{k_1}{1 - \varepsilon_1} = \frac{k_1}{1 - \varepsilon_1} + \frac{k_2}{1 - \varepsilon_{12}}, \quad (16)$$

$$k_1 \left(\frac{1}{\bar{\rho}_1} - \frac{\bar{\rho}_2}{(1 - \varepsilon_2)\bar{\rho}_1} \right) + \frac{k_2}{1 - \varepsilon_2} = \frac{k_1}{1 - \varepsilon_{12}} + \frac{k_2}{1 - \varepsilon_2}, \quad (17)$$

$$\max \left\{ \frac{k_1}{1 - \varepsilon_1} + \frac{k_2}{1 - \varepsilon_{12}}, \frac{k_1}{1 - \varepsilon_{12}} + \frac{k_2}{1 - \varepsilon_2} \right\} = \begin{cases} \frac{k_1}{1 - \varepsilon_1} + \frac{k_2}{1 - \varepsilon_{12}} & \text{if } \frac{k_1}{\bar{\rho}_1} \geq \frac{k_2}{\bar{\rho}_2} \\ \frac{k_1}{1 - \varepsilon_{12}} + \frac{k_2}{1 - \varepsilon_2} & \text{if } \frac{k_1}{\bar{\rho}_1} \leq \frac{k_2}{\bar{\rho}_2} \end{cases}. \quad (18)$$

Let

$$c = 4 \max \{ \bar{\rho}_1, \bar{\rho}_2 \} \left(\max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} + 1 \right). \quad (19)$$

For any ϵ such that

$$0 < \epsilon < \frac{1}{c}, \quad (20)$$

and for any k_1, k_2 such that

$$\left| \frac{k_2}{\bar{\rho}_2} - \frac{k_1}{\bar{\rho}_1} \right| \leq \epsilon (k_1 + k_2), \quad (21)$$

it holds,

$$\left| \frac{\min \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}}{\max \left\{ \frac{k_1}{1 - \varepsilon_1} + \frac{k_2}{1 - \varepsilon_{12}}, \frac{k_1}{1 - \varepsilon_{12}} + \frac{k_2}{1 - \varepsilon_2} \right\}} - 1 \right| \leq c\epsilon. \quad (22)$$

Proof: Based on the transition diagram in Figure 1, the rewards can be computed based on (5) and identities (14), (15) can be seen, from which (16)-(18) follow. To show (22) note first that

$$\begin{aligned} \max \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\} &= \max \left\{ \frac{k_1}{\bar{\rho}_1 (k_1 + k_2)}, \frac{k_2}{\bar{\rho}_2 (k_1 + k_2)} \right\} (k_1 + k_2) \\ &\geq \frac{1}{\max \{ \bar{\rho}_1, \bar{\rho}_2 \}} \max \left\{ \frac{k_1}{(k_1 + k_2)}, \frac{k_2}{(k_1 + k_2)} \right\} (k_1 + k_2) \\ &\geq \frac{(k_1 + k_2)}{2 \max \{ \bar{\rho}_1, \bar{\rho}_2 \}}, \end{aligned} \quad (23)$$

and similarly

$$\max \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\} \leq \frac{(k_1 + k_2)}{\min \{\bar{\rho}_1, \bar{\rho}_2\}}.$$

It follows then that

$$0 \leq \max \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\} - \epsilon(k_1 + k_2) \leq \min \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}, \quad (24)$$

where the left inequality follows from (20) and (23), and the right inequality from (21).

We also have from (16),

$$\begin{aligned} \frac{k_1}{1 - \varepsilon_1} + \frac{k_2}{1 - \varepsilon_{12}} &= k_2 \left(\frac{1}{\bar{\rho}_2} - \frac{\bar{\rho}_1}{(1 - \varepsilon_1)\bar{\rho}_2} \right) + \frac{k_1}{1 - \varepsilon_1} \\ &= \frac{k_2}{\bar{\rho}_2} - \frac{\bar{\rho}_1}{(1 - \varepsilon_1)} \left(\frac{k_2}{\bar{\rho}_2} - \frac{k_1}{\bar{\rho}_1} \right), \end{aligned}$$

hence using again (21) we have

$$\frac{k_2}{\bar{\rho}_2} - \frac{\bar{\rho}_1}{(1 - \varepsilon_1)} \epsilon(k_1 + k_2) \leq \frac{k_1}{1 - \varepsilon_1} + \frac{k_2}{1 - \varepsilon_{12}} \leq \frac{k_2}{\bar{\rho}_2} + \frac{\bar{\rho}_1}{(1 - \varepsilon_1)} \epsilon(k_1 + k_2)$$

and

$$\begin{aligned} \frac{k_2}{\bar{\rho}_2} - \max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} \epsilon(k_1 + k_2) &\leq \frac{k_1}{1 - \varepsilon_1} + \frac{k_2}{1 - \varepsilon_{12}} \\ &\leq \frac{k_2}{\bar{\rho}_2} + \max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} \epsilon(k_1 + k_2). \end{aligned}$$

Since a similar inequality holds if we interchange 1 and 2, we conclude,

$$\max \left\{ \frac{k_1}{1 - \varepsilon_1} + \frac{k_2}{1 - \varepsilon_{12}}, \frac{k_1}{1 - \varepsilon_{12}} + \frac{k_2}{1 - \varepsilon_2} \right\} \leq \max \left\{ \frac{k_2}{\bar{\rho}_2}, \frac{k_1}{\bar{\rho}_1} \right\} + \max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} \epsilon(k_1 + k_2), \quad (25)$$

$$\max \left\{ \frac{k_1}{1 - \varepsilon_1} + \frac{k_2}{1 - \varepsilon_{12}}, \frac{k_1}{1 - \varepsilon_{12}} + \frac{k_2}{1 - \varepsilon_2} \right\} \geq \max \left\{ \frac{k_2}{\bar{\rho}_2}, \frac{k_1}{\bar{\rho}_1} \right\} - \max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} \epsilon(k_1 + k_2), \quad (26)$$

From (20) and (23) we compute

$$\begin{aligned} \max \left\{ \frac{k_2}{\bar{\rho}_2}, \frac{k_1}{\bar{\rho}_1} \right\} - \max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} \epsilon(k_1 + k_2) &\geq \frac{(k_1 + k_2)}{2 \max \{\bar{\rho}_1, \bar{\rho}_2\}} - \frac{\max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} (k_1 + k_2)}{4 \max \{\bar{\rho}_1, \bar{\rho}_2\} \left(\max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} + 1 \right)} \\ &\geq \frac{(k_1 + k_2)}{2 \max \{\bar{\rho}_1, \bar{\rho}_2\}} - \frac{(k_1 + k_2)}{4 \max \{\bar{\rho}_1, \bar{\rho}_2\}} \\ &= \frac{(k_1 + k_2)}{4 \max \{\bar{\rho}_1, \bar{\rho}_2\}} \\ &> 0. \end{aligned} \quad (27)$$

Then, (26) implies,

$$\begin{aligned} \frac{\min \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}}{\max \left\{ \frac{k_1}{1 - \varepsilon_1} + \frac{k_2}{1 - \varepsilon_{12}}, \frac{k_1}{1 - \varepsilon_{12}} + \frac{k_2}{1 - \varepsilon_2} \right\}} &\leq \frac{\max \left\{ \frac{k_2}{\bar{\rho}_2}, \frac{k_1}{\bar{\rho}_1} \right\}}{\max \left\{ \frac{k_2}{\bar{\rho}_2}, \frac{k_1}{\bar{\rho}_1} \right\} - \max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} \epsilon(k_1 + k_2)} \\ &= 1 + \frac{\max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} \epsilon(k_1 + k_2)}{\max \left\{ \frac{k_2}{\bar{\rho}_2}, \frac{k_1}{\bar{\rho}_1} \right\} - \max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} \epsilon(k_1 + k_2)} \\ &\leq 1 + 4 \max \{\bar{\rho}_1, \bar{\rho}_2\} \max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} \epsilon, \end{aligned} \quad (28)$$

where the last inequality follows from (27). Also, by (24) and (25),

$$\begin{aligned} \frac{\min \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}}{\max \left\{ \frac{k_1}{1 - \varepsilon_1} + \frac{k_2}{1 - \varepsilon_{12}}, \frac{k_1}{1 - \varepsilon_{12}} + \frac{k_2}{1 - \varepsilon_2} \right\}} &\geq \frac{\max \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\} - \epsilon(k_1 + k_2)}{\max \left\{ \frac{k_2}{\bar{\rho}_2}, \frac{k_1}{\bar{\rho}_1} \right\} + \max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} \epsilon(k_1 + k_2)} \\ &= \left(1 - \frac{\epsilon(k_1 + k_2)}{\max \left\{ \frac{k_2}{\bar{\rho}_2}, \frac{k_1}{\bar{\rho}_1} \right\}} \right) \left(1 + \frac{\max \left\{ \frac{\bar{\rho}_1}{(1 - \varepsilon_1)}, \frac{\bar{\rho}_2}{(1 - \varepsilon_2)} \right\} \epsilon(k_1 + k_2)}{\max \left\{ \frac{k_2}{\bar{\rho}_2}, \frac{k_1}{\bar{\rho}_1} \right\}} \right)^{-1}. \end{aligned} \quad (29)$$

Using the inequality, $(1-a)(1+b)^{-1} \geq 1-(a+b)$, $a \leq 1$, $b \geq 0$, we conclude from (29) that,

$$\begin{aligned}
\frac{\min \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}}{\max \left\{ \frac{k_1}{1-\varepsilon_1} + \frac{k_2}{1-\varepsilon_{12}}, \frac{k_1}{1-\varepsilon_{12}} + \frac{k_2}{1-\varepsilon_2} \right\}} &\geq 1 - \frac{\max \left\{ \frac{\bar{\rho}_1}{(1-\varepsilon_1)}, \frac{\bar{\rho}_2}{(1-\varepsilon_2)} \right\} \epsilon (k_1 + k_2) + \epsilon (k_1 + k_2)}{\max \left\{ \frac{k_2}{\bar{\rho}_2}, \frac{k_1}{\bar{\rho}_2} \right\}} \\
&\geq 1 - \frac{\left(\max \left\{ \frac{\bar{\rho}_1}{(1-\varepsilon_1)}, \frac{\bar{\rho}_2}{(1-\varepsilon_2)} \right\} + 1 \right) \epsilon (k_1 + k_2)}{\frac{(k_1 + k_2)}{2 \max \{ \bar{\rho}_1, \bar{\rho}_2 \}}} \text{ by (23)} \\
&= 1 - 2 \max \{ \bar{\rho}_1, \bar{\rho}_2 \} \left(\max \left\{ \frac{\bar{\rho}_1}{(1-\varepsilon_1)}, \frac{\bar{\rho}_2}{(1-\varepsilon_2)} \right\} + 1 \right) \epsilon \\
&\geq 1 - 4 \max \{ \bar{\rho}_1, \bar{\rho}_2 \} \left(\max \left\{ \frac{\bar{\rho}_1}{(1-\varepsilon_1)}, \frac{\bar{\rho}_2}{(1-\varepsilon_2)} \right\} + 1 \right) \epsilon. \tag{30}
\end{aligned}$$

Inequalities (28) and (30) imply (22). ■

Consider now the Markov Chain describing Phase 1 of the algorithm and let $R_l^i \in \{0, 1\}$, $i = 1, 2$, be the number of packets successfully received (rewards) in slot l by receiver i . Let

$$\widehat{R}_l^i = \sum_{m=1}^l R_m^i, \quad i = 1, 2,$$

be the reward accumulated up to slot l , and define

$$T_k^i : \text{ first time slot such that } \widehat{R}_l^i = k, \quad i = 1, 2.$$

The following lemma follows from the underlying regenerative structure of the Markov Chain and the associated rewards.

Lemma 12. *The following hold*

$$\lim_{k \rightarrow \infty} \frac{T_k^i}{k} = \frac{1}{\bar{\rho}_i}, \quad \text{with probability 1,} \tag{31}$$

$$\lim_{k \rightarrow \infty} \mathbb{E} \left[\frac{T_k^i}{k} \right] = \frac{1}{\bar{\rho}_i}. \tag{32}$$

Proof: The Markov Chain starts from state A . Let $t_0 = 0$ and t_h , $h \geq 1$ be the first time after t_{h-1} that the Markov Chains returns to state A . Let also $I_h = t_h - t_{h-1}$, $h \geq 1$ and $\widehat{I}_h = \sum_{v=1}^h I_v$. It is well known that $\{I_h\}_{h=1}^\infty$ are i.i.d., i.e., they constitute a renewal process, and $\{\widehat{R}_l^i\}_{l=1}^\infty$, $i = 1, 2$, constitutes a regenerative process with respect to $\{I_h\}_{h=1}^\infty$.

Define the total reward accumulated up to renewal epoch h , as $\widetilde{R}_h^i = \widehat{R}_{t_h}^i$, where $\widetilde{R}_0^i = 0$. Define also,

$$\widehat{H}_k^i = \max \left\{ h \geq 0 : \widetilde{R}_h^i \leq k \right\}.$$

Since the Markov Chain is finite, it holds,

$$\mathbb{E} \left[\widetilde{R}_1^i \right] \leq \mathbb{E} [I_1] < \infty, \tag{33}$$

and hence, [29, Th6, p.164]

$$\lim_{l \rightarrow \infty} \mathbb{E} \left[\frac{\widehat{R}_l^i}{l} \right] = \lim_{l_i \rightarrow \infty} \frac{\widehat{R}_{l_i}^i}{l_i} = \frac{\mathbb{E} \left[\widetilde{R}_1^i \right]}{\mathbb{E} [I_1]} = \bar{\rho}_i < \infty. \tag{34}$$

Applying the Elementary Renewal Theorem [29, p. 59] (considering \widetilde{R}_h^i , $h = 1, 2, \dots$, as renewal intervals) we have,

$$\lim_{k \rightarrow \infty} \mathbb{E} \left[\frac{\widehat{H}_k^i}{k} \right] = \lim_{k \rightarrow \infty} \frac{\widehat{H}_k^i}{k} = \frac{1}{\mathbb{E} \left[\widetilde{R}_1^i \right]} > 0. \tag{35}$$

Also, since $\{I_h\}_{h=1}^\infty$ are i.i.d.,

$$\lim_{h \rightarrow \infty} \mathbb{E} \left[\frac{\widehat{I}_h}{h} \right] = \lim_{h \rightarrow \infty} \frac{\widehat{I}_h}{h} = \mathbb{E} [I_1]. \tag{36}$$

It follows from the definitions that

$$\widehat{I}_{\widehat{H}_k^i - 1} < T_k^i \leq \widehat{I}_{\widehat{H}_k^i}, \tag{37}$$

hence

$$\frac{\widehat{T}_{\widehat{H}_k^i-1}}{(\widehat{H}_k^i-1)} \frac{(\widehat{H}_k^i-1)}{k} < \frac{T_k^i}{k} \leq \frac{\widehat{T}_{\widehat{H}_k^i}}{\widehat{H}_k^i} \frac{\widehat{H}_k^i}{k}. \quad (38)$$

From (35) it follows that $\lim_{k \rightarrow \infty} \widehat{H}_k^i = \infty$ and hence, taking into account (36),

$$\lim_{l \rightarrow \infty} \frac{\widehat{T}_{\widehat{H}_l^i}}{\widehat{H}_l^i} = \lim_{l \rightarrow \infty} \frac{\widehat{T}_{\widehat{H}_l^i-1}}{\widehat{H}_l^i-1} = \mathbb{E}[I_1].$$

This, (34), (35) and (38) imply that,

$$\lim_{k \rightarrow \infty} \frac{T_k^i}{k} = \frac{\mathbb{E}[I_1]}{\mathbb{E}[\widetilde{R}_1^i]} = \frac{1}{\bar{\rho}_i}. \quad (39)$$

From the last equality and Fatou's Lemma [30, p. 209] it follows that,

$$\frac{1}{\bar{\rho}_i} \leq \liminf_{k \rightarrow \infty} \mathbb{E} \left[\frac{T_k^i}{k} \right]. \quad (40)$$

Observe next that \widehat{H}_k^i is a stopping time for $\{\widehat{T}_h, \widetilde{R}_h^i\}_{h=1}^{\infty}$, hence,

$$\mathbb{E}[\widehat{T}_{\widehat{H}_k^i}] = \mathbb{E}[\widehat{H}_k^i] \mathbb{E}[I_1].$$

From the last equality, (35) and (37) we get,

$$\limsup_{k \rightarrow \infty} \mathbb{E} \left[\frac{T_k^i}{k} \right] \leq \frac{\mathbb{E}[I_1]}{\mathbb{E}[\widetilde{R}_1^i]} = \frac{1}{\bar{\rho}_i}. \quad (41)$$

Equality (32) follows from (40) and (41). ■

According to Algorithm II, the number of slots needed for Phase 1 to complete is,

$$T_{k_1, k_2} = \min \{T_{k_1}^1, T_{k_2}^2\}.$$

By time T_{k_1, k_2} , there exists an $i \in \{1, 2\}$ such that k_i packets have been transmitted to receiver i . Let J_{k_1, k_2} be the receiver with packets remaining to be transmitted (if both receiver packets are transmitted by T_{k_1, k_2} , assign J_{k_1, k_2} arbitrarily to one of the receivers), i.e.,

$$J_{k_1, k_2} = \begin{cases} 2 & \text{if } T_{k_1, k_2} = T_{k_1}^1 \\ 1 & \text{otherwise} \end{cases}. \quad (42)$$

According to the definition, the number of packets successfully received by receiver i at the end of phase 1 is, $\widehat{R}_{T_{k_1, k_2}}^i$, $i = 1, 2$.

In the following, by

$$\lim_{\min \{k_1, k_2\} \rightarrow \infty} f(k_1, k_2),$$

$$ak_1 \geq bk_2$$

we denote the limit of $f(k_1, k_2)$ as $\min \{k_1, k_2\} \rightarrow \infty$ in the region $\{(k_1, k_2) : ak_1 \geq bk_2\}$.

Lemma 13. *The following hold with probability 1*

$$\lim_{k_1 \rightarrow \infty} T_{k_1, k_2} = T_{k_2}^2, \text{ for all } k_2, \quad (43)$$

$$\lim_{k_1 \rightarrow \infty} \widehat{R}_{T_{k_1, k_2}}^i = \widehat{R}_{T_{k_2}^2}^i, \text{ for all } k_2, \quad (44)$$

$$\lim_{k_1 \rightarrow \infty} J_{k_1, k_2} = 1, \text{ for all } k_2. \quad (45)$$

Similar limits hold by interchanging indices 1 and 2. Moreover,

$$\lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} = 1, \quad (46)$$

$$\lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{\widehat{R}_{T_{k_1, k_2}}^i}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} = \bar{\rho}_i, \quad (47)$$

$$\lim_{\substack{\min\{k_1, k_2\} \rightarrow \infty \\ \left(\frac{1}{\bar{\rho}_2} - \delta\right) k_2 \geq \left(\frac{1}{\bar{\rho}_1} + \delta\right) k_1}} J_{k_1, k_2} = 2, \text{ for any } \delta, 0 < \delta < \frac{1}{\bar{\rho}_2}, \quad (48)$$

$$\lim_{\substack{\min\{k_1, k_2\} \rightarrow \infty \\ \left(\frac{1}{\bar{\rho}_1} - \delta\right) k_1 \geq \left(\frac{1}{\bar{\rho}_2} + \delta\right) k_2}} J_{k_1, k_2} = 1, \text{ for any } \delta, 0 < \delta < \frac{1}{\bar{\rho}_1}. \quad (49)$$

All the above limits hold also in expectation.

Proof: The limits (43), (44) and (45) follow from the definition of T_{k_1, k_2} and the fact that by (31) $\lim_{k_1 \rightarrow \infty} T_{k_1}^1 = \infty$, while for any fixed k_2 , $T_{k_2}^2$ is finite. To show (46) it suffices to show that for any ϵ , $0 < \epsilon \leq \gamma$, γ being a constant, we can pick $K(\epsilon)$ large enough so that for all $k_1, k_2 \geq K(\epsilon)$,

$$1 - \epsilon/\gamma \leq \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} \leq 1 + \epsilon/\gamma. \quad (50)$$

We proceed as follows. Pick $\gamma = 1/8$ which by the fact that $\bar{\rho}_i \leq 1$ implies that

$$\gamma \leq \min\{1, 1/(2\bar{\rho}_1), 1/(2\bar{\rho}_2)\}, \quad (51)$$

$$\gamma \leq \frac{1}{2(\max\{\bar{\rho}_1, \bar{\rho}_2\} + \bar{\rho}_1 + \bar{\rho}_2 + \bar{\rho}_1\bar{\rho}_2)}. \quad (52)$$

Consider now any $0 < \epsilon \leq \gamma$. According to (31) we can pick $K_1(\epsilon)$, $K_2(\epsilon)$ such that for $i = 1, 2$,

$$\left(\frac{1}{\bar{\rho}_i} - \epsilon\right) k_i \leq T_{k_i}^i \leq \left(\frac{1}{\bar{\rho}_i} + \epsilon\right) k_i, \text{ for all } k_i \geq K(\epsilon) \triangleq \max_{i=1,2}\{K_i(\epsilon)\}. \quad (53)$$

Assume now that $k_i \geq K(\epsilon)$, $i = 1, 2$. We distinguish four cases

1. $\left(\frac{1}{\bar{\rho}_1} - \epsilon\right) k_1 \geq \left(\frac{1}{\bar{\rho}_2} + \epsilon\right) k_2$: Then $\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\} = \frac{k_2}{\bar{\rho}_2}$ and according to (53), $T_{k_1}^1 \geq T_{k_2}^2$. Hence,

$$\frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} = \frac{T_{k_2}^2}{\frac{k_2}{\bar{\rho}_2}},$$

which by (53) implies that

$$1 - \epsilon\bar{\rho}_2 \leq \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} \leq 1 + \epsilon\bar{\rho}_2, \quad (54)$$

$$1 - \epsilon/\gamma \leq \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} \leq 1 + \epsilon/\gamma, \text{ by (51).}$$

2. $\left(\frac{1}{\bar{\rho}_1} - \epsilon\right) k_1 < \left(\frac{1}{\bar{\rho}_2} + \epsilon\right) k_2 \leq \left(\frac{1}{\bar{\rho}_1} + \epsilon\right) k_1$: Equivalently,

$$\frac{\left(\frac{1}{\bar{\rho}_1} - \epsilon\right) k_1}{(1 + \epsilon\bar{\rho}_2)} < \frac{k_2}{\bar{\rho}_2} \leq \frac{\left(\frac{1}{\bar{\rho}_1} + \epsilon\right) k_1}{(1 + \epsilon\bar{\rho}_2)}. \quad (55)$$

In this case, both $T_{k_1}^1 \geq T_{k_2}^2$ or $T_{k_1}^1 \leq T_{k_2}^2$ may happen. If $T_{k_1}^1 \geq T_{k_2}^2$ and $\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\} = \frac{k_2}{\bar{\rho}_2}$, then as in case 1, (54) holds.

Let now $T_{k_1}^1 < T_{k_2}^2$ but $\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\} = \frac{k_2}{\bar{\rho}_2}$. Then we write,

$$\begin{aligned}
\frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} &= \frac{T_{k_1}^1}{\frac{k_2}{\bar{\rho}_2}}, \\
\frac{T_{k_1}^1}{\frac{k_1}{\bar{\rho}_1}} \frac{(1 + \epsilon \bar{\rho}_2)}{(1 + \epsilon \bar{\rho}_1)} &\leq \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} \leq \frac{T_{k_1}^1}{\frac{k_1}{\bar{\rho}_1}} \frac{(1 + \epsilon \bar{\rho}_2)}{(1 - \epsilon \bar{\rho}_1)}, \text{ by (55)} \\
\frac{(1 - \epsilon \bar{\rho}_1)(1 + \epsilon \bar{\rho}_2)}{(1 + \epsilon \bar{\rho}_1)} &\leq \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} \leq \frac{(1 + \epsilon \bar{\rho}_1)(1 + \epsilon \bar{\rho}_2)}{(1 - \epsilon \bar{\rho}_1)}, \text{ by (53)} \\
\frac{1 - \epsilon \bar{\rho}_1}{1 + \epsilon \bar{\rho}_1} &\leq \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} \leq \frac{1 + \epsilon(\bar{\rho}_1 + \bar{\rho}_2 + \bar{\rho}_1 \bar{\rho}_2)}{1 - \epsilon \bar{\rho}_1}, \text{ since } \epsilon \leq 1 \\
1 - \epsilon \frac{2\bar{\rho}_1}{1 + \epsilon \bar{\rho}_1} &\leq \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} \leq 1 + \epsilon \frac{(2\bar{\rho}_1 + \bar{\rho}_2 + \bar{\rho}_1 \bar{\rho}_2)}{1 - \epsilon \bar{\rho}_1}, \\
1 - 2\bar{\rho}_1 \epsilon &\leq \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} \leq 1 + 2(2\bar{\rho}_1 + \bar{\rho}_2 + \bar{\rho}_1 \bar{\rho}_2) \epsilon, \text{ since } 0 < \epsilon \leq 1/(2\bar{\rho}_1) \\
1 - \epsilon/\gamma &\leq \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} \leq 1 + \epsilon/\gamma, \text{ by (51) and (52).}
\end{aligned}$$

3. $\left(\frac{1}{\bar{\rho}_2} - \epsilon\right) k_2 < \left(\frac{1}{\bar{\rho}_1} + \epsilon\right) k_1 \leq \left(\frac{1}{\bar{\rho}_2} + \epsilon\right)$: Arguing as in case 2, we see that (50) holds.

4. $\left(\frac{1}{\bar{\rho}_2} - \epsilon\right) k_2 \geq \left(\frac{1}{\bar{\rho}_1} + \epsilon\right) k_1$: Arguing as in case 1, we see again that (50) holds.

Hence (50) holds for all 4 cases.

To prove (47) notice that since according to (46) we have

$$\lim_{\min\{k_1, k_2\} \rightarrow \infty} T_{k_1, k_2} = \infty,$$

it follows that

$$\begin{aligned}
\lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{\widehat{R}_{T_{k_1, k_2}}^i}{T_{k_1, k_2}} &= \lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{\sum_{m=1}^{T_{k_1, k_2}} R_m^i}{T_{k_1, k_2}} \\
&= \bar{\rho}_i.
\end{aligned}$$

Hence, taking also into account (46),

$$\begin{aligned}
\lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{\widehat{R}_{T_{k_1, k_2}}^i}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} &= \lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{\widehat{R}_{T_{k_1, k_2}}^i}{T_{k_1, k_2}} \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} \\
&= \bar{\rho}_i.
\end{aligned}$$

To show (48) pick in (53) $\epsilon < \min\{\delta, 1/8\}$. It follows that if $\min\{k_1, k_2\} \geq K(\epsilon)$ and

$$\left(\frac{1}{\bar{\rho}_2} - \delta\right) k_2 \geq \left(\frac{1}{\bar{\rho}_1} + \delta\right) k_1,$$

then $T_{k_2}^2 \geq T_{k_1}^1$, hence $T_{k_1, k_2} = T_{k_1}^1$ and according to definition (42), $J_{k_1, k_2} = 2$. In a similar fashion (49) can be shown.

It remains to show that the limits hold in expectation as well. For (43)-(44) this follows from

$$0 \leq \widehat{R}_{T_{k_1, k_2}}^i \leq T_{k_1, k_2} \leq T_{k_i^c}^i,$$

and the fact that $\mathbb{E}[T_{k_i^c}^i] < \infty$ [30, Problem 16.7]. For (45), (48), (49) convergence in expectation follows for the boundedness of the indices.

To show that the limits hold in expectation for (46), (47), observe that

$$0 \leq \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} \leq \max\left\{\frac{T_{k_1}^1}{\bar{\rho}_1} \frac{1}{k_1}, \bar{\rho}_2 \frac{T_{k_2}^2}{k_2}\right\}. \quad (56)$$

It follows from Lemma 12 that the sequences $\frac{T_{k_1}^1}{k_1}$ and $\frac{T_{k_2}^2}{k_2}$ are uniformly integrable [30, Corollary p. 218], hence (56) implies that $T_{k_1, k_2} / \min \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}$ is also uniformly integrable in the sense,

$$\lim_{a \rightarrow \infty} \sup_{k_1, k_2} \int_{\frac{T_{k_1, k_2}}{\min \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}} \geq a} \frac{T_{k_1, k_2}}{\min \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}} dP = 0. \quad (57)$$

To see this, set $\Phi(k_1, k_2) = T_{k_1, k_2} / \min \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}$, $\Theta(k_1, k_2) = \max \left\{ \bar{\rho}_1 \frac{T_{k_1}^1}{k_1}, \bar{\rho}_2 \frac{T_{k_2}^2}{k_2} \right\}$ and write,

$$\begin{aligned} 0 &\leq \sup_{k_1, k_2} \int_{\Phi(k_1, k_2) \geq a} \Phi(k_1, k_2) dP \leq \sup_{k_1, k_2} \int_{\Theta(k_1, k_2) \geq a} \Theta(k_1, k_2) dP \quad \text{by (56),} \\ &\leq \sup_{k_1, k_2} \left\{ \bar{\rho}_1 \int_{\Theta(k_1, k_2) \geq a, \bar{\rho}_1 \frac{T_{k_1}^1}{k_1} \geq \bar{\rho}_2 \frac{T_{k_2}^2}{k_2}} \frac{T_{k_1}^1}{k_1} dP + \bar{\rho}_2 \int_{\Theta(k_1, k_2) \geq a, \bar{\rho}_1 \frac{T_{k_1}^1}{k_1} \leq \bar{\rho}_2 \frac{T_{k_2}^2}{k_2}} \frac{T_{k_2}^2}{k_2} dP \right\} \\ &\leq \bar{\rho}_1 \sup_{k_1, k_2} \int_{\Theta(k_1, k_2) \geq a, \bar{\rho}_1 \frac{T_{k_1}^1}{k_1} \geq \bar{\rho}_2 \frac{T_{k_2}^2}{k_2}} \frac{T_{k_1}^1}{k_1} dP + \bar{\rho}_2 \sup_{k_1, k_2} \int_{\Theta(k_1, k_2) \geq a, \bar{\rho}_1 \frac{T_{k_1}^1}{k_1} \leq \bar{\rho}_2 \frac{T_{k_2}^2}{k_2}} \frac{T_{k_2}^2}{k_2} dP \\ &\leq \bar{\rho}_1 \sup_{k_1, k_2} \int_{\frac{T_{k_1}^1}{\bar{\rho}_1} \geq a} \frac{T_{k_1}^1}{k_1} dP + \bar{\rho}_2 \sup_{k_1, k_2} \int_{\frac{T_{k_2}^2}{\bar{\rho}_2} \geq a} \frac{T_{k_2}^2}{k_2} dP \\ &= \bar{\rho}_1 \sup_{k_1} \int_{\frac{T_{k_1}^1}{k_1} \geq \frac{a}{\bar{\rho}_1}} \frac{T_{k_1}^1}{k_1} dP + \bar{\rho}_2 \sup_{k_2} \int_{\frac{T_{k_2}^2}{k_2} \geq \frac{a}{\bar{\rho}_2}} \frac{T_{k_2}^2}{k_2} dP. \end{aligned} \quad (58)$$

Taking limits as $a \rightarrow \infty$ in (58) and using the uniform integrability of $\frac{T_{k_i}^i}{k_i}$, (57) follows.

Condition (57) implies convergence in expectation of (46). Since $\widehat{R}_{T_{k_1, k_2}}^i \leq T_{k_1, k_2}$,

$$0 \leq \frac{\widehat{R}_{T_{k_1, k_2}}^i}{\min \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}} \leq \frac{T_{k_1, k_2}}{\min \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}} \leq \max \left\{ \bar{\rho}_1 \frac{T_{k_1}^1}{k_1}, \bar{\rho}_2 \frac{T_{k_2}^2}{k_2} \right\}, \quad (59)$$

from which it follows again that (47) holds in expectation. ■

Consider now Phase 2 of the algorithm. The number of packets that need to be transmitted in this phase is

$$K_{k_1, k_2} = k_{J_{k_1, k_2}} - \widehat{R}_{T_{k_1, k_2}}^{J_{k_1, k_2}}.$$

Let $\{S_l^i\}_{l=1}^{\infty}$, $i = 1, 2$, be i.i.d. random variables, geometrically distributed with parameter ε_i , (i.e., $\Pr(S_l^i = k) = \varepsilon_i^{k-1} (1 - \varepsilon_i)$, $k \geq 1$) and independent of all processes during phase 1. Then Phase 2 lasts for

$$\hat{S}_{k_1, k_2} = \sum_{l=1}^{K_{k_1, k_2}} S_l^{J_{k_1, k_2}}$$

slots, where we used the convention $\sum_{m=l}^k x_m = 0$ if $k < l$. Since K_{k_1, k_2} , J_{k_1, k_2} are independent of $\{S_l^i\}_{l=1}^{\infty}$, $i = 1, 2$, we have

$$\begin{aligned} \mathbb{E} \left[\sum_{l=1}^{K_{k_1, k_2}} S_l^{J_{k_1, k_2}} \right] &= \mathbb{E} \left[\mathbb{E} \left[\sum_{l=1}^{K_{k_1, k_2}} S_l^{J_{k_1, k_2}} \middle| K_{k_1, k_2}, J_{k_1, k_2} \right] \right] \\ &= \mathbb{E} \left[\frac{K_{k_1, k_2}}{1 - \varepsilon_{J_{k_1, k_2}}} \right]. \end{aligned} \quad (60)$$

The total number of slots needed to complete both phases is

$$N_{k_1, k_2} = T_{k_1, k_2} + \hat{S}_{k_1, k_2}. \quad (61)$$

We can now proceed with the proof of Theorem 7. We repeat the theorem here for convenience

Theorem 7. *The following holds with probability 1.*

$$\lim_{|(k_1, k_2)| \rightarrow \infty} \frac{N_{k_1, k_2}}{v(k_1, k_2)} = 1.$$

The limit also holds in expectation.

Proof: It suffices to show that

$$\begin{aligned}\lim_{k_1 \rightarrow \infty} \frac{N_{k_1, k_2}}{v(k_1, k_2)} &= 1, \text{ for any fixed } k_2 \geq 0, \\ \lim_{k_2 \rightarrow \infty} \frac{N_{k_1, k_2}}{v(k_1, k_2)} &= 1, \text{ for any fixed } k_1 \geq 0, \\ \lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{N_{k_1, k_2}}{v(k_1, k_2)} &= 1,\end{aligned}$$

and that all the limits above hold also in expectation.

Let $k_1 \rightarrow \infty$. Then, according to Lemma 13,

$$\begin{aligned}\lim_{k_1 \rightarrow \infty} T_{k_1, k_2} &= T_{k_2}^2, \\ \lim_{k_1 \rightarrow \infty} \widehat{R}_{T_{k_1, k_2}}^{J_{k_1, k_2}} &= \widehat{R}_{T_{k_2}}^1, \\ \lim_{k_1 \rightarrow \infty} \frac{k_{J_{k_1, k_2}}}{k_1} &= 1, \\ \lim_{k_1 \rightarrow \infty} S_l^{J_{k_1, k_2}} &= S_l^1.\end{aligned}$$

Since the $\{S_l^1\}_{l=1}^\infty$ are i.i.d. and

$$\lim_{k_1 \rightarrow \infty} \frac{K_{k_1, k_2}}{k_1} = \lim_{k_1 \rightarrow \infty} \frac{k_{J_{k_1, k_2}} - \widehat{R}_{T_{k_1, k_2}}^{J_{k_1, k_2}}}{k_1} = 1,$$

we have

$$\begin{aligned}\lim_{k_1 \rightarrow \infty} \frac{\widehat{S}_{k_1, k_2}}{k_1} &= \lim_{k_1 \rightarrow \infty} \frac{\sum_{l=1}^{K_{k_1, k_2}} S_l^{J_{k_1, k_2}} K_{k_1, k_2}}{K_{k_1, k_2} k_1} \\ &= \frac{1}{1 - \varepsilon_1}.\end{aligned}$$

Hence for any $k_2 \geq 0$,

$$\begin{aligned}\lim_{k_1 \rightarrow \infty} \frac{N_{k_1, k_2}}{v(k_1, k_2)} &= \lim_{k_1 \rightarrow \infty} \left(\frac{T_{k_1, k_2}}{k_1} \frac{k_1}{v(k_1, k_2)} + \frac{\widehat{S}_{k_1, k_2}}{k_1} \frac{k_1}{v(k_1, k_2)} \right) \\ &= (1 - \varepsilon_1) \frac{1}{1 - \varepsilon_1} \\ &= 1.\end{aligned}\tag{62}$$

To show that (62) holds in expectation as well, note first that by Lemma 13 it holds

$$\lim_{k_1 \rightarrow \infty} \frac{\mathbb{E}[T_{k_1, k_2}]}{v(k_1, k_2)} = 0.$$

Hence it suffices to show that

$$\lim_{k_1 \rightarrow \infty} \frac{\mathbb{E}[\widehat{S}_{k_1, k_2}]}{v(k_1, k_2)} = 1.\tag{63}$$

To show the latter limit, note that

$$\lim_{k_1 \rightarrow \infty} \frac{K_{k_1, k_2}}{k_1 (1 - \varepsilon_{J_{k_1, k_2}})} = \frac{1}{1 - \varepsilon_1}.\tag{64}$$

Since

$$\begin{aligned}\frac{K_{k_1, k_2}}{k_1 (1 - \varepsilon_{J_{k_1, k_2}})} &\leq \frac{k_{J_{k_1, k_2}}}{k_1 (1 - \max\{\varepsilon_1, \varepsilon_2\})} \\ &\leq \frac{\max\{1, \frac{k_2}{k_1}\}}{(1 - \max\{\varepsilon_1, \varepsilon_2\})},\end{aligned}$$

for any fixed k_2 , the sequence

$$\frac{K_{k_1, k_2}}{k_1 (1 - \varepsilon_{J_{k_1, k_2}})}$$

is bounded. Convergence in expectation in (63) then follows from (60), (64) and the Bounded Convergence Theorem.

It remains to show that

$$\lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{N_{k_1, k_2}}{v(k_1, k_2)} = 1 \text{ a.e.} \quad (65)$$

$$\lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{\mathbb{E}[N_{k_1, k_2}]}{v(k_1, k_2)} = 1. \quad (66)$$

To analyze the behavior of these limits we consider their behavior in the following regions.

$$\mathcal{S}_1 = \left\{ (k_1, k_2) : \left(\frac{k_1}{\bar{\rho}_1} - \frac{k_2}{\bar{\rho}_2} \right) \geq (k_1 + k_2) \epsilon \right\},$$

$$\mathcal{S}_2 = \left\{ (k_1, k_2) : \left(\frac{k_2}{\bar{\rho}_2} - \frac{k_1}{\bar{\rho}_1} \right) \geq (k_1 + k_2) \epsilon \right\},$$

$$\mathcal{D} = \left\{ (k_1, k_2) : \left| \frac{k_1}{\bar{\rho}_1} - \frac{k_2}{\bar{\rho}_2} \right| < (k_1 + k_2) \epsilon \right\}.$$

Consider first the region \mathcal{S}_1 . In this region, it holds

$$k_1 - k_2 \frac{\bar{\rho}_1}{\bar{\rho}_2} \geq \bar{\rho}_1 (k_1 + k_2) \epsilon \geq 0, \quad (67)$$

and $\min \left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\} = \frac{k_2}{\bar{\rho}_2}$. Taking into account (47) and (49) we have for $(k_1, k_2) \in \mathcal{S}_1$,

$$\begin{aligned} \lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{K_{k_1, k_2}}{k_1 - k_2 \frac{\bar{\rho}_1}{\bar{\rho}_2}} &= \lim_{\min\{k_1, k_2\} \rightarrow \infty} \left(\frac{k_{J_{k_1, k_2}} - \frac{\widehat{R}_{T_{k_1, k_2}}^{J_{k_1, k_2}}}{\min\left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}} \frac{k_2}{\bar{\rho}_2}}{k_1 - k_2 \frac{\bar{\rho}_1}{\bar{\rho}_2}} \right) \\ &= 1. \end{aligned} \quad (68)$$

We conclude from (68) that $\lim_{\min\{k_1, k_2\} \rightarrow \infty} K_{k_1, k_2} = \infty$. Hence,

$$\begin{aligned} \lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{\hat{S}_{k_1, k_2}}{k_1 - k_2 \frac{\bar{\rho}_1}{\bar{\rho}_2}} &= \lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{\sum_{l=1}^{K_{k_1, k_2}} S_l^{J_{k_1, k_2}}}{K_{k_1, k_2}} \frac{K_{k_1, k_2}}{k_1 - k_2 \frac{\bar{\rho}_1}{\bar{\rho}_2}} \\ &= \frac{1}{1 - \varepsilon_1} \text{ by (68), (49),} \end{aligned} \quad (69)$$

and,

$$\begin{aligned} \lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{N_{k_1, k_2}}{v(k_1, k_2)} &= \lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{T_{k_1, k_2}}{\min\left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}} \frac{\min\left\{ \frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2} \right\}}{v(k_1, k_2)} + \frac{S_{k_1, k_2}}{k_1 - k_2 \frac{\bar{\rho}_1}{\bar{\rho}_2}} \frac{k_1 - k_2 \frac{\bar{\rho}_1}{\bar{\rho}_2}}{v(k_1, k_2)} \\ &= \lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{\frac{k_2}{\bar{\rho}_2}}{v(k_1, k_2)} + \frac{1}{1 - \varepsilon_1} \frac{k_1 - k_2 \frac{\bar{\rho}_1}{\bar{\rho}_2}}{v(k_1, k_2)}, \text{ by (46) and (69)} \\ &= \lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{\frac{k_1}{1 - \varepsilon_1} + \frac{k_2}{1 - \varepsilon_2}}{v(k_1, k_2)}, \text{ by (16)} \\ &= 1, \text{ by (18).} \end{aligned} \quad (70)$$

To show that (70) holds also in expectation in \mathcal{S}_1 , observe that in this region,

$$\begin{aligned} \frac{K_{k_1, k_2}}{\left(k_1 - k_2 \frac{\bar{\rho}_1}{\bar{\rho}_2} \right) (1 - \varepsilon_{J_{k_1, k_2}})} &\leq \frac{\max\{k_1, k_2\}}{\bar{\rho}_1 \epsilon (k_1 + k_2) (1 - \max\{\varepsilon_1, \varepsilon_2\})} \\ &\leq \frac{1}{\bar{\rho}_1 \epsilon (1 - \max\{\varepsilon_1, \varepsilon_2\})}, \end{aligned}$$

hence, using (60) and the Bounded Convergence Theorem,

$$\begin{aligned} \lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{\mathbb{E} [\hat{S}_{k_1, k_2}]}{\left(k_1 - k_2 \frac{\bar{\rho}_1}{\bar{\rho}_2}\right)} &= \mathbb{E} \left[\lim_{\min\{k_1, k_2\} \rightarrow \infty} \frac{K_{k_1, k_2}}{\left(k_1 - k_2 \frac{\bar{\rho}_1}{\bar{\rho}_2}\right) (1 - \varepsilon_{J_{k_1, k_2}})} \right] \\ &= \frac{1}{1 - \varepsilon_1}, \text{ by (68) and (49).} \end{aligned}$$

We can now repeat (70) by replacing the random variables with averages.

Similar arguments are used when we consider region \mathcal{S}_2 .

Finally, consider region \mathcal{D} . For (65) to hold in this region, it suffices to show that, for $\epsilon > 0$ we can pick $K(\epsilon)$ so that for all $(k_1, k_2) \in \mathcal{D}$ with $\min\{k_1, k_2\} \geq K(\epsilon)$, for some constants α, β , it holds,

$$1 - \alpha\epsilon \leq \frac{N_{k_1, k_2}}{v(k_1, k_2)} \leq 1 + \beta\epsilon,$$

For (66) to hold it suffices to show a similar inequality, but with N_{k_1, k_2} replaced by $\mathbb{E}[N_{k_1, k_2}]$.

In the following pick $0 < \epsilon \leq \min\{\bar{\rho}_1, \bar{\rho}_2, 1/3c\}$ where c is defined in (19). Writing

$$\frac{T_{k_1, k_2}}{v(k_1, k_2)} = \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} \frac{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}}{v(k_1, k_2)},$$

we get according to (22),

$$\frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} (1 + c\epsilon) \geq \frac{T_{k_1, k_2}}{v(k_1, k_2)} \geq \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} (1 - c\epsilon). \quad (71)$$

According to (46) we can pick $\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\} \geq K(\epsilon)$ so that

$$1 + c\epsilon \geq \frac{T_{k_1, k_2}}{\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}} \geq 1 - c\epsilon,$$

and then (71) becomes

$$1 + 3c\epsilon \geq \frac{T_{k_1, k_2}}{v(k_1, k_2)} \geq 1 - 3c\epsilon. \quad (72)$$

Note that since the convergence in (46) holds also in expectation, it also holds for $\min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\} \geq K(\epsilon)$,

$$1 + 3c\epsilon \geq \frac{\mathbb{E}[T_{k_1, k_2}]}{v(k_1, k_2)} \geq 1 - 3c\epsilon. \quad (73)$$

Consider now the second term in the total length (61), i.e.,

$$\begin{aligned} \hat{S}_{k_1, k_2} &= \sum_{l=1}^{K_{k_1, k_2}} S_l^{J_{k_1, k_2}}, \\ K_{k_1, k_2} &= k_{J_{k_1, k_2}} - \hat{R}_{T_{k_1, k_2}}^{J_{k_1, k_2}}. \end{aligned}$$

Pick $\min\{k_1, k_2\} \geq K(\epsilon)$ so that according to (47),

$$\hat{R}_{T_{k_1, k_2}}^{J_{k_1, k_2}} \geq (\bar{\rho}_1 - \epsilon) \min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}.$$

Note that if $J_{k_1, k_2} = 1$ then,

$$\begin{aligned} K_{k_1, k_2} &= k_{J_{k_1, k_2}} - \hat{R}_{T_{k_1, k_2}}^{J_{k_1, k_2}} \\ &\leq k_1 - (\bar{\rho}_1 - \epsilon) \min\left\{\frac{k_1}{\bar{\rho}_1}, \frac{k_2}{\bar{\rho}_2}\right\}, \text{ by (47)} \\ &= k_1 - (\bar{\rho}_1 - \epsilon) \frac{k_1}{\bar{\rho}_1} - (\bar{\rho}_1 - \epsilon) \min\left\{\frac{k_2}{\bar{\rho}_2} - \frac{k_1}{\bar{\rho}_1}, 0\right\} \\ &\leq \frac{k_1}{\bar{\rho}_1} \epsilon + (\bar{\rho}_1 - \epsilon) \epsilon (k_1 + k_2), \text{ since } (k_1, k_2) \in \mathcal{D} \\ &\leq \left(\frac{1}{\bar{\rho}_1} + \bar{\rho}_1\right) \epsilon (k_1 + k_2). \end{aligned}$$

A similar inequality holds when $J_{k_1, k_2} = 2$ and hence we conclude that in \mathcal{D} ,

$$0 \leq \frac{K_{k_1, k_2}}{k_1 + k_2} = \frac{k_{J_{k_1, k_2}} - \widehat{R}_{T_{k_1, k_2}}^{J_{k_1, k_2}}}{k_1 + k_2} \leq \max_{i=1,2} \left\{ \frac{1}{\bar{\rho}_i} + \bar{\rho}_i \right\} \epsilon. \quad (74)$$

Note that since (47) holds also in expectation, using similar steps we show that for $\min\{k_1, k_2\} \geq k(\epsilon)$ it holds

$$0 \leq \frac{\mathbb{E}[K_{k_1, k_2}]}{k_1 + k_2} \leq \max_{i=1,2} \left\{ \frac{1}{\bar{\rho}_i} + \bar{\rho}_i \right\} \epsilon. \quad (75)$$

Observe next that

$$\hat{S}_{k_1, k_2} \leq \sum_{l=1}^{K_{k_1, k_2}} \max\{S_l^1, S_l^2\} \triangleq U_{k_1, k_2}. \quad (76)$$

Since the sequence $\{\max\{S_l^1, S_l^2\}\}_{l=1}^\infty$ consists of i.i.d. variables with finite expectation, we conclude that with probability 1,

$$\sup_k \frac{\sum_{l=1}^k \max\{S_l^1, S_l^2\}}{k} \leq \infty,$$

and hence,

$$\sup_{k_1, k_2} \frac{U_{k_1, k_2}}{K_{k_1, k_2}} = C_1 < \infty.$$

Hence,

$$\begin{aligned} 0 &\leq \frac{S_{k_1, k_2}}{V(k_1, k_2)} \\ &\leq \frac{U_{k_1, k_2}}{K_{k_1, k_2}} \frac{K_{k_1, k_2}}{k_1 + k_2} \frac{k_1 + k_2}{v(k_1, k_2)} \\ &\leq C_1 \max_{i=1,2} \left\{ \frac{1}{\bar{\rho}_i} + \bar{\rho}_i \right\} \epsilon \max_{i=1,2} \{1 - \varepsilon_i\}. \end{aligned} \quad (77)$$

Combining (61), (72) and (77) we conclude that in \mathcal{D} and for $\min\{k_1, k_2\} \geq K(\epsilon)$,

$$1 - 3c\epsilon \leq \frac{N_{k_1, k_2}}{v(k_1, k_2)} \leq 1 + \left(3c + C_1 \max_{i=1,2} \left\{ \frac{1}{\bar{\rho}_i} + \bar{\rho}_i \right\} \max_{i=1,2} \{1 - \varepsilon_i\} \right) \epsilon.$$

This shows (65) in \mathcal{D} . To show that the same result holds in expectation, observe that (76) implies

$$0 \leq \mathbb{E}[\hat{S}_{k_1, k_2}] \leq \mathbb{E}[\max\{S_1^1, S_1^2\}] \mathbb{E}[K_{k_1, k_2}],$$

and hence for $\min\{k_1, k_2\} \geq k(\epsilon)$,

$$\begin{aligned} 0 &\leq \frac{\mathbb{E}[S_{k_1, k_2}]}{V(k_1, k_2)} \\ &\leq \mathbb{E}[\max\{S_1^1, S_1^2\}] \frac{\mathbb{E}[K_{k_1, k_2}]}{v(k_1, k_2)} \\ &= \mathbb{E}[\max\{S_1^1, S_1^2\}] \frac{\mathbb{E}[K_{k_1, k_2}]}{k_1 + k_2} \frac{k_1 + k_2}{v(k_1, k_2)} \\ &\leq \mathbb{E}[\max\{S_1^1, S_1^2\}] \max_{i=1,2} \left\{ \frac{1}{\bar{\rho}_i} + \bar{\rho}_i \right\} \epsilon \max_{i=1,2} \{1 - \varepsilon_i\}, \text{ by (75)}. \end{aligned}$$

Arguing again as above, we conclude that (66) holds. ■

BIOGRAPHIES

Yalin Evren Sagduyu

Yalin Evren Sagduyu (S'02, M'08) received the B.S. degree from Bogazici University, Turkey, in electrical and electronics engineering and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland at College Park. He has been a Graduate Research Assistant with the Institute for Systems Research at the University of Maryland and a Postdoctoral Fellow in the Department of Electrical Engineering and Computer Science at Northwestern University. He is currently a Senior Research Scientist in Intelligent Automation Inc. and a Visiting Assistant Research Scientist in the Institute for Systems Research at the University of Maryland, College Park. His research interests are in the areas of design and optimization of wireless networks, multi-user communications, network coding, information theory, network security and game theory.

Leonidas Georgiadis

Leonidas Georgiadis (S'76, M'78, SM'96) received the Diploma degree in electrical engineering from Aristotle University, Thessaloniki, Greece, in 1979, and his M.S. and Ph.D degrees both in electrical engineering from the University of Connecticut, in 1981 and 1986 respectively. From 1981 to 1983 he was with the Greek army. From 1986 to 1987 he was Research Assistant Professor at the University of Virginia, Charlottesville. In 1987 he joined IBM T. J. Watson Research Center, Yorktown Heights as a Research Staff Member. Since October 1995, he has been with the Telecommunications Department of Aristotle University, Thessaloniki, Greece. His interests are in the area of wireless networks, information theory and coding, high speed networks, distributed systems, routing, scheduling, congestion control, modeling and performance analysis. In 1992 he received the IBM Outstanding Innovation Award for his work on Goal-Oriented Workload Management for Multi-class systems.

Leandros Tassioulas

Leandros Tassioulas (S'89, M'91, SM'05, F'07) is Professor of Telecommunication Networks in the Department of Computer Engineering and Telecommunications at the University of Thessaly Greece since 2002 and Associate Director of the Informatics and Telematics Institute of the Center for Research and Technology Hellas (CERTH).

His research interests are in the field of computer and communication networks with emphasis on fundamental mathematical models and algorithms, architectures and protocols of wireless systems, sensor networks, novel internet architectures and satellite communications. He holds a Diploma in Electrical Engineering from the Aristotelian Univ. of Thessaloniki, Greece, in 1987, and a Ph.D. degree in Electrical Engineering from the Univ. of Maryland, College Park in 1991. He has held positions as Assistant Professor at Polytechnic University New York (1991-95), Assistant and Associate Professor Univ. of Maryland College Park (1995-2001) and professor Univ. of Ioannina Greece (1999-2001). He has been visiting researcher at IBM T.J.Watson research center in 1999 and in Flarion Technologies in 2003. Dr. Tassioulas has been the principal investigator in several research projects funded by industry, European Commission and USA government.

He is a Fellow of IEEE since 2007 while his research has been recognized by several awards including the inaugural INFOCOM 2007 Achievement Award "For fundamental contributions to resource allocation in communication networks", the INFOCOM 1994 best paper award, a National Science Foundation (NSF) Research Initiation Award in 1992, an NSF CAREER Award in 1995, an Office of Naval Research Young Investigator Award in 1997 and a Bodosaki Foundation award in 1999.

Anthony Ephremides

Anthony Ephremides (M'71, SM'77, F'84, LF'09) holds the Cynthia Kim Professorship of Information Technology at the Electrical and Computer Engineering Department of the University of Maryland in College Park where he holds a joint appointment at the Institute for Systems Research, of which he was among the founding members in 1986. He obtained his PhD in Electrical Engineering from Princeton University in 1971 and has been with the University of Maryland ever since.

He has held various visiting positions at other Institutions (including MIT, UC Berkeley, ETH Zurich, INRIA, etc) and co-founded and co-directed a NASA-funded Center on Satellite and Hybrid Communication Networks in 1991. He has been the President of Pontos, Inc, since 1980 and has served as President of the IEEE Information Theory Society in 1987 and as a member of the IEEE Board of Directors in 1989 and 1990. He has been the General Chair and/or the Technical Program Chair of several technical conferences (including the IEEE Information Theory Symposium in 1991 and 2000, the IEEE Conference on Decision and Control in 1986, the ACM Mobihoc in 2003, and the IEEE Infocom in 1999). He has served on the Editorial Board of numerous journals and was the Founding Director of the Fairchild Scholars and Doctoral Fellows Program, a University-Industry Partnership from 1981 to 1985.

He has received the IEEE Donald E. Fink Prize Paper Award in 1991 and the first ACM Achievement Award for Contributions to Wireless Networking in 1996, as well as the 2000 Fred W. Ellersick MILCOM Best Paper Award, the IEEE Third Millennium Medal, the 2000 Outstanding Systems Engineering Faculty Award from the Institute for Systems Research, and the Kirwan Faculty Research and Scholarship Prize from the University of Maryland in 2001, and a few other official recognitions of his

work. He also received the 2006 Aaron Wyner Award for Exceptional Service and Leadership to the IEEE Information Theory Society.

He is the author of several hundred papers, conference presentations, and patents, and his research interests lie in the areas of Communication Systems and Networks and all related disciplines, such as Information Theory, Control and Optimization, Satellite Systems, Queueing Models, Signal Processing, etc. He is especially interested in Wireless Networks and Energy Efficient Systems