Cache-Enabled Erasure Broadcast Channels with Feedback -Asymmetric User Memory Case

Asma Ghorbel, Mari Kobayashi, Sheng Yang

Cache-Enabled Erasure Broadcast Channels with Feedback - Asymmetric User Memory Case

Asma Ghorbel
CentraleSupélec
Gif-sur-Yvette, France
asma.ghorbel@centralesupelec.fr

Mari Kobayashi
CentraleSupélec
Gif-sur-Yvette, France
mari.kobayashi@centralesupelec.fr

Sheng Yang
CentraleSupélec
Gif-sur-Yvette, France
sheng.yang@centralesupelec.fr

ABSTRACT

We consider a cache-enabled K-user erasure broadcast channel in which a server with a library of N files wishes to deliver a requested file to each user k who is equipped with a cache of a finite memory M_k. Assuming that the transmitter has state feedback and user caches can be filled during off-peak hours reliably by decentralized cache placement, we characterize the achievable rate region as a function of the memory sizes and the erasure probabilities. The proposed delivery scheme, based on the broadcasting scheme proposed by Wang and Gatzianas et al., exploits the receiver side information established during the placement phase. A two-user toy example shows that the cache network with asymmetric memory sizes might achieve better sum rate performance than the network with symmetric memory sizes.

1. INTRODUCTION

The exponentially growing mobile data traffic is mainly due to video applications, e.g., content-based video streaming). Such video traffic has interesting features characterized by its asynchronous and skew nature. Namely, the user demands are highly asynchronous (since they request when and where they wish) and a few very popular files are requested over and over. The skewness of the video traffic together with the ever-growing cheap on-board storage memory suggests that the quality of experience can be boosted by caching popular contents at (or close to) end users in wireless networks. A number of recent works have studied such concept under different models and assumptions (see [1, 2, 3] and references therein). In most of these works, it is assumed that caching is performed in two phases: *placement phase* to prefetch users' caches under their memory constraints (typically during off-peak hours) prior to the actual demands; *delivery phase* to transmit codewords such that each user, based on the received signal and the contents of its cache, is able to decode the requested file. In this work, we focus on a coded caching model where a content-providing server is connected to many users, each equipped with a cache of finite memory [1]. By carefully choosing the sub-files to be distributed across users, coded caching exploits opportunistic multicasting such that a common signal is simultaneously useful for all users even with distinct file requests. A number of extensions of [1] have been developed including the case of decentralized placement phase [5], the case of non-uniform demands [4, 6, 9], the case of unequal file sizes [10], as well as the case of unequal memory sizes [15]. Although the potential merit of coded caching has been highlighted in these works, many of them have ignored the inherent features of wireless channels except a few recent works [7, 8].

In order to relax the unrealistic assumption of a perfect shared link, we model the bottleneck link as an erasure broadcast channel (EBC) to capture random failure or disconnection of any server-user link that a packet transmission may experience especially during high-traffic hours (delivery phase). The placement phase is performed either in a decentralized [5] or centralized manner [1] over the erasure-free shared link. We further assume that the channel at hand is memoryless and independently distributed across users and that the server acquires the channel states causally via feedback sent by users. Under this setting, we have characterized the achievable rate region of the cache-enabled EBC focusing on the special case with equal erasure probability, memory size in [16]. It is found that state feedback is useful to improve the performance of coded caching especially in the regime of a small memory size (with respect to the number of files) and with a large erasure probability. In this paper, we extend the results [16] to a general setup with asymmetric erasure probabilities and memory sizes. Our contribution is the characterization of the achievable rate region of the cache-enabled EBC under decentralized placement. We prove that a multi-phase delivery scheme extending the algorithm proposed by Wang and Gatzianas et al. [13, 14] to the case of receiver side information can achieve the rate region. A toy example shows that the asymmetry in user memory sizes is potentially beneficial to improve the sum rate performance. Finally, we note that a recent work in [15] has also studied the impact of asymmetric memory sizes in decentralized coded caching. Contrary to our observation, this work essentially shows that as the asymmetry in memory sizes grows, the gain of coded caching collapses. Such conclusion is somehow misleading because it builds on the assumptions of a perfect shared link as well as equal file sizes, which is seldom the case in practice.

The structure of the paper is as follows. Section 2 introduces first the system model and definitions, and then highlights the main results of this work. Section3 provides a sketch of achievability proof of the rate region of the cache-enabled EBC with feedback, while section 4 proves the corresponding converse part. Finally section 5 concludes the paper. Throughout the paper, we use the following notational conventions. The superscript notation X^n represents a sequence $(X_1, \dots, X_n)$ of variables. $X_{\mathcal{I}}$ is used to denote the set of variables $\{X_i\}_{i\in\mathcal{I}}$. The entropy of X is denoted by $H(X)$. We let $[k] = \{1, \dots, k\}$.

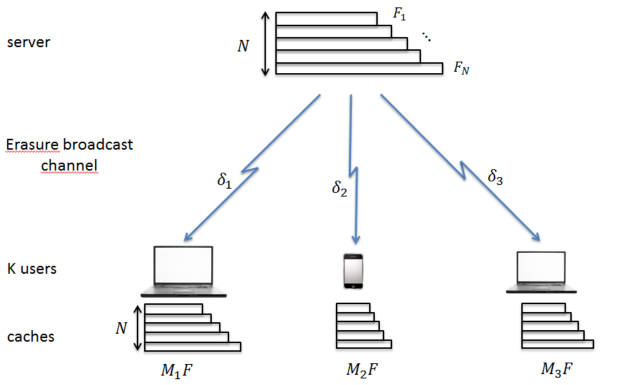

Figure 1: A cached-enabled erasure broadcast channel with $K = 3$.

Due to the space limitation, proofs for some lemmas are omitted and will be deferred to the full version of the paper.

2. SYSTEM MODEL AND MAIN RESULTS

2.1 System model and definitions

We consider a cache-enabled network depicted in Fig. 1 where a server is connected to K users through an erasure broadcast channel (EBC). The server has an access to N files $W_1, \dots, W_N$ where file i, i.e. W_i, consists of F_i packets of L bits each (F_iL bits). Each user k has a cache memory Z_k of M_kF packets for $M_k \in [0, N]$, where $F \triangleq \frac{1}{N}\sum_{i=1}^{N} F_i$ is the average size of the files. Under such a setting, consider a discrete time communication system where a packet is sent in each slot over the K-user EBC. The channel input $X_i \in \mathbb{F}_q$ belongs to the input alphabet of size $L = \log q$ bits. The channel is assumed to be memoryless and independently distributed across users so that in a given slot we have

$$\Pr(Y_1, Y_2, \dots, Y_K|X) = \prod_{k=1}^{K} \Pr(Y_k|X) \tag{1}$$

$$\Pr(Y_k|X) = \begin{cases} 1-\delta_k, & Y_k = X, \\ \delta_k, & Y_k = E \end{cases} \tag{2}$$

where Y_k denotes the channel output of receiver k, E stands for an erased output, δ_k denotes the erasure probability of user k. We let $S_i \in \mathcal{S} = 2^{\{1,\dots,K\}}$ denote the state of the channel in slot i which indicates the users who received correctly the packet. We assume that the transmitter obtains the state feedback S^{i-1} at the end of slot i while all the receivers know S^n at the end of the transmission.

The caching is performed in two phases: placement phase and delivery phase. In placement phase, the server fills the caches of all users $Z_1, \dots, Z_K$ up to the memory constraint. As in most works in the literature, we assume that the placement phase is done without error and neglect the cost, since it takes place usually during off-peak traffic hours. Once each user k makes a request d_k, the server sends codewords so that each user can decode its requested file as a function of its cache content and received signals during delivery phase. We provide a more formal definition below. A $(M_1, \dots, M_K, F_{d_1}, \dots, F_{d_K}, n)$ caching scheme consists of the following components.

- N message files $W_1, \dots, W_N$ are independently and uniformly distributed over $\mathcal{W}_1 \times \cdots \times \mathcal{W}_N$ with $\mathcal{W}_i = \mathbb{F}_q^{F_i}$ for all i.
- K caching functions are given by $\phi_k : \mathbb{F}_q^{\sum_{i=1}^N F_i} \to \mathbb{F}_q^{FM_k}$ map the files $W_1, \dots, W_N$ into the cache contents
$$Z_k \triangleq \phi_k(W_1, \dots, W_N) \tag{3}$$
for each user k.
- A sequence of encoding functions which transmit at slot i a symbol $X_i = f_i(W_{d_1}, \dots, W_{d_K}, S^{i-1}) \in \mathbb{F}_q$, based on the requested file set and the channel feedback up to slot $i-1$ for $i = 1, \dots, n$, where W_{d_k}, $d_k \in \{1, \dots, N\}$, denotes the message file requested by user k.
- A decoding function of user k is given by the mapping $\psi_k : \mathbb{F}_q^n \times \mathbb{F}_q^{FM_k} \times \mathcal{S}^n \to \mathbb{F}_q^{F_{d_k}}$ so that the decoded file is $\hat{W}_{d_k} = \psi_k(Y_k^n, Z_k, S^n)$ as a function of the received signals Y_k^n, the cache content Z_k, as well as the state information S^n.

A rate tuple $(R_1, \dots, R_K)$ is said to be achievable if, for every $\epsilon > 0$, there exists a $(M_1, \dots, M_K, F_{d_1}, \dots, F_{d_K}, n)$ caching strategy that satisfies the reliability condition

$$\max_{(d_1,\dots,d_K)\in\{1,\dots,N\}^K} \max_k \Pr(\psi_k(Y_k^n, Z_k, S^n) \neq W_{d_k}) < \epsilon$$

as well as rate condition

$$R_k < \frac{F_{d_k}}{n}. \tag{4}$$

Throughout the paper, we express the entropy and the rate in terms of packets in oder to avoid the constant factor $L = \log_2 q$. For simplicity reasons we denote for any parameter a_k for $k \in [K]$ and any subset $\mathcal{J} \subseteq [K]$, $\bar{a}_{\mathcal{J}} = \prod_{k\in\mathcal{J}}(1-a_k)$ and $a_{\mathcal{J}} = \prod_{k\in\mathcal{J}} a_k$

2.2 Decentralized cache placement

We mainly focus on decentralized cache placement proposed in [5] and adapt it to the packet-based broadcast channel (with no error). Under the memory constraint of M_kF packets, each user k independently caches a subset of p_kF packets of file i, chosen uniformly at random for $i = 1, \dots, N$, where $p_k = \frac{M_k}{N}$. By letting $\mathcal{L}_{\mathcal{K}}(W_i)$ denote the sub-file of W_i stored exclusively in the cache memories (known) of the users in $\mathcal{K}$, the cache memory Z_k of user k after decentralized placement is given by

$$Z_k = \{\mathcal{L}_{\mathcal{K}}(W_i) \ \ \forall k \subseteq \mathcal{K} \subseteq [K], \ \ \forall i = 1, \dots, N\} \tag{5}$$

which satisfies the memory constraint for user k

$$|Z_k| = p_k \sum_{i=1}^{N} F_i = Np_k \frac{\sum_{i=1}^N F_i}{N} = M_kF. \tag{6}$$

2.3 Main results

In order to present the main results, we specify two special cases.

Definition 1. *The cache-enabled EBC (or the network) is said symmetric if the erasure probabilities as well as the* *memory sizes are the same for all users, i.e.* $\delta_1 = \cdots = \delta_K$ *and* $p_1 = \cdots = p_K$.

DEFINITION 2. *The rate vector is said one-sided fair in the cache-enabled EBC if* $\delta_k \geq \delta_j$ *and for* $k \neq j$ *implies*

$$\begin{cases} \frac{\bar{p}_k}{p_k} R_k \geq \frac{\bar{p}_j}{p_j} R_j \\ \delta_k R_k \geq \delta_j R_j \end{cases} \tag{7}$$

For the special case where $p_k = 0 \; \forall k \in [K]$, it is reduced to $\delta_k R_k \geq \delta_j R_j$ which coincide with the definition of one-sided fair in [13, 14].

We focus on the case of most interest with $N \geq K$ and assume further that users' demands are all distinct.

THEOREM 1. *For* $K \leq 3$*, or for the symmetric network with* $K > 3$*, or for the one-sided fair rate with* $K > 3$*, the achievable rate region of the cached-enabled EBC with state feedback under decentralized cache placement is given by*

$$\sum_{k=1}^{K} \frac{\prod_{j=1}^{k}(1 - p_{\pi_j})}{1 - \prod_{j=1}^{k} \delta_{\pi_j}} R_{\pi_k} \leq 1 \tag{8}$$

for any permutation π *of* $\{1, \dots, K\}$.

The proof of Theorem 1 is provided in upcoming sections. The following corollary holds.

COROLLARY 1. *For* $K \leq 3$*, or for the symmetric network with* $K > 3$*, or for the one-sided fair rate with* $K > 3$*, the minimum number of transmissions to deliver a distinct requested file to each user in the cached-enabled EBC under decentralized cache placement is given by*

$$T_{\text{tot}} = \sum_{k=1}^{K} \frac{\Theta(F_{d_{\pi_k}}) \prod_{j=1}^{k}(1 - p_{\pi_j})}{1 - \prod_{j=1}^{k} \delta_{\pi_j}} \tag{9}$$

as $F_i \to \infty$ *for all* i*, for some permutation* π *determined by the parameters.*

The following remarks are in order. The results cover some special cases of interest. For the symmetric cache-enabled EBC ($p_k = p, \delta_k = \delta \;\; \forall k$), the above region simplifies to

$$\sum_{k=1}^{K} \frac{(1-p)^k}{1 - \delta^k} R_{\pi_k} \leq 1, \quad \forall \pi. \tag{10}$$

Exploiting a polyhedron structure, the achievability for a general K has been proved in [16]. For the case without cache memory ($p_k = 0 \;\; \forall k$), the region in Theorem 1 boils down to the rate region of the EBC with state feedback [13, 14] given by

$$\sum_{k=1}^{K} \frac{R_{\pi_k}}{1 - \prod_{j=1}^{k} \delta_j} \leq 1, \quad \forall \pi. \tag{11}$$

For the case of no erasure and equal file size ($\delta_k = 0, \forall k$, $F_i = F, \forall i$), the number of transmission in Corollary 1 scaled by F coincides with the rate-memory tradeoff under decentralized cache placement for asymmetric memory sizes [15] given by

$$\frac{T_{\text{tot}}}{F} = \sum_{k=1}^{K} \left[\prod_{j=1}^{k} \left(1 - \frac{M_j}{N}\right) \right]. \tag{12}$$

Further, if the memory size is equal for all users, i.e. $M_k = M$ for all k, the number of transmission (12) boils down into the rate-memory tradeoff under decentralized cache placement [5] given by

$$\frac{T_{\text{tot}}}{F} = \frac{N}{M}\left(1 - \frac{M}{N}\right)\left\{1 - \left(1 - \frac{M}{N}\right)^K\right\}. \tag{13}$$

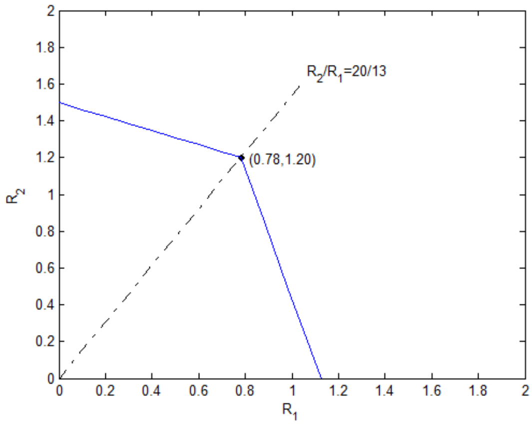

Figure 2: A two-user rate region for with $(p_1, p_2) = (\frac{1}{3}, \frac{2}{3})$, $(\delta_1, \delta_2) = (\frac{1}{4}, \frac{1}{2})$.

To illustrate the impact of the asymmetric memory sizes and erasure probabilities, let us consider a two-user example (case 1) with $(p_1, p_2) = (\frac{1}{3}, \frac{2}{3})$ and $(\delta_1, \delta_2) = (\frac{1}{4}, \frac{1}{2})$. Theorem 1 yields the rate region given by

$$\begin{aligned} \frac{8}{9}R_1 + \frac{16}{63}R_2 &\leq 1 \\ \frac{16}{63}R_1 + \frac{2}{3}R_2 &\leq 1. \end{aligned} \tag{14}$$

The region depicted in Fig. 2 is characterized by three vertices $(\frac{9}{8}, 0)$ $(0.78, 1.20)$, and $(0, \frac{63}{16})$. The vertex $(0.78, 1.20)$, achieving the sum rate of 1.98, corresponds to the case when the requested files satisfy the ratio $F_{d_2}/F_{d_1} = 20/13$. Let us consider another example where two users have both a memory size of $p_1 = p_2 = \frac{1}{2}$ while keeping the erasure probabilities as before. The maximum sum rate in the latter case is 1.87 smaller than the former case.

3. ACHIEVABILITY

We provide a sketch of achievability proof of Theorem 1 by focusing on the case of the one-sided fair rate vector. We first revisit the broadcasting scheme proposed in [13, 14] for the EBC with feedback, and then adapt it to the cached-enabled network.

3.1 Revisiting the broadcasting scheme [13, 14]

We provide a high-level description of the broadcasting scheme [13, 14] by assuming the number of private packets $\{F_k\}$ is arbitrarily large so that the length of each phase becomes deterministic. The broadcasting algorithm has two main roles: 1) broadcast new information packets and 2) multicast side information or overheard packets thanks to state feedback. From this reason, we can call phase 1 *broadcasting phase* and phases 2 to K *multicasting phase*. Phase j consists of $\binom{K}{j}$ sub-phases in each of which the transmitter sends packets intended to a subset of users $\mathcal{J}$ for $j = |\mathcal{J}|$. We let $\mathcal{L}_{\mathcal{J}}(V_{\mathcal{K}})$ denote the part of packet $V_{\mathcal{K}}$ received by users in $\mathcal{J}$ and erased by $[K] \setminus \mathcal{J}$. Here is a high-level description of the broadcasting algorithm:

1. Broadcasting phase (phase 1): send each message $V_k = W_k$ of F_k packets sequentially for $k = 1, \dots, K$. This phase generates overheard symbols $\{\mathcal{L}_{\mathcal{J}}(V_k)\}$ to be transmitted via linear combination in multicasting phase, where $\mathcal{J} \subseteq [K] \setminus k$ for all k.

2. Multicasting phase (phases $2 - K$): for a subset $\mathcal{J}$ of users, generate $V_{\mathcal{J}}$ as a linear combination of overheard packets such that
$$V_{\mathcal{J}} = \mathcal{F}_{\mathcal{J}}\left(\{\mathcal{L}_{\mathcal{J}\setminus\mathcal{I}\cup\mathcal{I}'}(V_{\mathcal{I}})\}_{\mathcal{I}'\subset\mathcal{I}\subset\mathcal{J}}\right) \tag{15}$$
where $\mathcal{F}_{\mathcal{J}}$ denotes a linear function. Send $V_{\mathcal{J}}$ sequentially for all $\mathcal{J} \subseteq [K]$ of the cardinality $|\mathcal{J}| = 2, \dots, K$.

The achievability result of [13, 14] implies the following lemma.

LEMMA 1. *For $K \le 3$, or for the symmetric channel with $K > 3$, or for the one-sided fair rate with $K > 3$, there exist linear functions $\{\mathcal{F}_{\mathcal{J}}\}$ as well as π, such that F_k independent linear functions of W_k can be obtained from $\{\mathcal{L}_{\mathcal{K}}(V_{\mathcal{J}})\}$ for $k \in \mathcal{K}$ and $\mathcal{J} \subseteq [K]$, for the total duration of*
$$T_{\rm tot} = \sum_{k=1}^{K} \frac{\Theta(F_{\pi_k})}{1 - \prod_{j=1}^{k} \delta_{\pi_j}}.$$

The proof is omitted because the proof in section 3.2 covers the case without user memories. In order to determine the total transmission duration, we need to introduce further some notions and parameters.

- A packet intended to $\mathcal{J}$ is consumed *for a given user* $k \in \mathcal{J}$ if this user or at least one user in $[K] \setminus \mathcal{J}$ receives it. The probability of such event is equal to $1 - \prod_{j\in[K]\setminus\mathcal{J}\cup\{k\}} \delta_j$.
- A packet intended to $\mathcal{I}$ creates a packet intended to users in $\mathcal{J}$ for user $k \in \mathcal{I} \subset \mathcal{J} \subseteq [K]$ if erased by user k and all users in $[K] \setminus \mathcal{J}$ but received by $\mathcal{J} \setminus \mathcal{I}$. The probability of such event is denoted by $\alpha^{\{k\}}_{\mathcal{I}\to\mathcal{J}} = \prod_{j'\in[K]\setminus\mathcal{J}\cup\{k\}} \delta_{j'} \prod_{j\in\mathcal{J}\setminus\mathcal{I}}(1-\delta_j)$. We let
$$N^{\{k\}}_{\mathcal{I}\to\mathcal{J}} = t^{\{k\}}_{\mathcal{J}} \alpha^{\{k\}}_{\mathcal{I}\to\mathcal{J}} \tag{16}$$
denote the number of such packets, where $t^{\{k\}}_{\mathcal{J}}$ will be defined shortly. We can also express $N^{\{k\}}_{\mathcal{I}\to\mathcal{J}}$ as
$$N^{\{k\}}_{\mathcal{I}\to\mathcal{J}} = \sum_{\mathcal{I}'\subseteq\mathcal{I}\setminus k} |\mathcal{L}_{\mathcal{J}\setminus\mathcal{I}\cup\mathcal{I}'}(V^{\{k\}}_{\mathcal{I}})| \tag{17}$$

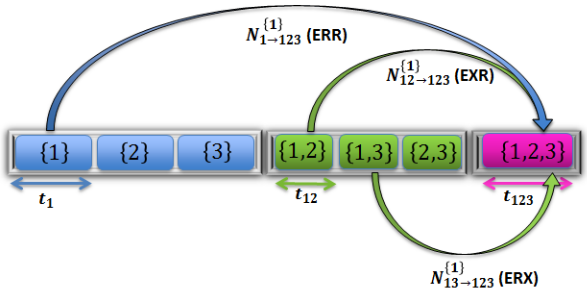

Figure 3: Phase organization for $K = 3$ and packet evolution viewed by user 1.

where we let $V^{\{k\}}_{\mathcal{I}}$ denotes the part of $V_{\mathcal{I}}$ required for user k.

- The duration $t_{\mathcal{J}}$ of sub-phase $\mathcal{J}$ is given by $t_{\mathcal{J}} = \max_{k\in\mathcal{J}} t^{\{k\}}_{\mathcal{J}}$ where
$$t^{\{k\}}_{\mathcal{J}} = \frac{\sum_{k\in\mathcal{I}\subset\mathcal{J}} N^{\{k\}}_{\mathcal{I}\to\mathcal{J}}}{1 - \prod_{j\in[K]\setminus\mathcal{J}\cup\{k\}} \delta_j} \tag{18}$$

The total duration is given by summing up all sub-phases, i.e. $T_{\rm tot} = \sum_{\mathcal{J}\subseteq[K]} t_{\mathcal{J}}$.

Fig. 3 illustrates the phase organization for $K = 3$ and the packet evolution viewed by user 1. The packets intended to $\{1, 2, 3\}$ are created both from phases 1 and 2. More precisely, sub-phase $\{1\}$ creates $\mathcal{L}_{23}(V_1)$ to be sent in phase 3 if erased by user 1 and received by others (ERR). The number of such packets is $N^{\{1\}}_{1\to123}$. Sub-phase $\{1, 2\}$ creates $\mathcal{L}_3(V_{12}), \mathcal{L}_{23}(V_{12})$ if erased by user 1 but received by user 3 (EXR), while sub-phase $\{1, 3\}$ creates $\mathcal{L}_2(V_{13}), \mathcal{L}_{23}(V_{13})$ if erased by user 1 and received by user 2 (ERX). The total number of packets intended to $\{1, 2, 3\}$ generated in phase 2 and required by user 1 is $N^{\{1\}}_{12\to123} + N^{\{1\}}_{13\to123}$.

3.2 Proposed delivery scheme

We describe the proposed delivery scheme by focusing on the one-sided fair rate vector. Namely, we assume without loss of generality $\delta_1 \ge \cdots \ge \delta_K$, $\delta_1 R_1 \ge \cdots \ge \delta_K R_K$, and $\frac{1-p_1}{p_1} R_1 \ge \cdots \ge \frac{1-p_2}{p_2} R_K$.

We assume also that user k requests file W_k of size F_k packets for $k = 1, \dots, K$. In the presence of users' caches, the packets to transmit in a given phase are composed by the packets created in the placement phase and previous phases of the algorithm.

Placement phase The placement phase creates packets $\{\mathcal{L}_{\mathcal{J}}(W_k)\}$ for $\mathcal{J} \subset [K]$ and for $k = 1, .., K$. The size of each sub-file given by
$$|\mathcal{L}_{\mathcal{J}}(W_k)| = \prod_{j\in\mathcal{J}} p_j \prod_{l\in[K]\setminus\mathcal{J}} (1 - p_l) F_k \tag{19}$$
Obviously, the sub-file $\mathcal{L}_{\mathcal{J}}(W_k)$ for $k \in \mathcal{J}$ are received by the destination and shall not be transmitted in delivery phase.

Phase 1 The transmitter sends $V_1, .., V_K$ sequentially until at least one user receives it, where $V_k = \mathcal{L}_{\emptyset}(W_k)$ corresponds to the order-1 packets created by placement phase

The length of sub-phase $\{k\}$ is given by

$$t_k = \frac{|V_k|}{1-\prod_{j\in[K]}\delta_j} = \frac{\prod_{j\in[K]}(1-p_j)F_k}{1-\prod_{j\in[K]}\delta_j}. \quad (20)$$

Phase 1 creates packets to be sent in phases 2 and 3. The sub-phase $\{k\}$ creates packets $\mathcal{L}_{\mathcal{J}}(V_k)$ to be sent in a sub-phase $\mathcal{J}\cup\{k\}$, where $k\notin\mathcal{J}$, whose number is given by

$$N^{\{k\}}_{k\to\mathcal{J}\cup\{k\}} = |\mathcal{L}_{\mathcal{J}}(V_k)| = t_k \prod_{k\in\mathcal{J}}(1-\delta_k) \prod_{k\in[K]\setminus\mathcal{J}}\delta_k \quad (21)$$

for $\mathcal{J}\subseteq[K]\setminus\{k\}$.

Phases $2\dots K$ For a subset $\mathcal{J}$ of users, generate $V_{\mathcal{J}}$ as a linear combination of overheard packets during the placement phase as well as during phases 1 to $j-1$.

$$V_{\mathcal{J}} = \mathcal{F}_{\mathcal{J}}\left(\{\mathcal{L}_{\mathcal{J}\setminus\mathcal{I}\cup\mathcal{I}'}(V_{\mathcal{I}})\}_{\mathcal{I}'\subset\mathcal{I}\subset\mathcal{J}}, \mathcal{L}_{\mathcal{J}\setminus\{k\}}(W_k)\right) \quad (22)$$

The duration $t_{\mathcal{J}}$ of sub-phase $\mathcal{J}$ is given by $t_{\mathcal{J}} = \max_{k\in\mathcal{J}} t^{\{k\}}_{\mathcal{J}}$ where

$$t^{\{k\}}_{\mathcal{J}} = \frac{\sum_{k\in\mathcal{I}\subset\mathcal{J}} N^{\{k\}}_{\mathcal{I}\to\mathcal{J}} + |\mathcal{L}_{\mathcal{J}\setminus\{k\}}(W_k)|}{1-\prod_{j\in[K]\setminus\mathcal{J}\cup\{k\}}\delta_j} \quad (23)$$

Following similar steps as [14, Appendix C], it is possible to prove that

$$\sum_{\mathcal{I}\subseteq\mathcal{J}} t^{\{k\}}_{\mathcal{I}} = \frac{\prod_{j\in[K]\setminus\mathcal{J}\cup\{k\}}(1-p_j)}{1-\prod_{j\in[K]\setminus\mathcal{J}\cup\{k\}}\delta_j}F_k \quad (24)$$

As a result we prove in Appendix A that it holds

$$t^{\{k\}}_{\mathcal{J}} = \sum_{\mathcal{H}\subseteq\mathcal{J}\setminus\{k\}} (-1)^{|\mathcal{H}|} \frac{\prod_{j\in[K]\setminus\mathcal{J}\cup\{k\}\cup\mathcal{H}}(1-p_j)}{1-\prod_{j\in[K]\setminus\mathcal{J}\cup\{k\}\cup\mathcal{H}}\delta_j}F_k \quad (25)$$

The length of a sub-phase $\mathcal{J}$ is given by the worst user among the subset $\mathcal{J}$. For the case of one-sided fair rate vectors, it is possible to prove that there exists $\tilde{\pi}$ satisfying

$$\arg\max_{k\in\mathcal{J}} t^{\{k\}}_{\mathcal{J}} = \min_{k\in\mathcal{J}}(\tilde{\pi}^{-1}(k)) \quad \forall\mathcal{J}\subseteq[K]. \quad (26)$$

The proof is provided in Appendix B. Moreover, such permutation corresponds to the identity permutation. This means the user permutation (which determines the sub-phase length) is preserved in all sub-phases for the one-sided fair rate vector of Definition 2. The worst user, user 1 with the largest erasure probability, yields the maximum $t^1_{\mathcal{J}}$ over all subsets $\mathcal{J}$ including 1. Proved in Appendix D, the total duration can be derived as follows

$$\begin{aligned} T_{\rm tot} &= \sum_{J\subseteq[K]} \max_{k\in J} t^{\{k\}}_{\mathcal{J}} = \sum_{J\subseteq[K]} t^{\{k^\star\}}_{\mathcal{J}} &(27)\\ &= \sum_{k=1}^{K} \frac{F_{d_k}\prod_{j=1}^{k}(1-p_j)}{1-\prod_{j=1}^{k}\delta_j} &(28)\end{aligned}$$

where we let $k^\star$ the smallest index in $\mathcal{J}$. Dividing both sides by $T_{\rm tot}$ and letting $R_k = \frac{F_{d_k}}{T_{\rm tot}}$, we readily obtain

$$\sum_{k=1}^{K} \frac{\prod_{j=1}^{k}(1-p_j)}{1-\prod_{j=1}^{k}\delta_j} R_k = 1 \quad (29)$$

Since under the one-sided fair rate constraint of Definition 2 the above inequality implies all the other $K!-1$ inequalities of the rate region as proved in Appendix C, this establishes the achieability.

4. OPTIMALITY OF DELIVERY PHASE

In this section, we provide the converse part of Theorem 1.First we provide two useful lemmas.

LEMMA 2. *[11, Lemma 5] For the erasure broadcast channel, if U is such that $X_i \leftrightarrow UY^{i-1}_{\mathcal{I}}S^{i-1} \leftrightarrow (S_{i+1},\dots,S_n)$, $\forall\mathcal{I}$,*

$$\frac{1}{1-\prod_{j\in\mathcal{I}}\delta_j} H(Y^n_{\mathcal{I}}\,|\,U,S^n) \le \frac{1}{1-\prod_{j\in\mathcal{J}}\delta_j} H(Y^n_{\mathcal{J}}\,|\,U,S^n), \quad (30)$$

for any sets $\mathcal{I},\mathcal{J}$ such that $\mathcal{J}\subseteq\mathcal{I}\subseteq\{1,\dots,K\}$.

LEMMA 3. *Under decentralized cache placement [5], the following equality holds for any i and $\mathcal{K}\subseteq[K]$*

$$H(W_i\,|\,\{Z_k\}_{k\in\mathcal{K}}) = \prod_{k\in\mathcal{K}}(1-p_k)\,H(W_i).$$

PROOF.

$$\begin{aligned} &H(W_i\,|\,\{Z_k\}_{k\in\mathcal{K}})\\ &\quad= H(W_i\,|\,\{\mathcal{L}_{\mathcal{J}}(W_l)\}_{\mathcal{J}\cap\mathcal{K}\neq\emptyset,\, l=1,\dots,N}) &(31)\\ &\quad= H(W_i\,|\,\{\mathcal{L}_{\mathcal{J}}(W_l)\}_{\mathcal{J}\cap\mathcal{K}\neq\emptyset}) &(32)\\ &\quad= H(\{\mathcal{L}_{\mathcal{J}}(W_i)\}_{\mathcal{J}\cap\mathcal{K}=\emptyset}) &(33)\\ &\quad= \sum_{\mathcal{J}\subseteq[K]\setminus\mathcal{K}} H(\mathcal{L}_{\mathcal{J}}(W_i)) &(34)\\ &\quad= \sum_{\mathcal{J}\subseteq[K]\setminus\mathcal{K}} \prod_{j\in\mathcal{J}} p_j \prod_{k\in[K]\setminus\mathcal{J}}(1-p_k)H(W_i) &(35)\\ &\quad= \prod_{k\in\mathcal{K}}(1-p_k) \sum_{\mathcal{J}\subseteq[K]\setminus\mathcal{K}} \prod_{j\in\mathcal{J}} p_j \prod_{k\in[K]\setminus\mathcal{K}\setminus\mathcal{J}}(1-p_k)H(W_i) &(36)\\ &\quad= \prod_{k\in\mathcal{K}}(1-p_k)\,H(W_i) &(37)\end{aligned}$$

where the first equality follows from (5); the second equality follows due to the independence between message files; the third equality follows by identifying the unknown parts of W_i given the cache memories of $\mathcal{K}$ and using the independence of all sub-files; (34) is again from the independence of the sub-files; (35) is from the law of large number similarly as in (6); finally, the last equality is obtained from the following basic property that we have $\sum_{\mathcal{J}\subseteq\mathcal{L}}\prod_{j\in\mathcal{J}}p_j\prod_{k\in\mathcal{L}\setminus\mathcal{J}}(1-p_k)=1$ for a subset $\mathcal{L}=[K]\setminus\mathcal{K}$. □

We apply genie-aided bounds to create a degraded erasure broadcast channel by providing the messages, the channel outputs, as well as the receiver side information (contents of cache memories) to enhanced receivers. Without loss of generality, we focus on the case without permutation and the demand $(d_1, \ldots, d_K) = (1, \ldots, K)$. We have for user k,

$$n \prod_{j=1}^{k}(1-p_j)R_k = \prod_{j=1}^{k}(1-p_j)H(W_k) \quad (38)$$

$$= H(W_k|Z^k S^n) \quad (39)$$

$$\leq I(W_k; Y_{[k]}^n \,|\, Z^k S^n) + n\epsilon'_{n,k} \quad (40)$$

$$\leq I(W_k; Y_{[k]}^n, W^{k-1} \,|\, Z^k S^n) + n\epsilon'_{n,k} \quad (41)$$

$$= I(W_k; Y_{[k]}^n \,|\, W^{k-1} Z^k S^n) + n\epsilon'_{n,k} \quad (42)$$

where the second equality is by applying Lemma 3 and noting that S^n is independent of others, (40) is from the Fano's inequality; the last equality is from $I(W_k; W^{k-1} \,|\, Z^k S^n) = 0$. Putting all the rate constraints together, and letting $\epsilon_{n,k} \triangleq \epsilon'_{n,k} / \prod_{j=1}^{k}(1-p_j)$,

$$n(1-p_1)(R_1 - \epsilon_{n,1}) \leq H(Y_1^n \,|\, Z_1 S^n) - H(Y_1^n \,|\, W_1 Z_1 S^n)$$

$$\vdots$$

$$n \prod_{j=1}^{K}(1-p_j)(R_K - \epsilon_{n,K}) \leq H(Y_{[K]}^n \,|\, W^{K-1} Z^K S^n) - H(Y_{[K]}^n \,|\, W^K Z^K S^n) \quad (43)$$

We now sum up the above inequalities with different weights, and applying Lemma 2 for $K-1$ times, namely, for $k = 1, \ldots, K-1$,

$$\frac{H(Y_{[k+1]}^n \,|\, W^k Z^{k+1} S^n)}{1 - \prod_{j \in [k+1]} \delta_j} \leq \frac{H(Y_{[k+1]}^n \,|\, W^k Z^k S^n)}{1 - \prod_{j \in [k+1]} \delta_j} \quad (44)$$

$$\leq \frac{H(Y_{[k]}^n \,|\, W^k Z^k S^n)}{1 - \prod_{j \in [k]} \delta_j} \quad (45)$$

where the first inequality follows because removing conditioning increases the entropy. Finally, we have

$$\sum_{k=1}^{K} \frac{\prod_{j \in [k]}(1-p_j)}{1 - \prod_{j \in [k]} \delta_j}(R_k - \epsilon_n)$$

$$\leq \frac{H(Y_1^n \,|\, Z_1 S^n)}{n(1-\delta_1)} - \frac{H(Y_{[K]}^n \,|\, W^K Z^K S^n)}{n(1 - \prod_{j \in [k]} \delta_j)} \quad (46)$$

$$\leq \frac{H(Y_1^n)}{n(1-\delta_1)} \leq 1 \quad (47)$$

which establishes the converse proof.

5. CONCLUSION

In this paper, we studied the decentralized coded caching in the erasure broadcast channel with state feedback by introducing the asymmetry in the user memory sizes as well as the channel statistic. Our contribution is the characterization of the achievable rate region of the channel under a general setting as a non-trivial extension of the work [16] focusing on the case of the equal memory size and erasure probability. A toy example with two users shows that the asymmetry in memory sizes potentially improves the sum rate performance. The detailed analysis as well as numerical examples for a large network dimension remain as future works.

6. REFERENCES

[1] M. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2856–2867, 2014

[2] N. Golrezaei, K. Shanmugam, A. G. Dimakis, A. F. Molisch, G. Caire, "FemtoCaching: Wireless Video Content Delivery through Distributed Caching Helpers" , *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8402–8413, 2013.

[3] M. Ji, G. Caire, A. Molisch, "Fundamental Limits of Distributed Caching in D2D Wireless Networks" , arXiv/1304.5856, 2013

[4] M. Ji, A. Tulino, J. Llorca, and G. Caire, "Order-optimal rate of caching and coded multicasting with random demands", arXiv:1502.03124, 2015

[5] M. Maddah-Ali and U. Niesen, "Decentralized coded caching attains order-optimal memory-rate tradeoff", *IEEE/ACM Trans. on Networking*, vol. 23, no. 4, pp. 1029–1040, 2015. http://arxiv.org/abs/1301.5848, 2013

[6] M. Maddah-Ali and U. Niesen, "Coded caching with nonuniform demands", http://arxiv.org/abs/1308.0178, 2013

[7] W. Huang, S.Wang, L. Ding, F. Yang, and W. Zhang, " The Performance Analysis of Coded Cache in Wireless Fading Channel", arXiv:1504.01452, 2015

[8] R. Timo and M. Wigger, "Joint Cache-Channel Coding over Erasure Broadcast Channels", arXiv:1505.01016, 2015

[9] J. Hachem, N. Karamchandani, and S. Diggavi, "Effect of Number of Users in Multi-level Coded Caching", in *Proceedings of the IEEE International Symposium on Information Theory (ISIT'2015)*, Hong-Kong, China, 2015

[10] J. Zhang, X. Lin, C. C. Wang, and X. Wang, "Coded Caching for Files with Distinct File Sizes", in *Proceedings of the IEEE International Symposium on Information Theory (ISIT'2015)*, Hong-Kong, China, 2015

[11] S. Yang and M. Kobayashi, "Secrecy Communications in K-user Multi-Antenna Broadcast Channel with State Feedback", in *Proceedings of the IEEE International Symposium on Information Theory (ISIT'2015)*, Hong-Kong, China, 2015

[12] M. A. Maddah-Ali and D. N. C. Tse, "Completely Stale Transmitter Channel State Information is Still Very Useful," *IEEE Trans. Inf. Theory* vol. 58, no. 7, pp. 4418–4431, July 2012.

[13] C. C. Wang, "On the Capacity of 1-to-Broadcast Packet Erasure Channels with Channel Output feedback", *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 931–956, February 2012.

[14] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Multiuser Broadcast Erasure Channel With Feedback-Capacity and Algorithms", *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5779–5804, September 2013.

[15] Wang, Sinong and Li, Wenxin and Tian, Xiaohua and Liu, Hui, "Fundamental limits of heterogenous cache", arXiv preprint arXiv:1504.01123, 2015

[16] A. Ghorbel, M. Kobayashi, and S. Yang, "Cache-Enabled Broadcast Packet Erasure Channels with State Feedback", to appear in the 53rd Annual

Allerton Conference on Communication, Control, and Computing (Allerton), 2015

APPENDIX

A. LENGTH OF SUB-PHASE

we first introduce a new variable $g_{\mathcal{J}}^{\{k\}} = \frac{t_{\mathcal{J}}^{\{k\}}}{F_k}$ for $k \in \mathcal{J} \subseteq [K]$. Using (24) we obtain

$$\sum_{\mathcal{I} \subseteq \mathcal{J}} g_{\mathcal{I}}^{\{k\}} = w_{j \in [K] \setminus \mathcal{J} \cup \{k\}} \tag{48}$$

where $w_{\mathcal{J}} = \frac{\prod_{j \in \mathcal{J}} (1-p_j)}{1 - \prod_{j \in \mathcal{J}} \delta_j}$. We first need to prove the following lemma

LEMMA 4. *For any nonempty set $[K]$ and $\mathcal{J} \subseteq [K]$. It holds*

$$\sum_{\mathcal{I} \subseteq \mathcal{J}} \sum_{\mathcal{H} \subseteq \mathcal{I}} (-1)^{|\mathcal{H}|} w_{[K] \setminus \mathcal{I} \cup \mathcal{H}} = w_{[K] \setminus \mathcal{J}} \tag{49}$$

PROOF.

$$\sum_{\mathcal{I} \subseteq \mathcal{J}} \sum_{\mathcal{H} \subseteq \mathcal{I}} (-1)^{|\mathcal{H}|} w_{[K] \setminus \mathcal{I} \cup \mathcal{H}} = \sum_{\mathcal{I} \subseteq \mathcal{J}} \sum_{\mathcal{H} \subseteq \mathcal{I}} (-1)^{|\mathcal{H}|} w_{[K] \setminus (\mathcal{I} \setminus \mathcal{H})} \tag{50}$$

$$= \sum_{\mathcal{I} \subseteq \mathcal{J}} \sum_{\mathcal{H}' \subseteq \mathcal{I}} (-1)^{|\mathcal{I} \setminus \mathcal{H}'|} w_{[K] \setminus \mathcal{H}'} \tag{51}$$

$$= \sum_{\mathcal{H}' \subseteq \mathcal{J}} \sum_{\mathcal{H}' \subseteq \mathcal{I} \subseteq \mathcal{J}} (-1)^{|\mathcal{I} \setminus \mathcal{H}'|} w_{[K] \setminus \mathcal{H}'} \tag{52}$$

$$= \sum_{\mathcal{H}' \subseteq \mathcal{J}} w_{[K] \setminus \mathcal{H}'} \sum_{\mathcal{H}' \subseteq \mathcal{I} \subseteq \mathcal{J}} (-1)^{|\mathcal{I} \setminus \mathcal{H}'|} \tag{53}$$

$$= \sum_{\mathcal{H}' \subseteq \mathcal{J}} w_{[K] \setminus \mathcal{H}'} \sum_{\mathcal{I}' \subseteq \mathcal{J} \setminus \mathcal{H}'} (-1)^{|\mathcal{I}'|} \tag{54}$$

$$= w_{[K] \setminus \mathcal{J}} \tag{55}$$

The last equality came from the fact that $\sum_{\mathcal{I} \subseteq \mathcal{J}} (-1)^{|\mathcal{I}|} = 0$ for all $\mathcal{J} \neq \emptyset$. $\square$

We prove (25) by induction on $|\mathcal{J}|$. for $\mathcal{J} = \{i\}$ we obtain $\sum_{i \in \mathcal{I} \subseteq \mathcal{J}} g_{\mathcal{I}}^{\{i\}} = g_{\mathcal{J}}^{\{i\}} = w_{[K] \setminus \mathcal{J} \cup \{i\}}$ and $\sum_{\mathcal{H} \subseteq \mathcal{J} \setminus \{i\}} (-1)^{|\mathcal{H}|} w_{[K] \setminus \mathcal{J} \cup \{i\} \cup \mathcal{H}} = w_{[K] \setminus \mathcal{J} \cup \{i\}}$. Thus for $|\mathcal{J}| = 1$

$$g_{\mathcal{J}}^{\{i\}} = \sum_{\mathcal{H} \subseteq \mathcal{J} \setminus \{i\}} (-1)^{|\mathcal{H}|} w_{[K] \setminus \mathcal{J} \cup \{i\} \cup \mathcal{H}} \tag{56}$$

Now suppose (56) holds for any $\mathcal{I} \subseteq [K]$ such that $|\mathcal{I}| < |\mathcal{J}|$ and we prove in the following that it holds for $\mathcal{J}$ too. We have

$$\sum_{i \in \mathcal{I} \subseteq \mathcal{J}} g_{\mathcal{I}}^{\{i\}} = w_{[K] \setminus \mathcal{J} \cup \{i\}} \tag{57}$$

$$= g_{\mathcal{J}}^{\{i\}} + \sum_{i \in \mathcal{I} \subset \mathcal{J}} g_{\mathcal{I}}^{\{i\}} \tag{58}$$

$$\tag{59}$$

thus, using Lemma 4 we obtain

$$g_{\mathcal{J}}^{\{i\}} = w_{[K] \setminus \mathcal{J} \cup \{i\}} - \sum_{i \in \mathcal{I} \subset \mathcal{J}} g_{\mathcal{I}}^{\{i\}} \tag{60}$$

$$= w_{[K] \setminus \mathcal{J} \cup \{i\}} - \sum_{i \in \mathcal{I} \subset \mathcal{J}} \sum_{\mathcal{H} \subseteq \mathcal{I} \setminus \{i\}} (-1)^{|\mathcal{H}|} w_{[K] \setminus \mathcal{I} \cup \{i\} \cup \mathcal{H}} \tag{61}$$

$$= w_{[K] \setminus \mathcal{J} \cup \{i\}} - \sum_{i \in \mathcal{I} \subseteq \mathcal{J}} \sum_{\mathcal{H} \subseteq \mathcal{I} \setminus \{i\}} (-1)^{|\mathcal{H}|} w_{[K] \setminus \mathcal{I} \cup \{i\} \cup \mathcal{H}} \tag{62}$$

$$+ \sum_{\mathcal{H} \subseteq \mathcal{J} \setminus \{i\}} (-1)^{|\mathcal{H}|} w_{[K] \setminus \mathcal{J} \cup \{i\} \cup \mathcal{H}} \tag{63}$$

$$= w_{[K] \setminus \mathcal{J} \cup \{i\}} - \sum_{\mathcal{I} \subseteq \mathcal{J} \setminus \{i\}} \sum_{\mathcal{H} \subseteq \mathcal{I}} (-1)^{|\mathcal{H}|} w_{[K] \setminus \mathcal{I} \cup \mathcal{H}} \tag{64}$$

$$+ \sum_{\mathcal{H} \subseteq \mathcal{J} \setminus \{i\}} (-1)^{|\mathcal{H}|} w_{[K] \setminus \mathcal{J} \cup \{i\} \cup \mathcal{H}} \tag{65}$$

$$= w_{[K] \setminus \mathcal{J} \cup \{i\}} - w_{[K] \setminus (\mathcal{J} \setminus \{i\})} + \sum_{\mathcal{H} \subseteq \mathcal{J} \setminus \{i\}} (-1)^{|\mathcal{H}|} w_{[K] \setminus \mathcal{J} \cup \{i\} \cup \mathcal{H}} \tag{66}$$

$$= \sum_{\mathcal{H} \subseteq \mathcal{J} \setminus \{i\}} (-1)^{|\mathcal{H}|} w_{[K] \setminus \mathcal{J} \cup \{i\} \cup \mathcal{H}} \tag{67}$$

B. EXISTENCE OF THE PERMUTATION

$\forall \quad \mathcal{J} \subseteq [K]$ s.t $|\mathcal{J}| \geq 2$ we set $m = \min(\mathcal{J})$. $\forall \quad i \in \mathcal{J}$ we have $\delta_m \geq \delta_i$; $\delta_m R_m \geq \delta_i R_i$ and $\frac{\bar{p}_m}{p_m} Rm \geq \frac{\bar{p}_i}{p_i} Ri$.

From (17), (19) and (23) we obtain

$$g_{\mathcal{J}}^{\{i\}} = \frac{1}{1 - \delta_{[K] \setminus \mathcal{J} \cup \{i\}}} \left[\sum_{i \in \mathcal{I} \subset \mathcal{J}} g_{\mathcal{I}}^{\{i\}} \bar{\delta}_{\mathcal{J} \setminus \mathcal{I}} \delta_{[K] \setminus \mathcal{J} \cup \{i\}} + p_{\mathcal{J} \setminus \{i\}} \bar{p}_{[K] \setminus \mathcal{J} \cup \{i\}} \right] \tag{68}$$

and

$$g_{\mathcal{J}}^{\{m\}} = \frac{1}{1 - \delta_{[K] \setminus \mathcal{J} \cup \{m\}}} \left[\sum_{m \in \mathcal{I} \subset \mathcal{J}} g_{\mathcal{I}}^{\{m\}} \bar{\delta}_{\mathcal{J} \setminus \mathcal{I}} \delta_{[K] \setminus \mathcal{J} \cup \{m\}} + p_{\mathcal{J} \setminus \{m\}} \bar{p}_{[K] \setminus \mathcal{J}} \right. \tag{69}$$

We prove by induction on $|\mathcal{J}|$ that $R_m g_{\mathcal{J}}^{\{m\}} \geq R_i g_{\mathcal{J}}^{\{i\}}$: For $|\mathcal{J}| = 2$, $\mathcal{J} = \{m, i\}$ so

$$g_{\mathcal{J}}^{\{i\}} = \frac{1}{1 - \delta_{[K] \setminus \mathcal{J} \cup \{i\}}} \left[g_i^{\{i\}} \bar{\delta}_m \delta_{[K] \setminus \mathcal{J} \cup \{i\}} + p_m \bar{p}_{[K] \setminus \mathcal{J} \cup \{i\}} \right] \tag{70}$$

and

$$g_{\mathcal{J}}^{\{m\}} = \frac{1}{1 - \delta_{[K] \setminus \mathcal{J} \cup \{m\}}} \left[g_m^{\{m\}} \bar{\delta}_i \delta_{[K] \setminus \mathcal{J} \cup \{m\}} + p_i \bar{p}_{[K] \setminus \mathcal{J} \cup \{m\}} \right] \tag{71}$$

Since $\delta_m \geq \delta_i$, it holds $\frac{1}{1 - \delta_{[K] \setminus \mathcal{J} \cup \{m\}}} \geq \frac{1}{1 - \delta_{[K] \setminus \mathcal{J} \cup \{i\}}}$ and $\bar{\delta}_i \geq \bar{\delta}_m$. Since $\frac{\bar{p}_m}{p_m} Rm \geq \frac{\bar{p}_i}{p_i} Ri$, it holds $p_i \bar{p}_{[K] \setminus \mathcal{J} \cup \{m\}} R_m \geq p_m \bar{p}_{[K] \setminus \mathcal{J} \cup \{i\}} R_i$. In addition we have from (25) : $g_m^{\{m\}} =$

$g_i^{\{i\}} = \frac{\bar{p}_{[K]}}{1-\delta_{[K]}}$ and $\delta_m R_m \geq \delta_i R_i$, thus we obtain $R_m g_{\mathcal{J}}^{\{m\}} \geq R_i g_{\mathcal{J}}^{\{i\}}$.

Suppose that it holds for any $\mathcal{I} \subseteq [K]$ s.t. $|\mathcal{I}| < |\mathcal{J}|$ and $m = min(\mathcal{I})$: $R_m g_{\mathcal{I}}^{\{m\}} \geq R_i g_{\mathcal{I}}^{\{i\}}$, and we prove it in the following that it holds also for $\mathcal{J}$.

Since $\delta_m \geq \delta_i$, it holds $\frac{1}{1-\delta_{[K]\setminus\mathcal{J}\cup\{m\}}} \geq \frac{1}{1-\delta_{[K]\setminus\mathcal{J}\cup\{i\}}}$. Since $\frac{\bar{p}_m}{p_m} Rm \geq \frac{\bar{p}_i}{p_i} Ri$, it holds $p_{\mathcal{J}\setminus\{m\}} \bar{p}_{[K]\setminus\mathcal{J}\cup\{m\}} R_m \geq p_{\mathcal{J}\setminus\{i\}} \bar{p}_{[K]\setminus\mathcal{J}\cup\{i\}} R_i$. It is left to prove that

$$R_m \sum_{m\in\mathcal{I}\subset\mathcal{J}} g_{\mathcal{I}}^{\{m\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}} \delta_{[K]\setminus\mathcal{J}\cup\{m\}} \geq R_i \sum_{i\in\mathcal{I}\subset\mathcal{J}} g_{\mathcal{I}}^{\{i\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}} \delta_{[K]\setminus\mathcal{J}\cup\{i\}} \tag{72}$$

We have

$$\begin{aligned}
&\sum_{m\in\mathcal{I}\subset\mathcal{J}} g_{\mathcal{I}}^{\{m\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}} \delta_{[K]\setminus\mathcal{J}\cup\{m\}} \\
&= \sum_{\{m,i\}\subseteq\mathcal{I}\subset\mathcal{J}} g_{\mathcal{I}}^{\{m\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}} \delta_{[K]\setminus\mathcal{J}\cup\{m\}} \\
&\quad + \sum_{m\in\mathcal{I}\subset\mathcal{J}\setminus\{i\}} g_{\mathcal{I}}^{\{m\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}} \delta_{[K]\setminus\mathcal{J}\cup\{m\}} &(73)\\
&= \sum_{\{m,i\}\subseteq\mathcal{I}\subset\mathcal{J}} g_{\mathcal{I}}^{\{m\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}} \delta_{[K]\setminus\mathcal{J}\cup\{m\}} \\
&\quad + \sum_{\mathcal{I}\subset\mathcal{J}\setminus\{i,m\}} g_{\mathcal{I}\cup\{m\}}^{\{m\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}\setminus\{m\}} \delta_{[K]\setminus\mathcal{J}\cup\{m\}} &(74)
\end{aligned}$$

on the other hand we have

$$\begin{aligned}
&\sum_{i\in\mathcal{I}\subset\mathcal{J}} g_{\mathcal{I}}^{\{i\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}} \delta_{[K]\setminus\mathcal{J}\cup\{i\}} \\
&= \sum_{\{m,i\}\subseteq\mathcal{I}\subset\mathcal{J}} g_{\mathcal{I}}^{\{i\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}} \delta_{[K]\setminus\mathcal{J}\cup\{i\}} \\
&\quad + \sum_{i\in\mathcal{I}\subset\mathcal{J}\setminus\{m\}} g_{\mathcal{I}}^{\{i\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}} \delta_{[K]\setminus\mathcal{J}\cup\{i\}} &(75)\\
&= \sum_{\{m,i\}\subseteq\mathcal{I}\subset\mathcal{J}} g_{\mathcal{I}}^{\{i\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}} \delta_{[K]\setminus\mathcal{J}\cup\{i\}} \\
&\quad + \sum_{\mathcal{I}\subset\mathcal{J}\setminus\{m,i\}} g_{\mathcal{I}\cup\{i\}}^{\{i\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}\setminus\{i\}} \delta_{[K]\setminus\mathcal{J}\cup\{i\}} &(76)
\end{aligned}$$

$\forall\ \{m,i\} \subseteq \mathcal{I} \subset \mathcal{J}$ $|\mathcal{I}| < |\mathcal{J}|$; $min(\mathcal{I}) = m$ and $i \in \mathcal{I}$ so by the hypothesis we have $g_{\mathcal{I}}^{\{m\}} R_m \geq g_{\mathcal{I}}^{\{i\}} R_i$. In addition we have $\delta_m \geq \delta_i$ thus

$$\sum_{\{m,i\}\subseteq\mathcal{I}\subset\mathcal{J}} g_{\mathcal{I}}^{\{m\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}} \delta_{[K]\setminus\mathcal{J}\cup\{m\}} R_m \geq \sum_{\{m,i\}\subseteq\mathcal{I}\subset\mathcal{J}} g_{\mathcal{I}}^{\{i\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}} \delta_{[K]\setminus\mathcal{J}\cup\{i\}} R_i \tag{77}$$

We have $\delta_m \geq \delta_i$ so $\bar{\delta}_i \geq \bar{\delta_m}$. Since $R_m \delta_m \geq R_i \delta_i$ then $\bar{\delta}_{\mathcal{J}\setminus\mathcal{I}\setminus\{m\}} \delta_{[K]\setminus\mathcal{J}\cup\{m\}} R_m \geq \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}\setminus\{i\}} \delta_{[K]\setminus\mathcal{J}\cup\{i\}} R_i$. $\forall\ \mathcal{I} \subset \mathcal{J} \setminus \{m,i\}$ we have from (25) $g_{\mathcal{I}\cup\{m\}}^{\{m\}} = g_{\mathcal{I}\cup\{i\}}^{\{i\}}$. Thus

$$R_m \sum_{\mathcal{I}\subset\mathcal{J}\setminus\{i,m\}} g_{\mathcal{I}\cup\{m\}}^{\{m\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}\setminus\{m\}} \delta_{[K]\setminus\mathcal{J}\cup\{m\}} \geq R_i \sum_{\mathcal{I}\subset\mathcal{J}\setminus\{m,i\}} g_{\mathcal{I}\cup\{i\}}^{\{i\}} \bar{\delta}_{\mathcal{J}\setminus\mathcal{I}\setminus\{i\}} \delta_{[K]\setminus\mathcal{J}\cup\{i\}} \tag{78}$$

and the proof is completed.

C. THE OUTERBOUND UNDER THE ONE-SIDED FAIR RATE

Suppose $\exists\ \pi_1$ such that $\sum_{j=1}^{K} R_{\pi_1(j)} w_{\pi_1(1)..\pi_1(j)} \leq 1$ and for a given $i \in [K]$ $\pi_1(i) \leq \pi_1(i+1)$. We prove that for any permutation π_2 that satisfies $\pi_2(i+1) = \pi_1(i) = k$; $\pi_2(i) = \pi_1(i+1) = k'$ and $\pi_1(j) = \pi_2(j)\ \forall\ j \in [K] \setminus \{i, i+1\}$, it holds $\sum_{j=1}^{K} R_{\pi_2(j)} w_{\pi_2(1)..\pi_2(j)} \leq 1$. It suffices to show that

$$\begin{aligned}
&w_{\pi_1(1)..\pi_1(i)} R_{\pi_1(i)} + w_{\pi_1(1)..\pi_1(i+1)} R_{\pi_1(i+1)} \\
&\quad \geq w_{\pi_2(1)..\pi_2(i)} R_{\pi_2(i)} + w_{\pi_2(1)..\pi_2(i+1)} R_{\pi_2(i+1)} \\
&\Leftrightarrow \\
&(w_{\pi_1(1)..\pi_1(i)} - w_{\pi_2(1)..\pi_2(i+1)}) R_{\pi_1(i)} \\
&\quad \geq (w_{\pi_2(1)..\pi_2(i)} - w_{\pi_1(1)..\pi_1(i+1)}) R_{\pi_1(i+1)} \\
&\Leftrightarrow \\
&(w_{\mathcal{I}k} - w_{\mathcal{I}kk'}) R_k \\
&\quad \geq (w_{\mathcal{I}k'} - w_{\mathcal{I}kk'}) R_{k'} &(79)
\end{aligned}$$

where $\mathcal{I} = \pi_1(1)..\pi_1(i-1)$. By replacing the weigh by its expression we obtain

$$\begin{aligned}
w_{\mathcal{I}k} - \tilde{w}_{\mathcal{I}kk'} &= \frac{\bar{p}_{\mathcal{I}k}}{1-\delta_{\mathcal{I}k}} - \frac{\bar{p}_{\mathcal{I}kk'}}{1-\delta_{\mathcal{I}kk'}} &(80)\\
&= \bar{p}_{\mathcal{I}k} \left[\frac{1}{1-\delta_{\mathcal{I}k}} - \frac{1}{1-\delta_{\mathcal{I}kk'}} + \frac{p_{k'}}{1-\delta_{\mathcal{I}kk'}} \right] &(81)\\
&= \bar{p}_{\mathcal{I}k} \left[\frac{(1-\delta_{\mathcal{I}kk'}) - (1-\delta_{\mathcal{I}k})}{(1-\delta_{\mathcal{I}k})(1-\delta_{\mathcal{I}kk'})} + \frac{p_{k'}}{1-\delta_{\mathcal{I}kk'}} \right] &(82)\\
&= \frac{\bar{p}_{\mathcal{I}k}}{1-\delta_{\mathcal{I}kk'}} \left[\frac{\delta_{\mathcal{I}k}(1-\delta_{k'})}{(1-\delta_{\mathcal{I}k})} + p_{k'} \right] &(83)
\end{aligned}$$

and similarly

$$w_{\mathcal{I}k'} - \tilde{w}_{\mathcal{I}kk'} = \frac{\bar{p}_{\mathcal{I}k'}}{1-\delta_{\mathcal{I}kk'}} \left[\frac{\delta_{\mathcal{I}k'}(1-\delta_k)}{(1-\delta_{\mathcal{I}k'})} + p_k \right] \tag{84}$$

thus (79) is equivalent to

$$\frac{\delta_{\mathcal{I}}(1-\delta_{k'})}{(1-\delta_{\mathcal{I}k})} \bar{p}_k \delta_k R_k - \frac{\delta_{\mathcal{I}}(1-\delta_k)}{(1-\delta_{\mathcal{I}k'})} \bar{p}_{k'} \delta_{k'} R_{k'} + (\bar{p}_k p_{k'} R_k - \bar{p}_{k'} p_k R_{k'}) \geq 0 \tag{85}$$

Since $k \leq k'$ then $\delta_k \geq \delta_{k'}$, so it is sufficient to prove that

$$\frac{\delta_{\mathcal{I}}(1-\delta_k)}{(1-\delta_{\mathcal{I}k'})} \underbrace{[\bar{p}_k \delta_k R_k - \bar{p}_{k'} \delta_{k'} R_{k'}]}_{A} + \underbrace{(\bar{p}_k p_{k'} R_k - \bar{p}_{k'} p_k R_{k'})}_{B} \geq 0 \tag{86}$$

This is satisfied if $A \geq 0$ and $B \geq 0$. The condition B is equivalent to

$$\frac{R_{k'}}{R_k} \leq \frac{\bar{p}_k p_{k'}}{\bar{p}_{k'} p_k} \triangleq \theta \tag{87}$$

We will examine condition A by considering the case $p_{k'} \geq p_k$ and $p_k \geq p_{k'}$ separately.

- Case $\theta > 1$
  In this case we have $p_k < p_{k'}$, or $\bar{p}_k > \bar{p}_{k'}$. Condition A reduces to:

  $$\delta_k R_k - \delta_{k'} R_{k'} \geq 0.$$

- Case $\theta < 1$
  In this case we have $p_k > p_{k'}$ or $\bar{p}_k < \bar{p}_{k'}$. Then we have

  $$\frac{R_{k'}}{R_k} \leq \frac{\bar{p}_k p_{k'}}{\bar{p}_{k'} p_k} \leq \frac{\bar{p}_k \delta_k}{\bar{p}_{k'} \delta_{k'}} \leq \frac{\delta_k}{\delta_{k'}}$$

  This means that B implies A so that the desired inequality holds once B holds. Since A is inactive, we can then consider a looser bounds

  $$\delta_k R_k - \delta_{k'} R_{k'} \geq 0.$$

So we obtain the result. Starting by π_1 as the identity we can obtain all the remaining $K! - 1$ permutations.

D. TOTAL DURATION

We have

$$\begin{aligned}
T_{tot} &= \sum_{\mathcal{J} \subseteq [K]} \max_{j \in \mathcal{J}} \left\{ t_{\mathcal{J}}^j \right\} &(88)\\
&= \sum_{\mathcal{J} \subseteq [K]} \max_{j \in \mathcal{J}} \left\{ g_{\mathcal{J}}^j F_j \right\} &(89)\\
&= \sum_{i=1}^{K} \sum_{\mathcal{J}: i = min(\mathcal{J})} \max_{j \in \mathcal{J}} \left\{ g_{\mathcal{J}}^j F_j \right\} &(90)\\
&= \sum_{i=1}^{K} \sum_{\mathcal{J}: i = min(\mathcal{J})} g_{\mathcal{J}}^i F_i &(91)\\
&= \sum_{i=1}^{K} F_i \sum_{i \in \mathcal{J} \subseteq \{i..K\}} g_{\mathcal{J}}^i &(92)\\
&= \sum_{i=1}^{K} F_i w_{[K] \setminus \{i..K\} \cup \{i\}} &(93)\\
&= \sum_{i=1}^{K} F_i w_{\{1..i\}} &(94)\\
&= \sum_{i=1}^{K} F_i \frac{\prod_{j=1}^{i} (1 - p_j)}{1 - \prod_{j=1}^{i} \delta_j} &(95)
\end{aligned}$$

Since $\boldsymbol{R}$ is one sided fair, (91) follows from the fact that $\arg\max_{j \in \mathcal{J}} (g_{\mathcal{J}}^j R_j) = \arg\max_{j \in \mathcal{J}} (g_{\mathcal{J}}^j F_j) = min(\mathcal{J})$. (93) follows from the equality $\sum_{i \in \mathcal{I} \subseteq \mathcal{J}} g_{\mathcal{I}}^i = w_{[K] \setminus \mathcal{J} \cup \{i\}}$.