

# Improved Converse and Gap Results for Coded Caching

Chien-Yi Wang, Shirin Saeedi Bidokhti *Member, IEEE*, and Michèle Wigger *Senior Member, IEEE*,

**Abstract**—Improved lower bounds are derived on the average and worst-case rate-memory tradeoffs of the Maddah-Ali&Niesen coded caching scenario. For any number of users and files and for arbitrary cache sizes, the multiplicative gap between the exact rate-memory tradeoff and the new lower bound is shown to be less than 2.315 in the worst-case scenario and 2.507 in the average-case scenario.

**Index Terms**—Caching, rate-memory tradeoff, source coding, index coding.

## I. INTRODUCTION

We consider the canonical coded caching scenario by Maddah-Ali and Niesen [1] with a single transmitter and  $K$  receivers, where each receiver is equipped with a cache memory of equal size (see in Figure 1). Communication takes place in two phases: a *caching phase* and a subsequent *delivery phase*. In the caching phase, the transmitter stores contents (arbitrary functions of files) at the receivers' cache memories. In the delivery phase, each receiver makes a demand and the transmitter accordingly conveys the desired files to each of the receivers. The main challenge in this configuration is that during the caching phase it is not known which receiver demands which specific file from the library. The cache contents thus need to be designed so as to be useful for many possible demands.

Traditional caching systems store a portion of the most popular files in each and every cache memory. This allows the receivers to retrieve these files locally without burdening the common communication link from the transmitter to the receivers. Recently in [1], it was shown that much larger gains, so called *global caching gains*, are possible if various receivers store different parts of the files in their cache memories. In this case, the transmitter can simultaneously serve multiple receivers during the delivery phase by sending coded data, and thus significantly reduce the delivery rate (latency) of communication.

The main quantity of interest in this work is the *rate-memory tradeoff* introduced in [1]—i.e., the minimum required delivery rate, as a function of the cache memories, so that all receivers reliably recover their demanded files. We

consider both the *worst-case* rate-memory tradeoff defined in [1], which is the common scenario in the coded-caching literature, as well as the *average-case* rate-memory tradeoff defined in [2]. In the latter case, the rate can adapt to the receivers' demands and the rate-memory tradeoff is defined as the average rate over all possible demand vectors. Upper bounds on the worst-case rate-memory tradeoff were presented for certain special cases in [3]–[10] and lower bounds (converse results) were presented in [11]–[13]. The previously best lower and upper bounds for the worst-case rate-memory tradeoff match up to a multiplicative gap of 4 [11]. The works in [5], [9] determined the exact rate-memory tradeoff assuming uncoded cache placements, i.e., assuming that *fractions of contents are cached*. This caching strategy is however known to be suboptimal in general. Upper and lower bounds on the rate-memory tradeoff for the average-case scenario were derived in [9], [13]. The previously best lower and upper bounds in this scenario have been shown to match up to a multiplicative gap of 4.7 [13].

In this paper we provide new lower bounds on the rate-memory tradeoff. The new lower bounds match the worst-case and average-case rate-memory tradeoffs up to multiplicative gaps of 2.315 and 2.507, respectively. More precisely, these gaps are with respect to the upper bounds on the rate-memory tradeoff under *decentralized* caching in [9]. An upper bound on the rate-memory tradeoff under decentralized caching is also an upper bound on the rate-memory tradeoff under *centralized* caching considered here, because in the decentralized caching the cache content at a given receiver has to be chosen according to a specific distribution, whereas in centralized caching any content can be cached that satisfies the cache memory constraints. In an independent and concurrent work, [14] presents slightly improved lower bounds for both the worst-case and average-case rate-memory tradeoffs. These bounds are within a gap of 2.00884 from the decentralized schemes. The proof in [14] is similar to the proof here, but includes an additional averaging step over all possible labelings of receivers.

Many other variations of the caching problem have recently been studied such as online caching [15]; caching with non-uniform demands [2], [16]–[18]; caching of correlated files [19]–[25] where [19] shows how Wyner's and Gac-Körner's common information play a key role; caching in noisy broadcast channels [26]–[40] where coding opportunities could be exploited through joint cache-channel coding in

C.-Y. Wang is with MediaTek Inc., Hsinchu 30078, Taiwan. He has previously been with LTCI, Telecom ParisTech, 75013 Paris, France. (e-mail: chien-yi.wang@mediatek.com) S. Saeedi Bidokhti is with the Department of Electrical and Systems Engineering at the University of Pennsylvania, USA. (e-mail: saeedi@seas.upenn.edu.) S. Saeedi Bidokhti was supported by the Swiss National Science Foundation fellowship no. 158487. M. Wigger is with LTCI, Telecom ParisTech, 75013 Paris, France (e-mail: michele.wigger@telecom-paristech.fr).

The material in this paper was presented at the 2017 *IEEE International Symposium on Information Theory*, Aachen, Germany.

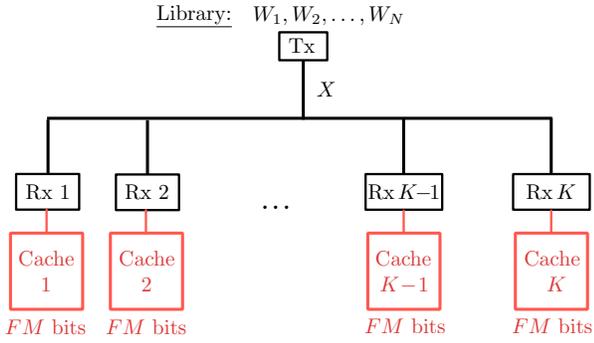


Fig. 1: Coded caching scenario with  $K$  receivers having equal cache size  $FM$  bits.

heterogeneous networks [27], [28], [30], [31]<sup>1</sup>, feedback and channel state information [34]–[37], and multiple antennas [39], [40]; caching in Gaussian interference networks [41]–[43]; hierarchical networks [44] and multi-server networks [45]; and cellular networks [46]–[53].

## II. DETAILED PROBLEM SETUP

Consider the communication scenario in Figure 1, which includes a single transmitter and  $K$  receivers that we term Receivers  $1, \dots, K$ . The transmitter has a library of  $N$  independent messages  $W_1, \dots, W_N$ . Each  $W_d$  is uniformly distributed over  $\{1, \dots, 2^F\}$  for  $F$  a positive integer. Every receiver is provided with a cache memory of size  $FM$  bits, and the range of interest for  $M$  is

$$0 \leq M \leq N. \quad (1)$$

Here,  $M = 0$  means that there is no cache memory in the system and  $M = N$  means that every receiver can store all the library in its cache memory.

Each receiver will demand exactly one message from the library. We denote the demand of Receiver  $k$  by

$$d_k \in \mathcal{N} := \{1, \dots, N\} \quad (2)$$

and thus the message demanded by Receiver  $k$  is  $W_{d_k}$ . Let

$$\mathbf{d} := (d_1, \dots, d_K) \quad (3)$$

denote the receivers' demand vector.

The communications process takes place in two phases, namely the caching phase and the delivery phase. Caching is done during a period of low network-congestion and before the receivers' demand vector  $\mathbf{d}$  is known. More specifically, for  $k \in \{1, \dots, K\}$ , the transmitter sends an individual cache message  $\mathbb{V}_k \in \{1, \dots, [2^{FM}]\}$  to Receiver  $k$ . Since  $\mathbf{d}$  is

unknown at this time, the cache messages will be functions of the entire library. For every  $k \in \{1, \dots, K\}$ , we have<sup>2</sup>

$$\mathbb{V}_k := g_k(W_1, \dots, W_N), \quad (4)$$

for some caching function

$$g_k : \{1, \dots, 2^F\}^N \rightarrow \{1, \dots, [2^{FM}]\}. \quad (5)$$

In the delivery phase, the transmitter is given the receivers' demands  $\mathbf{d} = (d_1, \dots, d_K)$ , and it generates the delivery-symbol  $X$  that is sent over the common noise-free bit-pipe:

$$X := f_{\mathbf{d}}(W_1, \dots, W_N), \quad (6)$$

for some encoding function

$$f_{\mathbf{d}} : \{1, \dots, 2^F\}^N \rightarrow \mathcal{X}, \quad (7)$$

where  $\mathcal{X}$  is the delivery alphabet that we will specify shortly. We assume that  $\mathbf{d}$  is known to all the receivers (e.g.,  $\mathbf{d}$  can be communicated to the receivers with asymptotically zero transmission rate<sup>3</sup>).

Receiver  $k$ ,  $k \in \{1, \dots, K\}$ , perfectly observes the delivery-symbol  $X$ , and thus recovers its desired message as

$$\hat{W}_k := \varphi_{k,\mathbf{d}}(X, \mathbb{V}_k) \quad (8)$$

using some decoding function

$$\varphi_{k,\mathbf{d}} : \mathcal{X} \times \{1, \dots, [2^{FM}]\} \rightarrow \{1, \dots, 2^F\}. \quad (9)$$

We are left with specifying the delivery alphabet  $\mathcal{X}$ . We distinguish the *worst-case* [1] and *average-case* [2] scenarios as follows:

- In the *worst-case scenario*, the delivery alphabet  $\mathcal{X}$  does not depend on the demand vector  $\mathbf{d}$ . In this scenario, the rate-memory pair  $(R, M)$  is *achievable* if for every  $\epsilon > 0$  and sufficiently large message lengths  $F$ , there exists a caching function (5), an encoding function (7), and decoding functions (9) for delivery alphabet

$$\mathcal{X} = \{1, \dots, [2^{F(R+\epsilon)}]\}, \quad (10)$$

so that for each demand vector  $\mathbf{d} \in \mathcal{N}^K$ , every Receiver  $k$ ,  $k \in \{1, \dots, K\}$ , can perfectly reconstruct its desired message:

$$\hat{W}_k = W_{d_k}. \quad (11)$$

- In the *average-case scenario*, the delivery alphabet  $\mathcal{X}$  depends on the demand vector  $\mathbf{d}$ . In this scenario, the rate-memory pair  $(R, M)$  is *achievable* if for each demand vector  $\mathbf{d} \in \mathcal{N}^K$ , any  $\epsilon > 0$ , and sufficiently large message lengths  $F$ , there exists a caching function (5), an encoding function (7), and decoding functions (9) for delivery alphabet

$$\mathcal{X}_{\mathbf{d}} = \{1, \dots, [2^{FR_{\mathbf{d}}}] \}, \quad (12)$$

<sup>2</sup>Alternatively, one could allow the caching functions to depend also on external randomness that does not depend on the library nor the receivers' demands. The rate-memory tradeoff, which is the focus of this paper, is the same under both assumptions. This can be proved in a similar way as proving that randomized encoding does not change the rate-distortion function of memoryless source coding problems.

<sup>3</sup>Alternatively, the desired information could also be sent from the server to the users as part of the subsequent delivery communication, see [54].

<sup>1</sup>Joint cache-channel coding is used for joint source-channel coding where part of the sources are actually given by the cache contents.

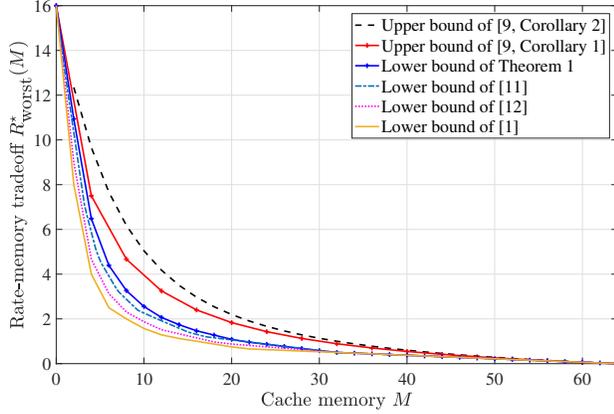


Fig. 2: Upper and lower bounds on  $R_{\text{worst}}^*(M)$  for  $K = 16$  and  $N = 64$ .

so that each Receiver  $k \in \{1, \dots, K\}$  can perfectly reconstruct its desired message (11) and

$$\frac{1}{NK} \sum_{\mathbf{d} \in \mathcal{N}^K} R_{\mathbf{d}} \leq R + \epsilon. \quad (13)$$

The main focus of this paper is on the *rate-memory tradeoffs* of the worst-case and the average-case scenarios.

*Definition 1:* Given the cache memory size  $M$ , we define the *rate-memory tradeoffs*  $R_{\text{worst}}^*(M)$  and  $R_{\text{avg}}^*(M)$  as the infimum of all rates  $R$  such that the rate-memory pair  $(R, M)$  is achievable for the worst-case and average-case scenarios, respectively.

### III. MAIN RESULTS

Define  $\bar{N} := \min\{K, N\}$  and  $\bar{\mathcal{N}} := \{1, 2, \dots, \bar{N}\}$ .

#### A. Worst-Case Scenario

Our first result is a lower bound on the rate-memory tradeoff in the worst-case scenario and it is proved in Section IV-B. The bound can also be extracted from the converse result for general degraded broadcast channels in [55].

*Theorem 1:* For all  $M \in [0, N]$ ,

$$R_{\text{worst}}^*(M) \geq R_{\text{worst}}^{\text{low}}(M), \quad (14)$$

where

$$R_{\text{worst}}^{\text{low}}(M) := \max \left\{ \max_{\ell \in \bar{\mathcal{N}}} \left[ \ell - M \frac{\ell^2}{N} \right], \max_{\ell \in \bar{\mathcal{N}}} \left[ \ell - M \sum_{j=1}^{\ell} \frac{j}{N-j+1} \right] \right\}. \quad (15)$$

Figure 2 compares our new lower bound on  $R_{\text{worst}}^*(M)$  with the existing lower bounds in [1], [11], [12]. The figure also shows upper bounds from [9]. The solid red upper bound is for centralized caching, as considered in this paper. The dashed black upper bound is for decentralized caching. For simplicity, the latter upper bound is used to derive gap results as stated in Theorem 2 below.

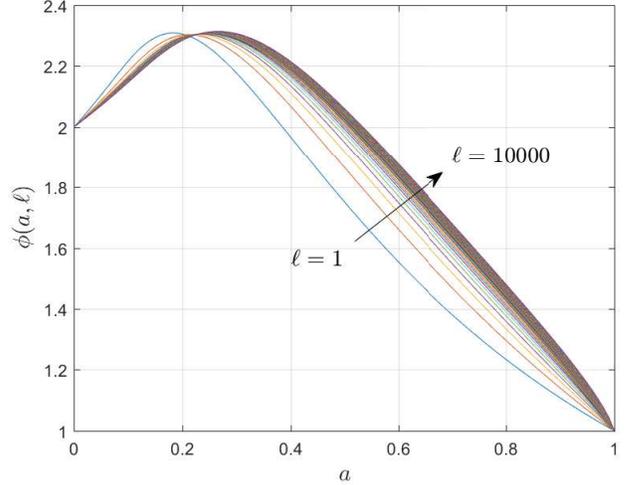


Fig. 3: The functions  $\phi(a, \ell)$  for  $a \in (0, 1)$  and  $\ell = 1, \dots, 10^4$ .

*Theorem 2:* Irrespective of the number of users  $K$ , the library size  $N$ , and the memory size  $M \in [0, N]$ :

$$\frac{R_{\text{worst}}^*(M)}{R_{\text{worst}}^{\text{low}}(M)} \leq \max_{\ell \in \mathbb{Z}^+} \max_{a \in (0, 1)} \phi(a, \ell), \quad (16)$$

where

$$\phi(a, \ell) := \frac{\frac{a(\ell+1)}{(1-a)\ell} \left( 1 - \left( \frac{\ell+a}{\ell+1} \right)^{\ell/a} \right)}{1 - (1-a) \sum_{j=0}^{\ell-1} \frac{1}{\ell-a-j}}. \quad (17)$$

*Proof:* See Section V. ■

*Remark 1:* For any  $\ell \in \mathbb{Z}^+$ , the function  $a \mapsto \phi(a, \ell)$  is continuous and bounded over  $(0, 1)$ , see also Figure 3. Numerical evaluations<sup>4</sup> show that for  $\ell \in \{1, \dots, 10^4\}$ :

$$\max_{a \in (0, 1)} \phi(a, \ell) \leq 2.315. \quad (18)$$

Moreover,

$$\max_{\ell \in \mathbb{Z}^+ : \ell > 10^4} \max_{a \in (0, 1)} \phi(a, \ell) \leq \max_{b \in (0, 10^{-4})} \max_{a \in (0, 1)} \psi(a, b), \quad (19)$$

where

$$\psi(a, b) := \frac{\frac{a(1+b)}{1-a} \left( 1 - \left( \frac{1+ab}{1+b} \right)^{\frac{1}{ab}} \right)}{1 - \frac{(1-a)b}{1-a+ab} + \frac{1-a}{a} \ln(1-a+ab)}. \quad (20)$$

The function  $(a, b) \mapsto \psi(a, b)$  is continuous and bounded over  $(0, 1) \times (0, 10^{-4})$ , see also Figure 4. Numerical evaluations show that

$$\max_{b \in (0, 10^{-4})} \max_{a \in (0, 1)} \psi(a, b) \leq 2.315. \quad (21)$$

*Proof:* Inequality (19) is proved in Section V-B. ■

<sup>4</sup>All numerical evaluations in this paper are performed by applying the MATLAB function `fmincon` with the sequential quadratic programming (SQP) method.

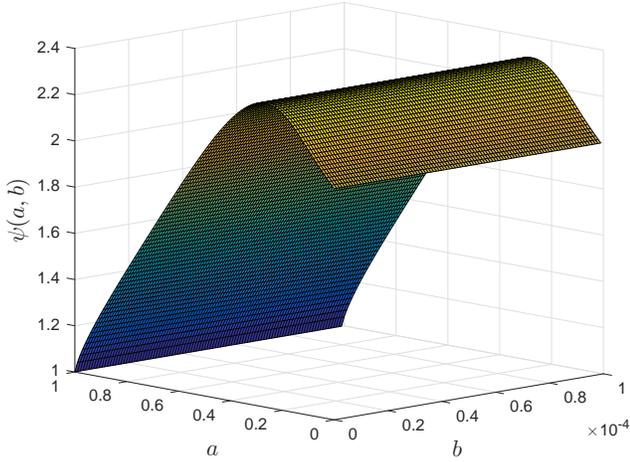


Fig. 4: The function  $\psi(a, b)$  over  $(a, b) \in (0, 1) \times (0, 10^{-4})$ .

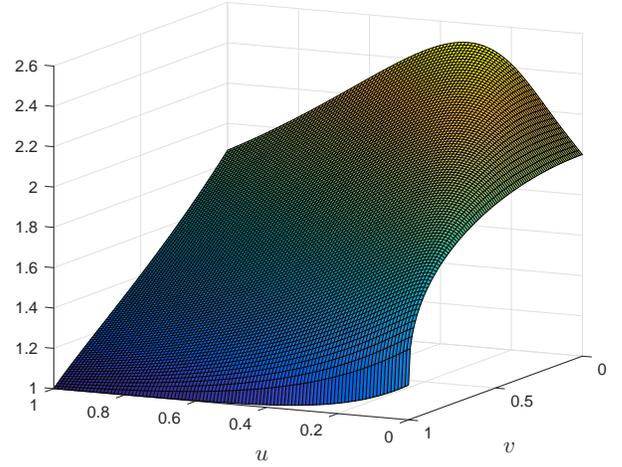


Fig. 6: The function  $\eta(u, v)$  over  $(u, v) \in (0, 1) \times (0, 1)$ .

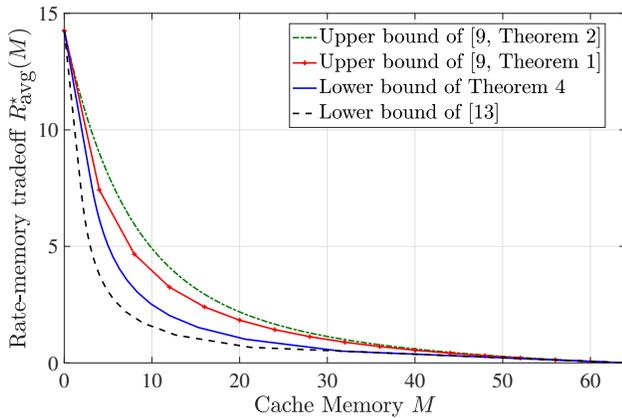


Fig. 5: Upper and lower bounds on  $R_{\text{avg}}^*(M)$  for  $K = 16$  and  $N = 64$ .

### B. Average-Case Scenario

*Theorem 3:* For all  $M \in [0, N]$ ,

$$R_{\text{avg}}^*(M) \geq R_{\text{avg}}^{\text{low}}(M), \quad (22)$$

where

$$R_{\text{avg}}^{\text{low}}(M) := \max \left\{ \begin{aligned} & \max_{\ell \in \{1, \dots, K\}} \left[ \left(1 - \left(1 - \frac{1}{N}\right)^\ell\right) (N - \ell M) \right], \\ & \max_{\ell \in \{1, \dots, K\}} \left[ \left(1 - \left(1 - \frac{1}{N}\right)^\ell\right) N - \frac{\ell(\ell+1)}{2N} M \right] \end{aligned} \right\}. \quad (23)$$

*Proof:* See Section IV-C. ■

Figure 5 compares this new lower bound on  $R_{\text{avg}}^*(M)$  with the existing lower bounds in [13] and the upper bounds in [9]. The solid red upper bound is for centralized caching, as considered in this paper. The dotted green upper bound is for decentralized caching and also from [9].

As the following theorem and remark show, the multiplicative gap between the lower bound of Theorem 3 and  $R_{\text{avg}}^*(M)$  is at most 2.507.

*Theorem 4:* Irrespective of the number of users  $K$ , the library size  $N$ , and the memory size  $M \in [0, N]$ :

$$\frac{R_{\text{avg}}^*(M)}{R_{\text{avg}}^{\text{low}}(M)} \leq \max_{u \in (0, 1]} \max_{v \in (0, 1/2]} \eta(u, v), \quad (24)$$

where

$$\eta(u, v) := \frac{\left(u + v - v(1-v)^{\frac{u}{v}}\right) \left(1 - \left(1 - \frac{v}{u+v}(1-v)^{\frac{u}{v}}\right)^{\frac{1}{v}}\right)}{(1-v)^{\frac{u}{v}} \left(1 - \left(1 + \frac{u}{2}\right)(1-v)^{\frac{u}{v}}\right)}. \quad (25)$$

*Proof:* See Section VI. ■

*Remark 2:* The function  $\eta(u, v)$  is continuous and bounded over  $(0, 1] \times (0, 1/2]$ . Numerical evaluations show that

$$\max_{u \in (0, 1]} \max_{v \in (0, 1/2]} \eta(u, v) \leq 2.507. \quad (26)$$

## IV. PROOF OF THEOREMS 1 AND 3

### A. Auxiliary Lemmas

The following two lemmas will be used in the proofs of Theorems 1 and 3.

Lemma 1 (see below) is stated for the average-case scenario. It also readily applies to the worst-case scenario if rate  $R_d$  is replaced by  $R$ .

*Lemma 1:* Fix a number  $\ell \in \bar{\mathcal{N}}$  and a demand vector  $\mathbf{d} \in \mathcal{N}^K$  whose first  $\ell$  entries are  $d_1, \dots, d_\ell$ . Fix also a small  $\epsilon > 0$  and assume a sufficiently large  $F$  with caching, encoding, and decoding functions so that (11) holds for all  $k \in \{1, \dots, K\}$ . Then,

$$R_d + \epsilon \geq \kappa_{\mathbf{d}}(\ell) - \frac{1}{F} \sum_{k=1}^{\ell} I(W_{d_k}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{d_1}, \dots, W_{d_{k-1}}), \quad (27)$$

where  $\kappa_{\mathbf{d}}(\ell)$  denotes the number of distinct demands for Receivers  $1, \dots, \ell$ :

$$\kappa_{\mathbf{d}}(\ell) := |\{d_1, \dots, d_\ell\}|. \quad (28)$$

*Proof:* For any  $k \in \{1, \dots, \ell\}$ :

$$\begin{aligned} I(X; W_{d_k} | \mathbb{V}_1, \dots, \mathbb{V}_k, W_{d_1}, \dots, W_{d_{k-1}}) & \\ \stackrel{(a)}{=} H(W_{d_k} | \mathbb{V}_1, \dots, \mathbb{V}_k, W_{d_1}, \dots, W_{d_{k-1}}) & \\ = H(W_{d_k} | W_{d_1}, \dots, W_{d_{k-1}}) & \\ - I(W_{d_k}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{d_1}, \dots, W_{d_{k-1}}) & \\ \stackrel{(b)}{=} F \cdot \mathbb{1}\{d_k \notin \{d_1, \dots, d_{k-1}\}\} & \\ - I(W_{d_k}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{d_1}, \dots, W_{d_{k-1}}), & \quad (29) \end{aligned}$$

where  $\mathbb{1}\{d_k \notin \{d_1, \dots, d_{k-1}\}\}$  denotes the indicator function that is 1 if  $d_k$  is not in  $\{d_1, \dots, d_{k-1}\}$  and is 0 otherwise. Moreover, (a) holds because (11) implies that  $H(W_{d_k} | X, \mathbb{V}_1, \dots, \mathbb{V}_k, W_{d_1}, \dots, W_{d_{k-1}}) = 0$  and (b) holds by the independence of the messages and because  $H(W_d) = F$  for any  $d \in \mathcal{N}$ .

On the other hand,

$$\begin{aligned} \sum_{k=1}^{\ell} I(X; W_{d_k} | \mathbb{V}_1, \dots, \mathbb{V}_k, W_{d_1}, \dots, W_{d_{k-1}}) & \\ \leq \sum_{k=1}^{\ell} I(X; W_{d_k}, \mathbb{V}_k | \mathbb{V}_1, \dots, \mathbb{V}_{k-1}, W_{d_1}, \dots, W_{d_{k-1}}) & \\ = I(X; W_{d_1}, \dots, W_{d_\ell}, \mathbb{V}_1, \dots, \mathbb{V}_\ell) & \\ \leq H(X) & \\ \stackrel{(a)}{\leq} F(R_{\mathbf{d}} + \epsilon), & \quad (30) \end{aligned}$$

where (a) holds by (10). Combining (29) and (30) establishes the lemma.  $\blacksquare$

*Lemma 2:* Let  $L$  be a positive integer,  $(A_1, \dots, A_L)$  be a random independent  $L$ -tuple, and  $\mathbb{V}$  be a random variable arbitrarily correlated with  $A_1, \dots, A_L$ . For any subset  $\mathcal{S} \subseteq \{1, \dots, L\}$ , denote the subset  $\{A_s, s \in \mathcal{S}\}$  by  $A_{\mathcal{S}}$ . Then, for all  $l \in \{1, \dots, L\}$ ,

$$\frac{1}{\binom{L}{l}} \sum_{\substack{\mathcal{S} \subseteq \{1, \dots, L\}: \\ |\mathcal{S}|=l}} I(A_{\mathcal{S}}; \mathbb{V}) \leq \frac{1}{L} I(A_1, \dots, A_L; \mathbb{V}). \quad (31)$$

*Proof:* Consider any  $l \in \{1, \dots, L\}$ . We have

$$\begin{aligned} \frac{1}{\binom{L}{l}} \sum_{\substack{\mathcal{S} \subseteq \{1, \dots, L\}: \\ |\mathcal{S}|=l}} I(A_{\mathcal{S}}; \mathbb{V}) & \\ \stackrel{(a)}{=} \frac{1}{\binom{L}{l}} \sum_{\substack{\mathcal{S} \subseteq \{1, \dots, L\}: \\ |\mathcal{S}|=l}} \sum_{j \in \mathcal{S}} H(A_j) & \\ - \frac{1}{\binom{L}{l}} \sum_{\substack{\mathcal{S} \subseteq \{1, \dots, L\}: \\ |\mathcal{S}|=l}} H(A_{\mathcal{S}} | \mathbb{V}) & \\ = \frac{\binom{L-1}{l-1}}{\binom{L}{l}} \sum_{j=1}^L H(A_j) - \frac{1}{\binom{L}{l}} \sum_{\substack{\mathcal{S} \subseteq \{1, \dots, L\}: \\ |\mathcal{S}|=l}} H(A_{\mathcal{S}} | \mathbb{V}) & \end{aligned}$$

$$\begin{aligned} & \stackrel{(b)}{\leq} \frac{l}{L} \sum_{j=1}^L H(A_j) - \frac{l}{L} H(A_1, \dots, A_L | \mathbb{V}) \\ & \stackrel{(c)}{=} \frac{l}{L} I(A_1, \dots, A_L; \mathbb{V}), \end{aligned} \quad (32)$$

where (a) and (c) follow since  $A_1, \dots, A_L$  are independent and (b) follows from the generalized Han Inequality (see [56, Theorem 17.6.1]).  $\blacksquare$

### B. Proof of Theorem 1

Fix  $\ell \in \bar{\mathcal{N}}$  and restrict attention to Receivers  $1, \dots, \ell$  and their cache memories. Let  $\mathcal{Q}_\ell^{\text{dist}}$  be the set of all ordered  $\ell$ -dimensional demand vectors  $(d_1, \dots, d_\ell)$  with all distinct entries. So,

$$|\mathcal{Q}_\ell^{\text{dist}}| = \binom{N}{\ell} \ell!. \quad (33)$$

Notice that for  $\mathbf{d} \in \mathcal{Q}_\ell^{\text{dist}}$ , we have  $\kappa_{\mathbf{d}}(\ell) = \ell$ , and averaging Inequality (27)<sup>5</sup> over all demand vectors  $\mathbf{d} \in \mathcal{Q}_\ell^{\text{dist}}$  yields the following inequality:

$$R + \epsilon \geq \ell - \sum_{k=1}^{\ell} \alpha_k, \quad (34)$$

where

$$\alpha_1 := \frac{1}{\binom{N}{\ell} \ell!} \sum_{\mathbf{d} \in \mathcal{Q}_\ell^{\text{dist}}} \frac{1}{F} I(W_{d_1}; \mathbb{V}_1), \quad (35a)$$

and for  $k = 2, \dots, \ell$ :

$$\alpha_k := \frac{1}{\binom{N}{\ell} \ell!} \sum_{\mathbf{d} \in \mathcal{Q}_\ell^{\text{dist}}} \frac{1}{F} I(W_{d_k}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{d_1}, \dots, W_{d_{k-1}}) \quad (35b)$$

We now upper bound the terms  $\sum_{k=1}^{\ell} \alpha_k$  that appear on the right hand side of (34). In particular, we prove the following lemma in Appendix B.

*Lemma 3:* Parameters  $\alpha_1, \dots, \alpha_\ell$  satisfy

$$\sum_{j=1}^{\ell} \alpha_k \leq \min \left\{ \frac{\ell^2}{N} M, \sum_{j=1}^{\ell} \frac{jM}{N-j+1} \right\}. \quad (36)$$

*Proof:* See Appendix B.  $\blacksquare$

Inserting (36) into (34), we obtain

$$R + \epsilon \geq \ell - \sum_{j=1}^{\ell} \frac{jM}{N-j+1}. \quad (37)$$

Finally, letting  $\epsilon \rightarrow 0$  concludes the proof.

### C. Proof of Theorem 3

For any  $\ell \in \{1, \dots, K\}$ , let  $\mathcal{Q}_\ell^{\text{rep}}$  be the set of all ordered length- $\ell$  vectors  $(d_1, \dots, d_\ell) \in \mathcal{N}^\ell$ , where repetitions are allowed. Notice that:

$$|\mathcal{Q}_\ell^{\text{rep}}| = N^\ell. \quad (38)$$

<sup>5</sup>In (27)  $R_{\mathbf{d}}$  needs to be replaced by  $R$  because here we consider a worst-case scenario.

Recall also that in the average-case scenario under investigation, the demand vector  $\mathbf{d} := (d_1, \dots, d_K)$  is uniform over  $\mathcal{Q}_K^{\text{rep}}$ . Let  $\mathbf{D} := (D_1, \dots, D_K) \sim \text{Uniform}(\mathcal{N}^K)$ .

Fix now an  $\ell \in \{1, \dots, K\}$ , and average Inequality (27) over all demand vectors  $\mathbf{d} \in \mathcal{Q}_K^{\text{rep}}$ . This yields:

$$\mathbb{E}_{\mathbf{D}}[R_{\mathbf{D}} + \epsilon] \geq \mathbb{E}_{\mathbf{D}}[\kappa_{\mathbf{D}}(\ell)] - \sum_{k=1}^{\ell} \beta_k, \quad (39)$$

where

$$\beta_1 := \frac{1}{F} I(W_{D_1}; \mathbb{V}_1 | \mathbf{D}), \quad (40a)$$

and for  $k = 2, \dots, \ell$ :

$$\beta_k := \frac{1}{F} I(W_{D_k}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{D_1}, \dots, W_{D_{k-1}}, \mathbf{D}). \quad (40b)$$

We can bound the terms in (39) with the following two lemmas.

*Lemma 4:*

$$\mathbb{E}_{\mathbf{D}}[\kappa_{\mathbf{D}}(\ell)] = N \left( 1 - \left( 1 - \frac{1}{N} \right)^\ell \right). \quad (41)$$

*Proof:*

$$\begin{aligned} & \mathbb{E}_{\mathbf{D}}[\kappa_{\mathbf{D}}(\ell)] \\ &= \mathbb{E}_{\mathbf{D}} \left[ \sum_{k=1}^{\ell} \mathbb{1}\{D_k \notin \{D_1, \dots, D_{k-1}\}\} \right] \\ &= \sum_{k=1}^{\ell} \mathbb{E}_{\mathbf{D}} \left[ \mathbb{1}\{D_k \notin \{D_1, \dots, D_{k-1}\}\} \right] \\ &\stackrel{(a)}{=} \sum_{k=1}^{\ell} \sum_{j=1}^N \frac{1}{N} \mathbb{E}_{\mathbf{D}} \left[ \mathbb{1}\{j \notin \{D_1, \dots, D_{k-1}\}\} \mid D_k = j \right] \\ &\stackrel{(b)}{=} \sum_{k=1}^{\ell} \left( 1 - \frac{1}{N} \right)^{k-1} \\ &\stackrel{(c)}{=} N \left( 1 - \left( 1 - \frac{1}{N} \right)^\ell \right), \end{aligned} \quad (42)$$

where (a) holds by the law of total expectation and because  $D_k$  is uniform over  $\{1, \dots, N\}$ ; (b) holds because  $D_1, \dots, D_{k-1}$  are i.i.d. and uniform over  $\{1, \dots, N\}$ ; and (c) follows by the formula of a geometric sum. ■

*Lemma 5:* Parameters  $\beta_1, \dots, \beta_\ell$  satisfy

$$\sum_{j=1}^{\ell} \beta_j \leq \min \left\{ \mathbb{E}_{\mathbf{D}}[\kappa_{\mathbf{D}}(\ell)] \cdot \frac{\ell M}{N}, \sum_{j=1}^{\ell} \frac{jM}{N} \right\}. \quad (43)$$

*Proof:* See Appendix C. ■

Combining the above two lemmas with (39), we obtain:

$$\begin{aligned} \mathbb{E}_{\mathbf{D}}[R_{\mathbf{D}} + \epsilon] &\geq N \left( 1 - \left( 1 - \frac{1}{N} \right)^\ell \right) \\ &\quad - \min \left\{ \mathbb{E}_{\mathbf{D}}[\kappa_{\mathbf{D}}(\ell)] \cdot \frac{\ell M}{N}, \sum_{j=1}^{\ell} \frac{jM}{N} \right\} \end{aligned} \quad (44)$$

Finally, letting  $\epsilon \rightarrow 0$  concludes the proof.

## V. PROOF OF THE GAP-RESULTS IN THEOREM 2 AND REMARK 1

### A. Proof of Theorem 2

We wish to uniformly bound the gap

$$\xi(K, N, M) := \frac{R_{\text{worst}}^*(M)}{R_{\text{worst}}^{\text{low}}(M)}, \quad (45)$$

irrespective of  $K, N \geq 1$  and  $M \in [0, N]$ .

Recall the achievable rate-memory tradeoff from [9, Corollary 2] and denote it by  $R_{\text{YMA}}(K, N, M)$ . For any pair of positive integers  $K, N \geq 1$ , we have

$$\begin{aligned} & R_{\text{YMA}}(K, N, M) \\ &:= \begin{cases} \bar{N} & \text{if } M = 0, \\ \frac{N-M}{M} \left( 1 - \left( 1 - \frac{M}{N} \right)^{\bar{N}} \right) & \text{if } M \in (0, N). \end{cases} \end{aligned} \quad (46)$$

Since  $R_{\text{YMA}}(K, N, M)$  upper bounds the rate-memory tradeoff under a *decentralized caching* assumption [9], it must also upper bound the rate-memory tradeoff under *centralized caching* as considered here. (In fact, decentralized caching imposes additional constraints on the caching functions  $g_k$  compared to our setup here.) Thus, for any number of users  $K$  and files  $N$ :

$$R_{\text{worst}}^*(M) \leq R_{\text{YMA}}(K, N, M), \quad M \in [0, N]. \quad (47)$$

We thus have

$$\begin{aligned} \xi(K, N, M) &\leq \frac{R_{\text{YMA}}(K, N, M)}{R_{\text{worst}}^{\text{low}}(M)} \\ &\leq \frac{R_{\text{YMA}}(K, N, M)}{R_{\text{worst}}(K, N, M)}, \end{aligned} \quad (48)$$

where we define

$$\underline{R}_{\text{worst}}(K, N, M) := \max_{\ell \in \mathcal{N}} \sum_{j=1}^{\ell} \left( 1 - \frac{jM}{N-j+1} \right) \quad (49)$$

and the second inequality holds because for all  $K, N, M$ :

$$\begin{aligned} R_{\text{worst}}^{\text{low}}(M) &\geq \max_{\ell \in \mathcal{N}} \left[ \ell - \sum_{j=1}^{\ell} \frac{jM}{N-j+1} \right] \\ &= \underline{R}_{\text{worst}}(K, N, M). \end{aligned} \quad (50)$$

We have a closer look at the function  $\underline{R}_{\text{worst}}(K, N, M)$ . Define

$$M_i := \begin{cases} \frac{N-i}{i+1} & \text{if } i \in \{0, 1, \dots, \bar{N}-1\}, \\ 0 & \text{if } i = \bar{N}. \end{cases}$$

and notice that

$$0 = M_{\bar{N}} < M_{\bar{N}-1} < \dots < M_0 = N. \quad (51)$$

In Appendix A it is shown that for each  $i \in \bar{N}$ :

$$\underline{R}_{\text{worst}}(K, N, M) = i - M \cdot \sum_{j=1}^i \frac{j}{N-j+1}, \quad M \in [M_i, M_{i-1}]. \quad (52)$$

So, for given  $K, N$ , the function  $\underline{R}_{\text{worst}}(K, N, M)$  is piecewise-linear with  $\bar{N}$  line segments over the intervals

$$[M_i, M_{i-1}], \quad i \in \{2, \dots, \bar{N}\}, \quad (53a)$$

$$[M_1, M_0]. \quad (53b)$$

We next upper bound  $R_{\text{YMA}}(K, N, M)$  by a function  $\bar{R}_{\text{YMA}}(K, N, M)$  that is piecewise-linear over the same intervals (53). Specifically, for every  $\ell \in \bar{N}$ , define for  $M \in [M_\ell, M_{\ell-1}]$ :

$$\begin{aligned} \bar{R}_{\text{YMA}}(K, N, M) &:= \frac{M_{\ell-1} - M}{M_{\ell-1} - M_\ell} R_{\text{YMA}}(K, N, M_\ell) \\ &+ \frac{M - M_\ell}{M_{\ell-1} - M_\ell} R_{\text{YMA}}(K, N, M_{\ell-1}). \end{aligned} \quad (54)$$

Notice that

$$\bar{R}_{\text{YMA}}(K, N, M_\ell) = R_{\text{YMA}}(K, N, M_\ell), \quad \forall \ell \in \{1, \dots, \bar{N}\}, \quad (55)$$

whereas for general  $M \in [0, N]$ :

$$\bar{R}_{\text{YMA}}(K, N, M) \geq R_{\text{YMA}}(K, N, M), \quad (56)$$

because for fixed values of  $K, N$  the function  $R_{\text{YMA}}(K, N, M)$  is convex in  $M$ .

Plugging (56) into (48), we obtain:

$$\xi(K, N, M) \leq \frac{\bar{R}_{\text{YMA}}(K, N, M)}{\underline{R}_{\text{worst}}(K, N, M)} =: \Xi(K, N, M). \quad (57)$$

Now, since the upper bound  $\Xi(K, N, M)$  is continuous and bounded in  $M \in [0, N]$  and because it is quasiconvex<sup>6</sup> in  $M$ , the maximum of  $\Xi(K, N, M)$  over each of the  $\bar{N} - 1$  closed intervals in (53a) is attained at one of the two boundary points of the interval. That means,

$$\max_{M \in [0, M_1]} \xi(K, N, M) \leq \max_{\ell \in \bar{N}} \Xi(K, N, M_\ell) \quad (58)$$

Consider now the half-open interval  $M \in [M_1, M_0]$ . Since  $\bar{R}_{\text{YMA}}(K, N, M_0) = \underline{R}_{\text{worst}}(K, N, M_0) = 0$ , for  $M \in [M_1, M_0]$ :

$$\begin{aligned} \Xi(K, N, M) &= \frac{M_0 - M}{M_0 - M_1} \bar{R}_{\text{YMA}}(K, N, M_1) + \frac{M - M_1}{M_0 - M_1} \cdot 0 \\ &= \frac{M_0 - M}{M_0 - M_1} \underline{R}_{\text{worst}}(K, N, M_1) + \frac{M - M_1}{M_0 - M_1} \cdot 0 \\ &= \frac{\bar{R}_{\text{YMA}}(K, N, M_1)}{\underline{R}_{\text{worst}}(K, N, M_1)} \\ &= \Xi(K, N, M_1), \end{aligned} \quad (59)$$

and  $\Xi(K, N, M)$  is constant over  $[M_1, M_0]$ . So, trivially the maximum is achieved for  $M = M_1$ . Combined with (58), this yields:

$$\max_{M \in [0, N]} \xi(K, N, M) \leq \max_{\ell \in \bar{N}} \Xi(K, N, M_\ell). \quad (60)$$

We continue to bound the right-hand side of (60). Irrespective of  $K, N \in \mathbb{Z}^+$ , we have:

$$\Xi(K, N, M_{\bar{N}}) = \frac{\bar{N}}{\bar{N}} = 1. \quad (61)$$

When  $\bar{N} = 1$  (i.e., only one file or only one user), Inequalities (60) and (61) imply that the gap  $\xi(K, N, M) = 1$  for all  $M \in [0, N]$ , and hence our lower bound is exact.

<sup>6</sup>A linear-fractional function is always quasiconvex [57].

We therefore assume in the following that  $\bar{N} \geq 2$ . For  $\ell \in \{1, \dots, \bar{N} - 1\}$ , we have

$$\begin{aligned} \Xi(K, N, M_\ell) &= \frac{\frac{N - \frac{N-\ell}{\ell+1}}{\frac{N-\ell}{\ell+1}} \left( 1 - \left( 1 - \frac{1}{N} \frac{N-\ell}{\ell+1} \right)^{\bar{N}} \right)}{\sum_{j=1}^{\ell} \left( 1 - \frac{j}{N-j+1} \frac{N-\ell}{\ell+1} \right)} \\ &= \frac{\frac{\ell(N+1)}{N-\ell} \left( 1 - \left( \frac{\ell(N+1)}{(\ell+1)N} \right)^{\bar{N}} \right)}{\sum_{j=1}^{\ell} \left( 1 + \left( 1 - \frac{N+1}{N-j+1} \frac{N-\ell}{\ell+1} \right) \right)} \\ &= \frac{\frac{\ell(N+1)}{N-\ell} \left( 1 - \left( \frac{\ell(N+1)}{(\ell+1)N} \right)^{\bar{N}} \right)}{\frac{\frac{\ell(N+1)}{\ell+1} - \frac{(N-\ell)(N+1)}{\ell+1} \sum_{j=1}^{\ell} \frac{1}{N-j+1}}{\ell - (N-\ell) \sum_{j=1}^{\ell} \frac{1}{N-j+1}}} \\ &= \frac{\frac{\ell(\ell+1)}{N-\ell} \left( 1 - \left( \frac{\ell(N+1)}{(\ell+1)N} \right)^{\bar{N}} \right)}{\ell - (N-\ell) \sum_{j=1}^{\ell} \frac{1}{N-j+1}} \\ &\leq \frac{\frac{\ell(\ell+1)}{N-\ell} \left( 1 - \left( \frac{\ell(N+1)}{(\ell+1)N} \right)^{\bar{N}} \right)}{\ell - (N-\ell) \sum_{j=1}^{\ell} \frac{1}{N-j+1}} \\ &\stackrel{a=\ell/N}{=} \frac{\frac{a(\ell+1)}{(1-a)\ell} \left( 1 - \left( \frac{\ell+a}{\ell+1} \right)^{\ell/a} \right)}{1 - (1-a) \sum_{j=0}^{\ell-1} \frac{1}{\ell-a+j}} =: \phi(a, \ell). \end{aligned} \quad (62)$$

Note that since  $\ell \in \{1, \dots, \bar{N} - 1\}$ ,

$$a \in [1/N, 1). \quad (63)$$

Therefore,

$$\max_{K \in \mathbb{Z}^+} \max_{N \in \mathbb{Z}^+} \max_{M \in [0, N]} \xi(K, N, M) \leq \max_{\ell \in \mathbb{Z}^+} \max_{a \in (0, 1)} \phi(a, \ell), \quad (64)$$

which concludes the proof.

## B. Proof of Inequality (19)

We have a closer look at the denominator of the function  $\phi(a, \ell)$ .

Notice that  $\frac{1}{n} \leq \int_{n-1}^n \frac{dt}{t}$  for all  $n \geq 2$ . Therefore,

$$\begin{aligned} \sum_{j=0}^{\ell-1} \frac{1}{\ell - aj} &= \sum_{j=0}^{\ell-1} \frac{1}{\ell(1-a) + a(\ell-j)} \\ &\stackrel{i=\ell-j}{=} \sum_{i=1}^{\ell} \frac{1}{\ell(1-a) + ai} \\ &= \frac{1}{\ell(1-a) + a} + \frac{1}{a} \sum_{i=2}^{\ell} \frac{1}{\ell(1-a)/a + i} \\ &\leq \frac{1}{\ell(1-a) + a} + \frac{1}{a} \sum_{i=2}^{\ell} \int_{\ell(1-a)/a + i-1}^{\ell(1-a)/a + i} \frac{1}{t} dt \\ &= \frac{1}{\ell(1-a) + a} + \frac{1}{a} \int_{\ell(1-a)/a+1}^{\ell(1-a)/a+\ell} \frac{1}{t} dt \\ &= \frac{1}{\ell(1-a) + a} + \frac{1}{a} \ln \left( \frac{\ell}{\ell(1-a) + a} \right) \end{aligned} \quad (65)$$

We use (65) to upper bound the function  $\phi(a, \ell)$ :

$$\begin{aligned} \phi(a, \ell) &= \frac{\frac{a(\ell+1)}{(1-a)^\ell} \left(1 - \left(\frac{\ell+a}{\ell+1}\right)^{\ell/a}\right)}{1 - (1-a) \sum_{j=0}^{\ell-1} \frac{1}{\ell-aj}} \\ &\leq \frac{\frac{a(\ell+1)}{(1-a)^\ell} \left(1 - \left(\frac{\ell+a}{\ell+1}\right)^{\ell/a}\right)}{1 - \frac{1-a}{\ell(1-a)+a} - \frac{1-a}{a} \ln \left(\frac{\ell}{\ell(1-a)+a}\right)} \\ &\stackrel{b=1/\ell}{=} \frac{\frac{a(1+b)}{1-a} \left(1 - \left(\frac{1+ab}{1+b}\right)^{\frac{1}{ab}}\right)}{1 - \frac{(1-a)y}{1-a+ab} + \frac{1-a}{a} \ln(1-a+ab)} \\ &= \psi(a, b). \end{aligned} \quad (66)$$

Noting also that if  $\ell > 10^4$ , then  $b < 10^{-4}$ , this concludes the proof of (19).

## VI. PROOF OF THE GAP-RESULT IN THEOREM 4

We wish to uniformly bound the gap

$$\theta(K, N, M) := \frac{R_{\text{avg}}^*(M)}{R_{\text{avg}}^{\text{low}}(M)}, \quad (67)$$

irrespective of  $K, N \geq 1$  and  $M \in [0, N)$ .

Since  $R_{\text{YMA}}(K, N, M)$  upper bounds the rate-memory tradeoff for the worst case, it must also upper bound the rate-memory tradeoff for the average case. Thus, for any number of users  $K$  and files  $N$ :

$$R_{\text{avg}}^*(M) \leq R_{\text{YMA}}(K, N, M), \quad M \in [0, N). \quad (68)$$

We thus have

$$\begin{aligned} \theta(K, N, M) &\leq \frac{R_{\text{YMA}}(K, N, M)}{R_{\text{avg}}^{\text{low}}(M)} \\ &\leq \frac{R_{\text{YMA}}(K, N, M)}{\underline{R}_{\text{avg}}(K, N, M)}, \end{aligned} \quad (69)$$

where we defined

$$\underline{R}_{\text{avg}}(K, N, M) := \max_{\ell \in \bar{N}} \sum_{k=1}^{\ell} \left[ \left(1 - \frac{1}{N}\right)^{k-1} - \frac{k}{N} M \right], \quad (70)$$

and where the second inequality holds because for all  $K, N, M$ :

$$\begin{aligned} R_{\text{avg}}^{\text{low}}(M) &\geq \max_{\ell \in \bar{N}} \left[ \left(1 - \left(1 - \frac{1}{N}\right)^\ell\right) N - \frac{\ell(\ell+1)}{2N} M \right] \\ &= \max_{\ell \in \bar{N}} \sum_{k=1}^{\ell} \left[ \left(1 - \frac{1}{N}\right)^{k-1} - \frac{k}{N} M \right]. \end{aligned} \quad (71)$$

Define

$$\tilde{M}_\ell := \begin{cases} \frac{N}{\ell+1} \left(1 - \frac{1}{N}\right)^\ell & \text{if } \ell \in \{0, 1, \dots, \bar{N}-1\}, \\ 0 & \text{if } \ell = \bar{N}. \end{cases} \quad (72)$$

and note that  $0 = \tilde{M}_{\bar{N}} < \tilde{M}_{\bar{N}-1} < \dots < \tilde{M}_0 = N$ . Using similar arguments as in Appendix A, it can be shown that function  $\underline{R}_{\text{avg}}(K, N, M)$  is piecewise-linear with  $\bar{N}$  line segments over the intervals

$$[\tilde{M}_\ell, \tilde{M}_{\ell-1}], \quad \ell \in \{2, \dots, \bar{N}\} \quad (73a)$$

$$[\tilde{M}_1, \tilde{M}_0). \quad (73b)$$

We next upper bound  $R_{\text{YMA}}(K, N, M)$  by a function  $\bar{R}_{\text{YMA}}(K, N, M)$  that is piecewise-linear over the same intervals (73). Specifically, for every  $\ell \in \bar{N}$ , define for  $M \in [\tilde{M}_\ell, \tilde{M}_{\ell-1})$ :

$$\begin{aligned} \bar{R}_{\text{YMA}}(K, N, M) &:= \frac{\tilde{M}_{\ell-1} - M}{\tilde{M}_{\ell-1} - \tilde{M}_\ell} R_{\text{YMA}}(K, N, \tilde{M}_\ell) \\ &\quad + \frac{M - \tilde{M}_\ell}{\tilde{M}_{\ell-1} - \tilde{M}_\ell} R_{\text{YMA}}(K, N, \tilde{M}_{\ell-1}). \end{aligned} \quad (74)$$

Notice that

$$\bar{R}_{\text{YMA}}(K, N, \tilde{M}_\ell) = R_{\text{YMA}}(K, N, \tilde{M}_\ell), \quad \forall \ell \in \{1, \dots, \bar{N}\},$$

whereas for general  $M \in [0, N)$ :

$$\bar{R}_{\text{YMA}}(K, N, M) \geq R_{\text{YMA}}(K, N, M), \quad (75)$$

because  $R_{\text{YMA}}(K, N, M)$  is convex in  $M$  for fixed  $K, N$ .

Plugging (75) into (69), we obtain:

$$\theta(K, N, M) \leq \frac{\bar{R}_{\text{YMA}}(K, N, M)}{\underline{R}_{\text{avg}}(K, N, M)} =: \Theta(K, N, M). \quad (76)$$

Following similar arguments as in the proof of Theorem 2, we have

$$\max_{M \in [0, N)} \theta(K, N, M) \leq \max_{\ell \in \bar{N}} \Theta(K, N, \tilde{M}_\ell). \quad (77)$$

We continue to bound the right-hand side of (77). Irrespective of  $K, N \in \mathbb{Z}^+$ , we have:

$$\begin{aligned} \Theta(K, N, \tilde{M}_{\bar{N}}) &= \frac{\bar{N}}{N(1 - (1 - 1/N)^{\bar{N}})} \\ &= \frac{x}{1 - ((1 - 1/N)^N)^x} \Big|_{x=\bar{N}/N} \\ &\stackrel{(a)}{\leq} \frac{x}{1 - e^{-x}} \Big|_{x=\bar{N}/N} \\ &\stackrel{(b)}{\leq} \frac{1}{1 - e^{-1}}, \end{aligned} \quad (78)$$

where (a) follows since  $(1 - 1/\zeta)^\zeta \leq e^{-1}$  for all  $\zeta > 1$  and (b) follows since  $x \mapsto \frac{x}{1 - e^{-x}}$  is an increasing function. This implies that when  $\bar{N} = 1$  (i.e., one file or one user), then  $\Theta(K, N, M) \leq \frac{1}{1 - e^{-1}} \leq 1.582$  for all  $M \in [0, N)$ .

In the following, we assume that  $\bar{N} \geq 2$ . As for  $\ell \in \{1, \dots, \bar{N}-1\}$ , we have

$$\begin{aligned} \Theta(K, N, \tilde{M}_\ell) &= \frac{N - \frac{N}{\ell+1} \left(1 - \frac{1}{N}\right)^\ell}{\frac{N}{\ell+1} \left(1 - \frac{1}{N}\right)^\ell} \left(1 - \left(1 - \frac{1}{N} \frac{N}{\ell+1} \left(1 - \frac{1}{N}\right)^\ell\right)^{\bar{N}}\right) \\ &= \frac{N - (N + \frac{\ell}{2}) \left(1 - \frac{1}{N}\right)^\ell}{\left(1 - \frac{1}{N}\right)^\ell \left(1 - \left(1 + \frac{\ell}{2N}\right) \left(1 - \frac{1}{N}\right)^\ell\right)} \\ &= \frac{\left(\frac{\ell+1}{N} - \frac{1}{N} \left(1 - \frac{1}{N}\right)^\ell\right) \left(1 - \left(1 - \frac{1}{\ell+1} \left(1 - \frac{1}{N}\right)^\ell\right)^{\bar{N}}\right)}{\left(1 - \frac{1}{N}\right)^\ell \left(1 - \left(1 + \frac{\ell}{2N}\right) \left(1 - \frac{1}{N}\right)^\ell\right)} \\ &\leq \frac{\left(\frac{\ell+1}{N} - \frac{1}{N} \left(1 - \frac{1}{N}\right)^\ell\right) \left(1 - \left(1 - \frac{1}{\ell+1} \left(1 - \frac{1}{N}\right)^\ell\right)^{\bar{N}}\right)}{\left(1 - \frac{1}{N}\right)^\ell \left(1 - \left(1 + \frac{\ell}{2N}\right) \left(1 - \frac{1}{N}\right)^\ell\right)} \end{aligned}$$

$$\stackrel{(a)}{=} \frac{\left(u + v - v(1-v)^{u/v}\right) \left(1 - \left(1 - \frac{v}{u+v}(1-v)^{u/v}\right)^{1/v}\right)}{(1-v)^{u/v} \left(1 - \left(1 + \frac{u}{2}\right)(1-v)^{u/v}\right)}$$

$$=: \eta(u, v), \quad (79)$$

where (a) follows by a change of variable  $u = \ell/N$  and  $v = 1/N$ . Note that since  $\ell \in \{1, \dots, \bar{N} - 1\}$  and assuming  $\bar{N} \geq 2$ ,

$$u \in \left[\frac{1}{\bar{N}}, \frac{\bar{N} - 1}{\bar{N}}\right] \text{ and } v \in (0, 1/2]. \quad (80)$$

Also, it holds that  $\frac{1}{1-e^{-1}} < \max_{u \in (0,1]} \max_{v \in (0,1/2]} \eta(u, v)$ . Therefore,

$$\max_{K \in \mathbb{Z}^+} \max_{N \in \mathbb{Z}^+} \max_{M \in [0, N]} \theta(K, N, M) \leq \max_{u \in (0,1]} \max_{v \in (0,1/2]} \eta(u, v), \quad (81)$$

which concludes the proof.

## VII. CONCLUSION

This paper derives new lower bounds on the rate-memory tradeoff under a worst-case or an average-case scenario. The obtained lower bounds are compared to upper bounds on the rate-memory tradeoffs in decentralized caching scenarios and shown to match these upper bounds up to a multiplicative gap of at most 2.315 (in the worst-case scenario) and 2.507 (in the average-case scenario). Previous bounds could establish a gap of 4.7 and 4, respectively. In a work that is parallel to this [14], improved upper bounds were presented which match the decentralized upper bounds up to factors of almost 2. The bounds in [14] are based on similar technical steps as used in this paper. The improvement is obtained through an additional averaging step over the labeling of the receivers.

The converse technique presented in this paper can be extended to setups where delivery communication takes place over a noisy broadcast channel (BC) (rather than a noise-free link as considered in this paper). Corresponding bounds on the capacity-memory tradeoff over general discrete memoryless BCs can be found in [55].

### APPENDIX A PROOF OF EQUATION (52)

Fix  $i \in \bar{\mathcal{N}}$  and  $M \in [M_i, M_{i-1}]$ . Define

$$\nu_j := \frac{1}{M_{j-1}} = \frac{k}{N - j + 1}, \quad j \in \{1, \dots, \bar{N}\}. \quad (82)$$

Notice that

$$\nu_1 \leq \nu_2 \leq \dots \leq \nu_{\bar{N}} \quad (83)$$

and

$$\frac{1}{\nu_{i+1}} \leq M \leq \frac{1}{\nu_i}. \quad (84)$$

Rewrite  $R_{\text{worst}}(K, N, M)$  as

$$R_{\text{worst}}(K, N, M) = \max_{\ell \in \bar{\mathcal{N}}} \left[ \sum_{j=1}^{\ell} (1 - M \cdot \nu_j) \right]. \quad (85)$$

By (83) and (84), the summands  $(1 - M \cdot \nu_j)$  are positive or zero for all  $j \leq i$  and they are negative  $j > i$ . The maximum in (85) is thus achieved by choosing  $\ell = i$ . This proves equation (52).

### APPENDIX B PROOF OF LEMMA 3

We first prove that for each  $k \in \{1, \dots, \ell\}$ :

$$\alpha_k \leq \frac{kM}{N - k + 1}, \quad (86)$$

which establishes the upper bound

$$\sum_{k=1}^{\ell} \alpha_k \leq \sum_{k=1}^{\ell} \frac{kM}{N - k + 1}. \quad (87)$$

For each partial demand vector  $\tilde{\mathbf{d}} = (d_1, \dots, d_{k-1})$ , let  $W_{\tilde{\mathbf{d}}} := \{W_{d_1}, \dots, W_{d_{k-1}}\}$ . We have:

$$\begin{aligned} & F\alpha_k \\ &= \frac{1}{\ell! \binom{N}{\ell}} \sum_{\mathbf{d} \in \mathcal{Q}_{\ell}^{\text{dist}}} I(W_{d_k}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{d_1}, \dots, W_{d_{k-1}}) \\ &= \frac{1}{\ell! \binom{N}{\ell}} \sum_{\tilde{\mathbf{d}} \in \mathcal{Q}_{k-1}^{\text{dist}}} \sum_{\substack{\mathbf{d} \in \mathcal{Q}_{\ell}^{\text{dist}} \\ (d_1, \dots, d_{k-1}) = \tilde{\mathbf{d}}}} I(W_{d_k}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{\tilde{\mathbf{d}}}) \\ &\stackrel{(a)}{=} \frac{1}{\ell! \binom{N}{\ell}} \sum_{\tilde{\mathbf{d}} \in \mathcal{Q}_{k-1}^{\text{dist}}} \sum_{j \in \mathcal{N} \setminus \tilde{\mathbf{d}}} I(W_j; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{\tilde{\mathbf{d}}}) \\ &\quad \cdot \binom{N-k}{\ell-k} (\ell-k)! \\ &= \frac{1}{k! \binom{N}{k}} \sum_{\tilde{\mathbf{d}} \in \mathcal{Q}_{k-1}^{\text{dist}}} \sum_{j \in \mathcal{N} \setminus \tilde{\mathbf{d}}} I(W_j; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{\tilde{\mathbf{d}}}) \\ &\stackrel{(b)}{=} \frac{1}{k! \binom{N}{k}} \left[ \sum_{\tilde{\mathbf{d}} \in \mathcal{Q}_{k-1}^{\text{dist}}} \left[ H(\{W_j : j \in \mathcal{N} \setminus \tilde{\mathbf{d}}\} | W_{\tilde{\mathbf{d}}}) - \right. \right. \\ &\quad \left. \left. \sum_{j \in \mathcal{N} \setminus \tilde{\mathbf{d}}} H(W_j | \mathbb{V}_1, \dots, \mathbb{V}_k, W_{\tilde{\mathbf{d}}}) \right] \right] \\ &\stackrel{(c)}{\leq} \frac{1}{k! \binom{N}{k}} \sum_{\tilde{\mathbf{d}} \in \mathcal{Q}_{k-1}^{\text{dist}}} I(\{W_j : j \in \mathcal{N} \setminus \tilde{\mathbf{d}}\}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{\tilde{\mathbf{d}}}) \\ &\stackrel{(d)}{\leq} \frac{(k-1)! \binom{N}{k-1}}{k! \binom{N}{k}} kFM \\ &= \frac{kFM}{N - k + 1}, \quad (88) \end{aligned}$$

where (a) holds because for each value of  $\ell$  and  $j$  there are  $\binom{N-k}{\ell-k} (\ell-k)!$  ordered demand vectors  $\mathbf{d} = (d_1, \dots, d_K) \in \mathcal{Q}_{\ell}^{\text{dist}}$  with  $(d_1, \dots, d_{k-1}) = \tilde{\mathbf{d}}$  and with  $d_k = j$ ; (b) holds by the independence of the messages; (c) holds because for any random tuple  $(A_1, \dots, A_L)$  it holds that  $\sum_{l=1}^L H(A_l) \geq H(A_1, \dots, A_L)$ ; and (d) holds because  $I(W_1, \dots, W_N; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{\tilde{\mathbf{d}}})$  cannot exceed  $kFM$ . This concludes the proof of (86) and thus of (87).

We now prove

$$\sum_{k=1}^{\ell} \alpha_k \leq \frac{\ell^2 M}{N}. \quad (89)$$

For each  $\mathbf{d} \in \mathcal{Q}_{\ell}^{\text{dist}}$ :

$$I(W_{d_1}; \mathbb{V}_1) + \sum_{k=2}^{\ell} I(W_{d_k}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{d_1}, W_{d_2}, \dots, W_{d_{k-1}})$$

$$\leq I(W_{d_1}, W_{d_2}, \dots, W_{d_\ell}; \mathbb{V}_1, \dots, \mathbb{V}_\ell). \quad (90)$$

So,

$$\begin{aligned} & F \binom{N}{\ell} \ell! \cdot \sum_{k=1}^{\ell} \alpha_k \\ &= \sum_{\mathbf{d} \in \mathcal{Q}_\ell^{\text{dist}}} \left[ I(W_{d_1}; \mathbb{V}_1) \right. \\ &\quad \left. + \sum_{k=2}^{\ell} I(W_{d_k}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{d_1}, W_{d_2}, \dots, W_{d_{k-1}}) \right] \\ &\leq \sum_{\mathbf{d} \in \mathcal{Q}_\ell^{\text{dist}}} I(W_{d_1}, W_{d_2}, \dots, W_{d_\ell}; \mathbb{V}_1, \dots, \mathbb{V}_\ell) \\ &= \ell! \sum_{\substack{\mathbf{d} \in \mathcal{Q}_\ell^{\text{dist}}: \\ d_1 < d_2 < \dots < d_\ell}} I(W_{d_1}, W_{d_2}, \dots, W_{d_\ell}; \mathbb{V}_1, \dots, \mathbb{V}_\ell) \\ &\stackrel{(a)}{\leq} \ell! \binom{N}{\ell} \frac{\ell}{N} I(W_1, \dots, W_N; \mathbb{V}_1, \dots, \mathbb{V}_\ell) \\ &\leq \frac{\ell}{N} \ell! \binom{N}{\ell} \ell FM, \end{aligned} \quad (91)$$

where (a) follows by Lemma 2.

#### APPENDIX C PROOF OF LEMMA 5

We first prove that for each  $k \in \{1, \dots, \ell\}$ :

$$\beta_k \leq \frac{kM}{N}, \quad (92)$$

which establishes the upper bound

$$\sum_{k=1}^{\ell} \beta_k \leq \sum_{k=1}^{\ell} \frac{kM}{N}. \quad (93)$$

Defining  $\mathbf{D}_k := (D_1, \dots, D_k)$ , we have:

$$\begin{aligned} F\beta_k &= I(W_{D_k}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{D_1}, \dots, W_{D_{k-1}}, \mathbf{D}) \\ &= I(W_{D_k}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{D_1}, \dots, W_{D_{k-1}}, \mathbf{D}_k) \\ &\stackrel{(a)}{=} \frac{1}{N^k} \sum_{\mathbf{d} \in \mathcal{Q}_k^{\text{rep}}} I(W_{d_k}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{d_1}, \dots, W_{d_{k-1}}) \\ &= \frac{1}{N^k} \sum_{\tilde{\mathbf{d}} \in \mathcal{Q}_{k-1}^{\text{rep}}} \sum_{j=1}^N I(W_j; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{\tilde{\mathbf{d}}}) \\ &\stackrel{(b)}{\leq} \frac{1}{N^k} \sum_{\tilde{\mathbf{d}} \in \mathcal{Q}_{k-1}^{\text{rep}}} I(W_1, \dots, W_N; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{\tilde{\mathbf{d}}}) \\ &\stackrel{(c)}{\leq} \frac{1}{N^k} \sum_{\tilde{\mathbf{d}} \in \mathcal{Q}_{k-1}^{\text{rep}}} kFM \\ &= \frac{kFM}{N}, \end{aligned} \quad (94)$$

where (a) holds by writing out the conditioning on  $\mathbf{D}_k$  in form of an expectation; (b) holds because the messages are independent and because  $H(A_1, \dots, A_L) \leq \sum_{l=1}^L H(A_l)$  for any random  $L$ -tuple  $(A_1, \dots, A_L)$ ; and (c) holds because

$I(W_1, \dots, W_N; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{\tilde{\mathbf{d}}})$  cannot exceed  $kFM$ . This concludes the proof of (92) and thus (93).

We now prove

$$\sum_{k=1}^{\ell} \beta_k \leq \mathbb{E}_{\mathbf{D}} [\kappa_{\mathbf{D}}(\ell)] \cdot \frac{\ell M}{N}. \quad (95)$$

Let  $\mathbf{D}_\ell^{\text{dist}}$  be a vector containing all distinct elements of  $\mathbf{D}_\ell := (D_1, \dots, D_\ell)$ . Notice that  $\mathbf{D}_\ell^{\text{dist}}$  is of length  $\kappa_{\mathbf{D}_\ell}(\ell)$ . Also, following the definition of the previous section,  $W_{\mathbf{D}_\ell} := \{W_{D_1}, \dots, W_{D_\ell}\} = W_{\mathbf{D}_\ell^{\text{dist}}}$ . We have:

$$\begin{aligned} & F \sum_{k=1}^{\ell} \beta_k \\ &= I(W_{D_1}; \mathbb{V}_1 | \mathbf{D}) \\ &\quad + \sum_{k=2}^{\ell} I(W_{D_k}; \mathbb{V}_1, \dots, \mathbb{V}_k | W_{D_1}, W_{D_2}, \dots, W_{D_{k-1}}, \mathbf{D}) \\ &\leq I(W_{\mathbf{D}_\ell}; \mathbb{V}_1, \dots, \mathbb{V}_\ell | \mathbf{D}) \\ &\leq I(W_{\mathbf{D}_\ell}; \mathbb{V}_1, \dots, \mathbb{V}_\ell | \mathbf{D}_\ell) \\ &= I(W_{\mathbf{D}_\ell}; \mathbb{V}_1, \dots, \mathbb{V}_\ell | \mathbf{D}_\ell, \kappa_{\mathbf{D}_\ell}(\ell)) \\ &= \sum_{i=1}^{\ell} \mathbb{P}(\kappa_{\mathbf{D}_\ell}(\ell) = i) I(W_{\mathbf{D}_\ell}; \mathbb{V}_1, \dots, \mathbb{V}_\ell | \mathbf{D}_\ell, \kappa_{\mathbf{D}_\ell}(\ell) = i) \\ &\stackrel{(a)}{=} \sum_{i=1}^{\ell} \mathbb{P}(\kappa_{\mathbf{D}_\ell}(\ell) = i) I(W_{\mathbf{D}_\ell^{\text{dist}}}; \mathbb{V}_1, \dots, \mathbb{V}_\ell | \mathbf{D}_\ell^{\text{dist}}, \kappa_{\mathbf{D}_\ell}(\ell) = i) \\ &\stackrel{(b)}{=} \sum_{i=1}^{\ell} \mathbb{P}(\kappa_{\mathbf{D}_\ell}(\ell) = i) \sum_{\substack{\tilde{\mathbf{d}} \in \mathcal{Q}_i^{\text{dist}}: \\ \tilde{d}_1 < \tilde{d}_2 < \dots < \tilde{d}_i}} \frac{1}{\binom{N}{i}} I(W_{\tilde{\mathbf{d}}}; \mathbb{V}_1, \dots, \mathbb{V}_\ell) \\ &\stackrel{(c)}{\leq} \sum_{i=1}^{\ell} \mathbb{P}(\kappa_{\mathbf{D}_\ell}(\ell) = i) \frac{i}{N} I(W_1, \dots, W_N; \mathbb{V}_1, \dots, \mathbb{V}_\ell) \\ &\stackrel{(d)}{\leq} \sum_{i=1}^{\ell} \mathbb{P}(\kappa_{\mathbf{D}_\ell}(\ell) = i) \cdot i \cdot \frac{\ell FM}{N} \\ &= \mathbb{E}_{\mathbf{D}} [\kappa_{\mathbf{D}}(\ell)] \cdot \frac{\ell FM}{N}. \end{aligned} \quad (96)$$

Notice that here (a) holds because for the involved mutual informations only the set of distinct demands matters and not the exact demand vector. Moreover, (b) holds because given  $\kappa_{\mathbf{D}_\ell}(\ell) = i$  the probability that  $\mathbf{D}_\ell^{\text{dist}}$  equals a specific vector  $\tilde{\mathbf{d}} \in \mathcal{Q}_i^{\text{dist}}$  equals  $\binom{N}{i}^{-1}$ ; (c) follows from Lemma 2; and (d) follows since  $I(W_1, \dots, W_N; \mathbb{V}_1, \dots, \mathbb{V}_\ell)$  cannot be larger than  $\ell FM$ .

#### REFERENCES

- [1] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.
- [2] U. Niesen and M. A. Maddah-Ali, "Coded caching with nonuniform demands," *IEEE Trans. Inf. Theory*, vol. 63, no. 2, pp. 1146–1158, Feb 2017.
- [3] Z. Chen, P. Fan, and K. B. Letaief, "Fundamental limits of caching: improved bounds for users with small buffers," *IET Commun.*, vol. 10, no. 17, pp. 2315–2318, 2016.
- [4] C. Tian, "A note on the fundamental limits of coded caching," *CoRR*, vol. abs/1503.00010, 2015. [Online]. Available: <http://arxiv.org/abs/1503.00010>

- [5] K. Wan, D. Tuninetti, and P. Piantanida, "On caching with more users than files," in *IEEE Int. Symp. Inf. Theory*, July 2016, pp. 135–139.
- [6] S. Sahraei and M. Gastpar, "K users caching two files: An improved achievable rate," in *2016 Annual Conf. Inf. Science and Syst.*, March 2016, pp. 620–624.
- [7] C. Tian and J. Chen, "Caching and delivery via interference elimination," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1548–1560, March 2018.
- [8] M. M. Amiri, Q. Yang, and D. Gündüz, "Coded caching for a large number of users," in *IEEE Inf. Theory Workshop*, Sept 2016, pp. 171–175.
- [9] Q. Yu, M. A. Maddah-Ali, and A. S. Avestimehr, "The exact rate-memory tradeoff for caching with uncoded prefetching," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 1281–1296, Feb 2018.
- [10] J. Gómez-Vilardebó, "Fundamental limits of caching: Improved bounds with coded prefetching," *IEEE Trans. Commun.*, vol. 66, 2018.
- [11] H. Ghasemi and A. Ramamoorthy, "Improved lower bounds for coded caching," *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4388–4413, July 2017.
- [12] A. Sengupta, R. Tandon, and T. C. Clancy, "Improved approximation of storage-rate tradeoff for caching via new outer bounds," in *IEEE Int. Symp. Inf. Theory*, June 2015, pp. 1691–1695.
- [13] C. Y. Wang, S. H. Lim, and M. Gastpar, "A new converse bound for coded caching," in *IEEE Inf. Theory and Applications Workshop*, Jan 2016, pp. 1–6.
- [14] Q. Yu, M. A. Maddah-Ali, and A. S. Avestimehr, "Characterizing the rate-memory tradeoff in cache networks within a factor of 2," in *IEEE Int. Symp. Inf. Theory*, June 2017, pp. 386–390.
- [15] R. Pedarsani, M. A. Maddah-Ali, and U. Niesen, "Online coded caching," *IEEE/ACM Trans. Netw.*, vol. 24, no. 2, pp. 836–845, April 2016.
- [16] M. Ji, A. M. Tulino, J. Llorca, and G. Caire, "Order-optimal rate of caching and coded multicasting with random demands," *IEEE Trans. Inf. Theory*, vol. 63, no. 6, pp. 3923–3949, June 2017.
- [17] J. Zhang, X. Lin, and X. Wang, "Coded caching under arbitrary popularity distributions," *IEEE Trans. Inf. Theory*, vol. 64, no. 1, pp. 349–366, Jan 2018.
- [18] J. Hachem, N. Karamchandani, and S. N. Diggavi, "Coded caching for multi-level popularity and access," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 3108–3141, May 2017.
- [19] R. Timo, S. Saeedi Bidokhti, M. Wigger, and B. C. Geiger, "A rate-distortion approach to caching," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1957–1976, March 2018.
- [20] P. Hassanzadeh, A. M. Tulino, J. Llorca, and E. Erkip, "Rate-memory trade-off for the two-user broadcast caching network with correlated sources," in *IEEE Int. Symp. Inf. Theory*, June 2017, pp. 1227–1231.
- [21] P. Hassanzadeh, A. Tulino, J. Llorca, and E. Erkip, "Correlation-aware distributed caching and coded delivery," in *IEEE Inf. Theory Workshop*, Sept 2016, pp. 166–170.
- [22] —, "Cache-aided coded multicast for correlated sources," in *Int. Symp. Turbo Codes and Iterative Inf. Process.*, Sept 2016, pp. 360–364.
- [23] C. Y. Wang, S. H. Lim, and M. Gastpar, "Information-theoretic caching: Sequential coding for computing," *IEEE Trans. Inf. Theory*, vol. 62, no. 11, pp. 6393–6406, Nov 2016.
- [24] Q. Yang and D. Gündüz, "Centralized coded caching of correlated contents," *CoRR*, vol. abs/1711.03798, 2017. [Online]. Available: <http://arxiv.org/abs/1711.03798>
- [25] Q. Yang, P. Hassanzadeh, D. Gündüz, and E. Erkip, "Centralized caching and delivery of correlated contents over a Gaussian broadcast channel," in *Int. Symp. on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, May 2018, pp. 1–6.
- [26] W. Huang, S. Wang, L. Ding, F. Yang, and W. Zhang, "The performance analysis of coded cache in wireless fading channel," *CoRR*, vol. abs/1504.01452, 2015. [Online]. Available: <http://arxiv.org/abs/1504.01452>
- [27] R. Timo and M. Wigger, "Joint cache-channel coding over erasure broadcast channels," in *Int. Symp. Wireless Commun. Syst.*, Aug 2015, pp. 201–205.
- [28] S. Saeedi Bidokhti, M. Wigger, and R. Timo, "Noisy broadcast networks with receiver caching," *IEEE Trans. Inf. Theory*, 2018, to appear.
- [29] S. Saeedi Bidokhti, M. Wigger, and R. Timo, "An upper bound on the capacity-memory tradeoff of degraded broadcast channels," in *Int. Symp. Turbo Codes and Iterative Inf. Process.*, Sept 2016, pp. 350–354.
- [30] A. S. Cacciapuoti, M. Caleffi, M. Ji, J. Llorca, and A. M. Tulino, "Speeding up future video distribution via channel-aware caching-aided coded multicast," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 8, pp. 2207–2218, Aug 2016.
- [31] M. M. Amiri and D. Gündüz, "Cache-aided data delivery over erasure broadcast channels," in *IEEE Int. Conf. Commun.*, May 2017, pp. 1–6.
- [32] —, "Decentralized caching and coded delivery over Gaussian broadcast channels," in *2017 IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 2785–2789.
- [33] P. Hassanzadeh, E. Erkip, J. Llorca, and A. Tulino, "Distortion-memory tradeoffs in cache-aided wireless video delivery," in *Allerton Conf. on Commun., Control, and Computing*, Sept 2015, pp. 1150–1157.
- [34] A. Ghorbel, M. Kobayashi, and S. Yang, "Content delivery in erasure broadcast channels with cache and feedback," *IEEE Trans. Inf. Theory*, vol. 62, no. 11, pp. 6407–6422, Nov 2016.
- [35] J. Zhang, F. Engelmann, and P. Elia, "Coded caching for reducing CSIT-feedback in wireless communications," in *Allerton Conf. on Commun., Control, and Computing*, Sept 2015, pp. 1099–1105.
- [36] J. Zhang and P. Elia, "Fundamental limits of cache-aided wireless BC: Interplay of coded-caching and CSIT feedback," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 3142–3160, May 2017.
- [37] —, "Feedback-aided coded caching for the MISO BC with small caches," in *IEEE Int. Conf. Commun.*, May 2017, pp. 1–6.
- [38] —, "Wireless coded caching: A topological perspective," in *IEEE Int. Symp. Inf. Theory*, June 2017, pp. 401–405.
- [39] S. Yang, K. H. Ngo, and M. Kobayashi, "Content delivery with coded caching and massive MIMO in 5G," in *Int. Symp. Turbo Codes and Iterative Inf. Process.*, Sept 2016, pp. 370–374.
- [40] S. P. Shariatpanahi, G. Caire, and B. H. Khalaj, "Physical-layer schemes for wireless coded caching," *CoRR*, vol. abs/1711.05969, 2017. [Online]. Available: <http://arxiv.org/abs/1711.05969>
- [41] M. A. Maddah-Ali and U. Niesen, "Cache-aided interference channels," in *IEEE Int. Symp. Inf. Theory*, June 2015, pp. 809–813.
- [42] N. Naderializadeh, M. A. Maddah-Ali, and A. S. Avestimehr, "Fundamental limits of cache-aided interference management," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 3092–3107, May 2017.
- [43] A. Sengupta, R. Tandon, and O. Simeone, "Cache aided wireless networks: Tradeoffs between storage and latency," in *2016 Annual Conf. Inf. Science and Syst.*, March 2016.
- [44] N. Karamchandani, U. Niesen, M. A. Maddah-Ali, and S. Diggavi, "Hierarchical coded caching," in *IEEE Int. Symp. Inf. Theory*, July 2014, pp. 2142–2146.
- [45] S. P. Shariatpanahi, S. A. Motahari, and B. H. Khalaj, "Multi-server coded caching," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7253–7271, Dec 2016.
- [46] K. Shanmugam, N. Golrezaei, A. G. Dimakis, A. F. Molisch, and G. Caire, "Femtocaching: Wireless content delivery through distributed caching helpers," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8402–8413, Dec 2013.
- [47] V. Ntranos, M. A. Maddah-Ali, and G. Caire, "Cellular interference alignment: Omni-directional antennas and asymmetric configurations," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6663–6679, Dec 2015.
- [48] Y. Ugur, Z. H. Awan, and A. Sezgin, "Cloud radio access networks with coded caching," in *Int. ITG Workshop Smart Antennas*, March 2016, pp. 1–5.
- [49] X. Yi and G. Caire, "Topological coded caching," in *IEEE Int. Symp. Inf. Theory*, July 2016, pp. 2039–2043.
- [50] S. H. Park, O. Simeone, and S. S. Shitz, "Joint optimization of cloud and edge processing for fog radio access networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7621–7632, Nov 2016.
- [51] R. Tandon and O. Simeone, "Cloud-aided wireless networks with edge caching: Fundamental latency trade-offs in fog radio access networks," in *IEEE Int. Symp. Inf. Theory*, July 2016, pp. 2029–2033.
- [52] B. Azari, O. Simeone, U. Spagnolini, and A. M. Tulino, "Hypergraph-based analysis of clustered co-operative beamforming with application to edge caching," *IEEE Wireless Commun. Letters*, vol. 5, no. 1, pp. 84–87, Feb 2016.
- [53] B. Blaszczyk and A. Giovanidis, "Optimal geographic caching in cellular networks," in *IEEE Int. Conf. Commun.*, June 2015, pp. 3358–3363.
- [54] Y. Fadlallah, A. M. Tulino, D. Barone, G. Vettigli, J. Llorca, and J. M. Gorce, "Coding for caching in 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 106–113, February 2017.
- [55] S. Saeedi Bidokhti, M. Wigger, and A. Yener, "Benefits of cache assignment on degraded broadcast channels," *CoRR*, vol. abs/1702.08044, 2017. [Online]. Available: <http://arxiv.org/abs/1702.08044>
- [56] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.

- [57] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

**Chien-Yi Wang** received the B.S. degree in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2007, the M.S. degree in electronics engineering from National Taiwan University, Taipei, Taiwan, in 2010, and the PhD degree in Computer and Communication Sciences from the École Polytechnique Fédérale (EPFL), Lausanne, Switzerland, in 2015. During the academic year 2009-2010, he was an exchange student at RWTH Aachen University, Aachen, Germany. In 2016, he was a postdoctoral researcher in the Communications and Electronics Department at Telecom ParisTech. Currently he is a senior engineer at MediaTek. His research interests include network information theory and wireless communications.

**Shirin Saeedi Bidokhti** (S'09-M'12) is a research assistant professor in the Department of Electrical and Systems Engineering at the University of Pennsylvania (UPenn). She received the B.Sc. degree in Electrical Engineering from University of Tehran in 2005, and the M.Sc and Ph.D. degrees in Communication Systems from École Polytechnique Fédérale de Lausanne (EPFL) in 2007 and 2012, respectively. Before joining UPenn in 2017, she was a postdoctoral fellow with the Institute for Communications Engineering at the Technische Universität München (2013-2015), a postdoctoral fellow with the Department of Electrical Engineering at Stanford University (2015-2017), and a postdoctoral research scholar with the Department of Electrical Engineering at the Pennsylvania State University (2017). She has been awarded an Advanced Postdoc Mobility Fellowship (2014) and a Prospective Researcher Fellowship (2012) both from the Swiss National Science Foundation. Her research interests include network information theory, network coding, and data compression.

**Michèle Wigger** (S'05, M'09, SM'14) received the M.Sc. degree in electrical engineering, with distinction, and the Ph.D. degree in electrical engineering both from ETH Zurich in 2003 and 2008, respectively. In 2009, she was first a post-doctoral fellow at the University of California, San Diego, USA, and then joined Telecom Paris Tech, Paris, France, where she is currently a Full Professor. Dr. Wigger has held visiting professor appointments at the Technion-Israel Institute of Technology and ETH Zurich. Dr. Wigger has previously served as an Associate Editor of the IEEE Communication Letters, and is now Associate Editor for Shannon Theory of the IEEE Transactions on Information Theory. She is currently also serving on the Board of Governors of the IEEE Information Theory Society. Dr. Wigger's research interests are in multi-terminal information theory, in particular in distributed source coding, capacities of networks, and distributed hypothesis testing.