Pareto Optimal Schemes in Coded Caching: Uncoded Prefetching

Vijith Kumar K P
Department of EEE
IIT Guwahati, INDIA
Email: vijith@iitg.ac.in

Brijesh Kumar Rai
Department of EEE
IIT Guwahati, INDIA
Email: brijesh.rai@gmail.com

Tony Jacob
Department of EEE
IIT Guwahati, INDIA
Email: tonyj@iitg.ac.in

***Abstract*—The problem of coded caching was introduced by Maddah-Ali and Niesen and has been extensively studied in recent years. The problem is fundamentally a multi-objective optimization problem where the rates achieved for each demand type is of interest and Pareto optimality is a natural framework. Under the constraint that the placement phase is uncoded, Yu et al. introduced the YMA scheme which was shown to be universal for all demand types. Vijith et al. showed that there are no universal schemes when coded placement is permitted and introduced the problem of finding Pareto optimal schemes. In this paper we study the possibility of finding schemes that dominate the YMA scheme and demonstrate, rather surprisingly, that they continue to operate at the Pareto optimal frontier of coded caching for (N, K) cache networks when $K \leq 3$. We introduce new lower bounds which partially characterize the tradeoffs between different demand types.**

***Index Terms*—Cache network, coded caching, uncoded prefetching, Pareto optimal.**

I. Introduction

The (N, K) canonical cache network, introduced by Maddah-Ali and Niesen [1], has a server with N files, $\{W_1, \ldots, W_N\}$, and K users connected to the server through a shared link. The files are assumed to be of same size, F bits, and each user is assumed to have a cache memory of size MF bits ($M \in [0, N]$), as depicted in Fig. 1. This network operates in two phases. In the first phase, called the placement phase, server fills each user's cache with coded versions of the files without knowing their future demands. After knowing the demands of each user, the server broadcasts packets to all users over the shared link and this phase is called the delivery phase. The objective of the delivery phase is to aid the recovery of the demanded files at each user, from the broadcast packets and the user's cache contents. Let the users' demands be represented by $\boldsymbol{d} = (W_{d_1}, \ldots, W_{d_K})$, where W_{d_l} is the file requested by user U_l. In response to the demand $\boldsymbol{d}$, the server broadcasts a set of packets $X_{\boldsymbol{d}}$. As Tian [2] noted, the problem of coded caching as introduced above is a symmetric one and he demonstrated that for any coded caching scheme there exists a symmetric scheme with the same or better performance. As the caches are filled in the placement phase without any knowledge of the future demands by the users, it is natural to consider that each cache's content is a symmetric function of the files. Given this symmetry, any permutation among the files demanded by the users could be satisfied at the same broadcast rate in the delivery phase. To formalize this, all demands which are permutations of each other are grouped into a demand type and the rate required for each demand type is studied. The performance of a coded caching scheme is evaluated based on the rates corresponding to each demand type. The design of symmetric schemes for coded caching is thus a multi-objective optimization problem where the rate for each demand type is an objective of interest.

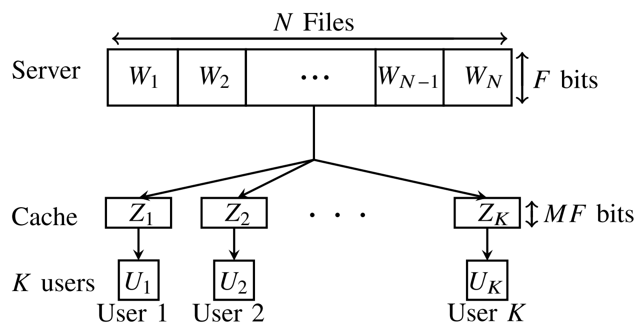

Fig. 1: The (N, K) cache network [1].

In [1], Maddah-Ali and Niesen focused on demands where each of the users request distinct files and introduced a scheme that was shown to be order optimal by deriving cut set bounds which demonstrated that the rate achieved by the scheme for these demands is within a multiplicative gap of 12 from the lower bounds. For these demands, several schemes that improve upon this scheme and several improvements to the lower bounds were obtained in the literature [4]–[10]. A modification to this scheme to address the problem of all demand types was proposed by Yu et al. and was shown to have a worst case performance within a factor of 2 from the best possible [8]. We refer to this scheme as the YMA scheme in this paper. The problem of coded caching with the placement phase restricted to be uncoded was studied by Yu et al. [11] who derived new lower bounds for each demand type under this constraint. A surprising consequence of these lower bounds was the conclusion that the YMA scheme was exactly optimal for all demand types, simultaneously. This demonstrated the existence of a universal scheme among coded caching schemes where the placement phase is uncoded. The non-existence of a universal scheme among all symmetric coded caching schemes was demonstrated by Vijith et al. [3]. In this context, they introduced the notion of dominance among coded caching schemes, where a scheme $\mathcal{A}$ said to dominate another scheme $\mathcal{B}$ if scheme $\mathcal{A}$ achieves the same or better performance as scheme $\mathcal{B}$ for all demand types and achieves strictly better performance for at least one demand type. If no scheme dominates a particular scheme, the scheme operates at the Pareto optimal frontier of the problem of coded caching. As an example, it was shown that the scheme introduced by Chen et al. [4] was Pareto optimal. In this paper, we investigate the possibility of obtaining a scheme that dominates the YMA scheme when the placement phase phase is not restricted to be uncoded. We demonstrate, rather surprisingly, that the YMA scheme continue to operate at the Pareto optimal frontier of coded caching for (N, K) cache networks when $K \leq 3$.

Consider the special case where the server has only one file, i.e., $N = 1$. The network has only one demand where all users request for the same file. Let $\boldsymbol{R}_1$ denote the rate that corresponds to this demand. We have the following result [1]:

Lemma 1. *For the $(1, K)$ cache network with cache size M, achievable rate R_1 must satisfy the following constraint*

$$M + R_1 \geq 1 \tag{1}$$

This situation happens to be trivial as there is no possibility of coding either in the placement phase or the delivery phase and the YMA scheme satisfies the constraint with equality and thus is optimal. In Section II we consider the $(N, 2)$ network for which new lower bounds and derived and are used to demonstrate that the YMA scheme is Pareto optimal. In Section III, we introduce another set of lower bounds for the $(N, 3)$ cache network and demonstrate that the YMA scheme is Pareto optimal. We conclude the paper in section IV. The key identities we use throughout the paper are

$$H(Z_k, X_{\boldsymbol{d}}) = H(W_{d_k}, Z_k, X_{\boldsymbol{d}}) \tag{2}$$

$$H(W_1, \ldots, W_N, Z_k, X_{\boldsymbol{d}}) = H(W_1, \ldots, W_N) \tag{3}$$

where (2) follows from the fact that user U_k can compute the file W_{d_k} from its cache contents Z_k and the received packet $X_{\boldsymbol{d}}$, and (3) follows from the fact that the cache contents Z_k and the broadcast packet $X_{\boldsymbol{d}}$ are functions of files $\{W_1, \ldots, W_N\}$. For $1 \leq L \leq N$, we use the notation $W_{[L]}$ to denote the set of files $\{W_1, \ldots, W_L\}$.

II. The $(N, 2)$ cache network

Consider the cache network where two users, $\{U_1, U_2\}$ are connected to a server with $N \geq 2$ files, $\{W_1, \ldots, W_N\}$. This network has two demand types, where $\boldsymbol{D}_1$ consist of the demands where both the users request the same file and $\boldsymbol{D}_2$ consists of the demands where both the users request distinct files. Let R_1 and R_2 denote the rates corresponding to $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ respectively. The rates achieved by the uncoded prefetching scheme YMA for the $(N, 2)$ cache network were characterized by Yu et al. [11] and is summarised in TABLE I where lower bounds shown in the table were derived under

Cache Size	R_1	R_2	YMA Lower Bounds
$0 \leq M \leq \frac{N}{2}$	$1 - \frac{1}{N}M$	$2 - \frac{3}{N}M$	$M + NR_1 \geq N$, $3M + NR_2 \geq 2N$
$\frac{N}{2} \leq M \leq N$	$1 - \frac{1}{N}M$	$1 - \frac{1}{N}M$	$M + NR_1 \geq N$, $M + NR_2 \geq N$

TABLE I: Rates achieved by the YMA scheme

the constraint that the placement phase is uncoded. They demonstrate that under such a constraint the YMA scheme simultaneously achieves the lower bounds for R_1 and R_2, and is thus universal. We now derive a new set of bounds, by relaxing this constraint of uncoded placement, which applies to all coded caching schemes.

Consider the set of demands $\{\boldsymbol{d}_{(p,k)} : 1 \leq p \leq 2, 1 \leq k \leq N - p + 1\}$, defined as

$$\boldsymbol{d}_{(1,k)} = (W_k, W_k) \tag{4}$$

$$\boldsymbol{d}_{(2,k)} = (W_k, W_{k+1}) \tag{5}$$

where in demand $\boldsymbol{d}_{(p,k)}$ users request p distinct files. Let $\boldsymbol{A}$ denote the set of demands

$$\boldsymbol{A} = \{\boldsymbol{d}_{(1,1)}, \boldsymbol{d}_{(1,2)}, \ldots, \boldsymbol{d}_{(1,N-2)}\} \tag{6}$$

and let $X_{\boldsymbol{A}}$ denote the set of all packets broadcast in response to these demands. As $|\boldsymbol{A}| = N - 2$, we have

$$M + (N-2)R_1 \geq H(Z_i) + H(X_{\boldsymbol{A}}) \geq H(Z_i, X_{\boldsymbol{A}}) \tag{7}$$

Note that $\boldsymbol{A} = \phi$ when $N = 2$. We now obtain the following result:

Lemma 2. *For the $(N, 2)$ cache network with cache size M, achievable rates R_1 and R_2 must satisfy the following constraints*

$$2M + (2N-3)R_1 + R_2 \geq 2N - 1 \tag{8}$$

$$M + (N-1)R_1 + R_2 \geq N \tag{9}$$

Proof. We have,

$$\begin{aligned}
&2M + (2N-3)R_1 + R_2 = 2(M + (N-2)R_1) + R_2 + R_1 \\
&\overset{(a)}{\geq} H(Z_1, X_{\boldsymbol{A}}) + H(Z_2, X_{\boldsymbol{A}}) + H(X_{\boldsymbol{d}_{(2,N-1)}}) + H(X_{\boldsymbol{d}_{(1,N-1)}}) \\
&\overset{(b)}{\geq} H(Z_1, X_{\boldsymbol{d}_{(2,N-1)}}, X_{\boldsymbol{A}}) + H(Z_2, X_{\boldsymbol{A}}, X_{\boldsymbol{d}_{(1,N-1)}}) \\
&\overset{(c)}{=} H(W_{[N-1]}, Z_1, X_{\boldsymbol{d}_{(2,N-1)}}, X_{\boldsymbol{A}}) + H(W_{[N-1]}, Z_2, X_{\boldsymbol{A}}, X_{\boldsymbol{d}_{(1,N-1)}}) \\
&\geq H(W_{[N-1]}, Z_1, X_{\boldsymbol{d}_{(2,N-1)}}) + H(W_{[N-1]}, Z_2) \\
&\overset{(b)}{\geq} H(W_{[N-1]}, Z_1, Z_2, X_{\boldsymbol{d}_{(2,N-1)}}) + H(W_{[N-1]}) \\
&\overset{(c)}{=} H(W_{[N]}, Z_1, Z_2, X_{\boldsymbol{d}_{(2,N-1)}}) + H(W_{[N-1]}) \\
&\overset{(d)}{=} H(W_{[N]}) + H(W_{[N-1]}) = 2N - 1
\end{aligned}$$

and

$$
\begin{aligned}
M + (N-1)R_1 + R_2 &= M + (N-2)R_1 + R_2 + R_1 \\
&\overset{(a)}{\geq} H(Z_1, X_{\boldsymbol{A}}) + H(X_{\boldsymbol{d}_{(2,N-1)}}) + H(X_{\boldsymbol{d}_{(1,N)}}) \\
&\overset{(b)}{\geq} H(Z_1, X_{\boldsymbol{A}}, X_{\boldsymbol{d}_{(2,N-1)}}, X_{\boldsymbol{d}_{(1,N)}}) \\
&\overset{(c)}{=} H(W_{[N]}, Z_1, X_{\boldsymbol{A}}, X_{\boldsymbol{d}_{(2,N-1)}}, W_N, X_{\boldsymbol{d}_{(1,N)}}) \\
&\overset{(d)}{=} H(W_{[N]}) = N
\end{aligned}
$$

where (a) follows from (7), (b) follows from the submodularity property of entropy, (c) follows from (2), and (d) follows from (3). □

As shown in TABLE II, the rates achieved by YMA scheme is tight with respect to the lower bounds presented in Lemma 2. Thus, even when coding is allowed in the placement phase, no code can dominate the YMA scheme which operates at the Pareto optimal frontier of the $(N,2)$ cache network. We summarize as:

Theorem 1. *For the* $(N,2)$ *cache network, the YMA uncoded prefetching scheme is Pareto optimal.*

Cache Size	R_1	R_2	Constraints
$0 \leq M \leq \frac{N}{2}$	$1 - \frac{1}{N}M$	$2 - \frac{3}{N}M$	$2M + (2N-3)R_1 + R_2 = 2N - 1$
$\frac{N}{2} \leq M \leq N$	$1 - \frac{1}{N}M$	$1 - \frac{1}{N}M$	$M + (N-1)R_1 + R_2 = N$

TABLE II: Constraints satisfied by the YMA scheme

III. The $(N,3)$ cache network

Consider the cache network where three users, $\{U_1, U_2, U_3\}$ are connected to a server with $N \geq 2$ files, $\{W_1, \ldots, W_N\}$. This network can have three demand types, where $\boldsymbol{D}_1$ consist of the demands where all three users request the same file, $\boldsymbol{D}_2$ consists of the demands where two of the users request the same file and the other user requests a different file and $\boldsymbol{D}_3$ consists of demands where the three users request different files (when $N \geq 3$). Let R_1, R_2 and R_3 denote the rates corresponding to $\boldsymbol{D}_1$, $\boldsymbol{D}_2$ and $\boldsymbol{D}_3$ respectively. The rates achieved by the uncoded prefetching scheme YMA for the $(N,3)$ cache network were characterized by Yu et al. [11] where lower bounds were derived under the constraint that the placement phase is uncoded. They demonstrate that under such a constraint the YMA scheme simultaneously achieves the lower bounds for R_1, R_2 and R_3, and is thus universal. We now derive a new set of bounds, by relaxing this constraint of uncoded placement, which applies to all coded caching schemes.

A. Case I: The $(2,3)$ cache network

Here, the server has two files $\{A, B\}$ and there are only two demand types, $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$. We now obtain the following result:

Lemma 3. *For the* $(2,3)$ *cache network with cache size* M*, achievable rates* R_2 *and* R_1 *must satisfy the following constraints*

$$2M + R_1 + R_2 \geq 3 \tag{10}$$

$$4M + 2R_1 + 3R_2 \geq 7 \tag{11}$$

$$M + R_1 + R_2 \geq 2 \tag{12}$$

Proof. We have,

$$
\begin{aligned}
2M + R_1 + R_2 &\geq H(Z_1) + H(X_{(A,B,A)}) + H(Z_2) + H(X_{(A,A,A)}) \\
&\overset{(a)}{\geq} H(Z_1, X_{(A,B,A)}) + H(Z_2, X_{(A,A,A)}) \\
&\overset{(b)}{=} H(A, Z_1, X_{(A,B,A)}) + H(A, Z_2, X_{(A,A,A)}) \\
&\overset{(a)}{\geq} H(A, Z_1, Z_2, X_{(A,A,A)}, X_{(A,B,A)}) + H(A) \\
&\overset{(b)}{=} H(A, B, Z_1, Z_2, X_{(A,A,A)}, X_{(A,B,A)}) + H(A) \\
&\overset{(c)}{=} H(A, B) + H(A) = 3
\end{aligned}
$$

We also have

$$
\begin{aligned}
&4M + 2R_1 + 3R_2 \\
&\geq H(Z_1) + H(X_{(A,A,B)}) + H(Z_1) + H(X_{(A,B,B)}) + H(X_{(B,B,B)}) \\
&\quad + H(Z_2) + H(X_{(A,A,B)}) + H(Z_3) + H(X_{(A,A,A)}) \\
&\overset{(a)}{\geq} H(Z_1, X_{(A,A,B)}) + H(Z_1, X_{(A,B,B)}) + H(X_{(B,B,B)}) \\
&\quad + H(Z_2, X_{(A,A,B)}) + H(Z_3, X_{(A,A,A)}) \\
&\overset{(b)}{=} H(A, Z_1, X_{(A,A,B)}) + H(A, Z_1, X_{(A,B,B)}) + H(X_{(B,B,B)}) \\
&\quad + H(A, Z_2, X_{(A,A,B)}) + H(A, Z_3, X_{(A,A,A)}) \\
&\overset{(a)}{\geq} H(A, Z_1, X_{(A,A,B)}, X_{(A,B,B)}) + H(A, Z_1) + H(X_{(B,B,B)}) \\
&\quad + H(A, Z_2, X_{(A,A,B)}) + H(A, Z_3, X_{(A,A,A)}) \\
&\overset{(a)}{\geq} H(A, Z_1, X_{(A,A,B)}, X_{(A,B,B)}) + H(A, Z_1, X_{(B,B,B)}) \\
&\quad + H(A, Z_2, X_{(A,A,B)}) + H(A, Z_3, X_{(A,A,A)}) \\
&\overset{(b)}{=} H(A, B, Z_1, X_{(B,B,B)}) + H(A, Z_1, X_{(A,A,B)}, X_{(A,B,B)}) \\
&\quad + H(A, Z_2, X_{(A,A,B)}) + H(A, Z_3, X_{(A,A,A)}) \\
&\overset{(c)}{=} H(A, B) + H(A, Z_1, X_{(A,A,B)}, X_{(A,B,B)}) \\
&\quad + H(A, Z_2, X_{(A,A,B)}) + H(A, Z_3, X_{(A,A,A)}) \\
&\overset{(a)}{\geq} H(A, B) + H(A, Z_1, Z_2, X_{(A,A,B)}, X_{(A,B,B)}) \\
&\quad + H(A, X_{(A,A,B)}) + H(A, Z_3, X_{(A,A,A)}) \\
&\overset{(b)}{=} H(A, B) + H(A, B, Z_1, Z_2, X_{(A,A,B)}, X_{(A,B,B)}) \\
&\quad + H(A, X_{(A,A,B)}) + H(A, Z_3, X_{(A,A,A)}) \\
&\overset{(c)}{=} 2H(A, B) + H(A, X_{(A,A,B)}) + H(A, Z_3, X_{(A,A,A)}) \\
&\overset{(a)}{\geq} 2H(A, B) + H(A, Z_3, X_{(A,A,A)}, X_{(A,A,B)}) + H(A) \\
&\overset{(b)}{=} 2H(A, B) + H(A, B, Z_3, X_{(A,A,A)}, X_{(A,A,B)}) + H(A) \\
&\overset{(c)}{=} 3H(A, B) + H(A) = 7
\end{aligned}
$$

We finally have,

$$\begin{aligned} M + R_1 + R_2 &\geq H(Z_1) + H(X_{A,A,A}) + H(X_{B,A,A}) \\ &\overset{(a)}{\geq} H(Z_1, X_{A,A,A}, X_{B,A,A}) \\ &\overset{(b)}{=} H(A, B, Z_1, X_{A,A,A}, X_{B,A,A}) \\ &\overset{(c)}{=} H(A, B) = 2 \end{aligned}$$

where (a) follows from the submodularity property of entropy, (b) follows from (2), and (c) follows from (3). □

The rates and the constraints satisfied by the YMA scheme for the $(2, 3)$ cache network is summarised in TABLE III.

Cache Size	R_1	R_2	Constraints
$0 \leq M \leq \frac{2}{3}$	$1 - \frac{1}{2}M$	$2 - \frac{3}{2}M$	$2M + R_1 + R_2 = 3$
$\frac{2}{3} \leq M \leq \frac{4}{3}$	$1 - \frac{1}{2}M$	$\frac{5}{3} - M$	$4M + 2R_1 + 3R_2 = 7$
$\frac{4}{3} \leq M \leq 2$	$1 - \frac{1}{2}M$	$1 - \frac{1}{2}M$	$M + R_1 + R_2 = 2$

TABLE III: Constraints satisfied by the YMA scheme

As seen in the table, the rates achieved by YMA scheme is tight with respect to the lower bounds presented in Lemma 3. Thus, even when coding is allowed in the placement phase, no code can dominate the YMA scheme which operates at the Pareto optimal frontier of the $(2, 3)$ cache network.

B. Case II: $N \geq 3$

Here, the server has N files $\{W_1, \ldots, W_N\}$. The network has three demand types $\boldsymbol{D}_1$, $\boldsymbol{D}_2$, and $\boldsymbol{D}_3$. Consider the set of demands $\{\boldsymbol{d}_{(p,k)} : 1 \leq p \leq 3, 1 \leq k \leq N - p + 1\}$, defined as

$$\boldsymbol{d}_{(1,k)} = (W_k, W_k, W_k) \tag{13}$$

$$\boldsymbol{d}_{(2,n)} = (W_n, W_{n+1}, W_{n+1}) \tag{14}$$

$$\boldsymbol{d}_{(3,n)} = (W_n, W_{n+1}, W_{n+2}) \tag{15}$$

where in demand $\boldsymbol{d}_{(p,k)}$ users request p distinct files. Let $\boldsymbol{B}$ denote the set of demands

$$\boldsymbol{B} = \{\boldsymbol{d}_{(1,1)}, \ldots, \boldsymbol{d}_{(1,N-3)}\} \tag{16}$$

and let $X_{\boldsymbol{B}}$ denote the set of all packets broadcast in response to these demands. As $|\boldsymbol{B}| = N - 3$, we have

$$M + (N-3)R_1 \geq H(Z_i) + H(X_{\boldsymbol{B}}) \geq H(Z_i, X_{\boldsymbol{B}}) \tag{17}$$

Note that $\boldsymbol{B} = \phi$ when $N = 3$. We now obtain the following result:

Lemma 4. *For the $(N, 3)$ cache network with cache size M, achievable rates R_1, R_2 and R_3 must satisfy the following constraints*

$$4M + R_3 + R_2 + (4N-9)R_1 \geq 4N - 4 \tag{18}$$

$$3M + (3N-6)R_1 + R_2 + 2R_3 \geq 3N - 1 \tag{19}$$

$$M + (N-2)R_1 + R_2 + R_3 \geq N \tag{20}$$

Proof. We have,

$$\begin{aligned}
&4M + R_3 + R_2 + (4N-9)R_1 \\
=&4(M + (N-3)R_1) + R_3 + R_2 + 3R_1 \\
\overset{(a)}{\geq}& 2H(Z_1, X_{\boldsymbol{B}}) + H(Z_2, X_{\boldsymbol{B}}) + H(Z_3, X_{\boldsymbol{B}}) + H(X_{\boldsymbol{d}_{(3,N-2)}}) \\
&+ H(X_{\boldsymbol{d}_{(2,N-2)}}) + 2H(X_{\boldsymbol{d}_{(1,N-2)}}) + H(X_{\boldsymbol{d}_{(1,N)}}) \\
\overset{(b)}{=}& 2H(W_{[N-3]}, Z_1, X_{\boldsymbol{B}}) + H(W_{[N-3]}, Z_2, X_{\boldsymbol{B}}) \\
&+ H(W_{[N-3]}, Z_3, X_{\boldsymbol{B}}) + H(X_{\boldsymbol{d}_{(3,N-2)}}) + H(X_{\boldsymbol{d}_{(2,N-2)}}) \\
&+ 2H(X_{\boldsymbol{d}_{(1,N-2)}}) + H(X_{\boldsymbol{d}_{(1,N)}}) \\
\geq& 2H(W_{[N-3]}, Z_1) + H(W_{[N-3]}, Z_2) + H(W_{[N-3]}, Z_3) \\
&+ H(X_{\boldsymbol{d}_{(3,N-2)}}) + H(X_{\boldsymbol{d}_{(2,N-2)}}) + 2H(X_{\boldsymbol{d}_{(1,N-2)}}) + H(X_{\boldsymbol{d}_{(1,N)}}) \\
\overset{(c)}{\geq}& H(W_{[N-3]}, Z_1, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-3]}, Z_2, X_{\boldsymbol{d}_{(1,N-2)}}) \\
&+ H(W_{[N-3]}, Z_1, X_{\boldsymbol{d}_{(2,N-2)}}) + H(W_{[N-3]}, Z_3, X_{\boldsymbol{d}_{(1,N-2)}}) \\
&+ H(X_{\boldsymbol{d}_{(1,N)}}) \\
\overset{(b)}{=}& H(W_{[N-2]}, Z_1, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-2]}, Z_2, X_{\boldsymbol{d}_{(1,N-2)}}) \\
&+ H(W_{[N-2]}, Z_1, X_{\boldsymbol{d}_{(2,N-2)}}) + H(W_{[N-2]}, Z_3, X_{\boldsymbol{d}_{(1,N-2)}}) \\
&+ H(X_{\boldsymbol{d}_{(1,N)}}) \\
\geq& H(W_{[N-2]}, Z_1, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-2]}, Z_2) \\
&+ H(W_{[N-2]}, Z_1, X_{\boldsymbol{d}_{(2,N-2)}}) + H(W_{[N-2]}, Z_3) + H(X_{\boldsymbol{d}_{(1,N)}}) \\
\overset{(c)}{\geq}& H(W_{[N-2]}, Z_1, Z_2, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-2]}) \\
&+ H(W_{[N-2]}, Z_1, Z_3, X_{\boldsymbol{d}_{(2,N-2)}}) + H(W_{[N-2]}) + H(X_{\boldsymbol{d}_{(1,N)}}) \\
\overset{(b)}{=}& H(W_{[N-1]}, Z_1, Z_2, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-1]}, Z_1, Z_3, X_{\boldsymbol{d}_{(2,N-2)}}) \\
&+ H(X_{\boldsymbol{d}_{(1,N)}}) + 2H(W_{[N-2]}) \\
\overset{(c)}{\geq}& H(W_{[N-1]}, Z_1, Z_2, Z_3, X_{\boldsymbol{d}_{(3,N-2)}}, X_{\boldsymbol{d}_{(2,N-2)}}) + H(W_{[N-1]}, Z_1) \\
&+ H(X_{\boldsymbol{d}_{(1,N)}}) + 2H(W_{[N-2]}) \\
\overset{(b)}{=}& H(W_{[N]}, Z_1, Z_2, Z_3, X_{\boldsymbol{d}_{(3,N-2)}}, X_{\boldsymbol{d}_{(2,N-2)}}) + H(W_{[N-1]}, Z_1) \\
&+ H(X_{\boldsymbol{d}_{(1,N)}}) + 2H(W_{[N-2]}) \\
\overset{(d)}{=}& H(W_{[N]}) + H(W_{[N-1]}, Z_1) + H(X_{\boldsymbol{d}_{(1,N)}}) + 2H(W_{[N-2]}) \\
\overset{(c)}{\geq}& H(W_{[N]}) + H(W_{[N-1]}, Z_1, X_{\boldsymbol{d}_{(1,N)}}) + 2H(W_{[N-2]}) \\
\overset{(b)}{=}& H(W_{[N]}) + H(W_{[N]}, Z_1, X_{\boldsymbol{d}_{(1,N)}}) + 2H(W_{[N-2]}) \\
\overset{(d)}{=}& 2H(W_{[N]}) + 2H(W_{[N-2]}) \geq 4N - 4
\end{aligned}$$

We also have,

$$\begin{aligned}
&3M + (3N-6)R_1 + R_2 + 2R_3 \\
&\quad =3(M + (N-3)R_1) + 3R_1 + R_2 + 2R_3 \\
&\quad \overset{(a)}{\geq} H(Z_1, X_{\boldsymbol{B}}) + H(Z_2, X_{\boldsymbol{B}}) + H(Z_3, X_{\boldsymbol{B}}) + 2H(X_{\boldsymbol{d}_{(1,N-2)}}) \\
&\quad\quad + H(X_{\boldsymbol{d}_{(1,N-1)}}) + H(X_{\boldsymbol{d}_{(2,N-1)}}) + 2H(X_{\boldsymbol{d}_{(3,N-2)}}) \\
&\quad \geq (\textit{see top of this page}) \\
&\quad \overset{(c)}{\geq} H(W_{[N-1]}, Z_1, Z_2, X_{\boldsymbol{d}_{(1,N-2)}}, X_{\boldsymbol{d}_{(2,N-1)}}, X_{\boldsymbol{d}_{(3,N-2)}}) \\
&\quad\quad + H(W_{[N-1]}, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-1]}, Z_3) \\
&\quad \overset{(b)}{=} H(W_{[N]}, Z_1, Z_2, X_{\boldsymbol{d}_{(1,N-2)}}, X_{\boldsymbol{d}_{(2,N-1)}}, X_{\boldsymbol{d}_{(3,N-2)}})
\end{aligned}$$

$$\overset{(b)}{=} H(W_{[N-3]}, Z_1, X_{\boldsymbol{B}}) + H(W_{[N-3]}, Z_2, X_{\boldsymbol{B}}) + H(W_{[N-3]}, Z_3, X_{\boldsymbol{B}}) + 2H(X_{\boldsymbol{d}_{(1,N-2)}}) + H(X_{\boldsymbol{d}_{(1,N-1)}}) + H(X_{\boldsymbol{d}_{(2,N-1)}}) + 2H(X_{\boldsymbol{d}_{(3,N-2)}})$$
$$\geq H(W_{[N-3]}, Z_1) + H(W_{[N-3]}, Z_2) + H(W_{[N-3]}, Z_3) + 2H(X_{\boldsymbol{d}_{(1,N-2)}}) + H(X_{\boldsymbol{d}_{(1,N-1)}}) + H(X_{\boldsymbol{d}_{(2,N-1)}}) + 2H(X_{\boldsymbol{d}_{(3,N-2)}})$$
$$\overset{(c)}{\geq} H(W_{[N-3]}, Z_1, X_{\boldsymbol{d}_{(2,N-1)}}, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-3]}, Z_2, X_{\boldsymbol{d}_{(1,N-2)}}, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-3]}, Z_3, X_{\boldsymbol{d}_{(1,N-2)}}, X_{\boldsymbol{d}_{(1,N-1)}})$$
$$\overset{(b)}{=} H(W_{[N-1]}, Z_1, X_{\boldsymbol{d}_{(2,N-1)}}, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-1]}, Z_2, X_{\boldsymbol{d}_{(1,N-2)}}, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-1]}, Z_3, X_{\boldsymbol{d}_{(1,N-2)}}, X_{\boldsymbol{d}_{(1,N-1)}})$$
$$\geq H(W_{[N-1]}, Z_1, X_{\boldsymbol{d}_{(2,N-1)}}, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-1]}, Z_2, X_{\boldsymbol{d}_{(1,N-2)}}, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-1]}, Z_3)$$

Cache Size	R_1	R_2	R_3	Constraints
$0 \leq M \leq \frac{N}{3}$	$1 - \frac{1}{N}M$	$2 - \frac{3}{N}M$	$3 - \frac{6}{N}M$	$4M + R_3 + R_2 + (4N-9)R_1 = 4N - 4$
$\frac{N}{3} \leq M \leq \frac{2N}{3}$	$1 - \frac{1}{N}M$	$\frac{5}{3} - \frac{2}{N}M$	$\frac{5}{3} - \frac{2}{N}M$	$3M + (3N-6)R_1 + R_2 + 2R_3 = 3N - 1$
$\frac{2N}{3} \leq M \leq N$	$1 - \frac{1}{N}M$	$1 - \frac{1}{N}M$	$1 - \frac{1}{N}M$	$M + (N-2)R_1 + R_2 + R_3 = N$

TABLE IV: Constraints satisfied by the YMA scheme

$$+ H(W_{[N-1]}, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-1]}, Z_3)$$
$$\overset{(d)}{=} H(W_{[N]}) + H(W_{[N-1]}, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-1]}, Z_3)$$
$$\overset{(c)}{\geq} H(W_{[N]}) + H(W_{[N-1]}, Z_3, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-1]})$$
$$\overset{(b)}{=} H(W_{[N]}) + H(W_{[N]}, Z_3, X_{\boldsymbol{d}_{(3,N-2)}}) + H(W_{[N-1]})$$
$$\overset{(d)}{=} 2H(W_{[N]}) + H(W_{[N-1]}) = 3N - 1$$

Finally we have,

$$M + (N-2)R_1 + R_2 + R_3$$
$$\overset{(a)}{\geq} H(Z_1, X_{\boldsymbol{B}}) + H(X_{\boldsymbol{d}_{(1,N)}}) + H(X_{\boldsymbol{d}_{(2,N-1)}}) + H(X_{\boldsymbol{d}_{(3,N-2)}})$$
$$\overset{(b)}{=} H(W_{[N-3]}, Z_1, X_{\boldsymbol{B}}) + H(X_{\boldsymbol{d}_{(1,N)}}) + H(X_{\boldsymbol{d}_{(2,N-1)}}) + H(X_{\boldsymbol{d}_{(3,N-2)}})$$
$$\overset{(c)}{\geq} H(W_{[N-3]}, Z_1, X_{\boldsymbol{B}}, X_{\boldsymbol{d}_{(1,N)}}, X_{\boldsymbol{d}_{(2,N-1)}}, X_{\boldsymbol{d}_{(3,N-2)}})$$
$$\overset{(b)}{=} H(W_{[N]}, Z_1, X_{\boldsymbol{B}}, X_{\boldsymbol{d}_{(1,N)}}, X_{\boldsymbol{d}_{(2,N-1)}}, X_{\boldsymbol{d}_{(3,N-2)}})$$
$$\overset{(d)}{=} H(W_{[N]}) = N$$

where (a) follows from (17), (b) follows from (2), (c) follows from the submodularity property of entropy and (d) follows from (3). □

As shown in TABLE IV, the rates achieved by YMA scheme is tight with respect to the lower bounds presented in Lemma 4. Thus, even when coding is allowed in the placement phase, no code can dominate the YMA scheme which operates at the Pareto optimal frontier of the $(N, 3)$ cache network. We summarize as:

Theorem 2. *For the $(N, 3)$ cache network, the YMA uncoded prefetching scheme is Pareto optimal.*

IV. CONCLUSIONS

In [11] Yu et al. studied the problem of coded caching under the constraint that the placement phase is uncoded. They introduced the YMA scheme which was shown to be universal by deriving lower bounds under the constraint of uncoded placement. In this paper we explore the possibility of finding schemes that dominate the YMA scheme when coded placement is permitted. By deriving new lower bounds, we show that the YMA scheme operates at the Pareto optimal frontier of the problem of coded caching for (N, K) networks when $K \leq 3$. A natural problem for further study is whether this result can be extended to general (N, K) cache networks.

REFERENCES

[1] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2856–2867, 2014.
[2] C. Tian, "Symmetry, outer bounds, and code constructions: A computer-aided investigation on the fundamental limits of caching," *MDPI Entropy*, vol. 20, no. 8, p. 603, 2018.
[3] K. P. Vijith Kumar, B. K. Rai, and T. Jacob, "Pareto optimal schemes in coded caching," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, July 2019, pp. 2629–2633.
[4] Z. Chen, P. Fan, and K. B. Letaief, "Fundamental limits of caching: Improved bounds for users with small buffers," *IET Communications*, vol. 10, no. 17, pp. 2315–2318, 2016.
[5] M. M. Amiri and D. Gündüz, "Fundamental limits of coded caching: Improved delivery rate-cache capacity tradeoff," *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 806–815, 2017.
[6] J. Gómez-Vilardebó, "Fundamental limits of caching: Improved rate-memory trade-off with coded prefetching," *IEEE Trans. Commun*, vol. 66, no. 10, pp. 4488–4497, 2018.
[7] K. P. Vijith Kumar, B. K. Rai, and T. Jacob, "Fundamental limits of coded caching: The memory rate pair (K-1-1/K, 1/(K-1))," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, July 2019, pp. 2624–2628.
[8] Q. Yu, M. A. Maddah-Ali, and A. S. Avestimehr, "Characterizing the rate-memory tradeoff in cache networks within a factor of 2," *IEEE Trans. Inf. Theory*, vol. 65, no. 1, pp. 647–663, 2018.
[9] K. P. Vijith Kumar, B. K. Rai, and T. Jacob, "Towards the exact rate memory tradeoff in coded caching," in *Proc. IEEE National Conf. Commun. (NCC)*, Bangalore, India, February 2019, pp. 1–6.
[10] K. P. Vijith Kumar, B. K. Rai, and T. Jacob, "Towards the Exact Memory Rate Tradeoff for the (4, 5) Cache Network," in *Proc. IEEE Int. Conf. Sig. Proce. Commun. (SPCOM)*, Bangalore, India, July 2020, pp. 1–5.
[11] Q. Yu, M. A. Maddah-Ali, and A. S. Avestimehr, "The exact rate-memory tradeoff for caching with uncoded pre-fetching," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 1281–1296, 2017.