Kamath, S., Ravi, J. & Dey, B. K. (2020). *Demand-Private Coded Caching and the Exact Trade-off for N=K=2*. In: 2020 National Conference on Communications (NCC), 21-23 Feb. 2020.

DOI: 10.1109/ncc48643.2020.9056026

Demand-Private Coded Caching and the Exact Trade-off for N=K=2

Sneha Kamath*, Jithin Ravi†, Bikash Kumar Dey‡
* Qualcomm, India. Email: snehkama@qti.qualcomm.com
† Universidad Carlos III de Madrid, Leganés, Spain. Email: rjithin@tsc.uc3m.es
‡ Indian Institute of Technology Bombay, Mumbai. Email: bikash@ee.iitb.ac.in

***Abstract*—The distributed coded caching problem has been studied extensively in the recent past. While the known coded caching schemes achieve an improved transmission rate, they violate the privacy of the users since in these schemes the demand of one user is revealed to others in the delivery phase. In this paper, we consider the coded caching problem under the constraint that the demands of the other users remain information theoretically secret from each user. We first show that the memory-rate pair $(M, \min\{N, K\}(1-M/N))$ is achievable under information theoretic demand privacy, while using broadcast transmissions. Using this, we show that perfectly demand-private coded caching rate is order optimal for all parameter regimes. We then show that a demand-private scheme for N files and K users can be obtained from a non-private scheme that satisfies only a restricted subset of demands of NK users for N files. We then focus on the demand-private coded caching problem for $K = 2$ users, $N = 2$ files. We characterize the exact memory-rate trade-off for this case. To show the achievability, we use our first result to construct a demand-private scheme from a non-private scheme satisfying a restricted demand subset that is known from an earlier work by Tian. Further, by giving a converse based on the extra requirement of privacy, we show that the obtained achievable region is the exact memory-rate trade-off.**

I. Introduction

In the seminal work [1], Maddah-Ali and Niesen demonstrated that significant gain in the transmission rate can be achieved in a noiseless broadcast network by clever design of caching and delivery schemes. The network studied in [1] consists of one server and K users, each user is equipped with a cache of uniform size. The server has N files and each user requests one of the N files in the delivery phase. By utilizing the broadcasting opportunity of this network, Maddah-Ali and Niesen provided a caching and delivery scheme which is shown to be order optimal within a factor of 12.

In this paper, we consider the coded caching problem under privacy requirement on the demands of the users, i.e., no user should learn anything about the demands of the other users. Recently, demand privacy for the coded caching setup has been studied from an information theoretic perspective [2], [3], [4]. In [2], it was studied under a setup where the delivery phase uses private multicasts to subsets of users, equivalently studying computational privacy guarantee (see Remark 1). Coded caching under perfect information theoretic privacy was studied first in [3]. In both [2], [3], construction techniques were proposed for deriving a demand-private scheme for N files and K users from a non-private coded caching scheme for N files and NK users. The achievable memory-rate pairs of the derived schemes are the same as that of the original non-private schemes. In [4], authors study the subpacketization requirement under information theoretic demand privacy constraint for $N = K = 2$. They have shown some lower bounds on the transmission rate for a given subpacketization when the caching scheme is constrained to be linear.

The non-private coded caching problem has been studied by many authors. The works [5], [6], [7], [8] focused on improving the achievable rates of Maddah-Ali and Niesen [1] by designing new schemes. Yu *et al.* [8] proposed a new caching scheme which was shown to be order optimal within a factor of 2. When the cache content is not allowed to be coded, the optimal rates were characterized in [8], [9]. Several works have obtained improved lower bounds on the rates, see for example [10], [11].

The coded caching schemes in the noiseless broadcast network is inherently prone to security and privacy issues since the broadcasted message is revealed to everyone. Information theoretic secrecy from an external adversary who can observe the broadcasted message was first studied by Sengupta *et al.* [12]. They proposed a scheme which prevents the adversary from getting any information about any file from the broadcasted message. Another privacy aspect was considered by Ravindrakumar *et al.* in [13] where each user should not get any information about any file other than the one requested by her. They proposed a scheme which achieves this constraint by distributing keys in the placement phase.

The contributions of this paper are as listed below.

1) We first show in Theorem 1 that the memory-rate pair $(M, \min\{N, K\}(1 - M/N))$ is achievable for coded caching under information theoretic demand privacy. Our achievable scheme uses broadcast transmissions in the delivery phase, and this complements a similar result in [2] for their model using private unicast transmissions in the delivery stage. We conclude in Theorem 2 that the optimal rates with and without demand privacy are always within a multiplicative factor, and this completes the order optimality [3] of information theoretically demand-private coded caching in all memory regimes.
2) We show in Theorem 3 that a demand-private scheme for N files and K users with the same memory-rate pair (M, R) can be obtained from a non-private scheme that serves only a subset of demands for N files and NK

users. This is a refinement of results of [2], [3], and the scheme uses the idea in [3]. However, the observation that the particular non-private scheme is required to serve only a subset of demands is new, and this is used later for the case of $N = K = 2$, discussed in the next item.

3) In Theorem 4, we characterize the exact memory-rate trade-off with demand privacy for $N = K = 2$. We note that the region given in Theorem 4 is strictly larger than achievable regions known from existing literature (See Fig. 2). To obtain this achievable region, we use two non-private caching schemes from [14] which are required to serve a restricted subset of demands. Proving converse for this problem is difficult in general, and the converse proof of Theorem 4 is a key contribution of this paper.
4) The achievability of the exact memory-rate trade-off in Theorem 4 is proved by showing that memory-rate pairs $(1/3, 4/3)$ and $(4/3, 1/3)$ are achievable. The caching and transmission schemes to achieve these points are linear with coded prefetching. Incidentally, these schemes also use a subpacketization of 3, which is the same as that of the schemes in [4] for the rate points $(2/3, 1)$ and $(1, 2/3)$. The question of whether the minimum required subpacketization is indeed 3 to achieve any memory-rate pair with demand privacy for $N = K = 2$ remains open.

We present the problem formulation and definitions in Sec. II. The results with proofs are presented in Sec. III.

II. Problem formulation and definitions

Consider a server with N files $W_0, W_1, \dots, W_{N-1}$ which are assumed to be independent and each of length F bits. File $W_i, i = 0, \dots, N-1$ takes values in the set $[2^F] := \{0, 1, \dots, 2^F - 1\}$ uniformly at random. The server is connected to K users via a noiseless broadcast link. Each user is equipped with a cache of size MF bits, where $M \in [0, N]$. There are two phases in a coded caching scheme. In the first phase, called the placement phase, the server fills the cache of each user. In the delivery phase, each user requests one file from the server. The index of the file requested by user k is denoted by D_k. We assume that all D_k are independent of each other, and each of them is uniformly distributed in the set $[N]$. Let the vector $\bar{D} = [D_0, D_1, ..., D_{K-1}]$ denote the demands of all users, and also let $\bar{D}_{\tilde{k}}$ denote all demands but D_k, i.e., $\bar{D}_{\tilde{k}} = \bar{D} \setminus \{D_k\}$. All users convey their demands secretly to the server. Then, the server broadcasts a message of size RF bits to serve the request of the users. The broadcasted message, denoted by X, consists of RF bits, where R is defined as the rate of transmission. User k decodes the requested file W_{D_k} using the received message, cache content, and D_k.

In a demand-private coded caching setup (see Fig. 1), we also have a privacy requirement on the demand in addition to the recovery requirement. The privacy constraint is such that user k should not gain any information about $\bar{D}_{\tilde{k}}$. To achieve this, we consider some *shared randomness* S_k which is shared between user k and the server, and it is not known to the other users. The shared randomness can be achieved during the placement phase since the placement is done secretly for each user. Random variables $S_0, \dots, S_{K-1}$ take values in some finite alphabets $\mathcal{S}_0, \dots, \mathcal{S}_{K-1}$, respectively. The set of random variables $(S_0, \dots, S_{K-1})$ is denoted by $\bar{S}$. Let $\mathcal{P}$ denote the set of values of a private randomness P available at the server. The random variables $P \cup \{S_k : k \in [K]\} \cup \{D_k : k \in [K]\} \cup \{W_i : i \in [N]\}$ are independent of each other.

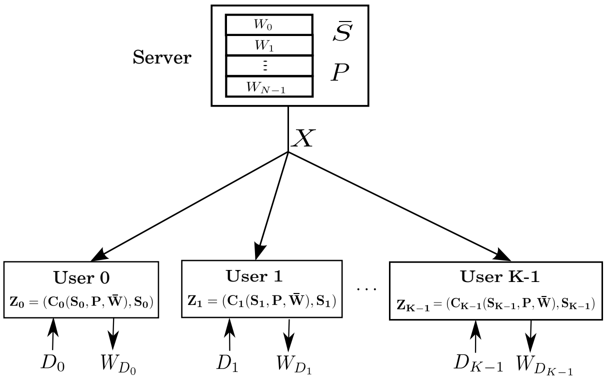
Server
W_0
W_1
$\vdots$
W_{N-1}
$\bar{S}$
P
X
User 0
$\mathbf{Z_0 = (C_0(S_0, P, \bar{W}), S_0)}$
User 1
$\mathbf{Z_1 = (C_1(S_1, P, \bar{W}), S_1)}$
$\cdots$
User K-1
$\mathbf{Z_{K-1} = (C_{K-1}(S_{K-1}, P, \bar{W}), S_{K-1})}$
D_0
W_{D_0}
D_1
W_{D_1}
D_{K-1}
$W_{D_{K-1}}$

Fig. 1: Demand-private coded caching model.

Non-private coded caching scheme: A *non-private coded caching scheme* consists of the following.

Cache encoding functions: For $k \in [K]$, the cache encoding function for the k-th user is a map

$$C_k : [2^F]^N \rightarrow [2^{MF}], \tag{1}$$

and the cache content Z_k is given by $Z_k = C_k(\bar{W})$.

Broadcast transmission encoding function: The transmission encoding is a map

$$E : [2^F]^N \times \mathcal{D}_0 \times \cdots \times \mathcal{D}_{K-1} \rightarrow [2^{RF}], \tag{2}$$

and the transmitted message is given by $X = (E(\bar{W}, \bar{D}), \bar{D})$.

Decoding functions: User k uses a decoding function

$$G_k : \mathcal{D}_0 \times \cdots \times \mathcal{D}_{K-1} \times [2^{RF}] \times [2^{MF}] \rightarrow [2^F]. \tag{3}$$

Let $\mathcal{C} = \{C_k : k = 0, \dots, K-1\}$ and $\mathcal{G} = \{G_k : k = 0, \dots, K-1\}$. Then the triple $(\mathcal{C}, E, \mathcal{G})$ is called an (N, K, M, R)-non-private scheme if it satisfies

$$W_{D_k} = G_k(\bar{D}, E(\bar{W}, \bar{D}), C_k(\bar{W})) \tag{4}$$

for all values of $\bar{D}, \bar{W}$. A memory-rate pair (M, R) is said to be *achievable* for the (N, K) coded caching problem if there exists an (N, K, M, R)-non-private scheme for some F.

Private coded caching scheme: A *private coded caching scheme* consists of the following.

Cache encoding functions: For $k \in [K]$, the cache encoding function for the k-th user is given by

$$C_k : \mathcal{S}_k \times \mathcal{P} \times [2^F]^N \rightarrow [2^{MF}], \tag{5}$$

and the cache content Z_k is given by $Z_k = (C_k(S_k, P, \bar{W}), S_k)$.

Broadcast transmission encoding function: The transmission encoding functions are

$$\begin{aligned} E : {[2^F]}^N &\times \mathcal{D}_0 \times \cdots \times \mathcal{D}_{K-1} \times \mathcal{P} \\ &\times \mathcal{S}_0 \times \cdots \times \mathcal{S}_{K-1} \to [2^{RF}], \\ J : \mathcal{D}_0 \times \cdots &\times \mathcal{D}_{K-1} \times \mathcal{P} \times \mathcal{S}_0 \times \cdots \times \mathcal{S}_{K-1} \to \mathcal{J}. \end{aligned}$$

The transmitted message X is given by

$$X = \big(E(\bar{W}, \bar{D}, P, \bar{S}), J(\bar{D}, P, \bar{S})\big).$$

Here $\log_2 |\mathcal{J}|$ is negligible[1] compared to file size F.

Decoding functions: User k has a decoding function

$$G_k : \mathcal{D}_k \times \mathcal{S}_k \times \mathcal{J} \times [2^{RF}] \times [2^{MF}] \to [2^F]. \qquad (6)$$

Let $\mathcal{C} = \{C_k : k = 0, \ldots, K-1\}$ and $\mathcal{G} = \{G_k : k = 0, \ldots, K-1\}$. The tuple $(\mathcal{C}, E, J, \mathcal{G})$ is called as an (N, K, M, R)-private scheme if it satisfies the following decoding and privacy conditions:

$$\begin{aligned} W_{D_k} = G_k\big(D_k, S_k, J(\bar{D}, P, \bar{S},), \\ E(\bar{W}, \bar{D}, P, \bar{S}), C_k(S_k, P, \bar{W})\big), \qquad (7) \\ I\left(\bar{D}_{\tilde{k}}; Z_k, X, D_k\right) = 0 \quad \text{for } k = 0, \ldots, K-1. \qquad (8) \end{aligned}$$

A memory-rate pair (M, R) is said to be *achievable with demand privacy* for the (N, K) coded caching problem if there exists an (N, K, M, R)-private scheme for some F.

The *memory-rate trade-off with demand privacy* is defined as

$$\begin{aligned} R^{*p}_{N,K}(M) = \inf\{R : (M, R) \text{ is achievable with demand} \\ \text{privacy for } (N, K) \text{ coded caching problem.}\} \quad (9) \end{aligned}$$

The *memory-rate trade-off* $R^*_{N,K}(M)$ for the non-private coded caching problem is defined similarly.

Remark 1 *The model studied in [2] assumed that the server can privately transmit to any subset of users by encryption using shared keys. The key length required for achieving such private multicast under information-theoretic privacy using broadcast transmissions is the same as the length of the multicast message. In that case, storing such keys in the cache will also contribute to the cache memory requirement. The required key rates are negligible only under computational privacy requirement, as noted in [2]. Since it was assumed that the shared keys are of negligible rates, the overall model in [2] does not ensure information-theoretic privacy under broadcast transmission. In contrast, we assume broadcast transmission in the delivery phase and we study perfect privacy in information-theoretic sense.*

III. Results

In [1, Example 1], it was shown that we can achieve rate $\min\{N, K\}(1 - M/N)$ for non-private scheme without any coding in cache placement or in broadcast transmission. Next we show that the same rate is achievable under perfect privacy of the demands under broadcast transmissions. The achievability of this rate using private unicast transmissions is simple [2, Theorem 1], and this implies the achievability under computational privacy guarantee using broadcast transmissions.

Theorem 1 *For any M, the memory-rate pair $(M, \min\{N, K\}(1 - M/N))$ is achievable in coded caching under information theoretic demand-privacy.*

Proof: Let each file W_i be split in two parts: cached part $W_i^{(c)}$ of length FM/N, and uncached part $W_i^{(u)}$ of length $F(1 - M/N)$. The cache contents of all the users are the same, and given by $Z_k = (Z^{(0)}, Z^{(1)}, \cdots, Z^{(N-1)})$, where $Z^{(i)} = W_i^{(c)}$ for each i. To describe the delivery phase, we consider two cases:

For $N < K$, the server broadcasts the remaining $(1 - M/N)$ fraction of each file. We now consider the case $N > K$. Let $D_0, D_1, \cdots, D_{K-1}$ be the demands of the users. The random variables $P_0, P_1, \cdots, P_{K-1}$ are defined inductively as

$$P_i = \begin{cases} P_j & \text{if } D_i = D_j \\ & \text{for some } j < i \\ \sim unif([K] \setminus \{P_0, P_1, \cdots, P_{i-1}\}) & \text{if } D_i \neq D_j \\ & \forall j < i. \end{cases}$$

The keys $S_0, S_1, \cdots, S_{K-1} \in [K]$ are chosen i.i.d. and uniformly distributed. The transmission X has two parts (X', J), where $X' = (X'_0, X'_1, \cdots, X'_{K-1})$ is the main payload, and J is the auxiliary transmission. The transmission is then given by

$$X'_j = \begin{cases} W^{(u)}_{D_i} & \text{if } j = P_i \\ & \text{for some } i \in [K] \\ \sim unif\left(\{0,1\}^{F(1-M/N)}\right) & \text{otherwise.} \end{cases}$$

and $J = (P_0 \oplus_K S_0, P_1 \oplus_K S_1, \cdots, P_{K-1} \oplus_K S_{K-1})$, where $\oplus_K$ denotes the addition modulo K operation. Since user k knows S_k, it can find P_k from J. It then can find $X'_{P_k} = W^{(u)}_{D_k}$, and thus $W_{D_k} = (Z^{(D_k)}, X'_{P_k})$.

Next we show that this scheme also satisfies the privacy condition. Let us denote $Q_i = P_i \oplus_K S_i$ for the ease of writing.

$$\begin{aligned} & I(\bar{D}_{\tilde{k}}; X, D_k, Z_k) \\ & = I(\bar{D}_{\tilde{k}}; X'_0, \cdots, X'_{K-1}, Q_0, Q_1, \cdots, Q_{K-1}, \\ & \qquad D_k, S_k, W_0^{(c)}, \ldots, W_{N-1}^{(c)}) \\ & = I(\bar{D}_{\tilde{k}}; Q_0, \cdots, Q_{K-1}, D_k, S_k) & (10) \\ & = I(\bar{D}_{\tilde{k}}; Q_0, \cdots, Q_{k-1}, Q_{k+1}, \cdots, Q_{K-1}, D_k, S_k, P_k) \\ & = 0 & (11) \end{aligned}$$

[1] The auxiliary transmission J essentially captures any additional transmission, that does not contribute any rate, in addition to the main payload. Such auxiliary transmissions of negligible rate are used even in non-private schemes without being formally stated in most work. For example, the scheme in [1] works only if the server additionally transmits the demand vector in the delivery phase. We have chosen to formally define such auxiliary transmission here.

where (10) follows because $(X'_0, \cdots, X'_{K-1}, W^{(c)}_0, \ldots, W^{(c)}_{N-1})$ is uniformly distributed in $\{0,1\}^{MF+FK(1-M/N)}$, and is independent of $(\bar{D}_{\tilde{k}}, Q_0, \cdots, Q_{K-1}, D_k, S_k)$, and (11) follows because all the random variables in the mutual information are independent. ■

One natural question that arises in demand-private coded caching is how much cost it incurs due to the extra constraint of demand privacy. The next theorem shows that the extra cost is always within a multiplicative factor of 8.

Theorem 2 *The optimal rates with and without privacy always satisfy the following:*

$$\frac{R^{*p}_{N,K}(M)}{R^{*}_{N,K}(M)} \leq 8. \tag{12}$$

The achievable memory-rate pair using the scheme given in [3] is shown [3, Theorem 2] to be within a factor of 8 from $R^*_{N,K}(M)$ for all memory regimes except for $0 \leq M \leq N/K$ when $N > K$. So, the result in Theorem 2 holds for all those memory regimes. It can be shown (see extended version [15]) that if $N > K$ and $0 \leq M \leq N/K$, then a combination of the scheme given in [3] and the scheme used to prove Theorem 1 gives an achievable memory-rate pair which is within a factor of 8 from $R^*_{N,K}(M)$. Thus, we have Theorem 2. This completes the order optimality result of demand-private coded caching. We also note that under computational privacy guarantee, the order optimality for all memory regimes was given in [2].

A demand-private scheme for N files and K users can be obtained using an existing non-private achievable scheme for N files and NK users as a blackbox. Here every user is associated with a stack of N users in the non-private caching problem. For example, demand-private schemes for $N = K = 2$ are obtained from the non-private schemes for $N = 2$ and $K = 4$. We use the ideas from the scheme presented in [3], where only certain types of demand vectors for the non-private scheme are used in the private scheme. Next we define this particular subset of demand vectors.

Consider a non-private coded caching problem with N files and NK users. A demand vector $\bar{d}$ in this problem is an NK-length vector, where the j^{th} component denotes the demand of user j. Then $\bar{d}$ can also be represented as K subvectors of length N, i.e.,

$$\bar{d} = [\bar{d}^{(0)}, \bar{d}^{(1)}, \ldots, \bar{d}^{(K-1)}],$$

where $\bar{d}^{(i)} \in [N]^N$ is an N-length vector for all $i \in [K]$. We now define a "restricted demand subset" $\mathcal{D}_{\mathcal{RS}}$.

Definition 1 (Restricted Demand Subset $\mathcal{D}_{\mathcal{RS}}$) *The restricted demand subset $\mathcal{D}_{\mathcal{RS}}$ for an (N, NK) coded caching problem is the set of all $\bar{d}$ such that $\bar{d}^{(i)}$ is a cyclic shift of the vector $(0, 1, \ldots, N-1)$ for all $i = 0, 1, \ldots, K-1$.*

Since N cyclic shifts are possible for each $\bar{d}^{(i)}$, there are a total of N^K such demand vectors in $\mathcal{D}_{\mathcal{RS}}$.

For a given $\bar{d} \in \mathcal{D}_{\mathcal{RS}}$ and $i \in [K]$, let c_i denote the number of right cyclic shifts of $(0, 1, \ldots, N-1)$ needed to get $\bar{d}^{(i)}$. Then, $\bar{d} \in \mathcal{D}_{\mathcal{RS}}$ is uniquely identified by the vector $\bar{c}(\bar{d}) := (c_1, \ldots, c_K)$. For $N = 2$ and $NK = 4$, the demands in $\mathcal{D}_{\mathcal{RS}}$ and their corresponding $\bar{c}(\bar{d}_s)$ are given in Table I.

D_0	D_1	D_2	D_3	$\bar{c}(d_s)$
0	1	0	1	$(0,0)$
0	1	1	0	$(0,1)$
1	0	0	1	$(1,0)$
1	0	1	0	$(1,1)$

TABLE I: Demand subset $\mathcal{D}_{\mathcal{RS}}$ for $N = 2$ and $NK = 4$.

A related concept is the "demand type" used in [14].

Definition 2 (Demand Types) *In (N, K)-non-private coded caching problem, for a given demand vector $\bar{d}$, let t_i denote the number of users requesting file i, where $i = 0, \ldots, N-1$. Demand type of $\bar{d}$, denoted by $T(\bar{d})$, is defined as the N-length vector $T(\bar{d}) := \bar{t} = (t_1, \ldots, t_N)$. The type class of $\bar{t}$ is defined as $\mathcal{D}_{\bar{t}} = \{\bar{d} | T(\bar{d}) = \bar{t}\}$.*

Clearly, the restricted demand subset $\mathcal{D}_{\mathcal{RS}}$ is a subset of the type class $(K, K, \ldots, K)$, i.e.,

$$\mathcal{D}_{\mathcal{RS}} \subseteq \mathcal{D}_{(K,K,\ldots,K)}. \tag{13}$$

A non-private scheme for an (N, K) coded caching problem that satisfies all demand vectors in a particular demand subset $\mathcal{D} \subset [N]^K$, is called a $\mathcal{D}$-non-private scheme. Clearly, for $\mathcal{D}_1 \subset \mathcal{D}_2$, a $\mathcal{D}_2$-non-private scheme is also a $\mathcal{D}_1$-non-private scheme. In particular, achievable rates for satisfying various demand type-classes were studied in [14], and their results are useful in our schemes for the type $(K, K, \cdots, K)$ due to the relation (13).

Theorem 3 *If there exists an (N, NK, M, R) $\mathcal{D}_{\mathcal{RS}}$-non-private scheme, then there exists an (N, K, M, R)-private scheme.*

The proof will construct an (N, K, M, R)-private scheme using an (N, NK, M, R) $\mathcal{D}_{\mathcal{RS}}$-non-private scheme as a blackbox using ideas from [3]. We first give an example to illustrate this construction for $N = 2, K = 2$ using only the restricted demand subset for a $(2, 4, \frac{1}{3}, \frac{4}{3})$ $\mathcal{D}_{(2,2)}$-non-private scheme from [14]. We will see that this allows a better achievable rate $(\frac{1}{3}, \frac{4}{3})$ for the $(N = 2, K = 2)$ demand-private coded caching problem than what can be achieved for the $N = 2, K = 4$ non-private caching problem.

Example 1 *We consider the demand-private coded caching problem for $N = 2, K = 2$. Using results from [3] and [2], we know that a demand-private scheme of the same rate-pair can be obtained from any non-private scheme for $N = 2, K = 4$. However, it was shown in [14] that for the memory $M = 1/3$, the optimum transmission rate $R^*_{2,4}(1/3) > 4/3$. It can be shown that other demand-private schemes in [2] also do not achieve $R = 4/3$ for $N = 2, K = 2$. See Fig. 2 for reference.*

Let A and B denote the two files. We will now give a scheme which achieves a rate $4/3$ for $M = 1/3$ with $F = 3l$. We denote the 3 segments of A and B by A_1, A_2, A_3 and B_1, B_2, B_3 respectively, of l bits each. First let us consider a $\mathcal{D}_{\mathcal{RS}}$-non-private scheme for $N = 2$ and $K = 4$ from [14]. Let $C_{i,j}(A,B)$, as shown in Table II, *correspond to the cache content of user $2i + j$ in the $\mathcal{D}_{\mathcal{RS}}$-non-private scheme. The transmission $T_{(i,j)}(A,B), i,j = 0,1$, as given in Table* II, *is chosen for the demand $\bar{d} \in \mathcal{D}_{\mathcal{RS}}$ such that $(i,j) = \bar{c}(\bar{d})$. Using Table* II, *it is easy to verify that the non-private scheme satisfies the decodability condition for demands in $\mathcal{D}_{\mathcal{RS}}$. From this scheme, we obtain a demand-private scheme for $N = 2, K = 2$ as follows. Let the shared key $S_k, k = 0,1$ of user k be a uniform binary random variable. The cache encoding functions and the transmission encoding function are denoted as*

$$C_k(S_k, A, B) = C_{k,S_k}(A,B) \text{ for } k = 0,1,$$
$$E(A,B,D_0,D_1,S_0,S_1) = T_{(D_0 \oplus S_0, D_1 \oplus S_1)}(A,B).$$

User k chooses $C_{k,S_k}(A,B)$ given in Table II *as the cache encoding function. The server broadcasts both $(D_0 \oplus S_0, D_1 \oplus S_1)$ and $T_{(D_0 \oplus S_0, D_1 \oplus S_1)}(A,B)$. This choice of transmission satisfies the decodability condition due to the way we have chosen the cache content and also since the chosen non-private scheme satisfies the decodability condition for demands in $\mathcal{D}_{\mathcal{RS}}$. Further, the broadcast transmission will not reveal any information about the demand of one user to the other user since one particular transmission $T_{(i,j)}(A,B)$ happens for all demand vectors (D_0, D_1), and also that S_i acts as one time pad for D_i for each $i = 0,1$. Here, all the transmissions consist of $4l$ bits (neglecting the 2 bits for $(D_0 \oplus S_0, D_1 \oplus S_1)$). Since $F = 3l$, this scheme achieves a rate $R = \frac{4}{3}$.*

(i,j)	Cache Content $C_{i,j}(A,B)$	Transmission $T_{(i,j)}(A,B)$
$(0,0)$	$A_1 \oplus B_1$	$B_1, B_2, A_3, A_1 \oplus A_2 \oplus A_3$
$(0,1)$	$A_3 \oplus B_3$	$A_2, A_3, B_1, B_1 \oplus B_2 \oplus B_3$
$(1,0)$	$A_2 \oplus B_2$	$B_2, B_3, A_1, A_1 \oplus A_2 \oplus A_3$
$(1,1)$	$A_1 \oplus A_2 \oplus A_3$ $\oplus B_1 \oplus B_2 \oplus B_3$	$A_1, A_2, B_3, B_1 \oplus B_2 \oplus B_3$

TABLE II: Cache contents and transmissions for $(2,2,\frac{1}{3},\frac{4}{3})$-private scheme.

Proof of Theorem 3: Let us consider any (N, NK, M, R) $\mathcal{D}_{\mathcal{RS}}$-non-private scheme. Let $C_k^{(np)}; k \in [NK]$ be the cache encoding functions, $E^{(np)}$ be the broadcast encoding function, and $G_k^{(np)}; k \in [NK]$ be the decoding functions for the given (N, NK, M, R) $\mathcal{D}_{\mathcal{RS}}$-non-private scheme. We will now present a construction of an (N, K, M, R)-private scheme from the given (N, NK, M, R) $\mathcal{D}_{\mathcal{RS}}$-non-private scheme.

Cache encoding: For $k \in [K]$ and $S_k \in [N]$, the k-th user's cache encoding function is given by

$$C_k(S_k, \bar{W}) := C^{(np)}_{kN+S_k}(\bar{W}), \tag{14}$$

The cache content is given by $Z_k = (C_k(S_k, \bar{W}), S_k)$.

Broadcast encoding: To define the broadcast encoding, we need some new notations and definitions. Let $\Psi : [N]^N \rightarrow [N]^N$ denote the cyclic shift operator, such that $\Psi(t_1, t_2, \cdots, t_N) = (t_N, t_1, \cdots, t_{N-1})$. Let us denote a vector $\mathbb{I} := (0, 1, \cdots, N-1)$. Let us also define

$$\bar{S} \ominus \bar{D} := (S_0 \ominus D_0, S_1 \ominus D_1, \cdots, S_{K-1} \ominus D_{K-1}),$$

where $S_k \ominus D_k$ denotes the difference of S_k and D_k modulo N. For a given $\bar{D} \in [N]^K$, we define an expanded demand vector for the non-private problem as:

$$\bar{D}^{(np)}(\bar{D}, \bar{S}) = (\Psi^{S_0 \ominus D_0}(\mathbb{I}), \cdots, \Psi^{S_{K-1} \ominus D_{K-1}}(\mathbb{I})),$$

where Ψ^i denotes the i-times cyclic shift operator.

The broadcast encoding function for the (N, K, M, R)-private scheme is defined by

$$E(\bar{W}, \bar{D}, \bar{S}) := E^{(np)}(\bar{W}, \bar{D}^{(np)}(\bar{D}, \bar{S})). \tag{15}$$

Let us denote $X_1 = E(\bar{W}, \bar{D}, \bar{S})$. The private scheme transmits the pair $X = (X_1, \bar{S} \ominus \bar{D})$.

Decoding: User k uses the decoding function of the $(kN + S_k)$-th user in the non-private scheme:

$$G_k(D_k, S_k, \bar{S} \ominus \bar{D}, X_1, Z_k) = G^{(np)}_{kN+S_k}(\bar{D}^{(np)}(\bar{D}, \bar{S}), X_1, Z_k)$$

Here the decoder computes $\bar{D}^{(np)}(\bar{D}, \bar{S})$ from $\bar{S} \ominus \bar{D}$.

Proof of decodability: The index of the output file is the $(kN + S_k)$-th component in $\bar{D}^{(np)}(\bar{D}, \bar{S})$, i.e., $S_k \ominus (S_k \ominus D_k) = D_k$. Thus the k-th user recovers its desired file.

Proof of privacy:

The proof of privacy essentially follows from the fact that S_i acts as one time pad for D_i which prevents any user $j \neq i$ getting any information about D_i. For a precise proof of $I(\bar{D}_{\tilde{k}}; Z_k, D_k, X | \bar{W}) = 0$, see the extended version [15].

It is easy to check that the memory-rate pair is close to (M, R) for the above private scheme for large F. ■

Corollary 1 *If there exists an (N, NK, M, R) $\mathcal{D}_{(K,K,\ldots,K)}$-non-private scheme, then there exists an (N, K, M, R)-private scheme.*

The converse proof of Theorem 4 uses the following lemma on some conditional distributions. The proof is elementary, and it can be found in [15].

Lemma 1 *Let $\tilde{k} = (k+1) \mod 2$ for $k = 0,1$. Then any demand-private scheme for $N = K = 2$ satisfies the following for user k, where $k = 0,1$, and for $j = 0,1$:*

$$\begin{aligned}(X, Z_k, W_j | D_k = j) &\sim (X, Z_k, W_j | D_{\tilde{k}} = 0, D_k = j) \\ &\sim (X, Z_k, W_j | D_{\tilde{k}} = 1, D_k = j). \end{aligned} \tag{16}$$

We present the optimal memory-rate region with demand privacy for $N = K = 2$ in Theorem 4. In Fig. 2, we plot the optimal trade-off for $N = K = 2$ along with the known achievable memory-rate pairs using different schemes in the literature.

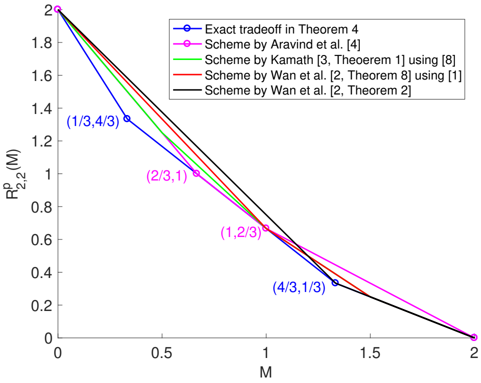

Fig. 2: Comparison of known private schemes for $N = K = 2$.

Theorem 4 *Any memory-rate pair (M, R) is achievable with demand privacy for $N = K = 2$ if and only if*

$$2M + R \geq 2, \quad 3M + 3R \geq 5, \quad M + 2R \geq 2. \tag{17}$$

Proof: It was shown in [14, Proposition 7] that the region given by (17) is an achievable rate region for Type (2,2) in $N = 2, K = 4$ coded caching problem. Then the achievability under demand-privacy for $N = K = 2$ follows from Corollary 1.

To show the converse, we only need to prove that any (M, R) pair satisfies $3M+3R \geq 5$. The other two inequalities in (17) are also necessary under no privacy requirement. So those hold under privacy requirement as well. We note that the bound $3M+3R \geq 5$ is given for 2 files and 3 users under no privacy requirement in [14, Proposition 5]. We obtain this bound for 2 users 2 files with privacy constraint, crucially using Lemma 1 below.

From the fact that the cache contents are independent of the demands and also that the transmission is independent of demands due to the privacy condition, it is easy to verify that $H(Z_0, X|D_0 = 0) + H(Z_1, X|D_1 = 0) + H(Z_1, X|D_0 = 1, D_1 = 0)$ is upper bounded by $3MF+3RF$. Next we lower bound the same quantity by $5F$ which proves the bound $3M+3R \geq 5$. First,

$$\begin{aligned}
&H(Z_0, X|D_0 = 0) + H(Z_1, X|D_1 = 0) \\
&= H(Z_0, W_0, X|D_0 = 0) + H(Z_1, W_0, X|D_1 = 0) && (18)\\
&= H(Z_0, W_0, X|D_0 = 0, D_1 = 0) \\
&\quad + H(Z_1, W_0, X|D_0 = 0, D_1 = 0) && (19)\\
&\geq H(Z_0, Z_1, W_0, X|D_0 = 0, D_1 = 0) \\
&\quad + H(W_0, X|D_0 = 0, D_1 = 0), && (20)
\end{aligned}$$

where (18) follows from the decodability condition, and (19) follows from Lemma 1.

Using the decodability condition, it can be shown (see [15]) that

$$\begin{aligned}
&H(Z_0, Z_1, W_0, X|D_0 = 0, D_1 = 0) \\
&\qquad + H(X, Z_1|D_0 = 1, D_1 = 0) \\
&\quad \geq H(W_1, W_0) + H(Z_1, W_0|D_0 = 1, D_1 = 0). && (21)
\end{aligned}$$

From (20) and (21), we obtain

$$\begin{aligned}
&H(Z_0, X|D_0 = 0) + H(Z_1, X|D_1 = 0) \\
&\quad + H(Z_1, X|D_0 = 1, D_1 = 0) \\
&\geq H(W_1, W_0) + H(Z_1, W_0|D_0 = 1, D_1 = 0) \\
&\quad + H(W_0, X|D_0 = 0, D_1 = 0) \\
&= H(W_1, W_0) + H(Z_1, W_0|D_0 = 0, D_1 = 1) \\
&\quad + H(W_0, X|D_0 = 0, D_1 = 1) && (22)\\
&\geq H(W_1, W_0) + H(Z_1, X, W_1|W_0, D_0 = 0, D_1 = 1) \\
&\quad + 2H(W_0|D_0 = 0, D_1 = 1) && (23)\\
&\geq 5F, && (24)
\end{aligned}$$

where in (22) we used Lemma 1, and in (23) we used the decodability condition. This completes the proof of Theorem 4. ■

IV. Acknowledgment

J. Ravi received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant No. 714161). The work of B. K. Dey was supported in part by the Bharti Centre for Communication, IIT Bombay.

References

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

[2] K. Wan and G. Caire, "On coded caching with private demands," arXiv:1908.10821 [cs.IT], Sep 2019.

[3] S. Kamath, "Demand private coded caching," arXiv:1909.03324 [cs.IT], Sep 2019.

[4] V. R. Aravind, P. Sarvepalli, and A. Thangaraj, "Subpacketization in coded caching with demand privacy," arXiv:1909.10471 [cs.IT], Sep 2019.

[5] M. Mohammadi Amiri and D. Gunduz, "Fundamental limits of coded caching: Improved delivery rate-cache capacity tradeoff," *IEEE Transactions on Communications*, vol. 65, no. 2, pp. 806–815, Feb 2017.

[6] K. Zhang and C. Tian, "Fundamental limits of coded caching: From uncoded prefetching to coded prefetching," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 6, pp. 1153–1164, June 2018.

[7] J. Gómez-Vilardebó, "Fundamental limits of caching: Improved rate-memory tradeoff with coded prefetching," *IEEE Transactions on Communications*, vol. 66, no. 10, pp. 4488–4497, Oct 2018.

[8] Q. Yu, M. A. Maddah-Ali, and A. S. Avestimehr, "The exact rate-memory tradeoff for caching with uncoded prefetching," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 1281–1296, Feb 2018.

[9] K. Wan, D. Tuninetti, and P. Piantanida, "On the optimality of uncoded cache placement," in *2016 IEEE Information Theory Workshop (ITW)*, Sep. 2016, pp. 161–165.

[10] H. Ghasemi and A. Ramamoorthy, "Improved lower bounds for coded caching," *IEEE Transactions on Information Theory*, vol. 63, no. 7, pp. 4388–4413, July 2017.

[11] C. Wang, S. Saeedi Bidokhti, and M. Wigger, "Improved converses and gap results for coded caching," *IEEE Transactions on Information Theory*, vol. 64, no. 11, pp. 7051–7062, Nov 2018.

[12] A. Sengupta, R. Tandon, and T. C. Clancy, "Fundamental limits of caching with secure delivery," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 2, pp. 355–370, Feb 2015.

[13] V. Ravindrakumar, P. Panda, N. Karamchandani, and V. M. Prabhakaran, "Private coded caching," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 3, pp. 685–694, March 2018.

[14] C. Tian, "Symmetry, outer bounds, and code constructions: A computer-aided investigation on the fundamental limits of caching," *Entropy*, vol. 20, no. 8, p. 603, 2018.

[15] S. Kamath, J. Ravi, and B. K. Dey, "Demand-private coded caching and the exact trade-off for N=K=2," arXiv:1911.06995 [cs.IT].