Coded Caching in the Presence of a Wire and a Cache Tapping Adversary of Type II

Mohamed Nafea, *Member, IEEE*, and Aylin Yener, *Fellow, IEEE*

***Abstract*—This paper introduces the notion of *cache-tapping* into the information theoretic models of coded caching. The wiretap channel II in the presence of multiple receivers equipped with fixed-size cache memories, and an adversary which selects symbols to tap into from cache placement and/or delivery is introduced. The legitimate terminals know neither whether placement, delivery, or both are tapped, nor the positions in which they are tapped. Only the size of the overall tapped set is known. For two receivers and two files, the strong secrecy capacity– the maximum achievable file rate while keeping the overall library strongly secure– is identified. Lower and upper bounds on the strong secrecy file rate are derived when the library has more than two files. Achievability relies on a code design which combines wiretap coding, security embedding codes, one-time pad keys, and coded caching. A genie-aided upper bound, in which the transmitter is provided with user demands before placement, establishes the converse for the two-files case. For more than two files, the upper bound is constructed by three successive channel transformations. Our results establish provable security guarantees against a powerful adversary which optimizes its tapping over both phases of communication in a cache-aided system.**

***Index Terms*—Secure coded caching, cache-tapping, wiretap channel II, strategic adversaries, strong secrecy, security embedding codes.**

I. Introduction

CACHING aims to reduce network traffic congestion by pro-actively storing partial contents at the cache memories of end users during off-peak times, providing local caching gain [3]–[5]. Seminal work in [6] has shown that, careful design of cache contents in a multi-receiver setting allows the transmitter to send delivery transmissions that are simultaneously useful for many users, providing a further gain termed as the *global caching gain*. This gain depends on the aggregate cache memory of the network and demonstrates the ability of coding over delivery transmission and/or cache contents.

In numerous works to date, coded caching has been studied under various modeling assumptions and network configurations, including decentralized caching [7], non-uniform demands [8], more users than files [9]–[11], heterogeneous cache sizes [12], improved bounds [13]–[15], hierarchical caching [16], interference networks [17]–[19], combination networks [20], [21], device-to-device communication [22], [23], coded placement [24], [25], and delivery over noisy channels [26]–[30].

Coded caching with security guarantees has been studied in [21], [31]–[38]. These, as they pertain to an external adversary, i.e., a wiretapper, assume secure cache placement; the adversary cannot tap into the cache nor the communication which performs cache placement. At the other extreme, if cache placement were to be public, i.e., if the adversary has perfect access to the cache contents, it follows from [39], [40] that the cache memories do not increase the secrecy capacity. This paper considers an intermediate setting between these two extremes in which the adversary may have *partial access* to cache placement.

The wiretap channel II (WTC-II) in [41] provides a model for an adversary with partial access to the legitimate communication; in the form of a threshold on the time fraction during which the adversary is able to tap into the communication. Specifically, the model considers a noiseless legitimate channel and an adversary which *selects* a *fixed-size* subset of the transmitted symbols to noiselessly observe. Reference [41] has shown that, despite this ability to choose the locations of the tapped symbols, with proper coding, the adversary can be made no more powerful than nature, i.e., the secrecy capacity of the WTC-II is identical to that of a binary erasure wiretapper channel with the same fraction of erasures. Reference [42] has generalized the WTC-II to one with a discrete memoryless– noisy– legitimate channel, and derived inner and outer bounds for its capacity-equivocation region. The secrecy capacity for this model has been identified in [43]. In [44], we have introduced a generalized wiretap model which includes both the classical wiretap [45] and wiretap II [42] channels as special cases. This generalized model has been extended to multi-transmitter and multi-receiver networks in [46]–[48]. In all these settings, the common theme is the robustness of stochastic wiretap encoding [45] against a type II adversary which can choose where to tap.

In this paper, we introduce an adversary model of type II to a cache-aided communication setting. The adversary noiselessly observes a partial subset of its choice of the transmitted symbols over cache placement and/or delivery. We term this model the caching broadcast channel with a *wire and cache tapping* adversary of type II (CBC-WCT II). The legitimate terminals do not know whether cache placement,

Manuscript received August 15, 2020; revised December 4, 2020; accepted January 14, 2021. Date of publication January 26, 2021; date of current version March 16, 2021. This work was supported in part by NSF under Grant CCF 2105872. This paper was presented in part at the 2018 IEEE Information Theory Workshop [1] and 2018 Allerton Conference on Communication, Control, and Computing [2]. *(Corresponding author: Aylin Yener.)*

Mohamed Nafea is with the Electrical and Computer Engineering Department, University of Detroit Mercy, Detroit, MI 48221 USA (e-mail: nafeamo@udmercy.edu).

Aylin Yener is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210 USA, also with the Department of Integrated Systems Engineering, Ohio State University, Columbus, OH 43210 USA, and also with the Department of Computer Science and Engineering, Ohio State University, Columbus, OH 43210 USA (e-mail: yener@ece.osu.edu).

Digital Object Identifier 10.1109/JSAIT.2021.3054663

delivery, or both phases are tapped; the relative fractions of tapped symbols in each, nor their positions. Only the overall size of the tapped set is known by the legitimate terminals.

The challenge in caching stems from the fact that the transmitter, which has access to a library of files, has no knowledge about future demands of end users when designing their cache contents. This remains to be the case when security against an external adversary is concerned. Further, for the new model introduced in this paper, the adversary might tap into cache placement, delivery, or both, and where the tapping occurs is *unknown* to the legitimate terminals. We show that even under these unfavorable conditions, strong secrecy guarantees that are invariant to the positions of the tapped symbols varying between cache placement, delivery, or both phases, can be provided.

In coded caching literature up to date, the physical communication which populates the cache memories at end users is not considered in the problem formulation, due to the assumption of secure cache placement. For the setting we propose in this work, in order to model cache placement that is tapped by an adversary, we consider a length-n communication block over a two-user broadcast WTC-II [47]. The sizes of cache memories at the receivers are fixed in our setting. Introducing variable memory sizes for which a rate-memory tradeoff can be characterized, as in the usual setup for caching, requires considering additional communication blocks for cache placement. Being of future interest, we discuss this extension to multiple communication blocks for cache placement in Section VII. We as well provide reasoning for our choice of the broadcast setting for cache placement.

The main contributions of this paper are summarized as follows:

1) We introduce the notion of *cache-tapping* into the information theoretic models of coded caching, in which an adversary of type II is able to tap into a fixed-size subset of its choice of the symbols transmitted during either cache placement, delivery, or both phases.
2) We characterize the strong secrecy capacity– the maximum achievable file rate which keeps the overall library strongly secure– for the instance of a transmitter's library with two files:
   - We devise an achievability scheme which integrates wiretap coding [41], security embedding codes [49], [50], one-time pad keys [39], *coded* cache placement and *uncoded* delivery [6].
   - We use a genie-aided upper bound in which the transmitter is provided with user demands before placement, rendering the model to a broadcast WTC-II [47], to establish the converse.
3) We derive lower and upper bounds on the strong secrecy file rate when the library has more than two files:
   - We use the same channel coding scheme as for the two files case. However, the cache placement and delivery schemes we employ to achieve the rates are different. In particular, we utilize here *uncoded* cache placement and a *partially coded* delivery.
   - We derive the upper bound in three steps: We (i) consider a transformed channel with an adversary which taps into a fraction of symbols equal to our model, but is only allowed to tap into the delivery phase. Since this adversary has a more restricted strategy space than the original one, its corresponding secrecy capacity is at least as large; (ii) use Sanov's theorem in method of types [51, Th. 11.4.1] to further upper bound the secrecy capacity of the restricted adversary model by the secrecy capacity when the adversary encounters a discrete memoryless binary erasure channel, and finally (iii) upper bound the secrecy capacity of the discrete memoryless model by that of a single receiver setting in which the receiver requests two files from the library.

The remainder of the paper is organized as follows. Section II describes the communication system proposed in this paper. Section III presents the main results. The proofs of these results are provided in Sections IV, V, and VI. Section VII provides a discussion about the communication model in question, the presented results, and the extension of our model to arbitrary number of users and to variable memory sizes. Section VIII concludes the paper.

II. System Model

We remark the notation we use throughout the paper. $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ denote the sets of natural, integer, real numbers, respectively. For $a, b \in \mathbb{R}$, $[a : b]$ denotes the set of integers $\{i \in \mathbb{N} : a \leq i \leq b\}$. $A_{[1:n]}$ denotes the sequence of variables $\{A_1, A_2, \ldots, A_n\}$. For two sets $\mathcal{A}_1$, $\mathcal{A}_2$; $\mathcal{A}_1 \times \mathcal{A}_2$ denotes their Cartesian product. $\mathcal{A}^T$ denotes the T-fold Cartesian product of the set $\mathcal{A}$. For $W_1, W_2 \in [1 : M]$, $W_1 \oplus W_2$ denotes the bit-wise XOR on the binary strings corresponding to W_1, W_2. $\mathbb{1}_{\mathcal{A}}$ denotes the indicator function for the event $\mathcal{A}$. $\mathbb{D}(p_x||q_x)$ denotes the Kullback-Leibler divergence between the probability distributions p_x, q_x, defined on the same probability space. $\{\epsilon_n\}_{n\geq 1}$ denotes a sequence of positive real numbers such that $\epsilon_n \to 0$ as $n \to \infty$.

Consider the communication system depicted in Fig. 1, in which the adversary is able to tap into both the cache placement and delivery transmissions. The transmitter observes $D \geq 2$ independent messages (files), $W_1, W_2, \ldots, W_D$, each of which is uniformly distributed over $[1 : 2^{nR_s}]$. Each receiver has a cache memory of size $\frac{n}{2}$ bits. The communication occurs over two phases: cache placement and delivery. The broadcast channel is noiseless during both phases. The communication model is described as follows:

Cache Placement Phase: The transmitter broadcasts a length-n binary signal, $\mathbf{X}_c^n \in \{0, 1\}^n$, to both receivers. The codeword $\mathbf{X}_c^n$ is a function of the library files; $\mathbf{X}_c^n \triangleq f_c(W_{[1:D]})$. The transmitter does not know the receiver demands during cache placement [6]. Each receiver has a cache memory of size $\frac{n}{2}$ bits in which they store a function of $\mathbf{X}_c^n$, $M_{c,j} \triangleq f_{c,j}(\mathbf{X}_c^n)$; $f_{c,j} : \{0, 1\}^n \mapsto [1 : 2^{\frac{n}{2}}]$, $j = 1, 2$.

Delivery Phase: At the beginning of this phase, the two receivers announce their demands $\mathbf{d} \triangleq (d_1, d_2) \in [1 : D]^2$

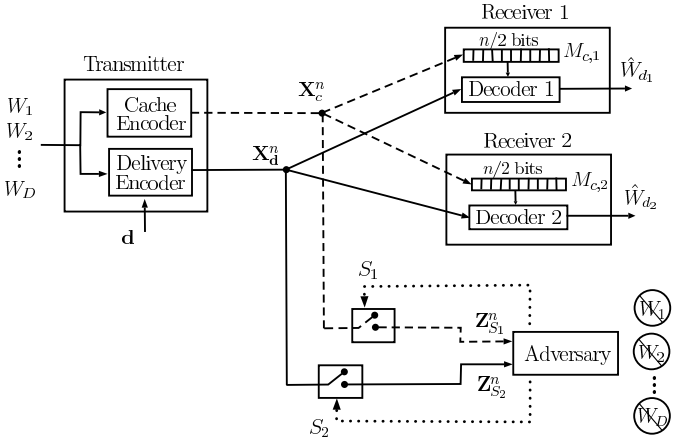

Fig. 1. The caching broadcast channel with a wire and cache tapping adversary of type II (CBC-WCT II). The adversary chooses tapping sets S_1 and S_2 in placement and delivery.

to the transmitter. To satisfy these demands, the transmitter encodes $W_{[1:D]}$ and $\mathbf{d}$ into a binary codeword $\mathbf{X}_\mathbf{d}^n \in \{0,1\}^n$: For each $\mathbf{d} \in [1{:}D]^2$, the transmitter uses the encoder $f_\mathbf{d} : [1{:}2^{nR_s}]^D \mapsto \{0,1\}^n$ and sends the codeword $\mathbf{X}_\mathbf{d}^n \triangleq f_\mathbf{d}(W_{[1:D]})$.

Decoding: Receiver j uses the decoder $g_{\mathbf{d},j} : [1 : 2^{\frac{n}{2}}] \times \{0,1\}^n \mapsto [1 : 2^{nR_s}]$ to output the estimate $\hat{W}_{d_j} \triangleq g_{\mathbf{d},j}(f_{c,j}(\mathbf{X}_c^n), \mathbf{X}_\mathbf{d}^n)$ of its desired message W_{d_j}; $j = 1, 2$.

Adversary Model: The adversary chooses two subsets $S_1, S_2 \subseteq [1 : n]$. The size of the sum of cardinalities of S_1 and S_2 is fixed: For $|S_1| = \mu_1$, $|S_2| = \mu_2$, $\mu_1, \mu_2 \leq n$, we have $\mu_1 + \mu_2 = \mu$. The subsets S_1, S_2 indicate the positions tapped by the adversary during cache placement and delivery, respectively. Over the two phases, the adversary observes the length-$2n$ sequence $\mathbf{Z}_S^{2n} = [\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n] \in \mathcal{Z}^{2n}$, where $\mathbf{Z}_{S_j}^n \triangleq [Z_{S_j,1}, \ldots, Z_{S_j,n}] \in \mathcal{Z}^n$, $j = 1, 2$,

$$Z_{S_1,i} = \begin{cases} X_{c,i}, & i \in S_1 \\ ?, & i \notin S_1 \end{cases}, \quad Z_{S_2,i} = \begin{cases} X_{\mathbf{d},i}, & i \in S_2 \\ ?, & i \notin S_2. \end{cases} \tag{1}$$

The alphabet is $\mathcal{Z} = \{0, 1, ?\}$, where "?" denotes an erasure.

The legitimate terminals know neither the realizations of S_1, S_2, nor the values of μ_1, μ_2. Only μ is known. Let $\alpha_1 = \frac{\mu_1}{n}$, $\alpha_2 = \frac{\mu_2}{n}$, be the fractions of the tapped symbols in the cache placement and delivery, and let $\alpha = \alpha_1 + \alpha_2$ be the overall tapped ratio. Note that $\alpha_1, \alpha_2 \in [0, 1]$ and $\alpha \in (0, 2]$.

Remark 1: We consider that $\alpha > 0$, i.e., the adversary is present. For $\alpha = 0$, i.e., no adversary, the problem considered in this paper has been extensively studied in the literature, see for example [6], [12], [52], [53].

A channel code $\mathcal{C}_{2n}$ for this model consists of

- D message sets; $\mathcal{W}_l \triangleq [1{:}2^{nR_s}]$, $l = 1, 2, \ldots, D$,
- Cache encoder; $f_c : [1 : 2^{nR_s}]^D \mapsto \{0,1\}^n$,
- Cache decoders; $f_{c,j} : \{0,1\}^n \mapsto [1 : 2^{\frac{n}{2}}]$, $j = 1, 2$,
- Delivery encoders; $f_\mathbf{d} : [1 : 2^{nR_s}]^D \mapsto \{0,1\}^n$; where $\mathbf{d} \in [1 : D]^2$,
- Decoders; $g_{\mathbf{d},j} : [1 : 2^{\frac{n}{2}}] \times \{0,1\}^n \mapsto [1 : 2^{nR_s}]$; where $j = 1, 2$, $\mathbf{d} \in [1 : D]^2$.

The file rate R_s is *achievable with strong secrecy* if there is a sequence of codes $\{\mathcal{C}_{2n}\}_{n\geq 1}$ such that

$$\lim_{n\to\infty} \max_{\mathbf{d}\in[1\,:\,D]^2} \mathbb{P}\left(\bigcup_{j=1,2} (\hat{W}_{d_j} \neq W_{d_j})\right) = 0 \quad \textbf{(Reliability)}, \tag{2}$$

$$\lim_{n\to\infty} \max_{\substack{S_1,S_2\subseteq[1\,:\,n]: \\ |S_1|+|S_2|\leq\mu}} I(W_{[1:D]}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n) = 0 \textbf{ (Strong Secrecy)}. \tag{3}$$

That is, R_s is the *symmetric secure* file rate, under any demand vector and adversarial strategy. The strong secrecy capacity C_s is the supremum of all achievable R_s.

Remark 2: While we consider the file rate R_s which guarantees reliability for the worst-case demand vector, the average rate for which there exists a prior distribution on the demands has been studied in coded caching literature as well; see for example [8], [13], [54].

Remark 3: The condition in (3) guarantees strong secrecy against all possible strategies for the adversary, i.e., choices of S_1, S_2 that satisfy the condition $|S_1| + |S_2| \leq \mu$.

III. Main Results

For clarity of exposition, we first study the model in Section II when the transmitter's library has two files; $D = 2$. We then extend the ideas and analysis to $D > 2$. For $D > 2$, we use a similar channel coding scheme to that we construct for $D = 2$, but the placement and delivery schemes that achieve the best rates are different. The following theorem presents the strong secrecy capacity for $D = 2$.

Theorem 1: For $0 < \alpha \leq 2$ and $D = 2$, the strong secrecy capacity for the caching broadcast channel with a wire and cache tapping adversary of type II (CBC-WCT II), described in Section II, is given by

$$C_s(\alpha) = 1 - \frac{\alpha}{2}. \tag{4}$$

Proof: The proof is provided in Section IV. ∎

Theorem 2 below presents an achievable strong secrecy file rate for $D > 2$.

Theorem 2: For $0 < \alpha \leq 2$, $D > 2$, an achievable strong secrecy file rate for the CBC-WCT II is

$$R_s(\alpha) \geq \begin{cases} \frac{1}{2} + \frac{3(1-\alpha)}{4D}, & 0 < \alpha < 1 \\ 1 - \frac{\alpha}{2}, & 1 \leq \alpha \leq 2. \end{cases} \tag{5}$$

Proof: The proof is provided in Section V. ∎

The following theorem upper bounds the secure file rate when $D > 2$.

Theorem 3: For $0 < \alpha \leq 2$, $D > 2$, the strong secrecy file rate for the CBC-WCT II is upper bounded as

$$R_s(\alpha) \leq \begin{cases} \frac{1}{2} + \frac{2D-1}{2D(D-1)}(1-\alpha), & 0 < \alpha < 1 \\ 1 - \frac{\alpha}{2}, & 1 \leq \alpha \leq 2. \end{cases} \tag{6}$$

Proof: The proof is provided in Section VI. ∎

The following corollary is immediate from Theorems 1–3.

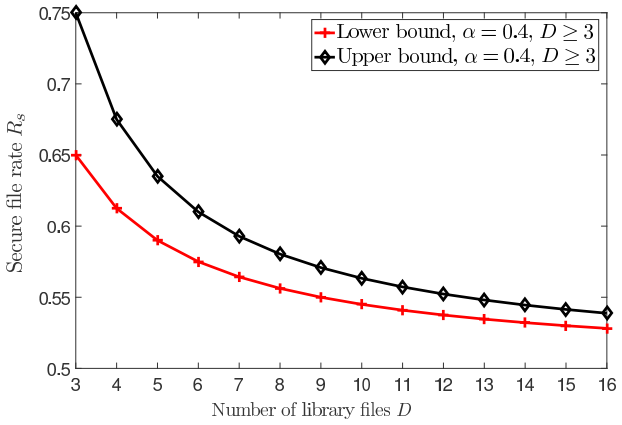

Fig. 2. Bounds on the achievable strong secrecy file rate R_s, when $\alpha = 0.4$ and $D \geq 3$.

Corollary 1: For $1 \leq \alpha \leq 2$, that is when the adversary can tap longer than one phase of communication, the strong secrecy capacity for the CBC-WCT II is

$$C_s(\alpha) = 1 - \frac{\alpha}{2}. \tag{7}$$

Remark 4: When $\alpha \in [1, 2]$ ($n \leq \mu \leq 2n$), two possible strategies for the adversary are $\{S_1 = [1 : n], S_2 \subset [1 : n]\}$ and $\{S_1 \subset [1 : n], S_2 = [1 : n]\}$; the adversary taps into all symbols in one phase and a subset of symbols in the other. Interestingly, the secrecy capacity for this range of α is $1 - \frac{\alpha}{2}$ for any library size. As we shall see in Sections IV-D and V, such an adversary limits the communication for cache placement to exchanging additional randomness (key bits) that allows for communicating a positive secure rate over the two phases. The cache memories are thus not used to store any information, and hence the lack of knowledge about user demands during cache placement is immaterial.

Unlike for $1 \leq \alpha \leq 2$, for $0 < \alpha < 1$, the lower and upper bounds in (5) and (6) have a gap. For illustration purposes, these bounds are plotted for $\alpha = 0.4$ in Fig. 2.

Remark 5: When $\alpha = 0$ (no adversary), our achievability scheme for $D > 2$ in Section V reduces to the achievability scheme in [6], which is shown to achieve the optimal rate-memory tradeoff for this case [52], [53]. However, the upper bound for $D > 2$ derived in this work is to address the intricacies of the adversarial model and is useful only when the adversary is present ($\alpha > 0$); (6) is loose when $\alpha = 0$.

IV. Proof of Theorem 1

In this section, we prove Theorem 1 which identifies the strong secrecy capacity when $D = 2$. Recall that the demand vector is denoted by $\mathbf{d} = (d_1, d_2)$, where $d_1, d_2 \in \{1, 2\}$.

A. Converse

When $\mathbf{d}$ is known to the transmitter during cache placement, the model in Theorem 1 reduces to a broadcast wiretap channel II (WTC-II), over a length-$2n$ communication block. The strong sum secrecy rate for that model, $2R_s$, is upper bounded by

$$2R_s \leq 2 - \alpha, \tag{8}$$

which follows from our recent work [47, Th. 1]. Note that (8) holds for any $\mathbf{d} = (d_1, d_2)$ such that $d_1 \neq d_2$, which represents the worst-case demands. Since $\mathbf{d}$ is unknown for the model in consideration, $1 - \frac{\alpha}{2}$ is an upper bound for its strong secrecy capacity.

B. Restricted Adversary Models as Building Blocks

Before proceeding with the achievability proof, it is relevant to take a step back and investigate the secrecy capacity when a known fraction of cache placement, a known fraction of delivery, or both, is tapped. In particular, we consider that the adversary taps into (i) cache placement only, (ii) delivery only, or (iii) both and the relative fractions of tapped symbols in each are known. For these three models, we show that the strong secrecy file rate in (4), i.e., $1 - \frac{\alpha}{2}$, is achievable, and hence determines their strong secrecy capacities. We then use these models as building blocks for when the relative fractions are *unknown*, and provide the achievability proof in Sections IV-C and IV-D.

1) Setting 1 (The Adversary Taps Into Cache Placement Only): This setting corresponds to $\alpha_1 = \alpha$ ($\alpha_2 = 0$) and $|S_1| = \mu$ ($S_2 = \varnothing$). The transmitter and receivers know that $\alpha_1 = \alpha$. We show that $1 - \frac{\alpha}{2}$ is an achievable strong secrecy file rate for this setting.

The transmitter divides W_l, $l = 1, 2$, into three independent messages, $W_l^{(1)}$, $W_l^{(2)}$, $W_{l,s}$; $W_l^{(1)}$, $W_l^{(2)}$ are uniform over $[1 : 2^{n\frac{1-\alpha-\epsilon_n}{2}}]$, $W_{l,s}$ is uniform over $[1 : 2^{n\frac{\alpha+\epsilon_n}{2}}]$. Define

$$M_c \triangleq \{M_{c,1}, M_{c,2}\};$$
$$M_{c,1} = W_1^{(1)} \oplus W_2^{(1)}, \quad M_{c,2} = W_1^{(2)} \oplus W_2^{(2)}, \tag{9}$$
$$M_{\mathbf{d}} \triangleq \left\{W_{d_1}^{(2)}, W_{d_2}^{(1)}, W_{d_1,s}, W_{d_2,s}\right\}, \tag{10}$$

where M_c and $M_{\mathbf{d}}$ are the messages sent by the transmitter during cache placement and delivery phases, respectively. Specifically, during placement, the transmitter maps M_c into $\mathbf{X}_c^n$ using stochastic wiretap coding [45]. Since the rate of M_c is less than $1 - \alpha$, M_c is strongly secure against the adversary that observes $n\alpha$ symbols of $\mathbf{X}_c^n$ [43], [44]. During delivery, the transmitter sends $\mathbf{X}_{\mathbf{d}}^n$ as the binary representation of $M_{\mathbf{d}}$ which is of length n bits, since the delivery phase is noiseless and secure.

Using $\mathbf{X}_c^n$, noiselessly received during placement, receiver j, $j = 1, 2$, recovers $M_{c,j}$ and stores it in its cache memory. The size of $M_{c,j}$ is smaller than $\frac{n}{2}$ bits, which is the cache size at each receiver. Using $\mathbf{X}_{\mathbf{d}}^n$, received noiselessly during delivery, both receivers recover $M_{\mathbf{d}}$. Using $M_{\mathbf{d}}$, along with its cache contents, $M_{c,j}$, and for n sufficiently large,[1] receiver j correctly recovers its desired message W_{d_j}, $j = 1, 2$.

For secrecy, we show in Appendix A that (3) is satisfied. Since $\epsilon_n \to 0$ as $n \to \infty$, the achievable strong secrecy file rate is given by

$$R_s(\alpha) = 2 \times \frac{1 - \alpha}{2} + \frac{\alpha}{2} = 1 - \frac{\alpha}{2}. \tag{11}$$

[1]Large block-length n is needed to ensure a valid subpacketization of the file W_l into the sub-files $\{W_l^{(1)}, W_l^{(2)}, W_{l,s}\}$, for $l = 1, 2$. That is, a bijective map between the file and its sub-files is preserved.

2) Setting 2 (The Adversary Taps Into the Delivery Only): This setting corresponds to $\alpha_1 = 0$ and $\alpha_2 = \alpha$, and the transmitter and receivers possess this knowledge. Once again, we show that $1 - \frac{\alpha}{2}$ is an achievable strong secrecy file rate. The transmitter (i) performs the same division of W_l, $l = 1, 2$, as in Setting 1, (ii) generates the keys K_1, K_2, each is uniform over $[1{:}2^{n\frac{\alpha+\epsilon_n}{2}}]$, independent from one another and from W_1, W_2. Define M_c, $M_{\mathbf{d}}$, $\tilde{M}_{\mathbf{d}}$, as follows:

$$M_c = \{M_{c,1}, M_{c,2}\};\ \ M_{c,1} = \left\{W_1^{(1)} \oplus W_2^{(1)}, K_1\right\},$$
$$M_{c,2} = \left\{W_1^{(2)} \oplus W_2^{(2)}, K_2\right\}, \tag{12}$$
$$M_{\mathbf{d}} = \left\{W_{d_1}^{(2)}, W_{d_2}^{(1)}\right\},\ \tilde{M}_{\mathbf{d}} = \left\{W_{d_1,s} \oplus K_1, W_{d_2,s} \oplus K_2\right\}. \tag{13}$$

During placement, the transmitter sends $\mathbf{X}_c^n$ as the binary representation of M_c, and receiver j, $j = 1, 2$, stores $M_{c,j}$ in its cache memory. During delivery, the transmitter encodes $M_{\mathbf{d}}$ into $\mathbf{X}_{\mathbf{d}}^n$ using wiretap coding, while using $\tilde{M}_{\mathbf{d}}$ as the randomization message. Receiver j recovers $M_{\mathbf{d}}$, $\tilde{M}_{\mathbf{d}}$, using which, along with $M_{c,j}$, it correctly decodes W_{d_j}, for sufficiently large n. By contrast, the adversary can only obtain $\tilde{M}_{\mathbf{d}}$ using which it can gain no information about W_1, W_2. In Appendix B, we show that (3) is satisfied. The achievable strong secrecy file rate is again $1 - \frac{\alpha}{2}$.

3) Setting 3 (The Legitimate Terminals Know the Values of α_1 and α_2): For this setting, neither $\alpha_1 = 0$ nor $\alpha_2 = 0$. However, the transmitter and receivers know the values of α_1, α_2. Under these assumptions, the scheme which achieves the secrecy rate $1 - \frac{\alpha}{2}$ depends on whether $\alpha_1 \geq \alpha_2$. For $\alpha_1 \geq \alpha_2$ ($\alpha_1 < \alpha_2$), we use an achievability scheme similar to Setting 1 (Setting 2).

Case 1 ($\alpha_1 \geq \alpha_2$): The transmitter divides W_l, $l = 1, 2$, into the independent messages $\{W_l^{(1)}, W_l^{(2)}, W_{l,s}\}$; $W_l^{(1)}, W_l^{(2)}$ are uniform over $\left[1 : 2^{n\frac{1-\alpha_1-\epsilon_n}{2}}\right]$ and $W_{l,s}$ is uniform over $\left[1 : 2^{n\frac{\alpha_1-\alpha_2}{2}}\right]$. The transmitter forms M_c, $M_{\mathbf{d}}$, as in (9), (10), and uses wiretap coding to map them into $\mathbf{X}_c^n$, $\mathbf{X}_{\mathbf{d}}^n$, respectively. As in setting 1, receiver j recovers W_{d_j}. For the secrecy constraint, note that M_c, $M_{\mathbf{d}}$ are independent, and their rates are $1-\alpha_1-\epsilon_n$, $1-\alpha_2-\epsilon_n$, respectively. Thus, wiretap coding strongly secures both M_c and $M_{\mathbf{d}}$ against the adversary. We show in Appendix C that (3) is satisfied. The achievable strong secrecy file rate is $R_s(\alpha) = 2 \times (\frac{1-\alpha_1}{2}) + \frac{\alpha_1-\alpha_2}{2} = 1 - \frac{\alpha}{2}$.

Case 2 ($\alpha_1 < \alpha_2$): The transmitter (i) divides W_l into $\{W_l^{(1)}, W_l^{(2)}, W_{l,s}\}$, where $W_l^{(1)}, W_l^{(2)}$ are uniform over $[1 : 2^{n\frac{1-\alpha_2-\epsilon_n}{2}}]$ and $W_{l,s}$ is uniform over $[1 : 2^{n\frac{\alpha_2-\alpha_1}{2}}]$, $l = 1, 2$; (ii) generates the keys K_1, K_2, uniformly over $[1 : 2^{n\frac{\alpha_2-\alpha_1}{2}}]$ and independently from W_1, W_2; (iii) forms M_c as in (12) and encodes it into $\mathbf{X}_c^n$ using wiretap coding; (iv) forms $M_{\mathbf{d}}$ as in (13) and forms $\tilde{M}_{\mathbf{d}}$ as

$$\tilde{M}_{\mathbf{d}} = \left\{W_{d_1,s} \oplus K_1, W_{d_2,s} \oplus K_2, \tilde{W}\right\}, \tag{14}$$

$\tilde{W}$ is independent from all other variables and uniform over $[1 : 2^{n(\alpha_1+\epsilon_n)}]$, and finally (v) encodes $M_{\mathbf{d}}$ into $\mathbf{X}_{\mathbf{d}}^n$ using wiretap coding, while using $\tilde{M}_{\mathbf{d}}$ as the randomization message.

As in Setting 2, receiver j correctly recovers W_{d_j}, and the adversary can only recover $\tilde{M}_{\mathbf{d}}$ using which it gains no information about W_1 and W_2. In Appendix D, we show (3) is satisfied. The achievable secrecy rate is $R_s(\alpha) = 2 \times (\frac{1-\alpha_2}{2}) + \frac{\alpha_2-\alpha_1}{2} = 1 - \frac{\alpha}{2}$.

With the aforementioned settings, we showed that the same secrecy rate, $1 - \frac{\alpha}{2}$, is achievable irrespective of where the adversary taps as long as α_1 and α_2 are known. The question then arises whether the lack of knowledge about relative fractions of tapped symbols would decrease the secrecy capacity. The following setting we propose provides a hint on the answer.

4) Setting 4 (Either $\alpha_1 = 0$ or $\alpha_2 = 0$, the Legitimate Terminals Do Not Know Which Is Zero): The adversary taps into either cache placement or delivery, but not both. The legitimate terminals *do not* know which phase is tapped. We show that the strong secrecy rate $1 - \frac{\alpha}{2}$ is again achievable.

The transmitter performs the same division of W_1 and W_2 as in Settings 1, 2, and generates independent keys K_1 and K_2 as in Setting 2. Define

$$M_c = \{M_{c,1}, M_{c,2}\};$$
$$M_{c,1} = W_1^{(1)} \oplus W_2^{(1)}, \quad M_{c,2} = W_1^{(2)} \oplus W_2^{(2)}, \tag{15}$$
$$\tilde{M}_c = \{\tilde{M}_{c,1}, \tilde{M}_{c,2}\};\ \ \tilde{M}_{c,1} = K_1,\ \tilde{M}_{c,2} = K_2 \tag{16}$$
$$M_{\mathbf{d}} = \left\{W_{d_1}^{(2)}, W_{d_2}^{(1)}\right\}, \tag{17}$$
$$\tilde{M}_{\mathbf{d}} = \left\{\tilde{M}_{\mathbf{d},1}, \tilde{M}_{\mathbf{d},2}\right\};\ \tilde{M}_{\mathbf{d},1} = W_{d_1,s} \oplus K_1,$$
$$\tilde{M}_{\mathbf{d},2} = W_{d_2,s} \oplus K_2. \tag{18}$$

During placement, the transmitter encodes M_c into $\mathbf{X}_c^n$ using wiretap coding, while using $\tilde{M}_c$ as the randomization message. Receiver j, $j = 1, 2$, stores $M_{c,j}$, $\tilde{M}_{c,j}$, in its cache. During delivery, the transmitter uses wiretap coding to encode $M_{\mathbf{d}}$ into $\mathbf{X}_{\mathbf{d}}^n$, while using $\tilde{M}_{\mathbf{d}}$ as the randomization message. Using its cache contents, and $M_{\mathbf{d}}$, $\tilde{M}_{\mathbf{d}}$, receiver j correctly decodes W_{d_j}. By contrast, the adversary can only recover either $\{K_1, K_2\}$ or $\{W_{d_1,s} \oplus K_1, W_{d_2,s} \oplus K_2\}$, but not both, using which it gains no information about W_1, W_2. We show in Appendix E that (3) is satisfied. The achievable strong secrecy rate is $1 - \frac{\alpha}{2}$.

The lack of knowledge about which phase is tapped is countered by encrypting pieces of information, $\{W_{d_1,s}, W_{d_2,s}\}$, with one-time pad keys, K_1, K_2, while ensuring the adversary only recovers either the keys or the encrypted bits but not both; using which it gains no information about W_1, W_2.

We next generalize this idea to tackle the case when the adversary taps into both phases, with no knowledge about the relative fractions of tapped symbols in each (the model in Fig. 1). Before continuing, let us first describe *security embedding codes* [49], [50]. The transmitted message is split into a number of layers, corresponding to different security levels. All layers of the message are encoded into a single codeword in an embedded fashion; each layer corresponds to one index identifying the codeword. Lower security-level layers serve as randomization (stochastic coding) for protecting higher security-level layers. The layers that can be securely transmitted are determined by the wiretapper's channel state.

Similar to [50] where the uncertainty about the wiretapper's channel is treated using security embedding codes, here, in each phase, we construct an embedding code in which $n\alpha$ single-bit layers are embedded into one another. Doing so, we

ensure that, no matter what the values for α_1, α_2 are, the adversary gets no more than $n\alpha_1$ bits from cache placement, and $n\alpha_2$ bits from delivery. By designing what the adversary recovers to be either a set of key bits and/or information bits encrypted with a distinct set of key bits, we guarantee no information on the messages is asymptotically leaked to the adversary. We thus prove that the lack of knowledge about relative fractions of tapped symbols *does not decrease* the secrecy capacity.

C. Achievability for $\alpha \in (0, 1)$

We are now ready to present the achievability for the general model when $D = 2$. Consider first $\alpha \in (0, 1)$. For simplicity, assume that $\frac{n\alpha_1}{2} = \frac{\mu_1}{2}$ and $\frac{n\alpha_2}{2} = \frac{\mu_2}{2}$ are integers. A minor modification to the analysis can be adopted otherwise. The transmitter (i) divides W_l, $l = 1, 2$, into the independent messages $W_l^{(1)}$, $W_l^{(2)}$, $W_{l,s}$; $W_l^{(1)}$, $W_l^{(2)}$ are uniform over $[1 : 2^{n\frac{1-\alpha}{2}}]$, and $W_{l,s}$ is uniform over $[1 : 2^{n\frac{\alpha}{2}}]$; (ii) generates the independent keys K_1, K_2, uniform over $[1 : 2^{n\frac{\alpha}{2}}]$, and independent from W_1, W_2. For simplicity of exposition, we have ignored the small rate reduction ϵ_n at this stage, as we will introduce this later into the security analysis. The main ideas of the achievability proof are:

1) The transmitter uses wiretap coding with a randomization message of size $n\alpha$ bits in *both placement and delivery*. As the adversary does not tap into more than $n\alpha$ bits in each phase, a secure transmission rate of $1-\alpha$ is achievable, as long as the randomization messages in the two phases are independent. Using coded placement for $\{W_l^{(1)}, W_l^{(2)}\}_{l=1,2}$, a secure file rate of $1 - \alpha$ can be achieved.
2) The randomization messages over the two phases can deliver additional secure information, of rate $\frac{\alpha}{2}$ per file, via encryption. The overall achievable file rate is $R_s = 1 - \frac{\alpha}{2}$. We use K_1, K_2, as the randomization message for placement. Along with wiretap coding, we employ a security embedding code [49], using bits of K_1, K_2, in a manner that allows the adversary to recover only the last $n\frac{\alpha_1}{2}$ bits from each. During delivery, we encrypt $W_{d_1,s}$, $W_{d_2,s}$, with K_1, K_2, and use this encrypted information as the randomization message. We employ again a security embedding code in the *reverse order* so that the adversary recovers only the first $n\frac{\alpha_2}{2}$ bits from each of $W_{d_1,s} \oplus K_1$, $W_{d_2,s} \oplus K_2$.
3) With the aforementioned construction, the adversary, for any values of α_1, α_2 it chooses, can only recover a set of key bits and/or a set of information bits encrypted with other key bits. Due to the *reversed embedding order*, the adversary does not obtain, during delivery, any message bits encrypted with key bits it has seen during placement. In addition, since $\{K_1, K_2\}$ is independent from $\{W_{d_1,s} \oplus K_1, W_{d_2,s} \oplus K_2\}$, and is an independent sequence, the adversary can not use the revealed key bits during cache placement to obtain any information about the bits of $W_{d_1,s} \oplus K_1$, $W_{d_2,s} \oplus K_2$ that need to be securely transmitted during delivery.

We now explain the achievability scheme in more detail. Let M_c, $\tilde{M}_c$ be as in (15), (16). M_c represents the message to be securely transmitted during placement, regardless of the adversary's choice of α_1. $\tilde{M}_c$ represents the randomization message used for wiretap coding. The transmitter further divides $\tilde{M}_{c,1}$, $\tilde{M}_{c,2}$ into sequences of independent bits, $\{\tilde{M}_{c,1}^{(1)}, \ldots, \tilde{M}_{c,1}^{(n\frac{\alpha}{2})}\}$, $\{\tilde{M}_{c,2}^{(1)}, \ldots, \tilde{M}_{c,2}^{(n\frac{\alpha}{2})}\}$, and generates $\mathbf{X}_c^n$ as:

Cache Placement Codebook Generation: Let m_c, $\tilde{m}_{c,1} = \{\tilde{m}_{c,1}^{(1)}, \ldots, \tilde{m}_{c,1}^{(n\frac{\alpha}{2})}\}$, $\tilde{m}_{c,2} = \{\tilde{m}_{c,2}^{(1)}, \ldots, \tilde{m}_{c,2}^{(n\frac{\alpha}{2})}\}$ be the realizations of M_c, $\tilde{M}_{c,1}$, $\tilde{M}_{c,2}$. We construct the codebook $\mathcal{C}_{c,n}$, from which $\mathbf{X}_c^n$ is drawn, as follows. We randomly and independently distribute all the possible 2^n length-n binary sequences into $2^{n(1-\alpha)}$ bins, indexed by $m_c \in [1{:}2^{n\frac{1-\alpha}{2}}]^2$, and each contains $2^{n\alpha}$ codewords. Further, we randomly and independently divide each bin m_c into two sub-bins, indexed by $\tilde{m}_{c,1}^{(1)}$, and each contains $2^{n\alpha-1}$ codewords. The two sub-bins $\tilde{m}_{c,1}^{(1)}$ are further divided into smaller bins, indexed by $\tilde{m}_{c,2}^{(1)}$, and each contains $2^{n\alpha-2}$ codewords. The process continues, going over $\tilde{m}_{c,1}^{(2)}$, $\tilde{m}_{c,2}^{(2)}$, $\cdots$, $\tilde{m}_{c,1}^{(n\frac{\alpha}{2}-1)}$, $\tilde{m}_{c,2}^{(n\frac{\alpha}{2}-1)}$, $\tilde{m}_{c,1}^{(n\frac{\alpha}{2})}$, until the remaining two codewords, after each sequence of divisions, are indexed by $\tilde{m}_{c,2}^{(n\frac{\alpha}{2})}$. $\mathcal{C}_{c,n}$ is described in Fig. 3.

Remark 6: An alternative representation of the former binning procedure is: Each of the $2^{n\alpha}$ codewords in the bin m_c, $m_c \in [1 : 2^{n\frac{1-\alpha}{2}}]^2$, is randomly assigned to an index $\{\tilde{m}_{c,1}^{(1)}, \tilde{m}_{c,2}^{(1)}, \ldots, \tilde{m}_{c,1}^{(n\frac{\alpha}{2})}, \tilde{m}_{c,2}^{(n\frac{\alpha}{2})}\}$. We chose to present the former description to provide a more detailed explanation of the embedding structure; in particular, the order of embedding which is a critical component in the achievability scheme.

Cache Encoder: Given w_1, w_2, the transmitter generates m_c, $\tilde{m}_c = \{\tilde{m}_{c,1}, \tilde{m}_{c,2}\}$ as in (15), (16), and sends $\mathbf{x}_c^n$, from $\mathcal{C}_{c,n}$, which corresponds to m_c, $\tilde{m}_{c,1}$, $\tilde{m}_{c,2}$, i.e., $\mathbf{x}_c^n(m_c, \tilde{m}_{c,1}^{(1)}, \tilde{m}_{c,2}^{(1)}, \ldots, \tilde{m}_{c,1}^{(n\frac{\alpha}{2})}, \tilde{m}_{c,2}^{(n\frac{\alpha}{2})})$.

For the delivery phase, define $M_{\mathbf{d}}$ and $\tilde{M}_{\mathbf{d}}$ as in (17) and (18). $M_{\mathbf{d}}$ represents the message to be securely transmitted during delivery no matter what the adversary's choice of α_2 is. $\tilde{M}_{\mathbf{d}}$ represents the randomization message. Similar to cache placement, the transmitter further divides $\tilde{M}_{\mathbf{d},1}$, $\tilde{M}_{\mathbf{d},2}$ into sequences of independent bits, $\{\tilde{M}_{\mathbf{d},1}^{(1)}, \ldots, \tilde{M}_{\mathbf{d},1}^{(n\frac{\alpha}{2})}\}$, $\{\tilde{M}_{\mathbf{d},2}^{(1)}, \ldots, \tilde{M}_{\mathbf{d},2}^{(n\frac{\alpha}{2})}\}$, and generates $\mathbf{X}_{\mathbf{d}}^n$ as follows.

Delivery Codebook Generation: Let $m_{\mathbf{d}}$, $\tilde{m}_{\mathbf{d},1} = \{\tilde{m}_{\mathbf{d},1}^{(1)}, \ldots, \tilde{m}_{\mathbf{d},1}^{(n\frac{\alpha}{2})}\}$, $\tilde{m}_{\mathbf{d},2} = \{\tilde{m}_{\mathbf{d},2}^{(1)}, \ldots, \tilde{m}_{\mathbf{d},2}^{(n\frac{\alpha}{2})}\}$ be the realizations of $M_{\mathbf{d}}$, $\tilde{M}_{\mathbf{d},1}$, $\tilde{M}_{\mathbf{d},2}$. We construct $\mathcal{C}_{\mathbf{d},n}$, from which $\mathbf{X}_{\mathbf{d}}^n$ is drawn, in a similar fashion as $\mathcal{C}_{c,n}$, but with a reversed indexing of the sub-bins. We randomly and independently divide all the 2^n binary sequences into $2^{n(1-\alpha)}$ bins, indexed by $m_{\mathbf{d}} \in [1{:}2^{n\frac{1-\alpha}{2}}]^2$, and each contains $2^{n\alpha}$ codewords. We further randomly and independently divide each bin $m_{\mathbf{d}}$ into two sub-bins, indexed by $\tilde{m}_{\mathbf{d},1}^{(n\frac{\alpha}{2})}$, and each contains $2^{n\alpha-1}$ codewords. The process continues, going in reverse order over $\tilde{m}_{\mathbf{d},2}^{(n\frac{\alpha}{2})}$, $\tilde{m}_{\mathbf{d},1}^{(n\frac{\alpha}{2}-1)}$, $\tilde{m}_{\mathbf{d},2}^{(n\frac{\alpha}{2}-1)}$, $\cdots$, $\tilde{m}_{\mathbf{d},1}^{(1)}$, until the remaining two codewords, after each sequence of divisions, are indexed by $\tilde{m}_{\mathbf{d},2}^{(1)}$. The codebook $\mathcal{C}_{\mathbf{d},n}$ is shown in Fig. 4.

Delivery Encoder: Given w_1, w_2 and $\mathbf{d} = (d_1, d_2)$, the transmitter generates $m_{\mathbf{d}}$, $\tilde{m}_{\mathbf{d}} = \{\tilde{m}_{\mathbf{d},1}, \tilde{m}_{\mathbf{d},2}\}$ as in (17), (18). The transmitter sends $\mathbf{x}_{\mathbf{d}}^n$, from $\mathcal{C}_{\mathbf{d},n}$, which corresponds to $m_{\mathbf{d}}$, $\tilde{m}_{\mathbf{d},1}$, and $\tilde{m}_{\mathbf{d},2}$, i.e., $\mathbf{x}_{\mathbf{d}}^n(m_{\mathbf{d}}, \tilde{m}_{\mathbf{d},1}^{(n\frac{\alpha}{2})}, \tilde{m}_{\mathbf{d},2}^{(n\frac{\alpha}{2})}, \ldots, \tilde{m}_{\mathbf{d},1}^{(1)}, \tilde{m}_{\mathbf{d},2}^{(1)})$.

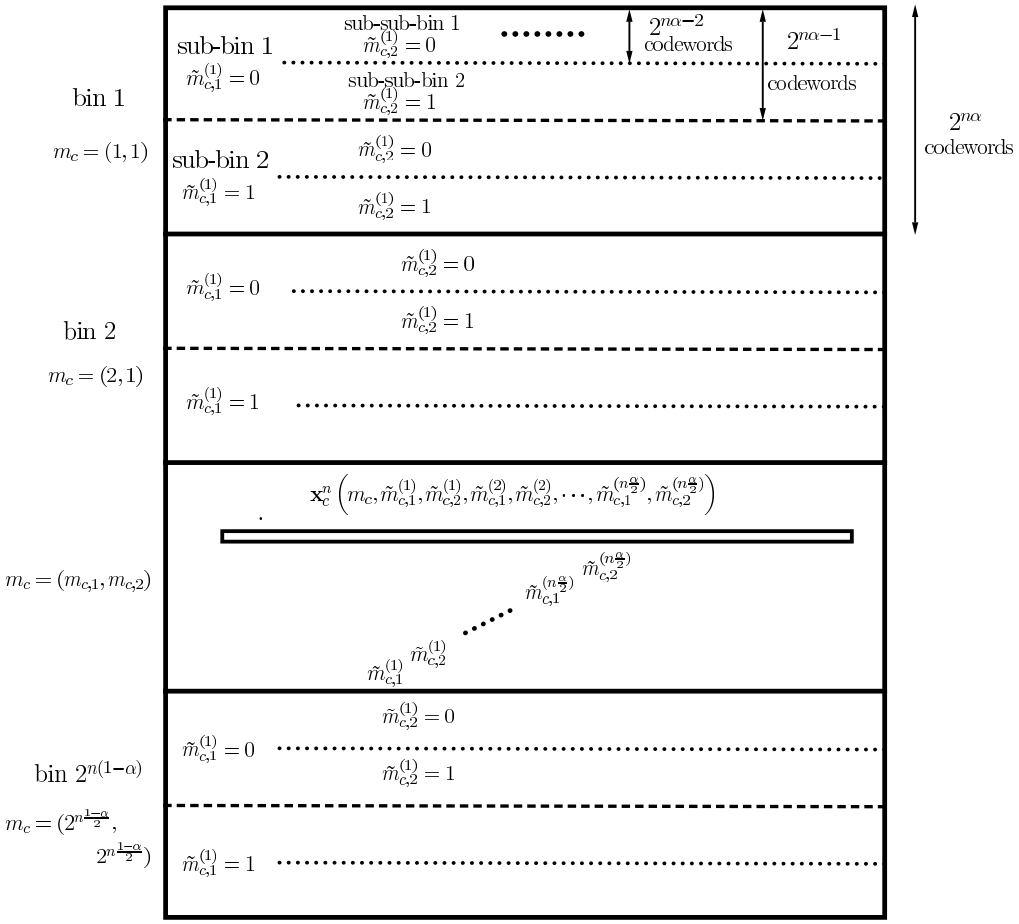

Fig. 3. Codebook construction for the cache placement phase, $\mathcal{C}_{c,n}$.

Decoding: Using $\mathbf{X}_c^n$, receiver j, $j = 1, 2$, recovers $M_{c,j}$, $\tilde{M}_{c,j}$, and stores them in its cache. For $j = 1, 2$, the combined size of $M_{c,j}$ and $\tilde{M}_{c,j}$ does not exceed $\frac{n}{2}$ bits. Using $\mathbf{X}_{\mathbf{d}}^n$, both receivers recover $M_{\mathbf{d}}$, $\tilde{M}_{\mathbf{d}}$. Using $M_{\mathbf{d}}$, $\tilde{M}_{\mathbf{d}}$, $M_{c,j}$, $\tilde{M}_{c,j}$, and for n sufficiently large, receiver j correctly decodes W_{d_j}.

Security Analysis: Let us first slightly modify the construction above as follows. Recall that $\{\epsilon_n\}_{n\geq 1}$ is a sequence of positive real numbers such that $\epsilon_n \to 0$ as $n \to \infty$. Define

$$\alpha_\epsilon = \alpha + 2\epsilon_n,\ \alpha_{1,\epsilon} = \alpha_1 + \epsilon_n,\ \alpha_{2,\epsilon} = \alpha_2 + \epsilon_n. \tag{19}$$

That is, $\alpha_{1,\epsilon} + \alpha_{2,\epsilon} = \alpha_\epsilon$. We increase the sizes of K_1, K_2, into $\frac{n\alpha_\epsilon}{2}$ bits, from $n\frac{\alpha}{2}$, and zero-pad the bit strings of $W_{d_1,s}$, $W_{d_2,s}$, accordingly. We also decrease the sizes of $W_l^{(1)}$, $W_l^{(2)}$, $l = 1, 2$, to $n\frac{1-\alpha_\epsilon}{2}$ bits, instead of $n\frac{1-\alpha}{2}$. We assume that $\frac{n\alpha_\epsilon}{2}$, $\frac{n\alpha_{1,\epsilon}}{2}$ are integers; as minor modifications can be adopted otherwise.

Fix $S_1, S_2 \subseteq [1:n]$. For the corresponding (fixed) values of α_1, α_2, the cache placement codebook $\mathcal{C}_{c,n}$ can be viewed as a wiretap code with $2^{n(1-\alpha_{1,\epsilon})}$ bins. Each bin is indexed by the message

$$w_c = \left(m_c, \tilde{m}_{c,1}^{(1)}, \tilde{m}_{c,2}^{(1)}, \tilde{m}_{c,1}^{(2)}, \tilde{m}_{c,2}^{(2)}, \ldots, \tilde{m}_{c,1}^{\left(n\frac{\alpha_{2,\epsilon}}{2}\right)}, \tilde{m}_{c,2}^{\left(n\frac{\alpha_{2,\epsilon}}{2}\right)}\right). \tag{20}$$

Each bin w_c contains $2^{n\alpha_{1,\epsilon}}$ binary codewords which are indexed by the randomization message

$$\tilde{w}_c = \left(\tilde{m}_{c,1}^{\left(n\frac{\alpha_{2,\epsilon}}{2}+1\right)}, \tilde{m}_{c,2}^{\left(n\frac{\alpha_{2,\epsilon}}{2}+1\right)}, \tilde{m}_{c,1}^{\left(n\frac{\alpha_{2,\epsilon}}{2}+2\right)}, \tilde{m}_{c,2}^{\left(n\frac{\alpha_{2,\epsilon}}{2}+2\right)}, \ldots, \tilde{m}_{c,1}^{\left(n\frac{\alpha_\epsilon}{2}\right)}, \tilde{m}_{c,2}^{\left(n\frac{\alpha_\epsilon}{2}\right)}\right). \tag{21}$$

Similarly, $\mathcal{C}_{\mathbf{d},n}$ can be seen as a wiretap code with $2^{n(1-\alpha_{2,\epsilon})}$ bins, each is indexed by the message

$$w_{\mathbf{d}} = \left(m_{\mathbf{d}}, \tilde{m}_{\mathbf{d},1}^{\left(n\frac{\alpha_\epsilon}{2}\right)}, \tilde{m}_{\mathbf{d},2}^{\left(n\frac{\alpha_\epsilon}{2}\right)}, \tilde{m}_{\mathbf{d},1}^{\left(n\frac{\alpha_\epsilon}{2}-1\right)}, \tilde{m}_{\mathbf{d},2}^{\left(n\frac{\alpha_\epsilon}{2}-1\right)}, \ldots, \tilde{m}_{\mathbf{d},1}^{\left(n\frac{\alpha_{2,\epsilon}}{2}+1\right)}, \tilde{m}_{\mathbf{d},2}^{\left(n\frac{\alpha_{2,\epsilon}}{2}+1\right)}\right). \tag{22}$$

Each bin $w_{\mathbf{d}}$ contains $2^{n\alpha_{2,\epsilon}}$ codewords, indexed by the randomization message

$$\tilde{w}_{\mathbf{d}} = \left(\tilde{m}_{\mathbf{d},1}^{\left(n\frac{\alpha_{2,\epsilon}}{2}\right)}, \tilde{m}_{\mathbf{d},2}^{\left(n\frac{\alpha_{2,\epsilon}}{2}\right)}, \tilde{m}_{\mathbf{d},1}^{\left(n\frac{\alpha_{2,\epsilon}}{2}-1\right)}, \tilde{m}_{\mathbf{d},2}^{\left(n\frac{\alpha_{2,\epsilon}}{2}-1\right)}, \ldots, \tilde{m}_{\mathbf{d},1}^{(1)}, \tilde{m}_{\mathbf{d},2}^{(1)}\right). \tag{23}$$

Let $\{\mathcal{B}_{w_c} : w_c \in [1:2^{n(1-\alpha_{1,\epsilon})}]\}$, $\{\mathcal{B}_{w_{\mathbf{d}}} : w_{\mathbf{d}} \in [1:2^{n(1-\alpha_{2,\epsilon})}]\}$ denote the partition (bins) of $\mathcal{C}_{c,n}$, $\mathcal{C}_{\mathbf{d},n}$, which correspond to

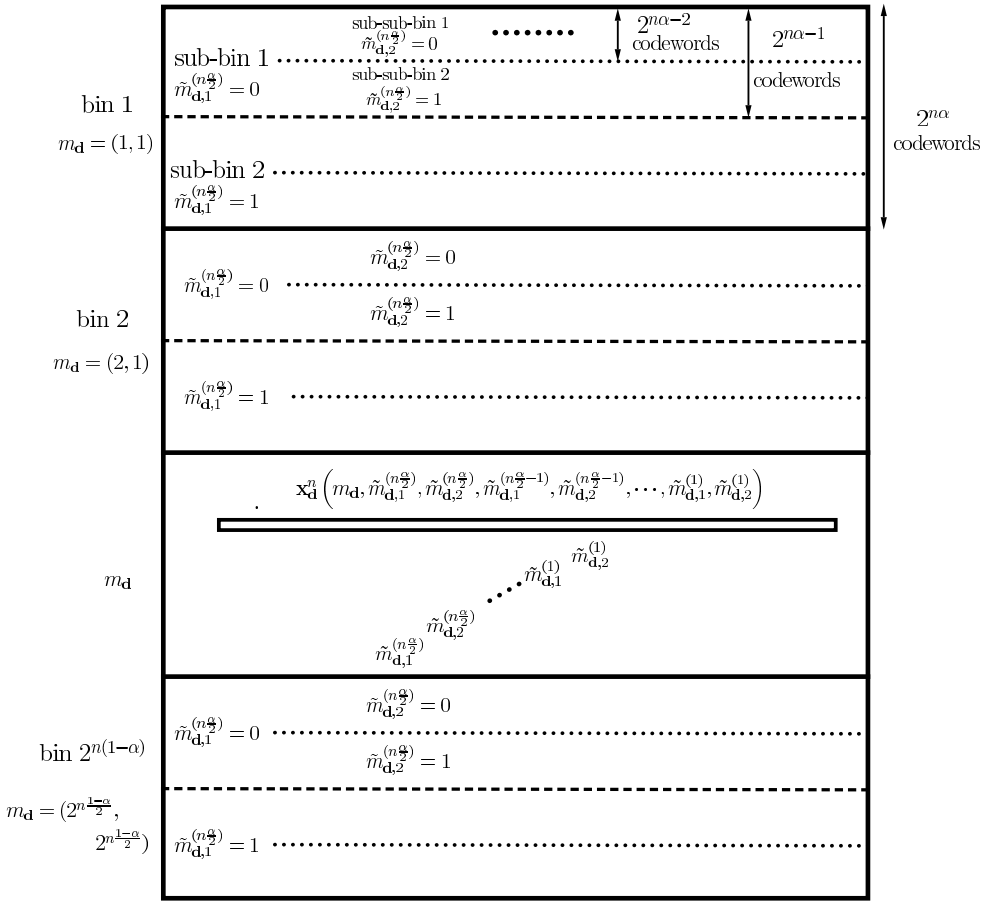

Fig. 4. Codebook construction for the delivery phase, $\mathcal{C}_{\mathbf{d},n}$.

w_c, $w_{\mathbf{d}}$ in (20), (22), respectively. Let $\mathbf{x}^{2n} \triangleq (\mathbf{x}_c^n, \mathbf{x}_{\mathbf{d}}^n)$ denote the concatenation of the two length-n binary codewords $\mathbf{x}_c^n$, $\mathbf{x}_{\mathbf{d}}^n$. Define the Cartesian product of the bins $\mathcal{B}_{w_c}$ and $\mathcal{B}_{w_{\mathbf{d}}}$, as

$$\mathcal{B}_{w_c,w_{\mathbf{d}}} \triangleq \left\{\mathbf{x}^{2n} = \left(\mathbf{x}_c^n, \mathbf{x}_{\mathbf{d}}^n\right) : \mathbf{x}_c^n \in \mathcal{B}_{w_c},\ \mathbf{x}_{\mathbf{d}}^n \in \mathcal{B}_{w_{\mathbf{d}}}\right\}. \quad (24)$$

Since the partitioning of $\mathcal{C}_{c,n}$ and $\mathcal{C}_{\mathbf{d},n}$ is random, for every w_c, $w_{\mathbf{d}}$; $\mathcal{B}_{w_c,w_{\mathbf{d}}}$ is a random codebook which results from the Cartesian product of the random bins $\mathcal{B}_{w_c}$ and $\mathcal{B}_{w_{\mathbf{d}}}$. Recall that $\mathcal{B}_{w_c}$ contains $2^{n\alpha_{1,\epsilon}}$ and $\mathcal{B}_{w_{\mathbf{d}}}$ contains $2^{n\alpha_{2,\epsilon}}$ length-n binary codewords. Thus, the product $\mathcal{B}_{w_c,w_{\mathbf{d}}}$ contains $2^{n\alpha_\epsilon}$ length-$2n$ binary codewords. Let $\{W_{d_l,s}^{(1)}, \ldots, W_{d_l,s}^{(n\frac{\alpha_\epsilon}{2})}\}$ and $\{K_l^{(1)}, \ldots, K_l^{(n\frac{\alpha_\epsilon}{2})}\}$ denote the bit strings of $W_{d_l,s}$ and K_l, $l = 1, 2$. In addition, for notational simplicity, define

$$\mathbf{W}_s^{(1)} = \left\{W_{d_1,s}^{(1)}, W_{d_2,s}^{(1)}, \ldots, W_{d_1,s}^{\left(n\frac{\alpha_{2,\epsilon}}{2}\right)}, W_{d_2,s}^{\left(n\frac{\alpha_{2,\epsilon}}{2}\right)}\right\} \quad (25)$$

$$\mathbf{W}_s^{(2)} = \left\{W_{d_1,s}^{\left(n\frac{\alpha_{2,\epsilon}}{2}+1\right)}, W_{d_2,s}^{\left(n\frac{\alpha_{2,\epsilon}}{2}+1\right)}, \ldots, W_{d_1,s}^{\left(n\frac{\alpha_\epsilon}{2}\right)}, W_{d_2,s}^{\left(n\frac{\alpha_\epsilon}{2}\right)}\right\} \quad (26)$$

$$\mathbf{K}^{(1)} = \left\{K_1^{(1)}, K_2^{(1)}, \ldots, K_1^{\left(n\frac{\alpha_{2,\epsilon}}{2}\right)}, K_2^{\left(n\frac{\alpha_{2,\epsilon}}{2}\right)}\right\} \quad (27)$$

$$\mathbf{K}^{(2)} = \left\{K_1^{\left(n\frac{\alpha_{2,\epsilon}}{2}+1\right)}, K_2^{\left(n\frac{\alpha_{2,\epsilon}}{2}+1\right)}, \ldots, K_1^{\left(n\frac{\alpha_\epsilon}{2}\right)}, K_2^{\left(n\frac{\alpha_\epsilon}{2}\right)}\right\} \quad (28)$$

$$\mathbf{W}_{\oplus\mathbf{K}}^{(1)} = \left\{W_{d_1,s}^{(i)} \oplus K_1^{(i)}, W_{d_2,s}^{(i)} \oplus K_2^{(i)} \colon i = 1, 2, \ldots, n\frac{\alpha_{2,\epsilon}}{2}\right\} \quad (29)$$

$$\mathbf{W}_{\oplus\mathbf{K}}^{(2)} = \left\{W_{d_1,s}^{(i)} \oplus K_1^{(i)}, W_{d_2,s}^{(i)} \oplus K_2^{(i)} \colon i = n\frac{\alpha_{2,\epsilon}}{2}+1, n\frac{\alpha_{2,\epsilon}}{2}+2, \ldots, n\frac{\alpha_\epsilon}{2}\right\}. \quad (30)$$

Let W_c, $\tilde{W}_c$, $W_{\mathbf{d}}$, and $\tilde{W}_{\mathbf{d}}$ denote the random variables that correspond to the realizations defined in (20)–(23). Using (15)–(18), (20)–(23), and (27)–(30), we have

$$W_c = \left\{M_c, \mathbf{K}^{(1)}\right\} = \left\{W_1^{(1)} \oplus W_2^{(1)}, W_1^{(2)} \oplus W_2^{(2)}, \mathbf{K}^{(1)}\right\}, \quad \tilde{W}_c = \mathbf{K}^{(2)} \quad (31)$$

$$W_{\mathbf{d}} = \left\{M_{\mathbf{d}}, \mathbf{W}_{\oplus\mathbf{K}}^{(2)}\right\} = \left\{W_{d_1}^{(2)}, W_{d_2}^{(1)}, \mathbf{W}_{\oplus\mathbf{K}}^{(2)}\right\}, \quad \tilde{W}_{\mathbf{d}} = \mathbf{W}_{\oplus\mathbf{K}}^{(1)}. \quad (32)$$

Notice that $\tilde{W}_c$ and $\tilde{W}_{\mathbf{d}}$ are independent, and each is uniformly distributed. $\{\tilde{W}_c, \tilde{W}_{\mathbf{d}}\}$ is thus jointly uniform. In addition, $\{\tilde{W}_c, \tilde{W}_{\mathbf{d}}\}$ is independent from $\{W_c, W_{\mathbf{d}}\}$. Thus, we can apply the analysis in [43, eqs. (94)-(103)] to show that, for every S_1, S_2, w_c, and $w_{\mathbf{d}}$, and every $\epsilon > 0$, there exists $\gamma(\epsilon) > 0$ such that

$$\mathbb{P}_{\mathcal{B}_{w_c,w_{\mathbf{d}}}}\left(\mathbb{D}\left(P_{\mathbf{Z}_{S_1}^n \mathbf{Z}_{S_2}^n \left| W_c=w_c, W_{\mathbf{d}}=w_{\mathbf{d}}\right.} || P_{\mathbf{Z}_{S_1}^n} P_{\mathbf{Z}_{S_2}^n}\right) > \epsilon\right) \leq \exp\left(-e^{n\gamma(\epsilon)}\right). \quad (33)$$

$P_{\mathbf{Z}_{S_1}^n \mathbf{Z}_{S_2}^n | W_c=w_c, W_{\mathbf{d}}=w_{\mathbf{d}}}$ is the induced distribution at the adversary when $\mathbf{x}_c^n(w_c, \tilde{w}_c)$, $\mathbf{x}_{\mathbf{d}}^n(w_{\mathbf{d}}, \tilde{w}_{\mathbf{d}})$ are the transmitted codewords over placement and delivery. $P_{\mathbf{Z}_{S_1}^n \mathbf{Z}_{S_2}^n}$ is the output distribution at the adversary when the placement and delivery codewords $\mathbf{X}_c^n$, $\mathbf{X}_{\mathbf{d}}^n$ are drawn independently and identically at random according to the input distribution, which is uniform over $\{0, 1\}$. Equation (33) states that the probability that these two distributions are not equal converges to zero doubly-exponentially fast with the block-length n.

The number of the messages $\{w_c, w_{\mathbf{d}}\}$ is $2^{n(2-\alpha_\epsilon)}$, and the number of possible choices for the subsets S_1, S_2, is $\binom{2n}{\alpha n} < 2^{2n}$. The combined number of the messages and the subsets is at most exponential in n. Using (33) and the union bound, as in [43], [44], we have

$$\lim_{n\to\infty} \max_{S_1,S_2} I\Big(W_c, W_{\mathbf{d}}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) = 0. \tag{34}$$

We also have, for any $\mathbf{d} = (d_1, d_2)$, $d_1, d_2 \in \{1, 2\}$,

$$\begin{aligned}
&I\Big(W_1, W_2; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) \\
&\quad = I\Big(W_1^{(1)}, W_1^{(2)}, W_2^{(1)}, W_2^{(2)}, W_{1,s}, W_{2,s}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) && (35)\\
&\quad = I\Big(W_1^{(1)}, W_1^{(2)}, W_2^{(1)}, W_2^{(2)}, \mathbf{W}_s^{(1)}, \mathbf{W}_s^{(2)}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) && (36)\\
&\quad = I\Big(W_1^{(1)} \oplus W_2^{(1)}, W_1^{(2)} \oplus W_2^{(2)}, W_{d_1}^{(2)}, W_{d_2}^{(1)}, \\
&\qquad\qquad \mathbf{W}_s^{(1)}, \mathbf{W}_s^{(2)}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) && (37)\\
&\quad = I\Big(M_c, M_{\mathbf{d}}, \mathbf{W}_s^{(1)}, \mathbf{W}_s^{(2)}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) && (38)\\
&\quad \le I\Big(M_c, M_{\mathbf{d}}, \mathbf{W}_s^{(1)}, \mathbf{W}_{\oplus\mathbf{K}}^{(2)}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) && (39)\\
&\quad = I\Big(M_c, \mathbf{W}_s^{(1)}, W_{\mathbf{d}}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) && (40)\\
&\quad = H\Big(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) - H\Big(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n \Big| M_c, \mathbf{W}_s^{(1)}, W_{\mathbf{d}}\Big), && (41)
\end{aligned}$$

where (36) follows because $\mathbf{Z}_{S_1}^n$, $\mathbf{Z}_{S_2}^n$ depend only on $\{W_1^{(1)}, W_1^{(2)}, W_2^{(1)}, W_2^{(2)}, W_{d_1,s}, W_{d_2,s}\}$, and by using (25), (26); (37) follows because there is a bijection between $\{W_1^{(1)}, W_1^{(2)}, W_2^{(1)}, W_2^{(2)}\}$ and $\{W_1^{(1)} \oplus W_2^{(1)}, W_1^{(2)} \oplus W_2^{(2)}, W_{d_1}^{(2)}, W_{d_2}^{(1)}\}$; (38) follows from (15), (17); (39) follows due to the Markov chain $\mathbf{W}_s^{(2)} - \{M_c, M_{\mathbf{d}}, \mathbf{W}_s^{(1)}, \mathbf{W}_{\oplus\mathbf{K}}^{(2)}\} - \{\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\}$, and the data processing inequality. This Markov chain holds because $\{M_c, M_{\mathbf{d}}, \mathbf{W}_s^{(1)}\}$ are independent from $\{\mathbf{W}_s^{(2)}, \mathbf{K}^{(2)}\}$, and only the encrypted information $\mathbf{W}_{\oplus\mathbf{K}}^{(2)}$ is transmitted. (40) follows from (32). The second term on the right hand side of (41) is lower bounded as

$$\begin{aligned}
&H\Big(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n \Big| M_c, \mathbf{W}_s^{(1)}, W_{\mathbf{d}}\Big) \\
&\quad = H\Big(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n, \mathbf{W}_s^{(1)} | M_c, W_{\mathbf{d}}\Big) - H\Big(\mathbf{W}_s^{(1)} | M_c, W_{\mathbf{d}}\Big) && (42)\\
&\quad = H\Big(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n, \mathbf{W}_s^{(1)}, \mathbf{W}_{\oplus K}^{(1)} | M_c, W_{\mathbf{d}}\Big) \\
&\qquad - H\Big(\mathbf{W}_{\oplus K}^{(1)} | M_c, W_{\mathbf{d}}, \mathbf{W}_s^{(1)}, \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) - H\Big(\mathbf{W}_s^{(1)}\Big) && (43)\\
&\quad = H\Big(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n, \mathbf{K}^{(1)}, \mathbf{W}_{\oplus K}^{(1)} | M_c, W_{\mathbf{d}}\Big) \\
&\qquad - H\Big(\mathbf{K}^{(1)} | M_c, W_{\mathbf{d}}, \mathbf{W}_s^{(1)}, \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) - H\Big(\mathbf{W}_s^{(1)}\Big) && (44)\\
&\quad \ge H\Big(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n, \mathbf{K}^{(1)}, \mathbf{W}_{\oplus K}^{(1)} | M_c, W_{\mathbf{d}}\Big) - H\Big(\mathbf{W}_s^{(1)}\Big) - \epsilon_n' && (45)\\
&\quad \ge H\Big(\mathbf{K}^{(1)} | M_c, W_{\mathbf{d}}\Big) + H\Big(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n | M_c, \mathbf{K}^{(1)}, W_{\mathbf{d}}\Big) \\
&\qquad\qquad - H\Big(\mathbf{W}_s^{(1)}\Big) - \epsilon_n' && (46)\\
&\quad = H\Big(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n | W_c, W_{\mathbf{d}}\Big) + H\Big(\mathbf{K}^{(1)}\Big) - H\Big(\mathbf{W}_s^{(1)}\Big) - \epsilon_n' && (47)\\
&\quad = H\Big(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n | W_c, W_{\mathbf{d}}\Big) - \epsilon_n', && (48)
\end{aligned}$$

where $\epsilon_n' \to 0$ as $n \to \infty$; (43) follows since $\mathbf{W}_s^{(1)}$ is independent from $\{M_c, W_{\mathbf{d}}\}$; (44) follows because there is a bijection between $\{\mathbf{W}_s^{(1)}, \mathbf{W}_{\oplus K}^{(1)}\}$ and $\{\mathbf{K}^{(1)}, \mathbf{W}_{\oplus K}^{(1)}\}$; (47) follows from (31), and since $\mathbf{K}^{(1)}$ is independent from $\{M_c, W_{\mathbf{d}}\}$; (48) follows since $\mathbf{K}^{(1)}$, $\mathbf{W}_s^{(1)}$ are independent and identically distributed.

The inequality in (45) follows since, given $\{M_c, \mathbf{W}_s^{(1)}, W_{\mathbf{d}}\}$, and for n sufficiently large, the adversary can decode $\mathbf{K}^{(1)}$ using its observations $\mathbf{Z}_{S_1}^n$, $\mathbf{Z}_{S_2}^n$. In particular, $\{M_c, \mathbf{W}_s^{(1)}, W_{\mathbf{d}}\}$ determine a partition of the codebook into bins, each of which contains $2^{n\alpha_\epsilon}$ codewords. For n large enough, and given the values of M_c, $\mathbf{W}_s^{(1)}$, $W_{\mathbf{d}}$, i.e., the bin index, the adversary can determine the codeword index inside the bin, and hence decode $\mathbf{K}^{(1)}$. We conclude that, $H(\mathbf{K}^{(1)} | M_c, W_{\mathbf{d}}, \mathbf{W}_s^{(1)}, \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n) \le \epsilon_n'$, where $\epsilon_n' \to 0$ as $n \to \infty$.

Substituting (48) in (41) gives

$$I\Big(W_1, W_2; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) \le I\Big(W_c, W_{\mathbf{d}}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) + \epsilon_n'. \tag{49}$$

Using (34) and (49), the secrecy constraint in (3) is satisfied. Since $\epsilon_n \to 0$ as $n \to \infty$, we conclude that, the achievable strong secrecy file rate is

$$R_s(\alpha) = 2 \times \frac{1-\alpha}{2} + \frac{\alpha}{2} = 1 - \frac{\alpha}{2}. \tag{50}$$

Remark 7: Although the codebooks $\mathcal{C}_{c,n}$, $\mathcal{C}_{\mathbf{d},n}$ are designed and generated disjointly, in the security analysis, we considered the Cartesian products of the individual bins of the two codebooks. We were able to do so since the input distributions for generating the codebooks are identical, i.e., uniform binary.

D. Achievability for $\alpha \in [1, 2]$

For $\alpha \in [1, 2]$, we adapt the scheme in Section IV-C as follows: W_1, W_2 are uniform over $[1 : 2^{n\frac{2-\alpha_\epsilon}{2}}]$; α_ϵ is as in (19). The transmitter (i) generates the independent keys K_1, K_2, uniform over $[1 : 2^{n\frac{2-\alpha_\epsilon}{2}}]$ and independent from W_1, W_2; (ii) generates the independent randomization messages $\tilde{W}$, $\tilde{W}_K$, uniform over $[1 : 2^{n(\alpha_\epsilon - 1)}]$ and independent from W_1, W_2, K_1, K_2. The messages for placement at receivers 1, 2 are

$$M_{c,1} = K_1, \qquad M_{c,2} = K_2, \tag{51}$$

i.e., receiver j stores the key K_j in its cache. The message to be securely transmitted during delivery is

$$M_{\mathbf{d}} = \big\{M_{\mathbf{d},1}, M_{\mathbf{d},2}\big\};\ M_{\mathbf{d},1} = W_{d_1} \oplus K_1, M_{\mathbf{d},2} = W_{d_2} \oplus K_2. \tag{52}$$

Note that, for $\alpha \in [1, 2]$, the adversary can see all symbols in at least one of the phases. Therefore, unlike Section IV-C,

we cannot utilize randomization messages, $\tilde{W}$, $\tilde{W}_K$, to carry any information; only keys are stored in the cache memories of the receivers. In addition, the placement and delivery codebooks for this case have a different embedding structure than for $\alpha \in (0, 1)$ in Section IV-C. In particular, the embedding here is performed on the bits of the messages M_c, $M_{\mathbf{d}}$, while the embedding in Section IV-C is performed on the bits of the randomization messages $\tilde{M}_c$, $\tilde{M}_{\mathbf{d}}$. Let $\{W_{d_l}^{(1)}, \ldots, W_{d_l}^{(n\frac{2-\alpha_\epsilon}{2})}\}$, $\{K_l^{(1)}, \ldots, K_l^{(n\frac{2-\alpha_\epsilon}{2})}\}$, $\{M_{\mathbf{d},l}^{(1)}, \ldots, M_{\mathbf{d},l}^{(n\frac{2-\alpha_\epsilon}{2})}\}$ denote the bit strings of W_{d_l}, K_l, $M_{\mathbf{d},l}$, $l = 1, 2$.

Cache Placement Codebook: The codebook $\mathcal{C}_{c,n}$ is generated as follows: The transmitter randomly and independently divides all the 2^n length-n binary sequences into 2 bins, indexed by $K_1^{(1)}$, and each contains 2^{n-1} codewords. These bins are further randomly and independently divided into two sub-bins, indexed by $K_2^{(1)}$, and each contains 2^{n-2} codewords. The process continues, going over $K_1^{(2)}$, $K_2^{(2)}, \ldots, K_1^{(n\frac{2-\alpha_\epsilon}{2})}$, $K_2^{(n\frac{2-\alpha_\epsilon}{2})}$, until the remaining $2^{n(\alpha_\epsilon-1)}$ codewords, after each sequence of divisions, are indexed by $\tilde{W}_K$.

Cache Encoder: The transmitter sends $\mathbf{X}_c^n$ which corresponds to the keys K_1, K_2, and the randomization message $\tilde{W}_K$, i.e., $\mathbf{X}_c^n(K_1^{(1)}, K_2^{(1)}, \ldots, K_1^{(n\frac{2-\alpha_\epsilon}{2})}, K_2^{(n\frac{2-\alpha_\epsilon}{2})}, \tilde{W}_K)$.

Delivery Codebook: The codebook $\mathcal{C}_{\mathbf{d},n}$ is generated as follows. The transmitter randomly and independently divides all the 2^n length-n binary sequences into two bins, indexed by $M_{\mathbf{d},1}^{(n\frac{2-\alpha_\epsilon}{2})}$, and each contains 2^{n-1} codewords. These bins are further randomly and independently divided into two sub-bins, indexed by $M_{\mathbf{d},2}^{(n\frac{2-\alpha_\epsilon}{2})}$, and each contains 2^{n-2} codewords. The process continues, going in reverse order over $M_{\mathbf{d},1}^{(n\frac{2-\alpha_\epsilon}{2}-1)}$, $M_{\mathbf{d},2}^{(n\frac{2-\alpha_\epsilon}{2}-1)}, \cdots, M_{\mathbf{d},1}^{(1)}, M_{\mathbf{d},2}^{(1)}$, until the remaining $2^{n(\alpha_\epsilon-1)}$ codewords, after each sequence of divisions, are indexed by the randomization message $\tilde{W}$.

Delivery Encoder: Given W_1, W_2, K_1, K_2, $\tilde{W}$, $\mathbf{d} = (d_1, d_2)$, the transmitter forms $M_{\mathbf{d},1}$, $M_{\mathbf{d},2}$, as in (52) and sends $\mathbf{X}_{\mathbf{d}}^n$ which corresponds to $M_{\mathbf{d},1}$, $M_{\mathbf{d},2}$, $\tilde{W}$, i.e., $\mathbf{X}_{\mathbf{d}}^n(M_{\mathbf{d},1}^{(n\frac{2-\alpha_\epsilon}{2})}, M_{\mathbf{d},2}^{(n\frac{2-\alpha_\epsilon}{2})}, \ldots, M_{\mathbf{d},1}^{(1)}, M_{\mathbf{d},2}^{(1)}, \tilde{W})$.

Decoding: Using $\mathbf{X}_c^n$, receiver $j = 1, 2$, recovers $M_{c,j} = K_j$ and stores it in its cache. Using $\mathbf{X}_{\mathbf{d}}^n$, both receivers recover $M_{\mathbf{d}} = \{M_{\mathbf{d},1}, M_{\mathbf{d},2}\}$. Using $M_{\mathbf{d},j}$, K_j, and for n large enough, receiver j recovers W_{d_j}.

Security Analysis: Fix S_1, S_2. Recall that $\alpha_1, \alpha_2 \leq 1$. Since $\alpha \geq 1$, $\alpha_1, \alpha_2 \geq \alpha - 1$. If $\alpha_1 = 1$, then $\alpha_2 = \alpha - 1$, and vice versa. In addition, notice that $1 - \alpha_1, 1 - \alpha_2 \leq 2 - \alpha$. As in Section IV-C, for a fixed value of α_1, the codebook $\mathcal{C}_{c,n}$ is a wiretap code with $2^{n(1-\alpha_{1,\epsilon})}$ bins, indexed by

$$W_c = \left(K_1^{(1)}, K_2^{(1)}, \ldots, K_1^{\left(n\frac{1-\alpha_{1,\epsilon}}{2}\right)}, K_2^{\left(n\frac{1-\alpha_{1,\epsilon}}{2}\right)}\right). \quad (53)$$

Each bin W_c contains $2^{n\alpha_{1,\epsilon}}$ binary codewords, indexed by

$$\tilde{W}_c = \left(K_1^{\left(n\frac{1-\alpha_{1,\epsilon}}{2}+1\right)}, K_2^{\left(n\frac{1-\alpha_{1,\epsilon}}{2}+1\right)}, \ldots, K_1^{\left(n\frac{2-\alpha_\epsilon}{2}\right)}, K_2^{\left(n\frac{2-\alpha_\epsilon}{2}\right)}, \tilde{W}_K\right). \quad (54)$$

Similarly, for a fixed value of α_2, $\mathcal{C}_{\mathbf{d},n}$ is a wiretap code with $2^{n(1-\alpha_{2,\epsilon})}$ bins, each is indexed by

$$W_{\mathbf{d}} = \left(\tilde{M}_{\mathbf{d},1}^{\left(n\frac{2-\alpha_\epsilon}{2}\right)}, \tilde{M}_{\mathbf{d},2}^{\left(n\frac{2-\alpha_\epsilon}{2}\right)}, \ldots, \tilde{M}_{\mathbf{d},1}^{\left(n\frac{1-\alpha_{1,\epsilon}}{2}+1\right)}, \tilde{M}_{\mathbf{d},2}^{\left(n\frac{1-\alpha_{1,\epsilon}}{2}+1\right)}\right). \quad (55)$$

Each bin $W_{\mathbf{d}}$ contains $2^{n\alpha_{2,\epsilon}}$ codewords, indexed by

$$\tilde{W}_{\mathbf{d}} = \left(\tilde{M}_{\mathbf{d},1}^{\left(n\frac{1-\alpha_{1,\epsilon}}{2}\right)}, \tilde{M}_{\mathbf{d},2}^{\left(n\frac{1-\alpha_{1,\epsilon}}{2}\right)}, \ldots, \tilde{M}_{\mathbf{d},1}^{(1)}, \tilde{M}_{\mathbf{d},2}^{(1)}, \tilde{W}\right). \quad (56)$$

Let us re-define $\mathbf{K}^{(1)}, \mathbf{K}^{(2)}, \mathbf{W}_{\oplus\mathbf{K}}^{(1)}$, and $\mathbf{W}_{\oplus\mathbf{K}}^{(2)}$ and define $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(2)}$ as

$$\mathbf{K}^{(1)} = \left\{K_1^{(i)}, K_2^{(i)} \colon i = 1, \ldots, n\frac{1-\alpha_{1,\epsilon}}{2}\right\} \quad (57)$$

$$\mathbf{K}^{(2)} = \left\{K_1^{(i)}, K_2^{(i)} \colon i = n\frac{1-\alpha_{1,\epsilon}}{2} + 1, \ldots, n\frac{2-\alpha_\epsilon}{2}\right\} \quad (58)$$

$$\mathbf{W}_{\oplus\mathbf{K}}^{(1)} = \left\{W_{d_1}^{(i)} \oplus K_1^{(i)}, W_{d_2}^{(i)} \oplus K_2^{(i)} \colon i = 1, \ldots, n\frac{1-\alpha_{1,\epsilon}}{2}\right\} \quad (59)$$

$$\mathbf{W}_{\oplus\mathbf{K}}^{(2)} = \left\{W_{d_1}^{(i)} \oplus K_1^{(i)}, W_{d_2}^{(i)} \oplus K_2^{(i)} \colon i = n\frac{1-\alpha_{1,\epsilon}}{2} + 1, \ldots, n\frac{2-\alpha_\epsilon}{2}\right\} \quad (60)$$

$$\mathbf{W}^{(1)} = \left\{W_{d_1}^{(i)}, W_{d_2}^{(i)} \colon i = 1, \ldots, n\frac{1-\alpha_{1,\epsilon}}{2}\right\} \quad (61)$$

$$\mathbf{W}^{(2)} = \left\{W_{d_1}^{(i)}, W_{d_2}^{(i)} \colon i = n\frac{1-\alpha_{1,\epsilon}}{2} + 1, \ldots, n\frac{2-\alpha_\epsilon}{2}\right\}. \quad (62)$$

From (53)-(60), we have

$$W_c = \mathbf{K}^{(1)}, \quad \tilde{W}_c = \left\{\mathbf{K}^{(2)}, \tilde{W}_K\right\},$$
$$W_{\mathbf{d}} = \mathbf{W}_{\oplus\mathbf{K}}^{(2)}, \quad \tilde{W}_{\mathbf{d}} = \left\{\mathbf{W}_{\oplus\mathbf{K}}^{(1)}, \tilde{W}\right\}. \quad (63)$$

Similar to Section IV-C, $\tilde{W}_c$, $\tilde{W}_{\mathbf{d}}$, are independent and uniform, and hence $\{\tilde{W}_c, \tilde{W}_{\mathbf{d}}\}$ is jointly uniform. Also, $\{\tilde{W}_c, \tilde{W}_{\mathbf{d}}\}$ is independent from $\{W_c, W_{\mathbf{d}}\}$. Thus, (34) is satisfied. For any $\mathbf{d} = (d_1, d_2)$, we have

$$I\left(W_1, W_2; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\right) = I\left(\mathbf{W}^{(1)}, \mathbf{W}^{(2)}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\right) \quad (64)$$
$$\leq I\left(\mathbf{W}^{(1)}, \mathbf{W}_{\oplus\mathbf{K}}^{(2)}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\right) \quad (65)$$
$$= I\left(\mathbf{W}^{(1)}, W_{\mathbf{d}}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\right) \quad (66)$$
$$= H\left(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\right) - H\left(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n | \mathbf{W}^{(1)}, W_{\mathbf{d}}\right) \quad (67)$$
$$\leq H\left(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\right) - H\left(\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n | \mathbf{K}^{(1)}, W_{\mathbf{d}}\right) + \epsilon_n' \quad (68)$$
$$= I\left(W_c, W_{\mathbf{d}}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\right) + \epsilon_n', \quad (69)$$

where (65) follows due to the Markov chain $\mathbf{W}^{(2)} - \{\mathbf{W}^{(1)}, \mathbf{W}_{\oplus\mathbf{K}}^{(2)}\} - \{\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\}$; (66), (69) follow from (63); (68) follows by using similar steps as in (42)-(48). Using (34)

and (69), the secrecy constraint in (3) is satisfied. Since $\epsilon_n \to 0$ as $n \to \infty$, the achievable strong secrecy file rate is

$$R_s(\alpha) = \frac{2-\alpha}{2} = 1 - \frac{\alpha}{2}. \tag{70}$$

This completes the proof for Theorem 1.

V. Proof of Theorem 2

In this section, we extend the achievability scheme in Section IV and provide a lower bound on the strong secrecy file rate when $D > 2$. The demand vector is $\mathbf{d} = (d_1, d_2)$, where $d_1, d_2 \in [1 : D]$. As in Section IV, we divide the proof into two cases for the ranges $\alpha \in (0, 1)$ and $\alpha \in [1, 2]$.

A. $\alpha \in [1, 2]$

For $\alpha \in [1, 2]$, we use the same scheme as in Section IV-D. For this range of α, only the keys K_1, K_2, are transmitted during placement, and stored in receivers 1 and 2 cache memories. That is, no information messages are stored in the caches, and the user demands are known during delivery. Hence, $D > 2$ is immaterial for this range of α. The achievable strong secrecy file rate is $1 - \frac{\alpha}{2}$.

B. $\alpha \in (0, 1)$

The achievability scheme for this case has the same channel coding structure as in Section IV-C. The difference however lies in generating the messages to be securely communicated over cache placement and delivery phases, i.e., M_c, $M_{\mathbf{d}}$. In particular, we use here uncoded placement for designing the cache contents, and a partially coded delivery transmission that is simultaneously useful for both receivers.

The transmitter (i) divides W_l, $l \in [1 : D]$, into the independent messages $\{W_l^{(1)}, W_l^{(2)}, W_{l,t}, W_{l,s}\}$, where $W_l^{(1)}$, $W_l^{(2)}$ are uniform over $[1 : 2^{n\frac{1-\alpha_\epsilon}{2D}}]$, α_ϵ is as in (19), $W_{l,t}$ is uniform over $[1 : 2^{n\frac{(2D-1)(1-\alpha_\epsilon)}{4D}}]$, and $W_{l,s}$ is uniform over $[1 : 2^{n\frac{\alpha_\epsilon}{2}}]$; (ii) generates the independent keys K_1, K_2, uniform over $[1 : 2^{n\frac{\alpha_\epsilon}{2}}]$ and independent from $W_{[1:D]}$.

Let $M_c = \{M_{c,1}, M_{c,2}\}$. Unlike (15), we use here *uncoded placement* for designing $M_{c,1}$, $M_{c,2}$:

$$M_{c,1} = \left\{W_1^{(1)}, W_2^{(1)}, \ldots, W_D^{(1)}\right\}, \tag{71}$$

$$M_{c,2} = \left\{W_1^{(2)}, W_2^{(2)}, \ldots, W_D^{(2)}\right\}. \tag{72}$$

The randomization message $\tilde{M}_c$ is identical to (16). Receiver j stores $M_{c,j}$, $\tilde{M}_{c,j}$ in its cache memory.

Unlike (17), we use *partially coded* delivery. The message to be securely transmitted in delivery is

$$M_{\mathbf{d}} = \left\{W_{d_2}^{(1)} \oplus W_{d_1}^{(2)}, W_{d_1,t}, W_{d_2,t}\right\}. \tag{73}$$

Notice that we use the term *partially coded* since part of $M_{\mathbf{d}}$ is coded, i.e., $W_{d_2}^{(1)} \oplus W_{d_1}^{(2)}$, while the other part is uncoded, i.e., $W_{d_1,t}$, $W_{d_2,t}$. The randomization message for delivery, $\tilde{M}_{\mathbf{d}}$, is identical to (18).

Remark 8: Note that the sizes of M_c, $M_{\mathbf{d}}$, $\tilde{M}_c$, and $\tilde{M}_{\mathbf{d}}$ are the same as in Section IV-C. In particular, the sizes of $\tilde{M}_c$ and $\tilde{M}_{\mathbf{d}}$ are both $n\alpha_\epsilon$ bits. The size of M_c is $2 \times D \times n\frac{1-\alpha_\epsilon}{2D} = n(1-\alpha_\epsilon)$ bits and the size of $M_{\mathbf{d}}$ is $n\frac{1-\alpha_\epsilon}{2D} + 2 \times n\frac{(2D-1)(1-\alpha_\epsilon)}{4D} = n(1-\alpha_\epsilon)$ bits.

Codebooks Generation and Encoders: For the messages M_c, $M_{\mathbf{d}}$, $\tilde{M}_c$, $\tilde{M}_{\mathbf{d}}$ defined above, the cache placement and delivery codebooks, $\mathcal{C}_{c,n}$ and $\mathcal{C}_{\mathbf{d},n}$, and the cache and delivery encoders, are designed in the same exact manner as in Section IV-C, see Figures 3 and 4.

Decoding: As in Section IV-C, using $M_{\mathbf{d}}$, $\tilde{M}_{\mathbf{d}}$, $M_{c,j}$, $\tilde{M}_{c,j}$, and for n sufficiently large, receiver j correctly decodes W_{d_j}, $j = 1, 2$.

Security analysis: Let W_c, $\tilde{W}_c$, $W_{\mathbf{d}}$, $\tilde{W}_{\mathbf{d}}$, be defined as in (20)-(23), (31), (32). Once again, $\tilde{W}_c$ and $\tilde{W}_{\mathbf{d}}$ are independent and uniform, and hence $\{\tilde{W}_c, \tilde{W}_{\mathbf{d}}\}$ is jointly uniform. In addition, $\{W_c, W_{\mathbf{d}}\}$ are independent from $\{\tilde{W}_c, \tilde{W}_{\mathbf{d}}\}$. Thus, (34) holds for this case. For any $\mathbf{d} = (d_1, d_2)$,

$$I\left(W_{[1:D]}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\right)$$

$$= I\left(\left\{W_l^{(1)}, W_l^{(2)}, W_{l,t}, W_{l,s}\right\}_{l=1}^D; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\right) \tag{74}$$

$$\le I\left(M_c, \left\{W_l^{(1)}, W_l^{(2)}, W_{l,t}, W_{l,s}\right\}_{l=1}^D; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\right) \tag{75}$$

$$\le I\left(M_c, W_{d_2}^{(1)} \oplus W_{d_1}^{(2)}, W_{d_1,t}, W_{d_2,t}, W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\right) \tag{76}$$

$$= I\left(M_c, M_{\mathbf{d}}, W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\right) \tag{77}$$

$$\le I\left(W_c, W_{\mathbf{d}}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\right) + \epsilon_n', \tag{78}$$

The inequality in (76) follows form the Markov chain $W_{[1:D]} - \{M_c, W_{d_2}^{(1)} \oplus W_{d_1}^{(2)}, W_{d_1,t}, W_{d_2,t}, W_{d_1,s}, W_{d_2,s}\} - \{\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\}$; (77) follows from (73); (78) follows using similar steps as in (37)-(48). Using (34), (78), the secrecy constraint in (3) is satisfied. With $\epsilon_n \to 0$ as $n \to \infty$, the achievable strong secrecy file rate is

$$R_s(\alpha) = \frac{(1-\alpha)}{D} + \frac{(2D-1)(1-\alpha)}{4D} + \frac{\alpha}{2} = \frac{1}{2} + \frac{3(1-\alpha)}{4D}. \tag{79}$$

This completes the proof for Theorem 2.

Remark 9: For $D = 2$, the achievable secrecy rate in (79) is strictly smaller than the secrecy rate obtained by *coded* placement and *uncoded* delivery in Section IV-C, i.e., $1 - \frac{\alpha}{2}$.

Remark 10: An achievability scheme which utilizes coded placement and uncoded delivery, as in Section IV-C, achieves the same secure file rate as (79) for $D = 3$. However, this scheme achieves a strictly smaller secure file rate for $D \ge 4$. In this scheme, $W_l^{(1)}$, $W_l^{(2)}$ are uniform over $[1 : 2^{n\frac{1-\alpha_\epsilon}{2(D-1)}}]$. $W_{l,t}$ is uniform over $[1 : 2^{n\frac{(D-2)(1-\alpha_\epsilon)}{2(D-1)}}]$. $W_{l,s}$, K_1, K_2, are uniform over $[1 : 2^{n\frac{\alpha_\epsilon}{2}}]$. $M_c = \{M_{c,1}, M_{c,2}\}$, where $M_{c,1} = \{W_1^{(1)} \oplus W_2^{(1)}, W_2^{(1)} \oplus W_3^{(1)}, \ldots, W_{D-1}^{(1)} \oplus W_D^{(1)}\}$ and $M_{c,2} = \{W_1^{(2)} \oplus W_2^{(2)}, W_2^{(2)} \oplus W_3^{(2)}, \ldots, W_{D-1}^{(2)} \oplus W_D^{(2)}\}$. $M_{\mathbf{d}} = \{W_{d_2}^{(1)}, W_{d_1}^{(2)}, W_{d_1,t}, W_{d_2,t}\}$. Without loss of generality, let $d_1 < d_2$. For any $\mathbf{d} = (d_1, d_2)$, using $M_{c,j}$, receiver j, can restore $W_{d_1}^{(j)} \oplus W_{d_2}^{(j)}$ by xor-ing $\{W_{d_1}^{(j)} \oplus W_{d_1+1}^{(j)}\}$, $\{W_{d_1+1}^{(j)} \oplus W_{d_1+2}^{(j)}\}, \cdots, \{W_{d_2-1}^{(j)} \oplus W_{d_2}^{(j)}\}$. The achievable strong secrecy file rate using this scheme is $R_s(\alpha) = \frac{1}{2} + \frac{1-\alpha}{2(D-1)}$.

VI. Proof of Theorem 3

When $\alpha \in [1, 2]$, the upper bound on R_s in Theorem 3 for $D > 2$ follows as in Section IV-A. It remains to prove the upper bound for $\alpha \in (0, 1)$. The proof is divided into the three following steps.

Step 1: We upper bound R_s by the secrecy capacity when the adversary is restricted to tap into the delivery phase only, denoted as C_s^{Res}. That is, C_s^{Res} is the maximum achievable file rate when $\alpha_1 = 0$, $\alpha_2 = \alpha$. Restricting the adversary to only tap into the delivery phase cannot decrease the secrecy capacity, i.e., $R_s \leq C_s^{\text{Res}}$, since this setting is included in its feasible strategy space. Cache placement is secure, and each receiver has a secure cache memory of size $\frac{n}{2}$ bits.

Step 2: We upper bound C_s^{Res} by the secrecy capacity when the delivery channel to the adversary is replaced by a discrete memoryless binary erasure channel, with erasure probability $1-\alpha$, denoted as C_s^{DM}. The proof for this step follows the same lines as in [44, Sec. V]. The idea is when the binary erasure channel produces a number of erasures greater than or equal to $(1-\alpha)n$, the adversary's channel in this discrete memoryless setup is worse than its channel in the former model, i.e., when it encounters exactly $(1-\alpha)n$ erasures and is able to select their positions. Hence, $C_s^{\text{Res}} \leq C_s^{\text{DM}}$ for this case. The result follows by using Sanov's theorem in method of types [51, Th. 11.4.1] to show that the probability of the binary erasure channel causing a number of erasures less than $(1-\alpha)n$ goes to zero as $n \to \infty$.

Step 3: From Step 1, each receiver has a secure cache of size $\frac{n}{2}$ bits. Since increasing the cache sizes cannot decrease the achievable file rate, we upper bound C_s^{DM} with the maximum achievable file rate when each receiver has a cache of size n bits, in which it stores $\mathbf{X}_c^n$. Receiver $j = 1, 2$, uses $\mathbf{X}_c^n$, $\mathbf{X}_{\mathbf{d}}^n$ to decode W_{d_j}, i.e., $\hat{W}_{d_j} = g_{\mathbf{d},j}(\mathbf{X}_c^n, \mathbf{X}_{\mathbf{d}}^n)$. This setup is equivalent to a single receiver, which has a cache of size n bits, demands two files W_{d_1}, W_{d_2}, and uses the decoder $g_{\mathbf{d}} \triangleq \{g_{\mathbf{d},1}, g_{\mathbf{d},2}\}$. Let C_s^{SR} be the maximum achievable file rate for this single receiver model. We have $C_s^{\text{DM}} \leq C_s^{\text{SR}}$. Next, we upper bound C_s^{SR}.

Let M_{D} denote the fraction of the size-n bits cache memory dedicated to store (coded or uncoded) information bits, and let M_{K} denote the fraction dedicated to store key bits. That is, $M_{\text{D}} + M_{\text{K}} = 1$. Let S_{D} denote the information bits stored in this memory, i.e., $S_{\text{D}} = f_{\text{D}}(W_{[1:D]})$ and $H(S_{\text{D}}) = nM_{\text{D}}$. We use the following lemma to upper bound C_s^{SR}.

Lemma 1 [37, Lemma 1]: For a fixed allocation of $M_{\text{D}}, M_{\text{K}}$, and a receiver who demands the files W_{d_1}, W_{d_2}, the secrecy rate for the single receiver model is upper bounded as

$$2R_s^{\text{SR}} \leq \min\{1, 1-\alpha+M_{\text{K}}\} + \frac{1}{n} I\big(W_{d_1}, W_{d_2}; S_{\text{D}}\big). \quad (80)$$

Notice that (80) holds for any demand pair $\mathbf{d} = (d_1, d_2)$ such that $d_1 \neq d_2$, i.e., the worst-case demands. Summing over all such demands, we have

$$2R_s^{\text{SR}} \leq \min\{1, 1-\alpha+M_{\text{K}}\} + \frac{1}{nD(D-1)} \sum_{d_1,d_2 \in [1\,:\,D],\ d_1 \neq d_2} I\big(W_{d_1}, W_{d_2}; S_{\text{D}}\big). \quad (81)$$

The second term on the right hand side of (81) can be written as

$$\begin{aligned}
&\frac{1}{nD(D-1)} \sum_{d_1,d_2 \in [1\,:\,D],\ d_1 \neq d_2} I\big(W_{d_1}, W_{d_2}; S_{\text{D}}\big) \\
&= \frac{1}{nD} \sum_{d_1 \in [1\,:\,D]} I\big(W_{d_1}; S_{\text{D}}\big) \\
&\quad + \frac{1}{nD(D-1)} \sum_{d_1,d_2 \in [1\,:\,D],\ d_1 \neq d_2} I\big(W_{d_2}; S_{\text{D}} | W_{d_1}\big) \quad (82) \\
&\leq \frac{1}{nD} \sum_{d_1 \in [1\,:\,D]} I(W_{d_1}; S_{\text{D}}) \\
&\quad + \frac{1}{nD(D-1)} \sum_{d_1 \in [1\,:\,D]} \left(\sum_{d_2 \in [1\,:\,D]} I\big(W_{d_2}; S_{\text{D}} | W_{d_1}\big) \right). \quad (83)
\end{aligned}$$

For any $d_1 \in [1 : D]$, we have

$$\begin{aligned}
&\sum_{d_2 \in [1\,:\,D]} I\big(W_{d_2}; S_{\text{D}} \big| W_{d_1}\big) \\
&= \sum_{d_2=1}^{D} \big[H\big(W_{d_2} \big| W_{d_1}\big) - H\big(W_{d_2} \big| W_{d_1}, S_{\text{D}}\big)\big] \quad (84) \\
&\leq \sum_{d_2=1}^{D} \big[H\big(W_{d_2} \big| W_1, W_2, \ldots, W_{d_2-1}, W_{d_1}\big) \\
&\qquad - H\big(W_{d_2} \big| W_1, W_2, \ldots, W_{d_2-1}, W_{d_1}, S_{\text{D}}\big)\big] \quad (85) \\
&= I\big(W_1, W_2, \ldots, W_{\text{D}}; S_{\text{D}} \big| W_{d_1}\big) \quad (86) \\
&\leq H(S_{\text{D}}) = nM_{\text{D}}, \quad (87)
\end{aligned}$$

where (85) follows because when $d_2 = d_1$, $H(W_{d_2}|W_{d_1}) = H(W_{d_2}|W_1, W_2, \ldots, W_{d_2-1}, W_{d_1}) = 0$, and when $d_2 \neq d_1$, $H(W_{d_2}|W_{d_1}) = H(W_{d_2}|W_1, W_2, \ldots, W_{d_2-1}, W_{d_1}) = H(W_{d_2})$. Similarly, we have

$$\sum_{d_1 \in [1\,:\,D]} I\big(W_{d_1}; S_{\text{D}}\big) \leq H(S_{\text{D}}) = nM_{\text{D}}. \quad (88)$$

Substituting (87) and (88) in (83) gives

$$\frac{1}{nD(D-1)} \sum_{\substack{d_1,d_2 \in [1\,:\,D] \\ d_1 \neq d_2}} I\big(W_{d_1}, W_{d_2}; S_{\text{D}}\big) \leq \frac{2D-1}{D(D-1)} M_{\text{D}}. \quad (89)$$

Thus, using (81) and (89), R_s^{SR} is further upper bounded as

$$R_s^{\text{SR}} \leq \frac{1}{2}\left[\min\{1, 1-\alpha+M_{\text{K}}\} + \frac{2D-1}{D(D-1)} M_{\text{D}}\right]. \quad (90)$$

Finally, by maximizing over all possible allocations for M_{D}, M_{K} such that $M_{\text{D}} + M_{\text{K}} = 1$, we get

$$\begin{aligned}
C_s^{\text{SR}} &\leq \frac{1}{2} \max_{\substack{M_{\text{D}}, M_{\text{K}}: \\ M_{\text{D}}+M_{\text{K}}=1}} \left\{\min\{1, 1-\alpha+M_{\text{K}}\} + \frac{2D-1}{D(D-1)} M_{\text{D}}\right\} \quad (91) \\
&= \frac{1}{2}\left[1 + \frac{2D-1}{D(D-1)}(1-\alpha)\right]. \quad (92)
\end{aligned}$$

Equation (92) follows because, for $D \geq 3$, the maximum occurs at $M_{\mathrm{K}} = \alpha$ and $M_{\mathrm{D}} = 1 - \alpha$. This completes the proof for Theorem 3.

Remark 11: An upper bound considering uncoded placement only can be derived as follows. The same analysis as in (80)-(92) carries through with $I(W_{d_2}; S_{\mathrm{D}}|W_{d_1})$ in (82) is equal to $I(W_{d_2}; S_{\mathrm{D}})$. Hence the right hand side of (89) is replaced by $\frac{2M_{\mathrm{D}}}{D}$. The resulting bound $R_s \leq \frac{1}{2} + \frac{(1-\alpha)}{D}$ is tighter than (92).

VII. Discussion

A. Variable Cache Memories

While the fixed-size cache memory setup considered in this work can be seen as a clean basic model for the intricate problem in consideration, it also allows us to obtain results and insights that are generalizable to more involved cache memory models. The extension to variable memory sizes can be done by considering multiple communication blocks for placement. Our results and coding scheme readily apply to an adversary model whose tapping capability during delivery is normalized with respect to tapping during placement; $\mu_1 + B\mu_2 \leq \mu$, B is the number of communication blocks for placement. This is a reasonable assumption given that cache placement generally takes place in a longer period than delivery. The problem turns to be more challenging when the adversary optimizes its tapping uniformly over the multiple blocks for cache placement as well as the delivery phase. This is left for future investigation.

B. Extension to More Than Two Users

The broadcast wiretap channel II in [47] can be generalized to more than two users. In [47], the achievability scheme does not depend on the number of receivers, and the converse proof can be extended to any broadcast setting of known secrecy capacity, see for example [55], [56]. It follows that the results in this work can also be generalized to more than two users. In particular, the key enabler for our achievability scheme is the proposed channel coding structure which involves using security embedding codes along with reversed embedding order across the two communication phases. By carefully choosing the messages and encryption keys to be transmitted over cache placement and delivery, the same channel coding structure can be applied to the case of more than two users.

C. The Broadcast Channel Model for Cache Placement

It is typical to model cache placement as a noiseless channel since placement is assumed to occur when networks are not congested and their rates are assumed to be large enough. Here however we model the cache placement as a broadcast channel communication. The broadcast model avails a clean and tractable solution without compromising its generalizability. A time division multiple access (TDMA) model for cache placement is a special case by imposing an additional constraint in which each receiver has to decode its desired file using only one half of the transmitted codeword. Additionally, the broadcast model is in line with the network information theory literature and it does not limit cache placement to occur over low rate traffic. With the ever-growing user demands, placement and delivery occurring in less asymmetric network loads is likely to be expected in the near future.

D. Larger Cache Sizes Lead to Simple Achievability

For a library with two files, if the receivers were to have cache memories of size n bits in which they store the transmitted signal during placement, the strong secrecy file rate in Theorem 1 is achievable using a simple wiretap code: The transmitter encodes $W = (W_1, W_2) \in [1 : 2^{n2R_s}]$ into a length-$2n$ binary codeword using a wiretap code, and sends the first n bits of this codeword during cache placement and the last n bits during delivery. Each receiver can thus decode both files, and the secrecy of W_1 and W_2 against the adversary follows by the results in [43], [44]. In caching problems, the relevant setup however is when the receivers have cache memories of limited size with respect to the overall transmission during cache placement. This calls for the limited size cache memory model considered in this paper, which in turn necessitates the use of the more elaborate coding scheme in Section IV.

VIII. Conclusion

We have introduced the caching broadcast channel with *a wire and cache* tapping adversary of type II. Each receiver is equipped with a fixed-size cache memory, and the adversary is able to tap into a subset of its choice of the transmitted symbols during cache placement, delivery, or both. The legitimate terminals have no knowledge about the fractions of the tapped symbols in each phase, nor their positions. Only the size of the overall tapped set is known. We have identified the strong secrecy capacity of this model– the maximum achievable file rate while keeping the overall library secure– when the transmitter's library has two files. We have derived lower and upper bounds for the strong secrecy file rate when the transmitter has more than two files in its library. We have devised an achievability scheme which combines wiretap coding, security embedding codes, one-time pad keys, and coded caching techniques.

The results presented in this paper highlight the robustness of (stochastic) coding in a cache-aided network, against a smart adversary who is able to perform a strategic attack jointly optimized over cache placement and delivery phases. Future directions that can build on this work include exploring variable cache memory sizes, models with more than two end-users, other network topologies, and models with noisy legitimate channels.

Appendix A
Secrecy Constraint for Setting 1

For every $S_1 \subseteq [1 : n]$ satisfying $|S_1| = \mu$, we have

$$\begin{aligned}
&\lim_{n\to\infty} I\Big(W_1, W_2; \mathbf{Z}^n_{S_1}\Big) \\
&\quad = \lim_{n\to\infty} I\Big(W_1^{(1)}, W_1^{(2)}, W_2^{(1)}, W_2^{(2)}, W_{1,s}, W_{2,s}; \mathbf{Z}^n_{S_1}\Big) \quad (93)\\
&\quad = \lim_{n\to\infty} I\Big(W_1^{(1)}, W_1^{(2)}, W_2^{(1)}, W_2^{(2)}; \mathbf{Z}^n_{S_1}\Big) \quad (94)
\end{aligned}$$

$$\leq \lim_{n\to\infty} I\Big(W_1^{(1)} \oplus W_2^{(1)}, W_1^{(2)} \oplus W_2^{(2)}; \mathbf{Z}_{S_1}^n\Big) \quad (95)$$

$$= \lim_{n\to\infty} I\Big(M_c; \mathbf{Z}_{S_1}^n\Big) = 0. \quad (96)$$

The adversary's observation over cache placement, $\mathbf{Z}_{S_1}^n$, results from sending $M_c = \{M_{c,1}, M_{c,2}\}$; $M_{c,1} = W_1^{(1)} \oplus W_2^{(1)}$. $M_{c,2} = W_1^{(2)} \oplus W_2^{(2)}$. Thus, (94) follows since $\mathbf{Z}_{S_1}^n$ does not depend on $\{W_{1,s}, W_{2,s}\}$ and (95) follows due to the Markov chain $\{W_1^{(1)}, W_1^{(2)}, W_2^{(1)}, W_2^{(2)}\} - \{W_1^{(1)} \oplus W_2^{(1)}, W_1^{(2)} \oplus W_2^{(2)}\} - \mathbf{Z}_{S_1}^n$. The second equality in (96) follows from [43, Th. 2], and since the rate of M_c is less than $1-\alpha$.

Appendix B
Secrecy Constraint for Setting 2

For every $S_2 \subseteq [1:n]$ satisfying $|S_2| = \mu$ and any $\mathbf{d} = (d_1, d_2)$, $d_1, d_2 \in \{1,2\}$,

$$I\Big(W_1, W_2; \mathbf{Z}_{S_2}^n\Big) = I\Big(W_{d_1}^{(2)}, W_{d_2}^{(1)}, W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_{S_2}^n\Big) \quad (97)$$

$$= I\Big(W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_{S_2}^n \big| W_{d_1}^{(2)}, W_{d_2}^{(1)}\Big) + I\Big(W_{d_1}^{(2)}, W_{d_2}^{(1)}; \mathbf{Z}_{S_2}^n\Big) \quad (98)$$

$$\leq I\Big(W_{d_1,s}, W_{d_2,s}; W_{d_1,s} \oplus K_1, W_{d_2,s} \oplus K_2 \big| W_{d_1}^{(2)}, W_{d_2}^{(1)}\Big) + I\Big(W_{d_1}^{(2)}, W_{d_2}^{(1)}; \mathbf{Z}_{S_2}^n\Big) \quad (99)$$

$$= I\big(W_{d_1,s}, W_{d_2,s}; W_{d_1,s} \oplus K_1, W_{d_2,s} \oplus K_2\big) + I\Big(W_{d_1}^{(2)}, W_{d_2}^{(1)}; \mathbf{Z}_{S_2}^n\Big) \quad (100)$$

$$= I\Big(W_{d_1}^{(2)}, W_{d_2}^{(1)}; \mathbf{Z}_{S_2}^n\Big) \quad (101)$$

$$= I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n\Big). \quad (102)$$

The adversary's observation over delivery, $\mathbf{Z}_{S_2}^n$, results from sending $M_{\mathbf{d}} = \{W_{d_1}^{(2)}, W_{d_2}^{(1)}\}$ and $\tilde{M}_{\mathbf{d}} = \{W_{d_1,s} \oplus K_1, W_{d_2,s} \oplus K_2\}$. Equation (97) follows because $\mathbf{Z}_{S_2}^n$ depends only on $W_{d_1}^{(2)}$, $W_{d_2}^{(1)}$, $W_{d_1,s}$, $W_{d_2,s}$. Equation (99) follows from the Markov chain $\{W_{d_1,s}, W_{d_2,s}\} - \{W_{d_1}^{(2)}, W_{d_2}^{(1)}, W_{d_1,s} \oplus K_1, W_{d_2,s} \oplus K_2\} - \mathbf{Z}_{S_2}^n$. Equation (100) follows because $\{W_{d_1}^{(2)}, W_{d_2}^{(1)}\}$, $\{W_{d_1,s}, W_{d_2,s}, K_1, K_2\}$ are independent. The randomization message for the wiretap code during delivery $\tilde{M}_{\mathbf{d}}$ is independent from the message $M_{\mathbf{d}}$. Using (102), [43, Th. 2],

$$\lim_{n\to\infty} \max_{S_2 \subseteq [1:n]\,:\,|S_2|=\mu} I\Big(W_1, W_2; \mathbf{Z}_{S_2}^n\Big) \leq \lim_{n\to\infty} \max_{S_2 \subseteq [1:n]\,:\,|S_2|=\mu} I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n\Big) = 0. \quad (103)$$

Appendix C
Secrecy Constraint for Setting 3 When $\alpha_1 \geq \alpha_2$

For a fixed choice of $S_1, S_2 \subseteq [1:n]$ such that $|S_1| + |S_2| = \mu$, and any $\mathbf{d} = (d_1, d_2)$, $d_1, d_2 \in \{1,2\}$,

$$I\Big(W_1, W_2; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) = I\Big(W_1^{(1)}, W_1^{(2)}, W_2^{(1)}, W_2^{(2)}, W_{1,s}, W_{2,s}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) \quad (104)$$

$$= I\Big(W_1^{(1)} \oplus W_2^{(1)}, W_1^{(2)} \oplus W_2^{(2)}, W_{d_1}^{(2)}, W_{d_2}^{(1)}, W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) \quad (105)$$

$$= I\Big(M_c, M_{\mathbf{d}}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) \quad (106)$$

$$= I\Big(M_c; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) + I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n \big| M_c\Big) \quad (107)$$

$$= I\Big(M_c; \mathbf{Z}_{S_1}^n\Big) + I\Big(M_c; \mathbf{Z}_{S_2}^n \big| \mathbf{Z}_{S_1}^n\Big) + I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n \big| M_c\Big) + I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_1}^n \big| M_c, \mathbf{Z}_{S_2}^n\Big). \quad (108)$$

Equation (105) follows because, for any $d_1, d_2 \in \{1,2\}$, there is a bijective map between $\{W_1^{(1)}, W_1^{(2)}, W_2^{(1)}, W_2^{(2)}\}$ and $\{W_1^{(1)} \oplus W_2^{(1)}, W_1^{(2)} \oplus W_2^{(2)}, W_{d_1}^{(2)}, W_{d_2}^{(1)}\}$.

From (9), (10); M_c and $M_{\mathbf{d}}$ are independent. $\mathbf{Z}_{S_1}^n$ results from sending M_c, while $\mathbf{Z}_{S_2}^n$ results from sending $M_{\mathbf{d}}$. For a fixed choice of S_1, S_2, $\{M_c, \mathbf{Z}_{S_1}^n\}$ are independent from $\mathbf{Z}_{S_2}^n$. Thus, we have

$$I\Big(M_c; \mathbf{Z}_{S_2}^n | \mathbf{Z}_{S_1}^n\Big) = 0. \quad (109)$$

In addition, $\{M_{\mathbf{d}}, \mathbf{Z}_{S_2}^n\}$ are independent from M_c. Thus,

$$I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n | M_c\Big) = H\Big(\mathbf{Z}_{S_2}^n | M_c\Big) - H\Big(\mathbf{Z}_{S_2}^n | M_c, M_{\mathbf{d}}\Big) \quad (110)$$

$$= H\Big(\mathbf{Z}_{S_2}^n | M_c\Big) - H\Big(\mathbf{Z}_{S_2}^n | M_{\mathbf{d}}\Big) \quad (111)$$

$$\leq I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n\Big). \quad (112)$$

Finally, using the Markov chain $\{M_{\mathbf{d}}, \mathbf{Z}_{S_2}^n\} - M_c - \mathbf{Z}_{S_1}^n$,

$$I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_1}^n | M_c, \mathbf{Z}_{S_2}^n\Big) = H\Big(\mathbf{Z}_{S_1}^n | M_c, \mathbf{Z}_{S_2}^n\Big) - H\Big(\mathbf{Z}_{S_1}^n | M_c, \mathbf{Z}_{S_2}^n, M_{\mathbf{d}}\Big) \quad (113)$$

$$\leq H\Big(\mathbf{Z}_{S_1}^n\Big) - H\Big(\mathbf{Z}_{S_1}^n | M_c\Big) = I\Big(M_c; \mathbf{Z}_{S_1}^n\Big). \quad (114)$$

Substituting (109), (112), and (114) in (108),

$$I\Big(W_1, W_2; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) \leq 2I\Big(M_c; \mathbf{Z}_{S_1}^n\Big) + I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n\Big). \quad (115)$$

The rates of M_c, $M_{\mathbf{d}}$ are $1-\alpha_1-\epsilon_n$, $1-\alpha_2-\epsilon_n$, respectively. By applying [43, Th. 2] to (115),

$$\lim_{n\to\infty} \max_{\substack{S_1,S_2 \subseteq [1:n]: \\ |S_1|+|S_2|=\mu}} I\Big(W_1, W_2; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) \leq 2 \lim_{n\to\infty} \max_{S_1 \subseteq [1:n]:\,|S_1|=\mu_1} I\Big(M_c; \mathbf{Z}_{S_1}^n\Big) + \lim_{n\to\infty} \max_{S_2 \subseteq [1:n]:\,|S_2|=\mu_2} I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n\Big) = 0. \quad (116)$$

Appendix D
Secrecy Constraint for Setting 3 When $\alpha_1 < \alpha_2$

For notational simplicity, let us define

$$M_{c,1\backslash K_1} = W_1^{(1)} \oplus W_2^{(1)},\ M_{c,2\backslash K_2} = W_1^{(2)} \oplus W_2^{(2)} \quad (117)$$

$$M_{c\backslash K} = \big\{M_{c,1\backslash K_1}, M_{c,2\backslash K_2}\big\}. \quad (118)$$

For fixed $S_1, S_2 \subseteq [1:n]$ such that $|S_1| + |S_2| = \mu$, and any $\mathbf{d} = (d_1, d_2)$, $d_1, d_2 \in \{1,2\}$,

$$I\Big(W_1, W_2; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) = I\Big(W_1^{(1)} \oplus W_2^{(1)}, W_1^{(2)} \oplus W_2^{(2)}, W_{d_1}^{(2)}, W_{d_2}^{(1)}, W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) \quad (119)$$

$$= I\Big(M_{c\backslash K}, M_{\mathbf{d}}, W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) \quad (120)$$

$$= I\Big(M_{c\backslash K}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) + I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n | M_{c\backslash K}\Big) + I\Big(W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n | M_{\mathbf{d}}, M_{c\backslash K}\Big). \quad (121)$$

From (12), (13), M_c is independent from $\{M_{\mathbf{d}}, \tilde{M}_{\mathbf{d}}\}$. $\mathbf{Z}_{S_1}^n$ results from sending $M_c = \{M_{c\backslash K}, K_1, K_2\}$, and $\mathbf{Z}_{S_2}^n$ results from sending $M_{\mathbf{d}} = \{W_{d_1}^{(2)}, W_{d_2}^{(1)}\}$, $\tilde{M}_{\mathbf{d}} = \{W_{d_1,s} \oplus K_1, W_{d_2,s} \oplus W_2\}$.

We upper bound each term on the right hand side of (121). For the third term, we have

$$I\Big(W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n \big| M_{\mathbf{d}}, M_{c\backslash K}\Big) \le I\big(W_{d_1,s}, W_{d_2,s}; \tilde{M}_{\mathbf{d}} \big| M_{\mathbf{d}}, M_{c\backslash K}\big) \quad (122)$$

$$= I\big(W_{d_1,s}, W_{d_2,s}; W_{d_1,s} \oplus K_1, W_{d_2,s} \oplus K_2\big) = 0, \quad (123)$$

where (122) follows due to the Markov chain $\{W_{d_1,s}, W_{d_2,s}\} - \{M_{c\backslash K}, M_{\mathbf{d}}, \tilde{M}_{\mathbf{d}}\} - \{\mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\}$, and (123) follows because $\tilde{M}_{\mathbf{d}}$ is independent from $\{W_{d_1,s}, W_{d_2,s}, M_{\mathbf{d}}, M_{c\backslash K}\}$.

For fixed S_1, S_2, $\mathbf{Z}_{S_2}^n$ is independent from $\{M_c, \mathbf{Z}_{S_1}^n\}$. Thus, the first term is bounded as

$$I\Big(M_{c\backslash K}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) \le I\Big(M_c; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) \quad (124)$$

$$= I\Big(M_c; \mathbf{Z}_{S_1}^n\Big) + I\Big(M_c; \mathbf{Z}_{S_2}^n \big| \mathbf{Z}_{S_1}^n\Big) = I\Big(M_c; \mathbf{Z}_{S_1}^n\Big). \quad (125)$$

For the second term on the right hand side of (121), we have

$$I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n \big| M_{c\backslash K}\Big) = I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n \big| M_{c\backslash K}\Big) + I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_1}^n \big| M_{c\backslash K}, \mathbf{Z}_{S_2}^n\Big). \quad (126)$$

Notice that $M_{c\backslash K}$ and $\mathbf{Z}_{S_2}^n$ are conditionally independent given $M_{\mathbf{d}}$. Thus,

$$I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n \big| M_{c\backslash K}\Big) = H\Big(\mathbf{Z}_{S_2}^n \big| M_{c\backslash K}\Big) - H\Big(\mathbf{Z}_{S_2}^n \big| M_{\mathbf{d}}\Big) \le I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n\Big). \quad (127)$$

In addition, using the independence between $\{M_{\mathbf{d}}, \mathbf{Z}_{S_2}^n\}$ and $\{M_c, \mathbf{Z}_{S_1}^n\}$, we have

$$I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_1}^n \big| M_{c\backslash K}, \mathbf{Z}_{S_2}^n\Big) = H\Big(\mathbf{Z}_{S_1}^n \big| M_{c\backslash K}, \mathbf{Z}_{S_2}^n\Big) - H\Big(\mathbf{Z}_{S_1}^n \big| M_{c\backslash K}, M_{\mathbf{d}}, \mathbf{Z}_{S_2}^n\Big) \quad (128)$$

$$\le H\Big(\mathbf{Z}_{S_1}^n\Big) - H\Big(\mathbf{Z}_{S_1}^n \big| M_{c\backslash K}, K_1, K_2, M_{\mathbf{d}}, \mathbf{Z}_{S_2}^n\Big) \quad (129)$$

$$= H\Big(\mathbf{Z}_{S_1}^n\Big) - H\Big(\mathbf{Z}_{S_1}^n \big| M_c\Big) = I\Big(M_c; \mathbf{Z}_{S_1}^n\Big). \quad (130)$$

Substituting (127) and (130) in (126) gives

$$I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n \big| M_{c\backslash K}\Big) \le I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n\Big) + I\Big(M_c; \mathbf{Z}_{S_1}^n\Big). \quad (131)$$

Finally, substituting (123), (125), (131) in (121), and applying [43, Th. 2], we have

$$\lim_{n\to\infty} \max_{\substack{S_1,S_2\subseteq[1\,:\,n]:\\ |S_1|+|S_2|=\mu}} I\Big(W_1, W_2; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) = 0, \quad (132)$$

since the rates of M_c and $M_{\mathbf{d}}$ are $1-\alpha_1-\epsilon_n$ and $1-\alpha_2-\epsilon_n$, respectively.

APPENDIX E
SECRECY CONSTRAINT FOR SETTING 4

From(15)–(18), M_c is independent from $\tilde{M}_c$; $M_{\mathbf{d}}$ is independent from $\tilde{M}_{\mathbf{d}}$, and $\{M_c, \tilde{M}_c\}$ are independent from $\{M_{\mathbf{d}}, \tilde{M}_{\mathbf{d}}\}$. Conditioned on a fixed choice of S_1, S_2, satisfying $\{|S_1| = \mu,\ |S_2| = 0\}$ or $\{|S_1| = 0,\ |S_2| = \mu\}$, define the random variable

$$\mathbf{Z}_S^n \triangleq \mathbf{Z}_{S_1}^n\ \mathbb{1}_{\{|S_2|=0\}} + \mathbf{Z}_{S_2}^n\ \mathbb{1}_{\{|S_1|=0\}}. \quad (133)$$

Note that $\mathbf{Z}_S^n$ only has a well-defined probability distribution when conditioned on the event $\{S_1, S_2\}$, since a prior distribution on these subsets is not defined. For this fixed choice of the subsets, and any $\mathbf{d} = (d_1, d_2)$, $d_1, d_2 \in \{1, 2\}$, we have

$$I\Big(W_1, W_2; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) = I\Big(W_1^{(1)} \oplus W_2^{(1)}, W_1^{(2)} \oplus W_2^{(2)}, W_{d_1}^{(2)}, W_{d_2}^{(1)}, W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) \quad (134)$$

$$= I\big(M_c, M_{\mathbf{d}}, W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_S^n\big) \quad (135)$$

$$= \mathbb{1}_{\{|S_2|=0\}}\, I\Big(M_c, M_{\mathbf{d}}, W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_S^n \Big| \{|S_2| = 0\}\Big) + \mathbb{1}_{\{|S_1|=0\}}\, I\Big(M_c, M_{\mathbf{d}}, W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_S^n \Big| \{|S_1| = 0\}\Big) \quad (136)$$

$$= \mathbb{1}_{\{|S_2|=0\}}\, I(M_c, M_{\mathbf{d}}, W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_{S_1}^n) + \mathbb{1}_{\{|S_1|=0\}}\, I\Big(M_c, M_{\mathbf{d}}, W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_{S_2}^n\Big) \quad (137)$$

$$= \mathbb{1}_{\{|S_2|=0\}}\, I\Big(M_c; \mathbf{Z}_{S_1}^n\Big) + \mathbb{1}_{\{|S_1|=0\}}\, I\Big(M_{\mathbf{d}}, W_{d_1,s}, W_{d_2,s}; \mathbf{Z}_{S_2}^n\Big) \quad (138)$$

$$\le \mathbb{1}_{\{|S_2|=0\}}\, I\Big(M_c; \mathbf{Z}_{S_1}^n\Big) + \mathbb{1}_{\{|S_1|=0\}}\, I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n\Big). \quad (139)$$

Equation (138) follows because (i) $\mathbf{Z}_{S_1}^n$ results from $\{M_c, \tilde{M}_c\}$ which are independent from $\{M_{\mathbf{d}}, W_{d_1,s}, W_{d_2,s}\}$, and (ii) $\mathbf{Z}_{S_2}^n$ is conditionally independent from M_c given $\{M_{\mathbf{d}}, W_{d_1,s}, W_{d_2,s}\}$, due to the Markov chain $M_c - \{M_{\mathbf{d}}, W_{d_1,s}, W_{d_2,s}\} - \{M_{\mathbf{d}}, \tilde{M}_{\mathbf{d}}\} - \mathbf{Z}_{S_2}^n$. Equation (139) follows using the same steps in (97)–(102).

Finally, since $\tilde{M}_c$ is independent from M_c; $\tilde{M}_{\mathbf{d}}$ is independent from $M_{\mathbf{d}}$, and the rates of M_c, $\tilde{M}_{\mathbf{d}}$ are both equal to $1-\alpha-\epsilon_n$, we have

$$\lim_{n\to\infty} \max_{\substack{S_1,S_2\subseteq[1\,:\,n]:\\ |S_1|+|S_2|=\mu}} I\Big(W_1, W_2; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) = \lim_{n\to\infty} \max_{\substack{S_1,S_2\subseteq[1\,:\,n]:\\ |S_i|=0,\ |S_j|=\mu\\ i,j\in\{1,2\},\ i\neq j}} I\Big(W_1, W_2; \mathbf{Z}_{S_1}^n, \mathbf{Z}_{S_2}^n\Big) \quad (140)$$

$$\le \lim_{n\to\infty} \max\Big\{ \max_{S_1\subseteq[1\,:\,n]:\,|S_1|=\mu} I(M_c; \mathbf{Z}_{S_1}^n), \max_{S_2\subseteq[1\,:\,n]:\,|S_2|=\mu} I(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n)\Big\} \quad (141)$$

$$= \max\Big\{ \lim_{n\to\infty} \max_{S_1\subseteq[1\,:\,n]:\,|S_1|=\mu} I(M_c; \mathbf{Z}_{S_1}^n), \lim_{n\to\infty} \max_{S_2\subseteq[1\,:\,n]:\,|S_2|=\mu} I\Big(M_{\mathbf{d}}; \mathbf{Z}_{S_2}^n\Big)\Big\} = 0, \quad (142)$$

Equation (141) follows from (139); (142) follows since both limits exist and equal to zero, using [43, Th. 2].

REFERENCES

[1] M. Nafea and A. Yener, "The caching broadcast channel with a wire and cache tapping adversary of type II," in *Proc. IEEE Inf. Theory Workshop*, Guangzhou, China, Nov. 2018, pp. 1–5.

[2] M. Nafea and A. Yener, "The caching broadcast channel with a wire and cache tapping adversary of type II: Multiple library files," in *Proc. 56th Allerton Conf. Commun. Control Comput.*, Monticello, IL, USA, Oct. 2018, pp. 989–996.

[3] L. W. Dowdy and D. V. Foster, "Comparative models of the file assignment problem," *ACM Comput. Surveys*, vol. 14, no. 2, pp. 287–313, 1982.

[4] K. C. Almeroth and M. H. Ammar, "The use of multicast delivery to provide a scalable and interactive video-on-demand service," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 6, pp. 1110–1122, Aug. 1996.

[5] S. Borst, V. Gupta, and A. Walid, "Distributed caching algorithms for content distribution networks," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, San Diego, CA, USA, Mar. 2010, pp. 1–9.

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

[7] M. A. Maddah-Ali and U. Niesen, "Decentralized coded caching attains order-optimal memory-rate tradeoff," *IEEE/ACM Trans. Netw.*, vol. 23, no. 4, pp. 1029–1040, Aug. 2015.

[8] 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.

[9] K. Wan, D. Tuninetti, and P. Piantanida, "On caching with more users than files," in *Proc. IEEE Int. Symp. Inf. Theory*, Barcelona, Spain, Jul. 2016, pp. 809–813.

[10] S. Sahraei and M. Gastpar, "*K* users caching two files: An improved achievable rate," in *Proc. IEEE Conf. Inf. Sci. Syst.*, Princeton, NJ, USA, Mar. 2016, pp. 620–624.

[11] M. M. Amiri, Q. Yang, and D. Gündüz, "Coded caching for a large number of users," May 2016. [Online]. Available: arXiv:1605.01993

[12] A. M. Ibrahim, A. A. Zewail, and A. Yener, "Coded caching for heterogeneous systems: An optimization perspective," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5321–5335, Aug. 2019.

[13] S. H. Lim, C.-Y. Wang, and M. Gastpar, "Information-theoretic caching: The multi-user case," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7018–7037, Nov. 2017.

[14] 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, Feb. 2017.

[15] 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.

[16] N. Karamchandani, U. Niesen, M. A. Maddah-Ali, and S. N. Diggavi, "Hierarchical coded caching," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3212–3229, Jun. 2016.

[17] M. A. Maddah-Ali and U. Niesen, "Cache-aided interference channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2015, pp. 809–813.

[18] J. Hachem, U. Niesen, and S. N. Diggavi, "Degrees of freedom of cache-aided wireless interference networks," *IEEE Trans. Inf. Theory*, vol. 64, no. 7, pp. 5359–5380, Jul. 2018.

[19] 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.

[20] M. Ji *et al.*, "On the fundamental limits of caching in combination networks," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun.*, Stockholm, Sweden, Jun. 2015, pp. 695–699.

[21] A. A. Zewail and A. Yener, "Combination networks with or without secrecy constraints: The impact of caching relays," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 6, pp. 1140–1152, Jun. 2018.

[22] M. Ji, G. Caire, and A. F. Molisch, "Fundamental limits of caching in wireless D2D networks," *IEEE Trans. Inf. Theory*, vol. 62, no. 2, pp. 849–869, Feb. 2016.

[23] A. M. Ibrahim, A. A. Zewail, and A. Yener, "Device-to-device coded-caching with distinct cache sizes," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 2748–2762, May 2020.

[24] K. Zhang and C. Tian, "Fundamental limits of coded caching: From uncoded prefetching to coded prefetching," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 6, pp. 1153–1164, Jun. 2018.

[25] A. M. Ibrahim, A. A. Zewail, and A. Yener, "Benefits of coded placement for networks with heterogeneous cache sizes," in *Proc. IEEE Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2018, pp. 1604–1608.

[26] S. S. Bidokhti, M. Wigger, and R. Timo, "Noisy broadcast networks with receiver caching," *IEEE Trans. Inf. Theory*, vol. 64, no. 11, pp. 6996–7016, Nov. 2018.

[27] 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.

[28] M. Amiri and D. Gündüz, "Cache-aided content delivery over erasure broadcast channels," *IEEE Trans. Commun.*, vol. 66, no. 1, pp. 370–381, Jan. 2018.

[29] 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.

[30] S. S. Bidokhti, M. Wigger, and A. Yener, "Benefits of cache assignment on degraded broadcast channels," *IEEE Trans. Inf. Theory*, vol. 65, no. 11, pp. 6999–7019, Nov. 2019.

[31] A. Sengupta, R. Tandon, and T. C. Clancy, "Fundamental limits of caching with secure delivery," *IEEE Trans. Inf. Forensics Security*, vol. 10, pp. 355–370, 2015.

[32] Z. H. Awan and A. Sezgin, "Fundamental limits of caching in D2D networks with secure delivery," in *Proc. IEEE Int. Conf. Commun. Workshop*, London, U.K., Jun. 2015, pp. 464–469.

[33] V. Ravindrakumar, P. Panda, N. Karamchandani, and V. M. Prabhakaran, "Private coded caching," *IEEE Trans. Inf. Forensics Security*, vol. 13, pp. 685–694, 2018.

[34] A. A. Zewail and A. Yener, "Device-to-device secure coded caching," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 1513–1524, 2020.

[35] A. A. Zewail and A. Yener, "Secure caching and delivery for combination networks with asymmetric connectivity," in *Proc. IEEE Inf. Theory Workshop*, Visby, Sweden, Aug. 2019, pp. 1–5.

[36] S. Kamel, M. Sarkiss, M. Wigger, and G. R.-B. Othman, "Secrecy capacity-memory tradeoff of erasure broadcast channels," *IEEE Trans. Inf. Theory*, vol. 65, no. 8, pp. 5094–5124, Aug. 2019.

[37] A. A. Zewail and A. Yener, "The wiretap channel with a cache," in *Proc. IEEE Int. Symp. Inf. Theory*, Vail, CO, USA, Jun. 2018, pp. 1720–1724.

[38] A. A. Zewail and A. Yener, "Untrusted caches in two-layer networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Paris, France, Jul. 2019, pp. 1–5.

[39] C. E. Shannon, "Communication theory of secrecy systems," *Bell Labs Techn. J.*, vol. 28, no. 4, pp. 656–715, Oct. 1949.

[40] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography. I. Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, Jul. 1993.

[41] L. H. Ozarow and A. D. Wyner, "Wire-tap channel II," *AT T Bell Lab. Techn. J.*, vol. 63, no. 10, pp. 2135–2157, Dec. 1984.

[42] M. Nafea and A. Yener, "Wiretap channel II with a noisy main channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, China, Jun. 2015, pp. 1159–1163.

[43] Z. Goldfeld, P. Cuff, and H. H. Permuter, "Semantic-security capacity for wiretap channels of type II," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 3863–3879, Jul. 2016.

[44] M. Nafea and A. Yener, "A new wiretap channel model and its strong secrecy capacity," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 2077–2092, Mar. 2018.

[45] A. D. Wyner, "The wire-tap channel," *Bell Syst. Techn. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[46] M. Nafea and A. Yener, "Generalizing multiple access wiretap and wiretap II channel models: Achievable rates and cost of strong secrecy," *IEEE Trans. Inf. Theory*, vol. 65, no. 8, pp. 5125–5143, Aug. 2019.

[47] M. Nafea and A. Yener, "A new broadcast wiretap channel model," in *Proc. IEEE Int. Symp. Inf. Theory*, Aachen, Germany, Jun. 2017, pp. 2128–2132.

[48] M. Nafea and A. Yener, "New models for interference and broadcast channels with confidential messages," in *Proc. IEEE Int. Symp. Inf. Theory*, Aachen, Germany, Jun. 2017, pp. 1808–1812.

[49] H. D. Ly, T. Liu, and Y. Blankenship, "Security embedding codes," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 1, pp. 148–159, Feb. 2012.

[50] Y. Liang, L. Lai, H. V. Poor, and S. Shamai, "A broadcast approach for fading wiretap channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 842–858, Feb. 2014.

[51] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: Wiley, 2006.

[52] 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.

[53] D. Cao, D. Zhang, P. Chen, N. Liu, W. Kang, and D. Gündüz, "Coded caching with asymmetric cache sizes and link qualities: The two-user case," *IEEE Trans. Commun.*, vol. 67, no. 9, pp. 6112–6126, Sep. 2019.
[54] 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, Jun. 2017.
[55] E. Ekrem and S. Ulukus, "Secrecy capacity of a class of broadcast channels with an eavesdropper," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, pp. 1–29, Aug. 2009.
[56] M. Benammar and P. Piantanida, "Secrecy capacity region of some classes of wiretap broadcast channels," *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5564–5582, Oct. 2015.

Mohamed Nafea (Member, IEEE) received the B.Sc. degree in electrical engineering from Alexandria University, Egypt, in 2010, the M.Sc. degree in wireless communications from Nile University, Egypt, in 2012, the M.A. degree in mathematics, and the Ph.D. degree in electrical engineering from Pennsylvania State University, University Park, in 2017 and 2018, respectively. He has been an Assistant Professor with the Electrical and Computer Engineering Department, University of Detroit Mercy, since Fall 2020. He spent a year as a Postdoctoral Scholar with the Electrical and Computer Engineering Department, Georgia Institute of Technology. His research interests include network information theory, information theoretic security and privacy, algorithmic fairness, statistical machine learning, and causal structure learning.

Aylin Yener (Fellow, IEEE) received the B.Sc. degree in electrical and electronics engineering and the B.Sc. degree in physics from Bogazici University, Istanbul, Turkey, and the M.S. and Ph.D. degrees in electrical and computer engineering from Rutgers University, New Brunswick, NJ, USA.

She holds the Roy and Lois Chope Chair in Engineering with The Ohio State University, Columbus, OH, USA, since 2020, where she is a Professor of Electrical and Computer Engineering, Integrated Systems Engineering, and Computer Science and Engineering. Until 2020, she was a University Distinguished Professor of Electrical Engineering and the Dean's Fellow with the Pennsylvania State University, University Park, PA, USA, where she joined the Faculty as an Assistant Professor in 2002. She was a Visiting Professor of electrical engineering with Stanford University from 2016 to 2018 and a Visiting Associate Professor with the same department from 2008 to 2009. Her current research interests are in information security, green communications, caching systems, 6G, and more generally in the fields of information theory, communication theory and networked systems. She received the NSF CAREER Award in 2003, the Best Paper Award in Communication Theory from the IEEE International Conference on Communications in 2010, the Penn State Engineering Alumni Society (PSEAS) Outstanding Research Award in 2010, the IEEE Marconi Prize Paper Award in 2014, the PSEAS Premier Research Award in 2014, the Leonard A. Doggett Award for Outstanding Writing in Electrical Engineering at Penn State in 2014, the IEEE Women in Communications Engineering Outstanding Achievement Award in 2018, the IEEE Communications Society Best Tutorial Paper Award in 2019, and the IEEE Communications Society Communication Theory Technical Achievement Award in 2020. She has been a Distinguished Lecturer for the IEEE Information Theory Society from 2019 to 2021, the IEEE Communications Society from 2018 to 2019, and the IEEE Vehicular Technology Society from 2017 to 2021. She is currently serving as the Junior Past President of the IEEE Information Theory Society. Previously, she was the President in 2020, the Vice President in 2019, the Second Vice President in 2018, an Elected Member of the Board of Governors from 2015 to 2018, and the Treasurer from 2012 to 2014, of the IEEE Information Theory Society. She served as the Student Committee Chair for the IEEE Information Theory Society from 2007 to 2011, and was the Co-Founder of the Annual School of Information Theory in North America in 2008. She was the Technical Co-Chair for various symposia/tracks at the IEEE ICC, PIMRC, VTC, WCNC, and Asilomar in 2005, 2008–2014, and 2018. Previously, she served as an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS from 2009 to 2012, an Editor for IEEE TRANSACTIONS ON MOBILE COMPUTING from 2017 to 2018, and an Editor and an Editorial Advisory Board Member for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2001 to 2012. She also served as a Guest Editor for IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY in 2011, and IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS in 2015. Currently, she serves as an Senior Editor for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS and is on the Senior Editorial Board of IEEE JOURNAL ON SELECTED AREAS IN INFORMATION THEORY.