Secrecy Capacity-Memory Tradeoff of Erasure Broadcast Channels

Sarah Kamel, Mireille Sarkiss, Michèle Wigger, and Ghaya Rekaya-Ben Othman

***Abstract*—This paper derives upper and lower bounds on the *secrecy capacity-memory tradeoff* of a wiretap erasure broadcast channel (BC) with K_w weak receivers and K_s strong receivers, where weak receivers, respectively strong receivers, have same erasure probabilities and cache sizes. The lower bounds are achieved by schemes that meticulously combine joint cache-channel coding with wiretap coding and key-aided one-time pads. The presented upper bound holds more generally for arbitrary degraded BCs and arbitrary cache sizes. When only weak receivers have cache memories, upper and lower bounds coincide for small and large cache memories, thus providing the exact secrecy capacity-memory tradeoff for this setup. The derived bounds further allow us to conclude that the secrecy capacity is positive even when the eavesdropper is stronger than all the legitimate receivers with cache memories. Moreover, they show that the secrecy capacity-memory tradeoff can be significantly smaller than its non-secure counterpart, but it grows much faster when cache memories are small.**

The paper also presents a lower bound on the *global secrecy capacity-memory tradeoff* where one is allowed to optimize the cache assignment subject to a total cache budget. It is close to the best known lower bound without secrecy constraint. For small total cache budget, the global secrecy capacity-memory tradeoff is achieved by assigning all the available cache memory uniformly over *all* receivers if the eavesdropper is stronger than all legitimate receivers, and it is achieved by assigning the cache memory uniformly only over the *weak* receivers if the eavesdropper is weaker than the strong receivers.

I. Introduction

Traffic load in communication systems varies tremendously during the day between busy periods where the network is highly congested causing packet loss, delivery delays and unsatisfied users and other periods where the network is barely used. Lately, caching has emerged as a promising technique to reduce the network load and latency in such dense wireless networks. In caching, the communication is divided into two phases: the caching phase and the delivery phase. The caching phase occurs during the off-peak periods of the network, where fragments of popular contents are stored in users' cache memories or on nearby servers. The delivery phase occurs when users request specific files during the peak-traffic periods of the network and they are served partly from their cache memories and partly from the server. The technical challenge in these networks is that in the caching phase the servers do not know exactly which files the receivers will demand during the delivery phase. They are thus obliged to store information about *all possibly requested files* in the receivers' cache memories.

Maddah-Ali and Niesen showed in their seminal work [1] that the delivery (high-traffic) communication can benefit from the cache memories more than the obvious local caching gain arising from locally retrieving parts of the requested files. The additional gain, termed *global caching gain*, is obtained through carefully designing the cached contents and applying a new coding scheme, termed *coded caching*, which allows the transmitter to simultaneously serve multiple receivers. In [1] and in many subsequent works, the delivery phase is modelled as an error-free broadcast channel and all receivers have equal cache sizes. Coded caching is straightforwardly extended to noisy BCs by means of a *separate cache-channel coding* architecture where a standard BC code, which ignores the cache contents, is used to transmit the *delivery-messages*[1] produced by the coded caching algorithm to the receivers. Improved global caching gains can be achieved by employing *joint cache-channel coding* where the encoder and the decoder simultaneously adapt to the cache contents and the channel statistics [2]–[5].

A different line of works has addressed secrecy issues in cache-aided BCs [6]–[8], where delivery communication takes place over a noiseless link Different secrecy requirements have been studied. In [6], [8], the entire library of messages needs to be kept secret from an external eavesdropper that has access to the outputs of the common bit-pipe but not to the cache memories. This is achieved by means of securing the XOR packets produced by coded caching with secret keys, which have been prestored at the receivers during the caching phase [6]. This approach has subsequently been extended to resolvable networks in [9] and to device-to-device communication models [10]. In [7], each legitimate receiver acts also as eavesdropper and is thus not allowed to learn anything about the files requested by the other receivers. In this case, uncoded fragments of the messages cannot be stored in the users' caches. Instead, random keys and combinations of messages XORed with these random keys are cached. In the delivery phase, messages (or combination of messages) XORed with random keys are transmitted in a way that each message can be

S. Kamel, M. Wigger, and G. Rekaya-Ben Othman are with LTCI, Telecom ParisTech, Université Paris-Saclay, 75013 Paris. Email: {sarah.kamel, michele.wigger, ghaya.rekaya}@telecom-paristech.fr. S. Kamel and M. Wigger were supported by the ERC under grant agreement 751111.

M. Sarkiss is with CEA, LIST, Communicating Systems Laboratory, BC 173, 91191 Gif-sur-Yvette, France. Email: mireille.sarkiss@cea.fr

Parts of the material in this paper have been presented at the *IEEE International Conference on Communications (ICC)*, Paris, France, May 2017 [16], and at the *IEEE Information Theory Workshop (ITW)*, Kaohsiung, Taiwan, Nov. 2017.

[1]Due to the presence of the cache memories the messages conveyed in the delivery phase, generally differ from the original messages in the library.

decoded only by its intended receiver. Secret communication has also been considered for cache-aided heterogeneous and multiantenna interference networks, see for example [11], [12], and [13]. A different angle of attack on this problem is taken in [14], which presents a privacy-preserving protocol that prevents eavesdroppers from learning which users are requesting which files as well as the statistical popularities of the files.

In this paper, we follow the secrecy requirement in [6] where an external eavesdropper is not allowed to learn anything about *all* the files in the library, but here delivery communication takes place over an erasure BC with one transmitter and $\mathsf{K} \geq 2$ receivers. The eavesdropper does not have access to the cache memories but overhears the delivery communication over the erasure BC. The main interest of this paper is the *secrecy capacity-memory tradeoff* under strong secrecy of such a system, i.e., the largest message rate for which it is possible to find encoding and decoding functions so that the mutual information

$$I(W_1, \ldots, W_{\mathsf{D}}; Z^n) \tag{1}$$

between all the messages $W_1, \ldots, W_{\mathsf{D}}$ of the library and the eavesdropper's observations Z^n vanishes asymptotically for increasing blocklengths n. In our previous work [15], we have addressed the weaker secrecy constraint where the eavesdropper is not allowed to learn any information about each of the actually demanded messages *individually*. Each of the K mutual informations $\frac{1}{n}I(W_{d_1}; Z^n), \ldots, \frac{1}{n}I(W_{d_{\mathsf{K}}}; Z^n)$ thus needs to vanish asymptotically, where $W_{d_1}, \ldots, W_{d_{\mathsf{K}}}$ denoting the files requested and delivered to Receivers $1, \ldots, \mathsf{K}$. Our conference publications [15] and [16] suggest that the ultimate performance limits under these two secrecy constraints are close.

There are two basic approaches to render cache-aided BCs secure: 1) introduce a randomization message to the non-secure coding schemes so as to transform them into wiretap codes [17]; and 2) store secret keys into the cache memories and apply one-time pads [18] to parts of the delivery-messages which can then serve as secret keys to wiretap codes [19]–[21]. As we will see, in the BC scenarios where different receivers have different cache sizes, combinations thereof should be employed. Moreover, by applying superposition coding or joint cache-channel coding, secret keys stored in caches can even be used to secure the communication to other receivers. Our superposition and joint cache-channel coding schemes share common elements with wiretap coding for broadcast channels [17], [22], wiretap coding with secret keys [19], [20], and wiretap coding with receiver side-information [23], [24]. The performances of our new schemes show that under secrecy constraints, further global caching gains are thus possible than the previously reported gains for non-secure communication [1], [3].

Based on the described coding ideas, we propose lower bounds on the secrecy capacity-memory tradeoff of the cache-aided erasure BC in Figure 1. The system consists of K_w weak receivers $1, \ldots, \mathsf{K}_w$ that have equal erasure probability $\delta_w \geq 0$ and cache size M_w, K_s strong receivers $\mathsf{K}_w + 1, \ldots, \mathsf{K}$ that have equal erasure probability $\delta_s \geq 0$ and cache size M_s, and a single eavesdropper. (There is no assumption on the strength of the eavesdropper. It can be weaker than the weak receivers, stronger than the strong receivers, or between weak and strong receivers.) We also provide a general upper bound on the secrecy capacity-memory tradeoff of an arbitrary degraded BC with receivers having arbitrary cache sizes. Upper and lower bounds match for the setup in Figure 1 in special cases, for example, when $\mathsf{M}_s = 0$ and M_w is sufficiently large or small. The proposed bounds moreover allow us to conclude the following:

- *Secrecy Capacity is Positive even when Eavesdropper Stronger than Some of the Legitimate Receivers with Cache Memories.* (But it needs to be weaker than the legitimate receivers without cache memories.)

- *Secrecy Constraint Can Significantly Harm Capacity:* The secrecy capacity-memory tradeoff can be significantly smaller than its non-secure counterpart, especially when only weak receivers have cache memories. One explanation is that when only weak receivers have cache memories, the communication to the strong receivers needs to be either secured through a randomization message as in wiretap coding or other secrecy mechanisms, which both significantly reduce the rate of communication.

- *Caching* Gains *are More Important under a Secrecy Constraint:* In the regime of small cache memories, the gains of caching (i.e., the slope of the capacity) are more pronounced in our system with secrecy constraint than in the standard non-secure system. Consider for example, $\mathsf{M}_s = 0$ and M_w sufficiently small. In this regime, when the eavesdropper's erasure probability δ_z is larger than erasure probability at the strong receivers δ_s, the slope γ_{sec} of the secrecy capacity-memory tradeoff satisfies (see Corollary 2)
$$\gamma_{\text{sec}} = \frac{\mathsf{K}_w(\delta_z - \delta_s)}{\mathsf{K}_w(\delta_z - \delta_s) + \mathsf{K}_s(\delta_z - \delta_w)^+}. \tag{2}$$
The slope γ of the standard non-secure capacity-memory tradeoff satisfies (see [2, Theorem 2])
$$\gamma \leq \frac{\mathsf{K}_w}{\mathsf{D}}. \tag{3}$$
This latter slope γ thus deteriorates with increasing library size D, which is not the case for γ_{sec}. The main reason for this behavior is that in a standard system the cache memories are filled with data, and intuitively each stored bit is useful only under some of the demands. In a secure system, a good option is to store secret keys in the cache memories of the receivers. These secret keys are helpful for all possible demands, and therefore the caching gain does not degrade with the library size D.

- *Optimal Cache Assignments for Small Total Cache Budgets:* For small total cache budgets, the global secrecy capacity-memory tradeoff is achieved by assigning all of the cache memory uniformly only over the *weak* receivers if the eavesdropper is weaker than strong receivers, and it is achieved by assigning all the cache memory uniformly over *all* receivers, if the eavesdropper is stronger than all receivers.

Paper Organization: The remainder of this paper is organized as follows. Section II formally defines the problem. Section III presents a general upper bound on the secrecy capacity-memory tradeoff and specializes it to the specific model of this work. In Sections IV and V, we present our results for the scenarios when only weak receivers have cache memories and when all receivers have cache memories, and we sketch the coding schemes achieving the proposed lower bounds. Section VI contains our results on the global secrecy capacity-memory tradeoff. Sections VIII and VII describe and analyze in detail all the coding schemes proposed in this paper. Finally, Section IX concludes the paper.

II. Problem Definition

A. Channel Model

We consider a wiretap erasure BC with a single transmitter, K receivers and one eavesdropper, as shown in Figure 1. The input alphabet of the BC is

$$\mathcal{X}=\{0,1\} \tag{4}$$

and all receivers and the eavesdropper have the same output alphabet

$$\mathcal{Z} = \mathcal{Y} = \mathcal{X} \cup \Delta. \tag{5}$$

The output erasure symbol Δ indicates the loss of a bit at the receiver. Let δ_k be the erasure probability of Receiver k's channel, for $k \in \mathcal{K} := \{1, \ldots, \mathsf{K}\}$, and δ_z be the erasure probability of the eavesdropper's channel. Then, for each $k \in \mathcal{K}$, the marginal transition law at Receiver k is described by

$$P_{Y_k|X}(y_k|x) = \begin{cases} 1-\delta_k & \text{if } \; y_k = x \\ \delta_k & \text{if } \; y_k = \Delta \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

and at the eavesdropper by

$$P_{Z|X}(z|x) = \begin{cases} 1-\delta_z & \text{if } \; z = x \\ \delta_z & \text{if } \; z = \Delta \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

for some parameters $0 \leq \delta_1, \ldots, \delta_K, \delta_z \leq 1$.

The K receivers are partitioned into two sets. The first set

$$\mathcal{K}_w := \{1, \ldots, \mathsf{K}_w\} \tag{8}$$

is formed by K_w weak receivers, which have bad channel conditions. The second set

$$\mathcal{K}_s := \{\mathsf{K}_w + 1, \ldots, \mathsf{K}\} \tag{9}$$

is formed by $\mathsf{K}_s = \mathsf{K} - \mathsf{K}_w$ strong receivers, which have good channel conditions.

In other words, we assume that

$$\delta_k = \begin{cases} \delta_w & \text{if } \; k \in \mathcal{K}_w \\ \delta_s & \text{if } \; k \in \mathcal{K}_s \end{cases} \tag{10}$$

with

$$0 \leq \delta_s \leq \delta_w \leq 1. \tag{11}$$

In a standard wiretap erasure BC, reliable communication with positive rates is only possible when the eavesdropper's erasure probability δ_z is larger than the erasure probabilities at all legitimate receivers. Here, any configuration of erasure probabilities is possible, because the legitimate receivers can prestore information in local cache memories and this information is not accessible by the eavesdropper. Specifically, we assume that each weak receiver has access to a local cache memory of size $n\mathsf{M}_w$ bits and each strong receiver has access to a local cache memory of size $n\mathsf{M}_s$ bits.

B. Library and Receiver Demands

The transmitter can access a library of $\mathsf{D} > \mathsf{K}$ independent messages

$$W_1, \ldots, W_{\mathsf{D}} \tag{12}$$

of rate $R \geq 0$ each. So for each $d \in \mathcal{D}$,

$$\mathcal{D} := \{1, \ldots, \mathsf{D}\}, \tag{13}$$

message W_d is uniformly distributed over the set $\{1, \ldots, \lfloor 2^{nR} \rfloor\}$, where n is the transmission blocklength.

Every receiver $k \in \mathcal{K}$ demands exactly one message W_{d_k} from the library. We denote the demand of Receiver k by $d_k \in \mathcal{D}$ and the *demand vector* of all receivers by

$$\mathbf{d} := (d_1, \ldots, d_{\mathsf{K}}) \in \mathcal{D}^{\mathsf{K}}. \tag{14}$$

Communication takes place in two phases: a first caching phase where the transmitter sends caching information to be stored in the receivers' cache memories and the subsequent delivery phase where the demanded messages W_{d_k}, for $k \in \mathcal{K}$, are conveyed to the receivers.

C. Placement Phase

The placement phase takes place during periods of low network-traffic. The transmitter therefore has all the available resources to transmit the cached contents in an error-free and secure fashion, and this first phase is only restricted by the storage constraints on the cache memories. However, since the caching phase takes place before the receivers demand their files, the cached content cannot depend on the demand vector $\mathbf{d}$, but only on the library and local randomness θ that is accessible by the transmitter. The cache content V_k stored at receiver $k \in \{1, \ldots, \mathsf{K}\}$ is thus of the form

$$V_k = g_k \left(W_1, \ldots, W_{\mathsf{D}}, \theta \right), \tag{15}$$

for some caching function

$$g_k \colon \left\{1, \ldots, \lfloor 2^{nR} \rfloor \right\}^{\mathsf{D}} \times \Theta \to \mathcal{V}_i \tag{16}$$

where for $k \in \mathcal{K}_w$

$$\mathcal{V}_k := \left\{1, \ldots, \lfloor 2^{n\mathsf{M}_w} \rfloor \right\}, \quad k \in \mathcal{K}_w \tag{17}$$

and for $k \in \mathcal{K}_s$

$$\mathcal{V}_k := \left\{1, \ldots, \lfloor 2^{n\mathsf{M}_s} \rfloor \right\}, \quad k \in \mathcal{K}_s. \tag{18}$$

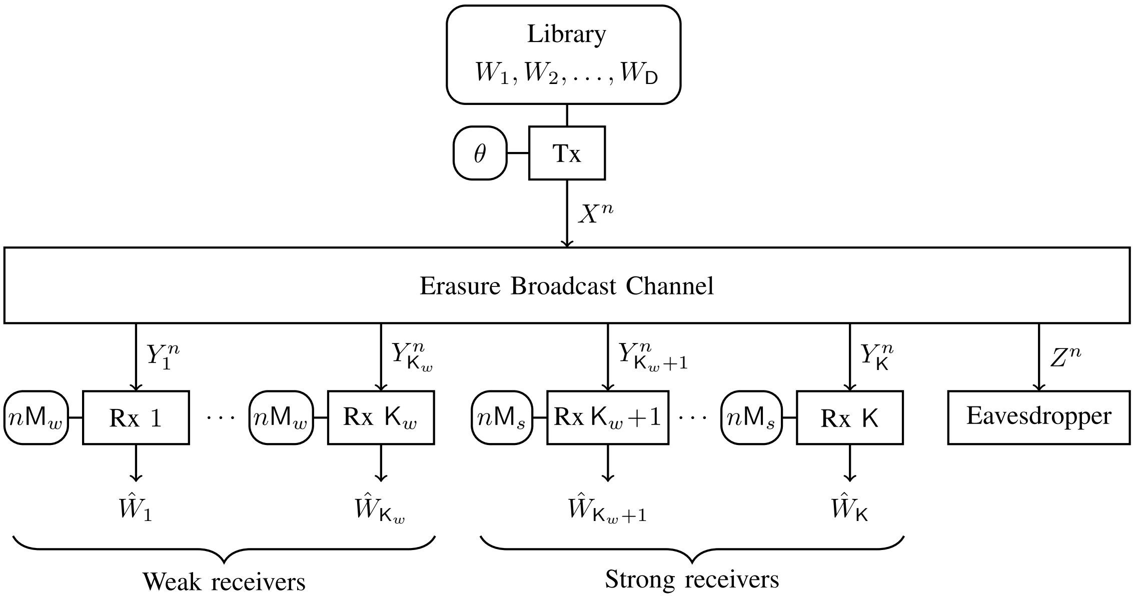


Fig. 1. Erasure BC with $\mathsf{K} = \mathsf{K}_w + \mathsf{K}_s$ legitimate receivers and an eavesdropper. The K_w weaker receivers have cache memories of size M_w and the K_s stronger receivers have cache memories of size M_s. The random variable θ models a source of randomness locally available at the transmitter.

D. Delivery Phase

Prior to the delivery phase, the demand vector $\mathbf{d}$ is learned by the transmitter and the legitimate receivers. The communication of the demand vector requires zero communication rate since it takes only $\mathsf{K} \cdot \lceil \log(\mathsf{D}) \rceil$ bits to describe $\mathbf{d}$.

Based on the demand vector $\mathbf{d}$, the transmitter sends

$$X^n = f_{\mathbf{d}}\left(W_1, \ldots, W_{\mathsf{D}}, \theta\right), \tag{19}$$

for some function

$$f_{\mathbf{d}} \colon \left\{1, \ldots, \lfloor 2^{nR} \rfloor\right\}^{\mathsf{D}} \times \Theta \to \mathcal{X}^n. \tag{20}$$

Each Receiver $k \in \mathcal{K}$, attempts to decode its demanded message W_{d_k} based on its observed outputs Y_k^n and its cache content V_k:

$$\hat{W}_k := \varphi_{k,\mathbf{d}}(Y_k^n, V_k), \quad k \in \mathcal{K}, \tag{21}$$

for some function

$$\varphi_{k,\mathbf{d}} \colon \mathcal{Y}^n \times \mathcal{V}_k \to \left\{1, \ldots, \lfloor 2^{nR} \rfloor\right\}. \tag{22}$$

E. Secrecy Capacity-Memory Tradeoff

A decoding error occurs whenever $\hat{W}_k \neq W_{d_k}$, for some $k \in \mathcal{K}$. We consider the worst-case probability of error over all feasible demand vectors

$$\mathrm{P}_e^{\text{Worst}} := \max_{\mathbf{d} \in \mathcal{D}^{\mathsf{K}}} \mathrm{P}\left[\bigcup_{k=1}^{\mathsf{K}} \left\{\hat{W}_k \neq W_{d_k}\right\}\right]. \tag{23}$$

The communication is considered secure if the eavesdropper's channel outputs Z^n during the delivery phase provide almost no information about the entire library. The mutual information $I\left(W_1, \ldots, W_{\mathsf{D}}; Z^n\right)$ is considered as a secrecy measure, i.e., we require:

$$I\left(W_1, \ldots, W_{\mathsf{D}}; Z^n\right) < \epsilon. \tag{24}$$

Since we assumed earlier that the caching phase is secure, the eavesdropped observations Z_n only concern the delivery phase.

Definition 1. *A rate-memory triple* $(R, \mathsf{M}_w, \mathsf{M}_s)$ *is* securely achievable *if for every* $\epsilon > 0$ *and sufficiently large blocklength* n*, there exist caching, encoding, and decoding functions as in* (16)*,* (20)*, and* (22) *so that*

$$\mathrm{P}_e^{\text{Worst}} \leq \epsilon \qquad \textit{and} \qquad I\left(W_1, \ldots, W_{\mathsf{D}}; Z^n\right) < \epsilon. \tag{25}$$

In our previous works [15] and [16], we called this secrecy constraint a *joint secrecy constraint* to distinguish it from the *individual secrecy constraint* in [15] where the second inequality in (25) is replaced by $\frac{1}{n} I\left(W_{d_k}; Z^n\right) < \epsilon, \forall k \in \mathcal{K}$.

Definition 2. *Given cache memory sizes* $(\mathsf{M}_w, \mathsf{M}_s)$*, the* secrecy capacity-memory tradeoff $C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s)$ *is the supremum of all rates* R *so that the triple* $(R, \mathsf{M}_w, \mathsf{M}_s)$ *is securely achievable:*

$$C_{\text{sec}}\left(\mathsf{M}_w, \mathsf{M}_s\right) := \sup\left\{R \colon \quad (R, \mathsf{M}_w, \mathsf{M}_s) \text{ securely achievable}\right\}. \tag{26}$$

Remark 1. *Without cache memories, i.e.,* $\mathsf{M}_w = \mathsf{M}_s = 0$*, the secrecy capacity-memory tradeoff* $C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s)$ *was determined in [17]:*

$$C_{\text{sec}}\left(\mathsf{M}_w = 0, \mathsf{M}_s = 0\right) = \left(\sum_{k=1}^{\mathsf{K}} \frac{1}{\delta_z - \delta_k}\right)^{-1}. \tag{27}$$

For comparison, we will also be interested in the standard (non-secure) capacity-memory tradeoff $C\left(\mathsf{M}_w, \mathsf{M}_s\right)$ as defined in [3], [25]. It is the largest rate R for given cache sizes M_w and M_s, for which there exist caching, encoding, and decoding functions as in (16), (20), and (22) so that

$$\mathrm{P}_e^{\text{Worst}} \leq \epsilon. \tag{28}$$

F. Preliminaries: A Mapping and A Lemma

We introduce a mapping $\mathsf{sec}(w,k)$ that represents the integers w and k as binary strings, zero-pads the shorter of the two strings to the length of the longer one, and takes the component-wise XOR function of the two strings. The output of the mapping is the integer corresponding to the produced binary string. Formally, the mapping is defined as follows. Assume that the positive real numbers n, R', R_{Key} are fixed and clear from the context. Then:

$$\begin{aligned}\mathsf{sec}(w,k)\colon \lfloor 2^{nR'}\rfloor \times \lfloor 2^{nR_{\text{Key}}}\rfloor &\to \lfloor 2^{nR'}\rfloor \\ (w,k) &\mapsto (w+k) \ \mathsf{mod}\ \lfloor 2^{nR'}\rfloor,\end{aligned} \tag{29}$$

where mod denotes the modulo operator. Notice that from $\mathsf{sec}(w,k)$ and k it is possible to recover w. Let $\mathsf{sec}_k^{-1}(\cdot)$ denote this inverse mapping so that

$$\mathsf{sec}_k^{-1}\left(\mathsf{sec}(w,k)\right) = w. \tag{30}$$

When $R' = R_{\text{Key}}$, we also write $\bigoplus$ instead of sec:

$$w_1 \bigoplus w_2 := \mathsf{sec}\big(w_1, w_2\big). \tag{31}$$

We mostly use sec when one of the arguments is a secret key and we use $\bigoplus$ when both arguments refer to messages.

The following lemma is essential in the secrecy analysis of our schemes.

Lemma 1. *Consider a blocklength n', the rates $\tilde{R} \geq R' > 0$, and the random codebook*

$$\mathcal{C} = \left\{ x^{n'}(\ell)\colon \ \ \ell \in \{1,\ldots,\lfloor 2^{n'\tilde{R}}\rfloor\}\right\} \tag{32}$$

with entries drawn i.i.d. according to some distribution P_X. Let the message W' be uniform over $\{1,\ldots,\lfloor 2^{n'R'}\rfloor\}$. To encode the message $W' = w'$, the transmitter picks uniformly at random*, for example as a function of an independent secret key, a codeword from a predetermined subset $\mathcal{S}(w') \subseteq \mathcal{C}$ and sends this codeword over the channel. If $\mathcal{S}(w')$ does not depend on $w' \in \{1,\ldots,\lfloor 2^{n'R'}\rfloor\}$,*

$$\mathcal{S}(1) = \ldots = \mathcal{S}\big(\lfloor 2^{n'R'}\rfloor\big), \tag{33}$$

or if the cardinalities of the subsets $\{\mathcal{S}(w')\}$ satisfy

$$\frac{1}{n'}\log_2|\mathcal{S}(w')| \geq I(X;Z), \quad \forall w' \in \{1,\ldots,\lfloor 2^{n'R'}\rfloor\}, \tag{34}$$

where $(X,Z) \sim P_X P_{Z|X}$, then

$$I\big(W'; Z^{n'}|\mathcal{C}\big) \to 0 \qquad \text{as} \qquad n' \to \infty. \tag{35}$$

Proof: If (33) holds, then the described encoding with a random choice over $\mathcal{S}(w')$ is equivalent to securing the message with a one-time pad [18] prior to encoding. In this case, $I(W'; Z^{n'}|\mathcal{C}) = 0$. If (34) holds, the limit (35) can be proved following the lines in [21, Theorem 7]. ∎

III. Upper Bounds on Secrecy Capacity-Memory Tradeoff

We start by presenting an upper bound on the secrecy capacity-memory tradeoff of a general degraded K-user BC with arbitrary cache sizes $\mathsf{M}_1,\ldots,\mathsf{M}_\mathsf{K}$ at the receivers. Subsequently, we specialize this bound to the erasure BC studied in this paper where all weak receivers and all strong receivers have equal cache sizes.

Consider an arbitrary degraded K-user discrete memoryless BC (not necessarily an erasure BC) with channel transition law $\Gamma(y_1,\ldots,y_\mathsf{K}|x)$. For simplicity, and because our result depends only on the conditional marginals $\Gamma_1(y_1|x),\ldots,\Gamma_\mathsf{K}(y_\mathsf{K}|x)$, we assume that the channel is physically degraded, so the Markov chain

$$X \to Y_\mathsf{K} \to Y_{\mathsf{K}-1} \to \ldots \to Y_1 \tag{36}$$

holds. In the same spirit, we also assume that the eavesdropper is degraded with respect to some of the legitimate receivers, and all other legitimate receivers are degraded with respect to the eavesdropper. Three scenarios can be considered:

a) The eavesdropper is degraded with respect to *all* legitimate receivers:
$$X \to Y_\mathsf{K} \to Y_{\mathsf{K}-1} \to \ldots \to Y_1 \to Z. \tag{37}$$
b) *All* legitimate receivers are degraded with respect to the eavesdropper:
$$X \to Z \to Y_\mathsf{K} \to Y_{\mathsf{K}-1} \to \ldots \to Y_1. \tag{38}$$
c) The eavesdropper is degraded with respect to the strongest $\mathsf{K} - \ell^*$ legitimate receivers, for some $\ell^* \in \{1,\ldots,\mathsf{K}-1\}$, and the remaining legitimate receivers are degraded with respect to the eavesdropper:
$$\begin{aligned}X \to Y_\mathsf{K} \to Y_{\mathsf{K}-1} \to \ldots \to Y_{\ell^*+1}& \\ \to Z \to Y_{\ell^*} \to \ldots \to Y_1&\end{aligned} \tag{39}$$

Let each Receiver $k \in \{1,\ldots,\mathsf{K}\}$ have cache size M_k. The following lemma holds.

Lemma 2 (Upper Bound for Arbitrary Degraded BCs and Cache Sizes)**.** *If a rate-memory tuple $(R,\mathsf{M}_1,\ldots,\mathsf{M}_\mathsf{K})$ is securely achievable, then for each receiver set $\mathcal{S} := \{j_1,\ldots,j_{|\mathcal{S}|}\} \subseteq \mathcal{K}$, there exist auxiliaries $(U_1,U_2,\ldots,U_{|\mathcal{S}|},Q)$ so that for each realization of $Q = q$ the following Markov chain holds:*

$$U_1 \to U_2 \to \ldots \to U_{|\mathcal{S}|} \to X \to (Y_{j_1},\ldots,Y_{j_{|\mathcal{S}|}},Z); \tag{40}$$

and the following $|\mathcal{S}|$ inequalities are satisfied:

$$R \leq \big[I(U_1;Y_{j_1}|Q) - I(U_1;Z|Q)\big]^+ + \mathsf{M}_{j_1}, \tag{41a}$$

and

$$kR \leq \sum_{\ell=1}^{k}\big[I(U_\ell;Y_{j_\ell}|U_{\ell-1},Q) - I(U_\ell;Z|U_{\ell-1},Q)\big]^+ + \sum_{\ell=1}^{k}\mathsf{M}_{j_\ell}, \quad k \in \{2,\ldots,|\mathcal{S}|\}, \tag{41b}$$

where $(\cdot)^+ := \max\{0,\cdot\}$.

Proof. See Appendix A. □

Turn back to the setup with weak and strong receivers in Figure 1. Based on the previous lemma and the upper bound on the standard (non-secure) capacity-memory tradeoff in [25, Theorem 5], the following Theorem 1 presents two upper bounds (Inequalities (42a) and (42b)) on the secrecy capacity-memory tradeoff for each choice of $k_w \in \{0,1,\ldots,\mathsf{K}_w\}$ and $k_s \in \{0,1,\ldots,\mathsf{K}_s\}$. Depending on the cache sizes M_w and M_s, a different choice of the parameters and of the bounds (42a) or (42b) is tightest.

Theorem 1 (Upper Bound on $C_{\text{sec}}(\mathsf{M}_w,\mathsf{M}_s)$)**.** *For each choice of $k_w \in \{0,1,\ldots,\mathsf{K}_w\}$ and $k_s \in \{0,1,\ldots,\mathsf{K}_s\}$, the secrecy capacity-memory tradeoff is upper bounded in the following two ways:*[2]

$$C_{\text{sec}}(\mathsf{M}_w,\mathsf{M}_s) \leq \max_{\beta\in[0,1]} \min\left\{ \frac{\beta(\delta_z-\delta_w)^+}{k_w} + \mathsf{M}_w,\; \frac{\beta(\delta_z-\delta_w)^+ + (1-\beta)(\delta_z-\delta_s)^+}{k_w+k_s} + \frac{k_w\mathsf{M}_w + k_s\mathsf{M}_s}{k_w+k_s} \right\} \tag{42a}$$

and

$$C_{\text{sec}}(\mathsf{M}_w,\mathsf{M}_s) \leq \min\left\{ \min_{i=1,\ldots,k_w} \{(1-\delta_w)\beta_i + \alpha_i\},\; \min_{j=1,\ldots,k_s}\left\{(1-\delta_s)\beta_{k_w+j} + \alpha_{k_w+j}\right\}\right\} \tag{42b}$$

for some tuple of nonnegative real numbers $\beta_1,\ldots,\beta_{k_w+k_s} \geq 0$ summing to 1, where for $i \in \{1,\ldots,k_w\}$:

$$\alpha_i := \min\left\{ \frac{i\mathsf{M}_w}{\mathsf{D}-i+1},\; \frac{1}{k_w+k_s-i+1} \cdot \left(\frac{(k_w+k_s)(k_w\mathsf{M}_w+k_s\mathsf{M}_s)}{\mathsf{D}} - \sum_{\ell=1}^{i-1}\alpha_\ell \right)\right\},$$

and for $j \in \{1,\ldots,k_s\}$:

$$\alpha_{k_w+j} := \min\left\{ \frac{k_w\mathsf{M}_w + j\mathsf{M}_s}{\mathsf{D}-k_w-j+1},\; \frac{1}{k_s-j+1} \cdot \left(\frac{(k_w+k_s)(k_w\mathsf{M}_w+k_s\mathsf{M}_s)}{\mathsf{D}} - \sum_{\ell=1}^{k_w+j-1}\alpha_\ell \right)\right\}.$$

[2]For any finite numbers a, b, we define $\min\{a/0, b\} = b$ and $\min\{a/0, b/0\} = \infty$. In the minimization (42b), a minimum over an empty set is defined as $+\infty$.

Each of the terms in the two upper bounds (42a) and (42b) is a sum of a first summand that depends only on the channel parameters and a second summand that depends also on the cache sizes. In (42a), the second summand equals the *average* cache size over a subset of users. In (42b), the second summand is approximately equal to the ratio between the *total* cache size of a subset of users and the total number of files D. As shown shortly by means of matching lower bounds or through numerical simulations, the bound in (42a) is tighter for small cache sizes and the bound in (42b) is tighter for moderate cache sizes. For large cache sizes, any of the two bounds can dominate depending on the scenario.

In fact, the results that we shall present shortly show that small cache memories should be used to exclusively store secret keys, resulting in a gain that only depends on the local cache size but not on the cache sizes of the other receivers nor on the total number of files in the system. This explains the form of the upper bound in (42a). Larger cache memories should be used to also store parts of each and every file in the library. On one hand, such a cache placement results in a caching gain that decays inversely proportional to the number of files D in the system. On the other hand, the cache placement creates multi-cast opportunities to many receivers, and thus can profit also from the cache sizes at other receivers. This explains why the upper bound in (42b) decay inversely proportional to D and depends on the total cache size across users.

Proof of Theorem 1. The first upper bound in (42a) is obtained by specializing Lemma 2 to the erasure BC in Figure 1. More specifically, setting $\beta_j := I(U_j; X|U_{j-1}, Q)$, constraints (41) can be rewritten as:

$$C_{\text{sec}}(\mathsf{M}_w,\mathsf{M}_s) \leq \beta_1(\delta_z-\delta_{j_1})^+ + \mathsf{M}_{j_1}, \tag{43a}$$

$$C_{\text{sec}}(\mathsf{M}_w,\mathsf{M}_s) \leq \frac{1}{k}\sum_{\ell=1}^{k}\left[\beta_\ell(\delta_z-\delta_{j_\ell})^+ + \mathsf{M}_{j_\ell}\right], \quad \forall k \in \{1,\ldots,|\mathcal{S}|\}. \tag{43b}$$

By well known properties on the mutual information, one finds that $\beta_1,\ldots,\beta_{|\mathcal{S}|} \geq 0$ and

$$\sum_{k=1}^{|\mathcal{S}|}\beta_k = I(U_1,\ldots,U_{|\mathcal{S}|}; X|Q) \leq H(X) \leq 1. \tag{44}$$

Upper bound (42a) is now obtained by specializing (43) to one of the subsets

$$\mathcal{S} = \{1,\ldots,k_w,\mathsf{K}_w+1,\ldots,\mathsf{K}_w+k_s\}, \qquad k_w,k_s>0, \tag{45a}$$

or

$$\mathcal{S} = \{1,\ldots,k_w\}, \tag{45b}$$

or

$$\mathcal{S} = \{\mathsf{K}_w+1,\ldots,\mathsf{K}_w+k_s\}, \tag{45c}$$

and by noticing that for the subset in (45a) one can restrict to

$$\beta_1 = \beta_2 = \ldots = \beta_{k_w} = \frac{\beta}{k_w} \tag{46}$$

and

$$\beta_{\mathsf{K}_w+1} = \beta_{\mathsf{K}_w+2} = \ldots = \beta_{\mathsf{K}_w+k_s} = \frac{1-\beta}{k_s}, \tag{47}$$

for some $\beta \in [0,1]$. For the subset in (45b) one can restrict to

$$\beta_1 = \beta_2 = \ldots = \beta_{k_w} = \frac{1}{k_w}, \tag{48}$$

and the subset in (45c) one can restrict to

$$\beta_{\mathsf{K}_w+1} = \beta_{\mathsf{K}_w+2} = \ldots = \beta_{\mathsf{K}_w+k_s} = \frac{1}{k_s}. \tag{49}$$

Constraint (42b) follows by ignoring the secrecy constraint and specializing [25, Theorem 5] to the erasure BC with weak and strong receivers considered in this paper. □

We simplify this upper bound for the special cases where only weak receivers have cache memories. More specifically, we replace the upper bound in (42b), which is obtained from [25, Theorem 5], by a simpler bound obtained by specializing the weaker upper bound in [3].

Corollary 1 (Upper Bound on $C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0)$). *For each choice of $k_w \in \{0, 1, \ldots, \mathsf{K}_w\}$, the secrecy capacity-memory tradeoff $C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0)$ is upper bounded in the following two ways:*

$$C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0) \leq \left(\frac{k_w}{1-\delta_w} + \frac{\mathsf{K}_s}{1-\delta_s} \right)^{-1} + \frac{k_w \mathsf{M}_w}{\mathsf{D}}, \tag{50a}$$

and

$$\begin{aligned} &C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0) \\ &\leq \max_{\beta \in [0,1]} \min \left\{ \frac{\beta(\delta_z - \delta_w)^+}{k_w} + \mathsf{M}_w, \right. \\ &\qquad \left. \frac{\beta(\delta_z - \delta_w)^+ + (1-\beta)(\delta_z - \delta_s)^+}{k_w + \mathsf{K}_s} + \frac{k_w}{k_w + \mathsf{K}_s} \mathsf{M}_w \right\}. \end{aligned} \tag{50b}$$

Proof. Constraint (50a) follows from [3] and by ignoring the secrecy constraint. Constraint (50b) is obtained by specializing (42a) to the case when $\mathsf{M}_s = 0$. We notice that in this case the constraint that (42a) generates for $j = \mathsf{K}_s$ is tighter than any constraint that it generates for $j < \mathsf{K}_s$. Thus, we can remove all constraints for $j < \mathsf{K}_s$ without affecting the result and we retain only the constraint in (50b). □

IV. Coding Schemes and Results when Only Weak Receivers have Cache Memories

Consider the special case where only weak receivers have cache memories, i.e.

$$\mathsf{M}_s = 0. \tag{51}$$

In this case, a positive secrecy rate can only be achieved if

$$\delta_z > \delta_s. \tag{52}$$

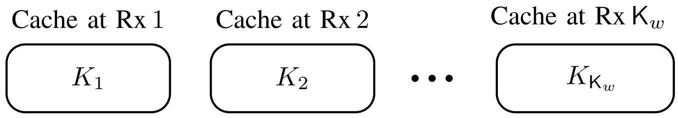
In the remainder of this section we assume that (52) holds.

A. Coding Schemes

We present four coding schemes in the order of increasing cache requirements. In the first two schemes, only random keys are placed in the cache memories. The third and fourth schemes also place parts of the messages in the cache memories and apply joint cache-channel coding for the delivery communication where the decoding operations at the receivers adapt at the same time to the channel statistics and the cache contents. For simplicity, and because time-sharing is optimal on an erasure BC to send independent messages to the various receivers, in some of our schemes communication is divided into subphases. When applied to general discrete memoryless BCs, the schemes can be improved by superposing various subphases on each other.

1) Wiretap and Cached Keys:

Placement phase: Store an independent secret key K_i in Receiver i's cache memory, for $i \in \mathsf{K}_w$.



Delivery phase: Time-sharing is applied over two subphases, where transmission in the first subphase is to all the weak receivers and transmission in the second subphase is to all strong receivers. In Subphase 1, the transmitter uses a standard (non-secure) broadcast code to send the secured message tuple

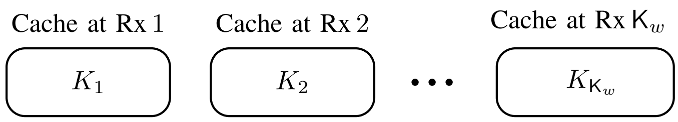
$$\begin{aligned} \mathbf{W}_{\text{sec}} := \Big(&\mathsf{sec}(W_{d_1}, K_1),\ \mathsf{sec}(W_{d_2}, K_2), \\ &\mathsf{sec}(W_{d_3}, K_3),\ \ldots,\ \mathsf{sec}(W_{d_{\mathsf{K}_w}}, K_{\mathsf{K}_w}) \Big), \end{aligned} \tag{53}$$

to weak receivers $1, \ldots, \mathsf{K}_w$, respectively. With the secret key K_i stored in its cache memory, each weak receiver $i \in \mathcal{K}_w$ can then recover a guess of its desired message W_{d_i}. In Subphase 2, the transmitter uses a wiretap broadcast code [17] to send messages $W_{d_{\mathsf{K}_w+1}}, \ldots, W_{d_\mathsf{K}}$ to the strong receivers $\mathsf{K}_w + 1, \ldots, \mathsf{K}$, respectively.

For a detailed analysis, see Section VII-A.

2) Cache-Aided Superposition Jamming:

Placement phase: As in the previous subsection, store an independent secret key K_i in Receiver i's cache memory, for $i \in \mathsf{K}_w$.



Delivery phase: The transmitter uses a superposition code to send the secured message tuple

$$\begin{aligned} \mathbf{W}_{\text{sec}} := \big(&\mathsf{sec}(W_{d_1}, K_1),\ \ \mathsf{sec}(W_{d_2}, K_2), \\ &\mathsf{sec}(W_{d_3}, K_3),\ \ldots,\ \mathsf{sec}(W_{d_{\mathsf{K}_w}}, K_{\mathsf{K}_w}) \big), \end{aligned} \tag{54}$$

in the cloud center and the non-secure message tuple

$$\mathbf{W}_{\text{sat}} := \left(W_{d_{\mathsf{K}_w+1}}, \ldots, W_{d_\mathsf{K}} \right) \tag{55}$$

in the satellite. The message $\mathbf{W}_{\text{sec}}$ sent in the cloud center is secured by the secret keys $K_1, \ldots, K_{\mathsf{K}_w}$. If the key rate exceeds $I(X;Z)$, these secret keys can secure also the message $\mathbf{W}_{\text{sat}}$ sent in the satellite. This holds by Lemma 1 and because the codeword used to send $\mathbf{W}_{\text{sat}}$ is chosen uniformly at random (depending on the secret keys $K_1, \ldots, K_{\mathsf{K}_w}$) over a subset of the codebook of rate exceeding $I(X;Z)$. If the key rate is below $I(X;Z)$, then we send an additional randomization message in the satellite to achieve the desired secrecy. The code construction in this latter case is depicted in Figure 2. Weak receivers decode only the cloud center and strong receivers the cloud center and the satellite codeword. From this decoding operation, each strong receiver $j \in \mathsf{K}_s$ directly obtains a guess of W_{d_j}. Each weak receiver $i \in \mathsf{K}_w$ uses the secret key K_i stored in its cache memory to recover a guess of its intended message W_{d_i}. Section VII-B presents the details of the scheme and its analysis.

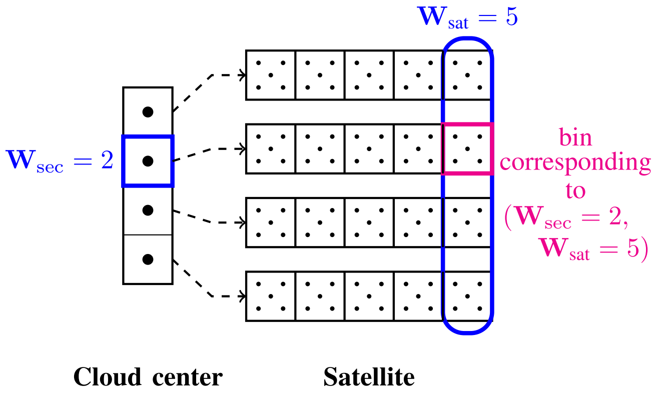


Fig. 2. Superposition codebook with randomization messages in the satellites.

3) Secure Cache-Aided Piggyback Coding I:

The scheme builds on the nested piggyback coding scheme in [3], which is rendered secure by applying secret keys to the produced XOR-messages and by introducing a randomization message as in wiretap coding. During the placement phase, each of these secret keys is stored in the cache memories of the weak receivers that decode the corresponding XOR-message. We outline the scheme for $\mathsf{K}_w = 3$ weak receivers and $\mathsf{K}_s = 1$ strong receiver.

Divide for each $d \in \mathcal{D}$, the message W_d into six submessages

$$W_d = \left(W^{(A)}_{d,\{1\}}, W^{(A)}_{d,\{2\}}, W^{(A)}_{d,\{3\}}, W^{(B)}_{d,\{1,2\}}, W^{(B)}_{d,\{1,3\}}, W^{(B)}_{d,\{2,3\}}\right), \tag{56}$$

where the first three are of equal rate and the latter three are of equal rate. Let $K_{\{1,2,3\}}$, $K_{\{1,2\}}, K_{\{2,3\}}, K_{\{1,3\}}$ be independent secret keys generated at the transmitter.

Placement phase: Placement is as described in the following table:

Cache at Rx 1	Cache at Rx 2	Cache at Rx 3
$\left\{W^{(A)}_{d,\{1\}}\right\}_{d=1}^{\mathsf{D}}$	$\left\{W^{(A)}_{d,\{2\}}\right\}_{d=1}^{\mathsf{D}}$	$\left\{W^{(A)}_{d,\{3\}}\right\}_{d=1}^{\mathsf{D}}$
$\left\{W^{(B)}_{d,\{1,2\}}, W^{(B)}_{d,\{1,3\}}\right\}_{d=1}^{\mathsf{D}}$	$\left\{W^{(B)}_{d,\{1,2\}}, W^{(B)}_{d,\{2,3\}}\right\}_{d=1}^{\mathsf{D}}$	$\left\{W^{(B)}_{d,\{1,3\}}, W^{(B)}_{d,\{2,3\}}\right\}_{d=1}^{\mathsf{D}}$
$K_{\{1,2,3\}}$	$K_{\{1,2,3\}}$	$K_{\{1,2,3\}}$
$K_{\{1,2\}}, K_{\{1,3\}}$	$K_{\{1,2\}}, K_{\{2,3\}}$	$K_{\{1,3\}}, K_{\{2,3\}}$

Delivery phase: Time-sharing is applied over three subphases and Subphase 2 is further divided into 3 periods.

In Subphase 1, the secured message

$$\mathsf{sec}\left(W^{(B)}_{d_1,\{2,3\}} \oplus W^{(B)}_{d_2,\{1,3\}} \oplus W^{(B)}_{d_3,\{1,2\}},\ K_{\{1,2,3\}}\right) \tag{57}$$

is sent to all three weak receivers using a standard point-to-point code. With their cache contents, each weak receiver i can decode the submessage of W_{d_i} sent in this subphase. Communication is secured when the key $K_{\{1,2,3\}}$ is sufficiently long. In Subphase 3, the non-secure message $W^{(A)}_{d_4}$ is sent to the strong receiver 4 using a standard wiretap code.

In the first period of Subphase 2, the transmitter uses the secure piggyback codebook in Figure 3 to transmit the secure message

$$\mathbf{W}^{(A)}_{\text{sec},\{1,2\}} = \mathsf{sec}\left(W^{(A)}_{d_1,\{2\}} \oplus W^{(A)}_{d_2,\{1\}},\ K_{\{1,2\}}\right) \tag{58}$$

to Receivers 1 and 2 and the non-secure message $W^{(B)}_{d_4,\{1,2\}}$ to Receiver 4. It randomly chooses a codeword in the wiretap bin indicated by $\mathbf{W}_{\text{sec},\{1,2\}}$ and $W^{(B)}_{d_4,\{1,2\}}$ and sends the chosen codeword over the channel. Weak receivers 1 and 2 have stored $W^{(B)}_{d_4,\{1,2\}}$ in their cache memories and can decode based on a restricted codebook consisting only of the bins in the column indicated by $W^{(B)}_{d_4,\{1,2\}}$. Their decoding performance is thus the same as if this message $W^{(B)}_{d_4,\{1,2\}}$ had not been sent at all. The strong receiver 4 has no cache memory and decodes both messages based on the entire codebook. Notice that the secured message $\mathbf{W}_{\text{sec},\{1,2\}}$ also acts as randomization message to secure the transmission of $W^{(B)}_{d_4,\{1,2\}}$ to Receiver 4. If this mechanism suffices to secure $W^{(B)}_{d_4,\{1,2\}}$, then no additional randomization message is needed in the satellite, i.e., the magenta bin in Figure 3 can be chosen of size 1.

Similar secure piggyback codebooks are also used during the second and third periods of Subphase 2 to send messages $\mathsf{sec}\left(W^{(A)}_{d_1,\{3\}} \oplus W^{(A)}_{d_3,\{1\}},\ K_{\{1,3\}}\right)$ and $W_{d_4,\{1,3\}}$ and messages $\mathsf{sec}\left(W^{(A)}_{d_2,\{3\}} \oplus W^{(A)}_{d_3,\{2\}},\ K_{\{2,3\}}\right)$ and $W_{d_4,\{2,3\}}$, respectively.

At the end of the delivery phase, each Receiver $k \in \mathcal{K}$ assembles all the guesses pertaining to its desired message

W_{d_k} and (in case of the weak receivers) all the parts of this message stored in its cache memory to form the final guess $\hat{W}_k$.

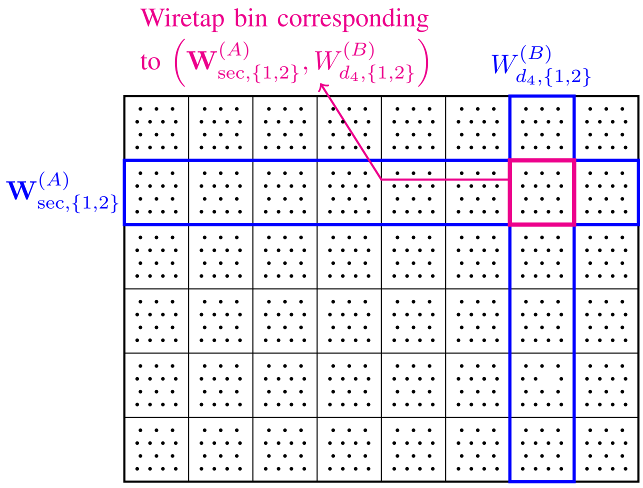


Fig. 3. Secure piggyback codebook with subcodebooks arranged in an array. Dots indicate codewords.

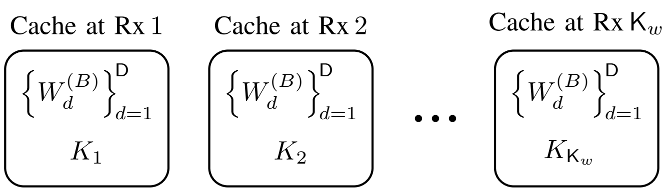
Remark 2. *The secure piggyback codebook used in Subphase 2 is inspired by the non-secure piggyback coding and Tuncel coding in [3] and [26], and by the secure coding scheme for BCs with complementary side-information in [23]. In fact, the main difference to the scheme in [23] is that here one of the receivers and the transmitter share a common secret key, which makes it possible to reduce the size of the wiretap bins or even eliminate them completely.*

Interestingly, in this construction, the secret keys $K_{\{1,2\}}, K_{\{1,3\}}, K_{\{2,3\}}$ *stored at the weak receivers can be used to "remotely secure" the transmission from the transmitter to the strong receiver 4.*

4) Secure Cache-Aided Piggyback Coding II:

This scheme is similar to the scheme in the previous section, but simpler. Divide each message W_d into 2 submessages $W_d = (W_d^{(A)}, W_d^{(B)})$ and let $K_1, \ldots, K_{\mathsf{K}_w}$ be independent secret keys.

Placement phase: Placement is as depicted in the following. In particular, each weak receiver $i \in \mathsf{K}_w$ caches the secret key K_i and all submessages $W_d^{(B)}$, for $d \in \mathcal{D}$.



Delivery phase: Transmission is in two subphases. In Subphase 1, the secure piggyback codebook is used to send the secured message tuple

$$\mathbf{W}^{(A)}_{\text{sec},w} := \left(\mathsf{sec}(W^{(A)}_{d_1}, K_1),\ \ldots,\ \mathsf{sec}(W^{(A)}_{d_{\mathsf{K}_w}}, K_{\mathsf{K}_w})\right) \quad (59)$$

to all weak receivers and the non-secure message tuple

$$\mathbf{W}_s^{(B)} := \left(W^{(B)}_{d_{\mathsf{K}_w+1}},\ \ldots,\ W^{(B)}_{d_{\mathsf{K}}}\right) \quad (60)$$

to the strong receivers. The codebook is depicted in Figure 3 where $\mathbf{W}^{(A)}_{\text{sec},\{1,2\}}$ needs to be replaced by $\mathbf{W}^{(A)}_{\text{sec},w}$ and $W^{(B)}_{d_4,\{1,2\}}$ by $\mathbf{W}_s^{(B)}$. The weak receivers can reconstruct $\mathbf{W}_s^{(B)}$ from their cache contents, and thus decode their desired message tuple $\mathbf{W}^{(A)}_{\text{sec},w}$ based on the single column of the codebook indicated by $\mathbf{W}_s^{(B)}$. From this decoded tuple and the secret key K_i stored in its cache memory, each weak receiver $i \in \mathcal{K}_w$ can then produce a guess of its desired message part $W^{(A)}_{d_i}$. The strong receivers decode both message tuples $\mathbf{W}^{(A)}_{\text{sec},w}$ and $\mathbf{W}_s^{(B)}$. Strong receiver $j \in \mathcal{K}_s$ keeps only its guess of $W^{(B)}_{d_j}$ and discards the rest.

Communication in this first subphase is secured because messages $W^{(A)}_{d_1}, \ldots, W^{(A)}_{d_{\mathsf{K}_w}}$ are perfectly secured by one-time pads and these one-time pads act as random bin indices to protect the messages $W^{(B)}_{d_{\mathsf{K}_w+1}}, \ldots, W^{(B)}_{d_{\mathsf{K}}}$ as in wiretap coding.

In Subphase 2, the message tuple

$$\mathbf{W}_s^{(A)} = \left(W^{(A)}_{d_{\mathsf{K}_w+1}},\ \ldots,\ W^{(A)}_{d_{\mathsf{K}}}\right) \quad (61)$$

is sent to all the strong receivers using a point-to-point wiretap code.

The choice of the rates and the lengths of the subphases are explained in Section VII-D, where the scheme is also analyzed.

B. Results on the Secrecy Capacity-Memory Tradeoff

Consider the following four rate-memory pairs:

- $$R^{(0)} := \frac{(\delta_z - \delta_s)(\delta_z - \delta_w)^+}{\mathsf{K}_w(\delta_z - \delta_s) + \mathsf{K}_s(\delta_z - \delta_w)^+}, \quad (62a)$$
  $$\mathsf{M}^{(0)} := 0; \quad (62b)$$
- $$R^{(1)} := \frac{(1-\delta_w)(\delta_z - \delta_s)}{\mathsf{K}_s(1-\delta_w) + \mathsf{K}_w(\delta_z - \delta_s)}, \quad (62c)$$
  $$\mathsf{M}^{(1)} := \frac{(\delta_z - \delta_s)\min\{1-\delta_z, 1-\delta_w\}}{\mathsf{K}_s(1-\delta_w) + \mathsf{K}_w(\delta_z - \delta_s)}; \quad (62d)$$
- $$R^{(2)} := \min\left\{\frac{(1-\delta_w)(1-\delta_s)}{\mathsf{K}_s(1-\delta_w) + \mathsf{K}_w(1-\delta_s)},\right.$$
  $$\left.\frac{(1-\delta_w)(\delta_z - \delta_s)}{\mathsf{K}_s(1-\delta_w) + \mathsf{K}_w(\delta_w - \delta_s)}\right\}, \quad (62e)$$
  $$\mathsf{M}^{(2)} := \min\left\{\frac{1-\delta_z}{\mathsf{K}_w}, \frac{(1-\delta_w)(\delta_z - \delta_s)}{\mathsf{K}_s(1-\delta_w) + \mathsf{K}_w(\delta_w - \delta_s)}\right\}; \quad (62f)$$
- $$R^{(\mathsf{K}_w+2)} := \frac{\delta_z - \delta_s}{\mathsf{K}_s}, \quad (62g)$$
  $$\mathsf{M}^{(\mathsf{K}_w+2)} := \frac{\mathsf{D}\cdot\mathsf{K}_w(\delta_z - \delta_s)^2}{\mathsf{K}_s^2\min\{1-\delta_z, 1-\delta_w\} + \mathsf{K}_s\mathsf{K}_w(\delta_z - \delta_s)}$$
  $$+ \frac{\mathsf{K}_s(\delta_z - \delta_s)\min\{1-\delta_z, 1-\delta_w\}}{\mathsf{K}_s^2\min\{1-\delta_z, 1-\delta_w\} + \mathsf{K}_s\mathsf{K}_w(\delta_z - \delta_s)};$$

(62h)

- $R^{(\mathsf{K}_w+3)} := \dfrac{\delta_z - \delta_s}{\mathsf{K}_s}$, (62i)

$$\mathsf{M}^{(\mathsf{K}_w+3)} := \frac{\mathsf{D}\cdot(\delta_z - \delta_s)}{\mathsf{K}_s}, \tag{62j}$$

and the K_w rate-memory pairs $\left\{(R^{(t+2)}, \mathsf{M}^{(t+2)})\colon t \in \{1,\ldots,\mathsf{K}_w - 1\}\right\}$ defined in (62k) and (62l) on top of the next page.

Theorem 2 (Lower Bound on $C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0)$)**.**

$$C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0) \geq \text{upper hull}\left\{\left(R^{(\ell)}, \mathsf{M}^{(\ell)}\right) : \ \ell \in \{0,\ldots,\mathsf{K}_w + 3\}\right\}. \tag{63}$$

Proof. It suffices to prove achievability of the $\mathsf{K}_w + 4$ rate-memory pairs $\{(R^{(\ell)}, \mathsf{M}^{(\ell)})\colon \ \ell = 0,\ldots,\mathsf{K}_w + 3\}$. Achievability of the upper convex hull follows by time/memory sharing arguments as in [1]. The pair $(R^{(0)}, \mathsf{M}^{(0)})$ is achievable by Remark 1. The pair $(R^{(1)}, \mathsf{M}^{(1)})$ is achieved by the "wiretap and cached keys" scheme described and analyzed in Sections IV-A1 and VII-A. The pair $(R^{(2)}, \mathsf{M}^{(2)})$ is achieved by the "cache-aided superposition jamming" scheme described and analyzed in Sections IV-A2 and VII-B. The pairs $(R^{(t+2)}, \mathsf{M}^{(t+2)})$, for $t \in \{1,\ldots,\mathsf{K}_w - 1\}$, are achieved by the "secure cache-aided piggyback coding I" scheme described and analyzed in Sections IV-A3 and VII-C. The pair $(R^{(\mathsf{K}_w+2)}, \mathsf{M}^{(\mathsf{K}_w+2)})$ is achieved by the "secure cache-aided piggyback coding II" scheme described and analyzed in Sections IV-A4 and VII-D. The pair $(R^{(\mathsf{K}_w+3)}, \mathsf{M}^{(\mathsf{K}_w+3)})$ is achieved by storing the entire library in the cache memory of each weak receiver and by applying a standard wiretap BC code [17] to send the requested messages to the strong receivers. □

Interestingly, upper and lower bounds in Corollary 1 and Theorem 2 match for small and large M_w irrespective of the number of weak and strong receivers K_w and K_s. In the absence of a secrecy constraint, the best upper and lower bounds for small M_w match only when $\mathsf{K}_w = 1$, irrespective of the value of K_s [3], [5].

Corollary 2. *For small cache memories* $\mathsf{M}_w \in [0, \mathsf{M}^{(1)}]$

$$C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0) = R^{(0)} + \frac{\mathsf{K}_w(\delta_z - \delta_s)}{\mathsf{K}_w(\delta_z - \delta_s) + \mathsf{K}_s(\delta_z - \delta_w)^+}\mathsf{M}_w, \tag{64}$$

where $R^{(0)}$ *is defined in* (62a) *and* $\mathsf{M}^{(1)}$ *is defined in* (62d).

Proof: Achievability follows from the two achievable rate-memory pairs $(R^{(0)}, \mathsf{M}^{(0)})$ and $(R^{(1)}, \mathsf{M}^{(1)})$ in (62a)–(62d) and by time/memory-sharing arguments. The converse follows from upper bound (50b) in Corollary 1 when specialized to $k_w = \mathsf{K}_w$. In fact, for $k_w = \mathsf{K}_w$ and cache size $\mathsf{M}_w \in [0, \mathsf{M}^{(1)}]$, the maximizing β is:

$$\beta = \frac{(\delta_z - \delta_s) - \mathsf{K}_s\mathsf{M}_w}{\mathsf{K}_w(\delta_z - \delta_s) + \mathsf{K}_s(\delta_z - \delta_w)^+}\mathsf{K}_w. \tag{65}$$

This choice of β makes the two terms in the minimization (50b) equal. ■

Notice that when $\delta_z \leq \delta_w$, then (64) specializes to

$$C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0) = \mathsf{M}_w, \qquad 0 \leq \mathsf{M}_w \leq \mathsf{M}^{(1)}. \tag{66}$$

The secrecy capacity thus grows in the same way as the cache size at weak receivers. This is achieved with the "Wiretap and Cached Keys" scheme of Subsection VII-A.

Notice that in this case the secrecy capacity-memory tradeoff $C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0)$ grows much faster in the cache size M_w than its non-secure counterpart $C(\mathsf{M}_w, \mathsf{M}_s = 0)$. In fact, by the upper bound in [3], the maximum slope of the standard capacity-memory tradeoff

$$\gamma := \max_{m \geq 0}\left\{ \left.\frac{\mathsf{d}C(\mathsf{M}_w, \mathsf{M}_s = 0)}{\mathsf{d}\mathsf{M}_w}\right|_{\mathsf{M}_w = m} \right\} \tag{67}$$

is at most

$$\gamma \leq \frac{\mathsf{K}_w\mathsf{M}_w}{\mathsf{D}}. \tag{68}$$

By the above Corollary 2 and the concavity of $C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0)$ in M_w, the maximum slope of the secrecy capacity-memory tradeoff

$$\gamma_{\text{sec}} := \max_{m \geq 0}\left\{ \left.\frac{\mathsf{d}C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0)}{\mathsf{d}\mathsf{M}_w}\right|_{\mathsf{M}_w = m} \right\} \tag{69}$$

is

$$\gamma_{\text{sec}} = \lim_{m \to 0}\left\{ \left.\frac{\mathsf{d}C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0)}{\mathsf{d}\mathsf{M}_w}\right|_{\mathsf{M}_w = m} \right\} = \frac{\mathsf{K}_w(\delta_z - \delta_s)}{\mathsf{K}_w(\delta_z - \delta_s) + \mathsf{K}_s(\delta_z - \delta_w)^+}\mathsf{M}_w. \tag{70}$$

So in contrast to the maximum slope of the standard capacity-memory tradeoff γ, the maximum slope of the secrecy capacity-memory tradeoff γ_{sec} does not deteriorate with the size of the library D. The reason for this discrepancy is that in the setup with secrecy constraint an optimal strategy for small cache memories is to exclusively place secret keys in the cache memories. In this case, each bit of the cache content is useful irrespective of the specifically demanded files. In a setup without secrecy constraint, only data is placed in the cache memories. So, at least on an intuitive level, each bit of cache memory is useful only under some of the demands.

We turn to the regime of large cache memories.

Corollary 3. *When the cache memory* M_w *is large:*

$$C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0) = \frac{\delta_z - \delta_s}{\mathsf{K}_s}, \quad \mathsf{M}_w \geq \mathsf{M}^{(\mathsf{K}_w+2)}, \tag{71}$$

where $\mathsf{M}^{(\mathsf{K}_w+2)}$ *is defined in* (62h).

The rate-memory pairs (62k)–(62h) are attained by means of joint cache-channel coding where the decoders simultaneously adapt to the cache contents and the channel statistics. To emphasize the strength of the joint coding approach, we characterize the rates that are securely achievable under a separate cache-channel coding approach.

Define the following rate-memory pair:

$$R^{(1)}_{\text{sep}} := \min\left\{ \frac{(1-\delta_w)(1-\delta_s)}{\mathsf{K}_s(1-\delta_w) + \mathsf{K}_w(1-\delta_s)},\right.$$

$$R^{(t+2)} := \frac{(t+1)(1-\delta_w)(\delta_z-\delta_s)\big[\mathsf{K}_s t(1-\delta_w)+(\mathsf{K}_w-t+1)\min\{\delta_w-\delta_s,\delta_z-\delta_s\}\big]}{(\mathsf{K}_w-t+1)(\delta_z-\delta_s)\big[\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)\min\{\delta_w-\delta_s,\delta_z-\delta_s\}\big]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2}, \tag{62k}$$

$$\begin{aligned}\mathsf{M}^{(t+2)} := {} & \frac{\mathsf{D}\cdot t(t+1)(1-\delta_w)(\delta_z-\delta_s)\big[\mathsf{K}_s(t-1)(1-\delta_w)+(\mathsf{K}_w-t+1)\min\{\delta_w-\delta_s,\delta_z-\delta_s\}\big]}{\mathsf{K}_w\big[(\mathsf{K}_w-t+1)(\delta_z-\delta_s)[\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)\min\{\delta_w-\delta_s,\delta_z-\delta_s\}]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2\big]} \\ & + \frac{(t+1)(\mathsf{K}_w-t+1)(\delta_z-\delta_s)\min\{1-\delta_z,1-\delta_w\}\big[\mathsf{K}_s t(1-\delta_w)+(\mathsf{K}_w-t)\min\{\delta_w-\delta_s,\delta_z-\delta_s\}\big]}{\mathsf{K}_w\big[(\mathsf{K}_w-t+1)(\delta_z-\delta_s)[\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)\min\{\delta_w-\delta_s,\delta_z-\delta_s\}]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2\big]}\end{aligned} \tag{62l}$$

$$\left.\frac{(1-\delta_w)(\delta_z-\delta_s)}{\mathsf{K}_s(1-\delta_w)+\mathsf{K}_w(\delta_w-\delta_s)}\right\}, \tag{72a}$$

$$\mathsf{M}_{\text{sep}}^{(1)} := \min\left\{\frac{(1-\delta_w)(1-\delta_s)}{\mathsf{K}_s(1-\delta_w)+\mathsf{K}_w(\delta_w-\delta_s)},\ \frac{(1-\delta_w)(\delta_z-\delta_s)}{\mathsf{K}_s(1-\delta_w)+\mathsf{K}_w(\delta_w-\delta_s)}\right\}, \tag{72b}$$

and for $t \in \{0,1,\ldots,\mathsf{K}_w-1\}$, define the following rate-memory pairs:

$$R_{\text{sep}}^{(t+2)} := \frac{(t+1)(1-\delta_w)(\delta_z-\delta_s)}{\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)(\delta_z-\delta_s)}, \tag{72c}$$

$$\mathsf{M}_{\text{sep}}^{(t+2)} := \frac{\mathsf{D}t+(\mathsf{K}_w-t)}{\mathsf{K}_w} \cdot \frac{(t+1)(1-\delta_w)(\delta_z-\delta_s)}{\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)(\delta_z-\delta_s)]}. \tag{72d}$$

Proposition 1. *Any rate $R > 0$ is achievable by means of separate cache-channel coding, if it satisfies*

$$R \le \text{upper hull}\left\{\left\{\left(R^{(0)},\mathsf{M}^{(0)}\right),\left(R^{(\mathsf{K}_w+3)},\mathsf{M}^{(\mathsf{K}_w+3)}\right)\right\} \bigcup \left\{\left(R_{\text{sep}}^{(j)},\mathsf{M}_{\text{sep}}^{(j)}\right)\right\}_{j=1}^{\mathsf{K}_w+1}\right\}. \tag{73}$$

Proof: The rate-memory tuples $\big\{(R^{(\ell)},\mathsf{M}^{(\ell)})\colon \ell = 0, \mathsf{K}_w+3\big\}$ in (62a)–(62b) and (62i)–(62j) correspond to the capacity-memory tradeoffs without cache memory and the capacity-memory tradeoffs when the entire library is stored in each weak receiver's cache memory, respectively. It can be verified that these schemes apply a separate cache-channel coding architecture. Rate-memory pair $\big(R_{\text{sep}}^{(1)},\mathsf{M}_{\text{sep}}^{(1)}\big)$ is achieved by the "cache-aided superposition jamming" scheme in Subsection VII-B, but with a key size equal to the message rate R. The key size here is larger than in the scheme in Subsection VII-B, because we insist on separate cache-channel coding where keys stored in cache memories cannot be combined with wiretap coding. Rate-memory pairs $\big\{\big(R_{\text{sep}}^{(t+2)},\mathsf{M}_{\text{sep}}^{(t+2)}\big)\colon t = 0,1,\ldots,\mathsf{K}_w-1\big\}$ are achieved by a scheme that communicates to weak and strong receivers in two independent phases: in the first phase the Sengupta-Clancy-Tandon secure coded caching scheme [6] is combined with a standard optimal BC code to communicate to the weak receivers, and in the second phase a standard wiretap BC code is used to communicate to the strong receivers. ∎

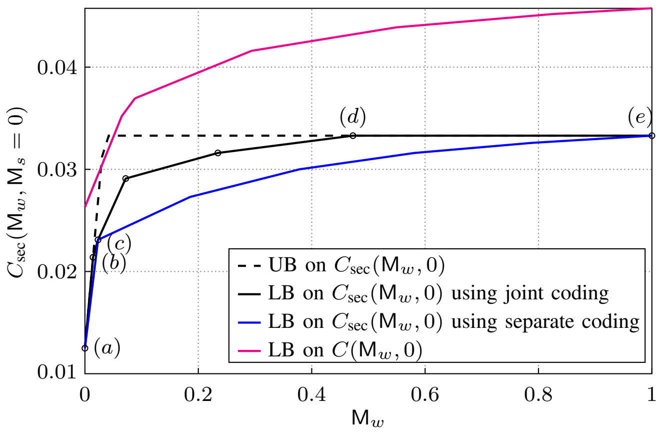


Fig. 4. Upper and lower bounds on $C_{\text{sec}}(\mathsf{M}_w,\mathsf{M}_s=0)$ for $\delta_w=0.7$, $\delta_s=0.3$, $\delta_z=0.8$, $\mathsf{D}=30$, $\mathsf{K}_w=5$, and $\mathsf{K}_s=15$. For the upper bound, the bound in (50a) dominates over (50b) only when $0.03 \le \mathsf{M}_w \le 0.04$. For other values of M_w, the bound in (50b) dominates over (50a). Notice that the eavesdropper is weaker than all legitimate receivers.

We notice from this Proposition 1 and from Theorem 2 that for certain channel parameters and cache sizes, separate cache-channel coding achieves the same performance as our joint source-channel coding schemes. For example, when each of the six minimizations in (62c)–(62f) and in (72a)–(72b) is attained by the second term (which is the case, e.g., when $\delta_z < \delta_w$), then

$$R^{(1)} = R_{\text{sep}}^{(2)} \tag{74}$$

$$\mathsf{M}^{(1)} = \mathsf{M}_{\text{sep}}^{(2)} \tag{75}$$

and

$$R^{(2)} = R_{\text{sep}}^{(1)} \tag{76}$$

$$\mathsf{M}^{(2)} = \mathsf{M}_{\text{sep}}^{(1)}, \tag{77}$$

and thus for all $\mathsf{M}_w \le \max\{\mathsf{M}^{(1)},\mathsf{M}^{(2)}\}$ separate cache-channel coding achieves the performance of our joint cache-channel coding schemes. For all other cache sizes $\mathsf{M}_w > \max\{\mathsf{M}^{(1)},\mathsf{M}^{(2)}\}$, the rates achieved by our joint cache-channel coding schemes of Theorem 2 are generally higher than the rates in Proposition 1.

C. Numerical Comparisons

In Figure 4, we compare the presented bounds at hand of an example with 5 weak and 15 strong receivers and where the eavesdropper is degraded with respect to all receivers.

The figure shows the upper and lower bounds on the secrecy capacity-memory tradeoff $C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0)$ in Corollary 1 and Theorem 2. It also shows the rates achieved by the separate cache-channel coding scheme leading to Proposition 1. Finally, the figure presents the lower bound on the standard capacity-memory tradeoff in [3].

The presented lower bound of Theorem 2 (see the black solid line in Figure 4) is piece-wise linear with the end points of the pieces corresponding to the points in (62). The left-most point (a) corresponds to the capacity in the absence of cache memories. The second and third left-most points (b) and (c) are obtained by storing only secret keys in the cache memories. The right-most point (e) corresponds to the point where the messages can be sent at the same rate as if only strong receivers were present in the system. This performance is trivially achieved by storing all messages in each of the weak receivers' cache memories and holding the delivery communication only to strong receivers. After this point, the capacity cannot be increased further because strong receivers do not have cache memories. Through coding, the same rate can also be achieved without storing the entire library at each weak receiver, see the second right-most point (d).

For small and large cache memories, our upper and lower bounds are exact. This shows that in the regime of small cache memories, it is optimal to place only secret keys in the weak receivers' cache memories. In this regime, the slope of $C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0)$ in M_w is steep (see Corollary 2 and Equation (70)) because the secret keys stored in the cache memories are always helpful, irrespective of the specific demands $\mathbf{d}$. In particular, the slope is not divided by the library size D as is the case in the traditional caching setup without secrecy constraint.

In the regime of moderate or large cache memories, the proposed placement strategies also store information about the messages in the cache memories. In this regime, the slope C_{sec} is smaller and proportional to $\frac{1}{\mathsf{D}}$, because only a fraction of the cache content is effectively helpful for a specific demand $\mathbf{d}$.

In the presented example, the proposed separate cache-channel coding performs strictly worse than joint cache-channel coding except for the single rate-memory pair (c). (To see that equality holds in (c), notice that for the present example, the second term in the $\min$ is active in all four expressions (62e), (62f), (72a), and (72b).)

Figure 5 shows the bounds for an example where the eavesdropper is stronger than the weak receivers but not the strong receivers:

$$\delta_s < \delta_z < \delta_w. \tag{78}$$

It shows that positive rates can be achieved even if $\delta_z \leq \delta_w$, because messages sent to weak receivers can be specially secured by means of one-time pads using the secret keys stored in their cache memories. In this case, our separate and joint cache-channel coding schemes achieve the same rates for small cache memories. (See also the paragraph following Proposition 1.)

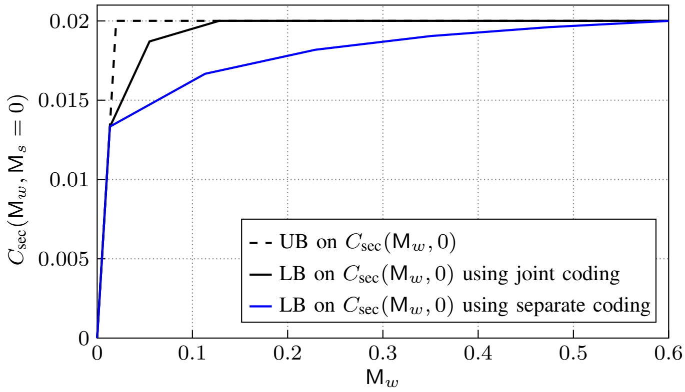


Fig. 5. Upper and lower bounds on $C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s = 0)$ for $\delta_w = 0.8$, $\delta_s = 0.3$, $\delta_z = 0.6$, $\mathsf{D} = 30$, $\mathsf{K}_w = 5$, and $\mathsf{K}_s = 15$. For the upper bound, the bound in (50b) dominates over (50a) for all values of M_w. Notice that the eavesdropper is stronger than weak receivers.

V. Coding Schemes and Results when All Receivers have Cache Memories

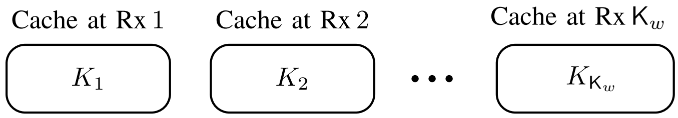
We turn to the case where all receivers have cache memories, so $\mathsf{M}_w, \mathsf{M}_s > 0$. In this case, we do not impose any constraint on the eavesdropper's channel, so δ_z can be larger or smaller than δ_s, δ_w.

A. Coding Schemes

We present four coding schemes. The first one only stores secret keys in all the cache memories, the second one stores keys in all cache memories and data at weak receivers, and the last two schemes store keys and data at all the receivers.

1) Cached Keys:

Placement phase: Store independent secret keys $K_1, \ldots, K_{\mathsf{K}}$ in the cache memories of Receivers $1, \ldots, \mathsf{K}$:



Delivery phase: Apply a standard (non-secure) broadcast code to send the secured message tuple

$$\mathbf{W}_{\text{sec}} := \big(\mathsf{sec}(W_{d_1}, K_1),\ \mathsf{sec}(W_{d_2}, K_2),\ \mathsf{sec}(W_{d_3}, K_3),\ \ldots,\ \mathsf{sec}(W_{d_{\mathsf{K}_w}}, K_{\mathsf{K}_w})\big), \tag{79}$$

to Receivers $1, \ldots, \mathsf{K}$, respectively. With the secret key K_k stored in its cache memory, each Receiver $k \in \mathsf{K}$ recovers a guess of its desired message W_{d_k}. See Section VIII-A on how to choose the parameters of the scheme.

2) Secure Cache-Aided Piggyback Coding with Keys at All Receivers:

The difference between this scheme and the secure cache-aided piggyback coding scheme of Sections IV-A3 and VII-C is that additional secret keys are placed in the cache memories of weak and strong receivers so that communications in Subphases 2 and 3 can entirely be secured with these keys, i.e., no wiretap binning is required.

We outline the scheme for $\mathsf{K}_w = 3$ weak receivers and $\mathsf{K}_s = 1$ strong receiver. For a detailed description and an analysis in the general case, see Section VIII-B.
Divide for each $d \in \mathcal{D}$ the message W_d into six submessages

$$W_d = \left(W^{(A)}_{d,\{1\}}, W^{(A)}_{d,\{2\}}, W^{(A)}_{d,\{3\}}, W^{(B)}_{d,\{1,2\}}, W^{(B)}_{d,\{1,3\}}, W^{(B)}_{d,\{2,3\}}\right), \tag{80}$$

where the first three are of equal rate and the latter three are of equal rate.

Placement phase: Placement is as described in the following table, where $K_{\{1,2,3\}}$, $K_{\{1,2\}}$, $K_{\{2,3\}}$, $K_{\{1,3\}}$, $K_{4,\{1,2\}}$, $K_{4,\{2,3\}}$, $K_{4,\{1,3\}}$, $K_{\{4\}}$ denote independent secret keys.

Cache at Rx 1

$$\left\{W^{(A)}_{d,\{1\}}\right\}_{d=1}^{\mathsf{D}}$$
$$\left\{W^{(B)}_{d,\{1,2\}}, W^{(B)}_{d,\{1,3\}}\right\}_{d=1}^{\mathsf{D}}$$
$$K_{\{1,2,3\}}$$
$$K_{\{1,2\}}, K_{\{1,3\}}$$
$$K_{4,\{1,2\}}, K_{4,\{1,3\}}$$

Cache at Rx 2

$$\left\{W^{(A)}_{d,\{2\}}\right\}_{d=1}^{\mathsf{D}}$$
$$\left\{W^{(B)}_{d,\{1,2\}}, W^{(B)}_{d,\{2,3\}}\right\}_{d=1}^{\mathsf{D}}$$
$$K_{\{1,2,3\}}$$
$$K_{\{1,2\}}, K_{\{2,3\}}$$
$$K_{4,\{1,2\}}, K_{4,\{2,3\}}$$

Cache at Rx 3

$$\left\{W^{(A)}_{d,\{3\}}\right\}_{d=1}^{\mathsf{D}}$$
$$\left\{W^{(B)}_{d,\{1,3\}}, W^{(B)}_{d,\{2,3\}}\right\}_{d=1}^{\mathsf{D}}$$
$$K_{\{1,2,3\}}$$
$$K_{\{1,3\}}, K_{\{2,3\}}$$
$$K_{4,\{1,3\}}, K_{4,\{2,3\}}$$

Cache at Rx 4

$$K_4$$
$$K_{4,\{1,2\}}$$
$$K_{4,\{1,3\}}$$
$$K_{4,\{2,3\}}$$

Delivery phase: Time-sharing is applied over three subphases and Subphase 2 is further divided into 3 periods. Transmission in Subphase 1 is as described in Section IV-A3. In Subphase 3, the secured message $\mathsf{sec}(W^{(A)}_{d_4}, K_4)$ is sent to the strong receiver 4 using a standard point-to-point code. This transmission is secure if the key K_4 is chosen sufficiently long.

In the first period of Subphase 2, the transmitter uses the standard piggyback codebook in Figure 6 to transmit the secure message

$$\mathbf{W}^{(A)}_{\text{sec},\{1,2\}} = \mathsf{sec}\left(W^{(A)}_{d_1,\{2\}} \oplus W^{(A)}_{d_2,\{1\}},\ K_{\{1,2\}}\right) \tag{81}$$

to Receivers 1 and 2 and the secure message

$$W^{(B)}_{\text{sec},\{4\}} = \mathsf{sec}\left(W^{(B)}_{d_4,\{1,2\}}, K_{4,\{1,2\}}\right) \tag{82}$$

to Receiver 4. Receivers 1 and 2 can reconstruct $W^{(B)}_{\text{sec},\{4\}}$ from their cache contents, and thus decode message $\mathbf{W}^{(A)}_{\text{sec},\{1,2\}}$ based solely on the column of the codebook that corresponds to $W^{(B)}_{\text{sec},\{4\}}$. Receiver 4 decodes both messages $\mathbf{W}^{(A)}_{\text{sec},\{1,2\}}$ and $W^{(B)}_{\text{sec},\{4\}}$. From the decoded secured messages and the keys stored in their cache memories, Receivers $1, 2$, and 4 can recover their desired message parts $W^{(A)}_{d_1,\{2\}}, W^{(A)}_{d_2,\{1\}}$ and $W^{(B)}_{d_4,\{1,2\}}$.

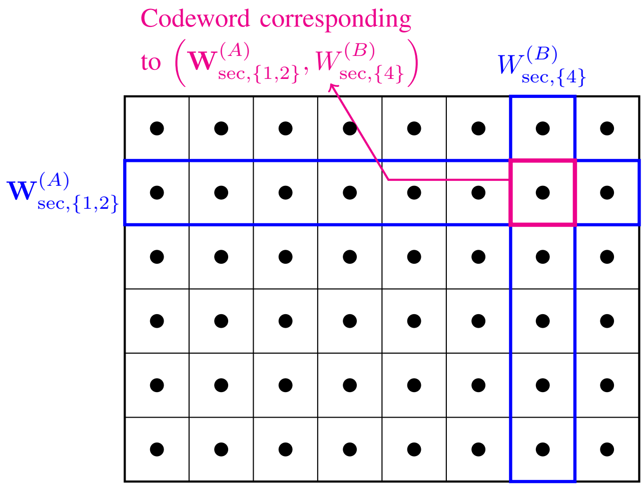


Fig. 6. Standard piggyback codebook where only a single codeword (indicated by a single dot) is assigned to each pair of messages.

In the same way, using a standard piggyback codebook, the secured messages $\mathsf{sec}\left(W^{(A)}_{d_1,\{3\}} \oplus W^{(A)}_{d_3,\{1\}},\ K_{\{1,3\}}\right)$ and $\mathsf{sec}\left(W^{(B)}_{d_4,\{1,3\}}, K_{4,\{1,3\}}\right)$ are transmitted in Period 2 to Receivers $1, 3$, and 4, and the secured messages $\mathsf{sec}\left(W^{(A)}_{d_2,\{3\}} \oplus W^{(A)}_{d_3,\{2\}},\ K_{\{2,3\}}\right)$ and $\mathsf{sec}\left(W^{(B)}_{d_4,\{2,3\}}, K_{4,\{2,3\}}\right)$ are transmitted in Period 3 to Receivers $2, 3$, and 4.

3) Symmetric Secure Piggyback Coding:
Each message is split into two submessages $W_d = (W^{(A)}_d, W^{(B)}_d)$, and communication is in three subphases. Submessages of $\{W^{(A)}_d\}$ and corresponding secret keys are placed in weak receivers' cache memories according to the Sengupta et al. secure coded caching placement algorithm [6]. Submessages of $\{W^{(B)}_d\}$ and corresponding secret keys are placed in strong receivers' cache memories according to the same placement algorithm. In Subphase 1, the Sengupta et al. delivery scheme for submessages $\{W^{(A)}_d\}$ is combined with a standard BC code to transmit only to weak receivers and in Subphase 3 it is combined with a standard BC code to transmit only to strong receivers. Transmission in Subphase 2 is divided into into $\mathsf{K}_w\mathsf{K}_s$ periods, each dedicated to a pair of weak and strong receivers $i \in \mathsf{K}_w$ and $j \in \mathsf{K}_s$. A standard piggyback codebook is used in each of these periods to send secured messages to the corresponding pair of receivers. The secret keys securing these messages have been pre-placed in the appropriate cache memories.

We now describe the scheme in more detail for the special case $\mathsf{K}_w = 3$ and $\mathsf{K}_s = 2$, and for parameters $t_w = 2$ and $t_s = 1$. The general scheme is described and analyzed in Section VIII-C.

Divide each W_d into six submessages

$$W_d = \left(W^{(A)}_{d,\{1,2\}}, W^{(A)}_{d,\{1,3\}}, W^{(A)}_{d,\{2,3\}}, W^{(B)}_{d,\{4\}}, W^{(B)}_{d,\{5\}}\right),\ d \in \mathcal{D}, \tag{83}$$

where the first three are of equal rate and the latter two are of equal rate.

Placement phase: Generate independent secret keys $K_{\{1,2,3\}}$, $K_{w,\{1,4\}}$, $K_{w,\{1,5\}}$, $K_{w,\{2,4\}}$, $K_{w,\{2,5\}}$, $K_{w,\{3,4\}}$, $K_{w,\{3,5\}}$, $K_{s,\{1,4\}}$, $K_{s,\{1,5\}}$, $K_{s,\{2,4\}}$, $K_{s,\{2,5\}}$, $K_{s,\{3,4\}}$, $K_{s,\{3,5\}}$, $K_{\{4,5\}}$.

Placement of information in the cache memories is as indicated in the following table.

Cache at Rx 1

$$\left\{W^{(A)}_{d,\{1,2\}}, W^{(A)}_{d,\{1,3\}}\right\}_{d=1}^{\mathsf{D}}$$
$$K_{\{1,2,3\}}$$
$$K_{w,\{1,4\}}, K_{w,\{1,5\}}$$
$$K_{s,\{1,4\}}, K_{s,\{1,5\}}$$

Cache at Rx 4

$$\left\{W^{(B)}_{d,\{4\}}\right\}_{d=1}^{\mathsf{D}}$$
$$K_{w,\{1,4\}}, K_{w,\{2,4\}}, K_{w,\{3,4\}}$$
$$K_{s,\{1,4\}}, K_{s,\{2,4\}}, K_{s,\{3,4\}}$$
$$K_{\{4,5\}}$$

Cache at Rx 2

$$\left\{W^{(A)}_{d,\{1,2\}}, W^{(A)}_{d,\{2,3\}}\right\}_{d=1}^{\mathsf{D}}$$
$$K_{\{1,2,3\}}$$
$$K_{w,\{2,4\}}, K_{w,\{2,5\}}$$
$$K_{s,\{2,4\}}, K_{s,\{2,5\}}$$

Cache at Rx 5

$$\left\{W^{(B)}_{d,\{5\}}\right\}_{d=1}^{\mathsf{D}}$$
$$K_{w,\{1,5\}}, K_{w,\{2,5\}}, K_{w,\{3,5\}}$$
$$K_{s,\{1,5\}}, K_{s,\{2,5\}}, K_{s,\{3,5\}}$$
$$K_{\{4,5\}}$$

Cache at Rx 3

$$\left\{W^{(A)}_{d,\{1,3\}}, W^{(A)}_{d,\{2,3\}}\right\}_{d=1}^{\mathsf{D}}$$
$$K_{\{1,2,3\}}$$
$$K_{w,\{3,4\}}, K_{w,\{3,5\}}$$
$$K_{s,\{3,4\}}, K_{s,\{3,5\}}$$

Delivery phase: The delivery phase is divided into three subphases, where Subphase 2 is further divided into 6 periods.

Subphase 1 is intended only for weak receivers. The transmitter sends the secured message

$$W^{(A)}_{\text{sec},\{1,2,3\}} = \mathsf{sec}\big(W^{(A)}_{d_1,\{2,3\}} \oplus W^{(A)}_{d_2,\{1,3\}} \oplus W^{(A)}_{d_3,\{1,2\}},\ K_{\{1,2,3\}}\big) \quad (84)$$

using a capacity-achieving code to the weak receivers $1, 2$, and 3. With their cache contents, each of these weak receivers can decode its desired submessage.

The first period of Subphase 2 is dedicated to Receivers 1 and 4. The transmitter sends the secured messages

$$W^{(B)}_{\text{sec},1,4} = \mathsf{sec}\big(W^{(B)}_{d_1,\{4\}}, K_{w,\{1,4\}}\big) \quad (85)$$

and

$$W^{(A)}_{\text{sec},4,1} = \mathsf{sec}\big(W^{(A)}_{d_4,\{1,2\}}, K_{s,\{1,4\}}\big) \quad (86)$$

using a standard piggyback codebook where rows encode $W^{(B)}_{\text{sec},1,4}$ and columns encode $W^{(A)}_{\text{sec},4,1}$. (This corresponds to the piggyback codebook in Figure 6 where $\mathbf{W}^{(A)}_{\text{sec},\{1,2\}}$ needs to be replaced by $W^{(B)}_{\text{sec},1,4}$ and $W^{(B)}_{\text{sec},\{4\}}$by $W^{(A)}_{\text{sec},4,1}$.) Since Receiver 1 can reconstruct $W^{(A)}_{\text{sec},4,1}$ from its cache content, it decodes $W^{(B)}_{\text{sec},1,4}$ based on the *single column* of the piggyback codebook indicated by $W^{(A)}_{\text{sec},4,1}$. Similarly, since Receiver 4 can reconstruct $W^{(B)}_{\text{sec},1,4}$ from its cache content, it decodes $W^{(A)}_{\text{sec},4,1}$ based on the *single row* corresponding to $W^{(B)}_{\text{sec},1,4}$. With their decoded secure messages and the secret keys in their cache memories, Receiver 1 can recover $W^{(B)}_{d_1,\{4\}}$ and Receiver 4 can recover $W^{(A)}_{d_4,\{1,2\}}$.

A similar scheme is used in the subsequent periods to convey the parts $W^{(B)}_{d_1}, W^{(B)}_{d_2}$, $W^{(B)}_{d_3}$, $W^{(A)}_{d_4}$, $W^{(A)}_{d_5}$ that are not stored in their cache memories to Receivers 1–5. Table I shows the concerned receivers, the conveyed messages and the keys used in each period of Subphase 2. (Notice that there are different choices on how to fill in the last row. For example, a cyclic shift (to the right) between the first three elements and a cyclic shift (to the right) between the last three elements, is possible as well. The performance of the scheme would be unchanged. The important aspect is that for each column, the element indicated in the last row is stored in the cache memory of the weak receiver indicated in the first row.)

In Subphase 3, the transmitter sends

$$W^{(B)}_{\text{sec},\{4,5\}} = \mathsf{sec}\big(W^{(B)}_{d_4,\{5\}} \oplus W^{(B)}_{d_5,\{4\}},\ K_{\{4,5\}}\big) \quad (87)$$

using a capacity-achieving code to the strong receivers 4 and 5. With their cache contents, each receiver can decode its desired submessage sent in this subphase.

Choosing all keys sufficiently long ensures that the delivery communication satisfies the secrecy constraint (24).

4) Secure Generalized Coded Caching:

The scheme is based on the *generalized coded caching* algorithms of [25], but where the produced zero-padded XORs are secured with independent secret keys and these keys are placed in the cache memories of the receivers that decode the XORs. Choosing the secret keys sufficient long, ensures than the secrecy constraint (24) is satisfied.

B. *Results on the Secrecy Capacity-Memory Tradeoff*

Consider the following $\mathsf{K}+\mathsf{K}_w+\mathsf{K}_w\mathsf{K}_s$ rate-memory tuples. Let $\tilde{R}^{(0)} = R^{(0)}$ and $\tilde{\mathsf{M}}_w^{(0)} = \tilde{\mathsf{M}}_s^{(0)} = 0$.

- Let

$$\tilde{R}^{(1)} := \frac{(1-\delta_s)(1-\delta_w)}{\mathsf{K}_w(1-\delta_s)+\mathsf{K}_s(1-\delta_w)}, \quad (88a)$$

$$\tilde{\mathsf{M}}_w^{(1)} := \frac{(1-\delta_s)\min\left\{1-\delta_z, 1-\delta_w\right\}}{\mathsf{K}_w(1-\delta_s)+\mathsf{K}_s(1-\delta_w)}, \quad (88b)$$

$$\tilde{\mathsf{M}}_s^{(1)} := \frac{(1-\delta_w)\min\left\{1-\delta_z, 1-\delta_s\right\}}{\mathsf{K}_w(1-\delta_s)+\mathsf{K}_s(1-\delta_w)}; \quad (88c)$$

- For $t \in \{1, \ldots, \mathsf{K}_w - 1\}$, let $\tilde{R}^{(t+1)}$, $\tilde{\mathsf{M}}_w^{(t+1)}$ and $\tilde{\mathsf{M}}_s^{(t+1)}$ defined in (88d), (88e) and (88f) on top of the next page.
- For each pair $t_w \in \{1, \ldots, \mathsf{K}_w\}$ and $t_s \in \{1, \ldots, \mathsf{K}_s\}$, let $\tilde{R}^{(\mathsf{K}_w+(t_w-1)\mathsf{K}_s+t_s)}$, $\tilde{\mathsf{M}}_w^{(\mathsf{K}_w+(t_w-1)\mathsf{K}_s+t_s)}$ and $\tilde{\mathsf{M}}_s^{(\mathsf{K}_w+(t_w-1)\mathsf{K}_s+t_s)}$ be defined as in (88g), (88h) and (88i) on top of the next page.
- For $t \in \{1, \ldots, \mathsf{K} - 1\}$, let $\tilde{R}^{(\mathsf{K}_w+\mathsf{K}_w\mathsf{K}_s+t)}$, $\tilde{\mathsf{M}}_w^{(\mathsf{K}_w+\mathsf{K}_w\mathsf{K}_s+t)}$ and $\tilde{\mathsf{M}}_s^{(\mathsf{K}_w+\mathsf{K}_w\mathsf{K}_s+t)}$ be defined as in (88j), (88k) and (88l) on top of the next page.

TABLE I
MESSAGES SENT AND KEYS USED IN THE SIX PERIODS OF SUBPHASE 2 FOR THE EXAMPLE WITH $\mathsf{K}_w = 3$ WEAK RECEIVERS AND $\mathsf{K}_s = 2$ STRONG RECEIVERS.

	Subphase 2					
	Period 1	**Period 2**	**Period 3**	**Period 4**	**Period 5**	**Period 6**
Receivers	$1,4$	$2,4$	$3,4$	$1,5$	$2,5$	$3,5$
Keys	$K_{w,\{1,4\}}, K_{s,\{1,4\}}$	$K_{w,\{2,4\}}, K_{s,\{2,4\}}$	$K_{w,\{3,4\}}, K_{s,\{3,4\}}$	$K_{w,\{1,5\}}, K_{s,\{1,5\}}$	$K_{w,\{2,5\}}, K_{s,\{2,5\}}$	$K_{w,\{3,5\}}, K_{s,\{3,5\}}$
Messages for weak receivers	$W^{(B)}_{d_1,\{4\}}$	$W^{(B)}_{d_2,\{4\}}$	$W^{(B)}_{d_3,\{4\}}$	$W^{(B)}_{d_1,\{5\}}$	$W^{(B)}_{d_2,\{5\}}$	$W^{(B)}_{d_3,\{5\}}$
Messages for strong receivers	$W^{(A)}_{d_4,\{1,2\}}$	$W^{(A)}_{d_4,\{2,3\}}$	$W^{(A)}_{d_4,\{1,3\}}$	$W^{(A)}_{d_5,\{1,2\}}$	$W^{(A)}_{d_5,\{2,3\}}$	$W^{(A)}_{d_5,\{1,3\}}$

Theorem 3 (Lower bound on $C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s)$)**.**

$$C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s) \geq \text{upper hull}\Big\{ \left(\tilde{R}^{(\ell)}, \tilde{\mathsf{M}}_w^{(\ell)}, \tilde{\mathsf{M}}_s^{(\ell)}\right) : \ell \in \{0, 1, \ldots, \mathsf{K} + \mathsf{K}_w + \mathsf{K}_w\mathsf{K}_s - 1\}\Big\}. \quad (89)$$

Proof. By time/memory-sharing arguments, it suffices to prove the achievability of the rate-memory triples $\{(\tilde{R}^{(\ell)}, \tilde{\mathsf{M}}_w^{(\ell)}, \tilde{\mathsf{M}}_s^{(\ell)}) : \ell \in \{0, 1, \ldots, \mathsf{K} + \mathsf{K}_w + \mathsf{K}_w\mathsf{K}_s - 1\}\}$. The triple $\{(\tilde{R}^{(1)}, \tilde{\mathsf{M}}_w^{(1)}, \tilde{\mathsf{M}}_s^{(1)})$ is achieved by the "cached keys" scheme, see Subsections V-A1 and VIII-A. The triples $(\tilde{R}^{(\ell)}, \tilde{\mathsf{M}}_w^{(\ell)}, \tilde{\mathsf{M}}_s^{(\ell)})$, for $\ell \in \{2, \ldots, \mathsf{K}_w\}\}$, are achieved by the "secure piggyback coding scheme with keys at all receivers", see Subsections V-A2 and VIII-B. The triples $(\tilde{R}^{(\ell)}, \tilde{\mathsf{M}}_w^{(\ell)}, \tilde{\mathsf{M}}_s^{(\ell)})$, for $\ell \in \{\mathsf{K}_w + 1, \ldots, \mathsf{K}_w + \mathsf{K}_w\mathsf{K}_s\}\}$, are achieved by the "symmetric secure piggyback coding" scheme, see Subsections V-A3 and VIII-C. Finally, the triples $(\tilde{R}^{(\ell)}, \tilde{\mathsf{M}}_w^{(\ell)}, \tilde{\mathsf{M}}_s^{(\ell)})$, for $\ell \in \{\mathsf{K}_w + \mathsf{K}_w\mathsf{K}_s + 1, \ldots, \mathsf{K} + \mathsf{K}_w + \mathsf{K}_w\mathsf{K}_s - 1\}$, are achieved by the "secure generalized coded caching" scheme sketched in Section V-A4. □

Corollary 4. *The rate-memory tradeoff* $\left(\tilde{R}^{(1)}, \tilde{\mathsf{M}}_w^{(1)}, \tilde{\mathsf{M}}_s^{(1)}\right)$ *is optimal, i.e.,*

$$C_{\text{sec}}\left(\mathsf{M}_w = \tilde{\mathsf{M}}_w^{(1)}, \mathsf{M}_s = \tilde{\mathsf{M}}_s^{(1)}\right) = \tilde{R}^{(1)}. \quad (90)$$

Proof: Achievability follows from the two achievable rate-memory triples $(\tilde{R}^{(0)}, \tilde{\mathsf{M}}_w^{(0)} = 0, \tilde{\mathsf{M}}_s^{(0)} = 0)$ and $(\tilde{R}^{(1)}, \tilde{\mathsf{M}}_w^{(1)}, \tilde{\mathsf{M}}_s^{(1)})$ in (88a)–(88c) and by time/memory-sharing arguments. The converse follows by specializing upper bound (42a) in Theorem 1 to $k_w = \mathsf{K}_w$ and $k_s = \mathsf{K}_s$. In fact, for $k_w = \mathsf{K}_w$, $k_s = \mathsf{K}_s$, and cache sizes $\mathsf{M}_w = \tilde{\mathsf{M}}_w^{(1)}$ and $\mathsf{M}_s = \tilde{\mathsf{M}}_s^{(1)}$, the maximizing β equals

$$\beta = \frac{\mathsf{K}_w(1-\delta_s)}{\mathsf{K}_w(1-\delta_s) + \mathsf{K}_s(1-\delta_w)} \quad (91)$$

when $\delta_z < \delta_w$, and it equals $\beta = 0$ when $\delta_z \geq \delta_w$. ■

Notice that when $\delta_z \leq \delta_s$, then

$$\tilde{R}^{(1)} = \tilde{\mathsf{M}}_w^{(1)} = \tilde{\mathsf{M}}_s^{(1)}. \quad (92)$$

This rate is achieved by XORing the messages with secret keys stored in the cache memories, and by sending the resulting bits using a traditional non-secure code for the erasure BC.

VI. GLOBAL SECRECY CAPACITY-MEMORY TRADEOFF

In the preceding sections, we considered scenarios with unequal cache sizes at the receivers and showed that in these scenarios joint cache-channel coding schemes can significantly improve over the traditional separation-based schemes with their typical uniform cache assignment. In this section, we emphasize the importance of unequal cache sizes that depend on the receivers' channel conditions by focusing on the *global secrecy capacity-memory tradeoff* $C_{\text{sec,glob}}$, which is the largest secrecy capacity-memory tradeoff that is possible given a total cache budget

$$\mathsf{K}_w\mathsf{M}_w + \mathsf{K}_s\mathsf{M}_s \leq \mathsf{M}_{\text{tot}}. \quad (93)$$

We assign the same cache memory size M_w to all weak receivers and the same cache memory size M_s to all strong receivers. Using simple time/memory-sharing arguments, it can be shown that this assumption is without loss in optimality.

So, the main quantity of interest in this section is

$$C_{\text{sec,glob}}(\mathsf{M}_{\text{tot}}) := \max_{\substack{\mathsf{M}_w, \mathsf{M}_s \geq 0: \\ \mathsf{K}_w\mathsf{M}_w + \mathsf{K}_s\mathsf{M}_s \leq \mathsf{M}_{\text{tot}}}} C_{\text{sec}}(\mathsf{M}_w, \mathsf{M}_s). \quad (94)$$

A. Results

Using the achievability results in Theorems 2 and 3 combined with an appropriate cache assignment and time/memory-sharing arguments, yields the following lower bound on $C_{\text{sec,glob}}(\mathsf{M}_{\text{tot}})$.

Corollary 5. *(Lower bound on* $C_{\text{sec,glob}}(\mathsf{M}_{\text{tot}})$*) The global secrecy capacity-memory tradeoff* $C_{\text{sec,glob}}$ *is lower bounded as*

$$C_{\text{sec,glob}}(\mathsf{M}_{\text{tot}}) \geq \text{upper hull}\Big\{\Big\{\left(R^{(\ell)}, \mathsf{M}_{\text{tot}} = \mathsf{K}_w\mathsf{M}_s^{(\ell)}\right)\Big\}_{\ell=0}^{\mathsf{K}_w+3}, \bigcup \Big\{\left(\tilde{R}^{(\ell')}, \mathsf{M}_{\text{tot}}^{\ell'}\right)\Big\}_{\ell'=1}^{\mathsf{K}+\mathsf{K}_w+\mathsf{K}_w\mathsf{K}_s-1}\Big\}, \quad (95)$$

$$\tilde{R}^{(t+1)} := \frac{(t+1)(1-\delta_w)(1-\delta_s)\big[\mathsf{K}_s t(1-\delta_w)+(\mathsf{K}_w-t+1)(\delta_w-\delta_s)\big]}{(\mathsf{K}_w-t+1)(1-\delta_s)\big[\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)(\delta_w-\delta_s)\big]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2}, \tag{88d}$$

$$\begin{aligned}\tilde{\mathsf{M}}_w^{(t+1)} :=\;& \frac{\mathsf{D}\cdot t(t+1)(1-\delta_w)(1-\delta_s)\big[\mathsf{K}_s(t-1)(1-\delta_w)+(\mathsf{K}_w-t+1)(\delta_w-\delta_s)\big]}{\mathsf{K}_w\big[(\mathsf{K}_w-t+1)(1-\delta_s)[\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)(\delta_w-\delta_s)]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2\big]}\\ &+\frac{\mathsf{K}_s t(t+1)(\mathsf{K}_w-t+1)(1-\delta_w)(1-\delta_s)\min\{1-\delta_z,1-\delta_s\}}{\mathsf{K}_w\big[(\mathsf{K}_w-t+1)(1-\delta_s)[\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)(\delta_w-\delta_s)]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2\big]}\\ &+\frac{(t+1)(\mathsf{K}_w-t)(\mathsf{K}_w-t+1)(1-\delta_s)(\delta_w-\delta_s)\min\{1-\delta_z,1-\delta_w\}\,]}{\mathsf{K}_w\big[(\mathsf{K}_w-t+1)(1-\delta_s)[\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)(\delta_w-\delta_s)]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2\big]},\end{aligned} \tag{88e}$$

$$\begin{aligned}\tilde{\mathsf{M}}_s^{(t+1)} :=\;& \frac{\mathsf{K}_s t(t+1)(1-\delta_w)^2\min\{1-\delta_z,1-\delta_s\}}{(\mathsf{K}_w-t+1)(1-\delta_s)\big[\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)(\delta_w-\delta_s)\big]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2}\\ &+\frac{(t+1)(\mathsf{K}_w-t+1)(1-\delta_w)(1-\delta_s)\min\{(\delta_w-\delta_z)^+,\delta_w-\delta_s\}}{(\mathsf{K}_w-t+1)(1-\delta_s)\big[\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)(\delta_w-\delta_s)\big]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2};\end{aligned} \tag{88f}$$

$$\tilde{R}^{(\mathsf{K}_w+(t_w-1)\mathsf{K}_s+t_s)} := \frac{(t_w+1)(t_s+1)(1-\delta_w)(1-\delta_s)\big[\mathsf{K}_s(1-\delta_w)+\mathsf{K}_w(1-\delta_s)\big]}{\mathsf{K}_w(\mathsf{K}_w-t_w)(t_s+1)(1-\delta_s)^2+\mathsf{K}_s(t_w+1)(1-\delta_w)\big[(\mathsf{K}_s-t_s)(1-\delta_w)+\mathsf{K}_w(t_s+1)(1-\delta_s)\big]}, \tag{88g}$$

$$\begin{aligned}\tilde{\mathsf{M}}_w^{(\mathsf{K}_w+(t_w-1)\mathsf{K}_s+t_s)} :=\;& \frac{(t_w+1)(t_s+1)(1-\delta_s)^2\big[\mathsf{D}\cdot t_w(1-\delta_w)+(\mathsf{K}_w-t_w)\min\{1-\delta_z,1-\delta_w\}\big]}{\mathsf{K}_w(\mathsf{K}_w-t_w)(t_s+1)(1-\delta_s)^2+\mathsf{K}_s(t_w+1)(1-\delta_w)\big[(\mathsf{K}_s-t_s)(1-\delta_w)+\mathsf{K}_w(t_s+1)(1-\delta_s)\big]}\\ &+\frac{\mathsf{K}_s(t_w+1)(t_s+1)(1-\delta_w)(1-\delta_s)\min\{1-\delta_z,2-\delta_w-\delta_s\}}{\mathsf{K}_w(\mathsf{K}_w-t_w)(t_s+1)(1-\delta_s)^2+\mathsf{K}_s(t_w+1)(1-\delta_w)\big[(\mathsf{K}_s-t_s)(1-\delta_w)+\mathsf{K}_w(t_s+1)(1-\delta_s)\big]},\end{aligned} \tag{88h}$$

$$\begin{aligned}\tilde{\mathsf{M}}_s^{(\mathsf{K}_w+(t_w-1)\mathsf{K}_s+t_s)} :=\;& \frac{(t_w+1)(t_s+1)(1-\delta_w)^2\big[\mathsf{D}\cdot t_s(1-\delta_s)+(\mathsf{K}_s-t_s)\min\{1-\delta_z,1-\delta_s\}\big]}{\mathsf{K}_w(\mathsf{K}_w-t_w)(t_s+1)(1-\delta_s)^2+\mathsf{K}_s(t_w+1)(1-\delta_w)\big[(\mathsf{K}_s-t_s)(1-\delta_w)+\mathsf{K}_w(t_s+1)(1-\delta_s)\big]}\\ &+\frac{\mathsf{K}_w(t_w+1)(t_s+1)(1-\delta_w)(1-\delta_s)\min\{1-\delta_z,2-\delta_w-\delta_s\}}{\mathsf{K}_w(\mathsf{K}_w-t_w)(t_s+1)(1-\delta_s)^2+\mathsf{K}_s(t_w+1)(1-\delta_w)\big[(\mathsf{K}_s-t_s)(1-\delta_w)+\mathsf{K}_w(t_s+1)(1-\delta_s)\big]};\end{aligned} \tag{88i}$$

$$\tilde{R}^{(\mathsf{K}_w+\mathsf{K}_w\mathsf{K}_s+t)} := \frac{\sum_{t_w=\max\{0,t-\mathsf{K}_s\}}^{\min\{t,\mathsf{K}_w\}}\binom{\mathsf{K}_w}{t_w}\binom{\mathsf{K}_s}{t-t_w}(1-\delta_w)^{-t_w}(1-\delta_s)^{t_w}}{\sum_{t_w=\max\{0,t+1-\mathsf{K}_s\}}^{\min\{t+1,\mathsf{K}_w\}}\binom{\mathsf{K}_w}{t_w}\binom{\mathsf{K}_s}{t+1-t_w}(1-\delta_w)^{-t_w}(1-\delta_s)^{t_w-1}}, \tag{88j}$$

$$\begin{aligned}\tilde{\mathsf{M}}_w^{(\mathsf{K}_w+\mathsf{K}_w\mathsf{K}_s+t)} :=\;& \frac{\mathsf{D}\cdot\sum_{t_w=\max\{1,t-\mathsf{K}_s\}}^{\min\{t,\mathsf{K}_w\}}\binom{\mathsf{K}_w-1}{t_w-1}\binom{\mathsf{K}_s}{t-t_w}(1-\delta_w)^{-t_w}(1-\delta_s)^{t_w}}{\sum_{t_w=\max\{0,t+1-\mathsf{K}_s\}}^{\min\{t+1,\mathsf{K}_w\}}\binom{\mathsf{K}_w}{t_w}\binom{\mathsf{K}_s}{t+1-t_w}(1-\delta_w)^{-t_w}(1-\delta_s)^{t_w-1}}\\ &+\min\left\{\frac{(t+1)(1-\delta_z)}{\mathsf{K}},\frac{(1-\delta_w)^{-t_w}(1-\delta_s)^{t_w}\left[\binom{\mathsf{K}-1}{t}(1-\delta_s)-\binom{\mathsf{K}_w-1}{t}(\delta_w-\delta_s)\right]}{\sum_{t_w=\max\{0,t+1-\mathsf{K}_s\}}^{\min\{t+1,\mathsf{K}_w\}}\binom{\mathsf{K}_w}{t_w}\binom{\mathsf{K}_s}{t+1-t_w}(1-\delta_w)^{-t_w}(1-\delta_s)^{t_w}}\right\},\end{aligned} \tag{88k}$$

$$\begin{aligned}\tilde{\mathsf{M}}_s^{(\mathsf{K}_w+\mathsf{K}_w\mathsf{K}_s+t)} :=\;& \frac{\mathsf{D}\cdot\sum_{t_w=\max\{0,t-\mathsf{K}_s\}}^{\min\{t-1,\mathsf{K}_w\}}\binom{\mathsf{K}_w}{t_w}\binom{\mathsf{K}_s-1}{t-t_w-1}(1-\delta_w)^{-t_w}(1-\delta_s)^{t_w}}{\sum_{t_w=\max\{0,t+1-\mathsf{K}_s\}}^{\min\{t+1,\mathsf{K}_w\}}\binom{\mathsf{K}_w}{t_w}\binom{\mathsf{K}_s}{t+1-t_w}(1-\delta_w)^{-t_w}(1-\delta_s)^{t_w-1}}\\ &+\min\left\{\frac{(t+1)(1-\delta_z)}{\mathsf{K}},\frac{\binom{\mathsf{K}-1}{t}(1-\delta_w)^{-t_w}(1-\delta_s)^{t_w+1}}{\sum_{t_w=\max\{0,t+1-\mathsf{K}_s\}}^{\min\{t+1,\mathsf{K}_w\}}\binom{\mathsf{K}_w}{t_w}\binom{\mathsf{K}_s}{t+1-t_w}(1-\delta_w)^{-t_w}(1-\delta_s)^{t_w}}\right\}.\end{aligned} \tag{88l}$$

---

where $\mathsf{M}-\mathrm{tot}^{\ell'} := \mathsf{K}_w\tilde{\mathsf{M}}_w^{(\ell')}+\mathsf{K}_s\tilde{\mathsf{M}}_s^{(\ell')}$.

Proposition 2. *If the eavesdropper is weaker than the strong receivers, i.e.,*

$$\delta_z > \delta_s, \tag{96}$$

then for small cache sizes $\mathsf{M}_{\text{tot}} \in [0, \mathsf{K}_w\mathsf{M}^{(1)}]$*:*

$$C_{\text{sec,glob}}(\mathsf{M}_{\text{tot}}) = R^{(0)} + \frac{(\delta_z-\delta_s)^+}{\mathsf{K}_w(\delta_z-\delta_s)^+ + \mathsf{K}_s(\delta_z-\delta_w)^+}\mathsf{M}_{\text{tot}}. \tag{97}$$

If the eavesdropper is at least as strong as the strong receivers, i.e.,

$$\delta_z \le \delta_s, \tag{98}$$

then for small cache sizes $\mathsf{M}_{\text{tot}} \in \Big[0, \mathsf{K}\cdot\frac{(1-\delta_w)(1-\delta_s)}{\mathsf{K}_w(1-\delta_s)+\mathsf{K}_s(1-\delta_w)}\Big]$*:*

$$C_{\text{sec,glob}}(\mathsf{M}_{\text{tot}}) = \frac{\mathsf{M}_{\text{tot}}}{\mathsf{K}}. \tag{99}$$

Proof: Achievability of (97) follows from Corollary 2 and by assigning all available cache memory uniformly across weak receivers, so $\mathsf{M}_w = \frac{\mathsf{M}_{\text{tot}}}{\mathsf{K}_w}$ and $\mathsf{M}_s = 0$. The converse to (97) can be proved by specializing Lemma 2 to the set of all users $\mathcal{S} = \mathcal{K}$. For a given choice of the cache assignment $\mathsf{M}_w, \mathsf{M}_s \geq 0$, Lemma 2 yields (amongst others) that for some $\beta \in [0,1]$, the secrecy capacity-memory tradeoff is upper bounded as

$$\mathsf{K}_w C_{\text{sec,glob}}(\mathsf{M}_{\text{tot}}) \leq \beta(\delta_z - \delta_w)^+ + \mathsf{K}_w \mathsf{M}_w \tag{100}$$

and

$$\mathsf{K} C_{\text{sec,glob}}(\mathsf{M}_{\text{tot}}) \leq \beta(\delta_z - \delta_w)^+ + (1-\beta)(\delta_z - \delta_s)^+ + \mathsf{M}_{\text{tot}}. \tag{101}$$

Upper bounding now $\mathsf{K}_w \mathsf{M}_w$ by M_{tot} and combining (100) and (101) into a single bound yields:

$$\begin{aligned} &C_{\text{sec,glob}}(\mathsf{M}_{\text{tot}}) \\ &\leq \max_{\beta \in [0,1]} \min \Bigg\{ \frac{\beta(\delta_z - \delta_w)^+ + \mathsf{M}_{\text{tot}}}{\mathsf{K}_w}, \\ &\qquad \frac{\beta(\delta_z - \delta_w)^+ + (1-\beta)(\delta_z - \delta_s)^+}{\mathsf{K}} + \frac{\mathsf{M}_{\text{tot}}}{\mathsf{K}} \Bigg\}. \end{aligned} \tag{102}$$

If $\delta_z > \delta_w$, then the maximizing β in the above upper bound is as in (65) when M_w is replaced by M_{tot}; otherwise the maximizing β equals 0. Plugging these values into (102) establishes the desired upper bound.

Achievability of (99) follows from time/memory-sharing arguments and from the achievability of the rate-memory triple $(\tilde{R}^{(1)}, \tilde{M}_w^{(1)}, \tilde{M}_s^{(1)})$ in (88a)–(88c), which under (98) specializes to the rate-memory triple:

$$\tilde{R}^{(1)} = \tilde{\mathsf{M}}_w^{(1)} = \tilde{\mathsf{M}}_s^{(1)} = \frac{(1-\delta_w)(1-\delta_s)}{\mathsf{K}_w(1-\delta_s) + \mathsf{K}_s(1-\delta_w)}. \tag{103}$$

In fact, the rate-memory triples under consideration are achieved by storing an independent secret key at each receiver and securing the messages with these keys by means of one-time pads. This requires a uniform cache assignment across *all* receivers, i.e., $\mathsf{M}_w = \mathsf{M}_s = \frac{\mathsf{M}_{\text{tot}}}{\mathsf{K}}$. The converse to (99) follows from upper bound (101), which under (98) specializes to $C_{\text{sec,glob}}(\mathsf{M}_{\text{tot}}) \leq \mathsf{M}_{\text{tot}}/\mathsf{K}$. ■

Figure 7 illustrates our choices of the cache contents when $\delta_z > \delta_s$ in the order of increasing total cache memory M_{tot}. When the total cache budget M_{tot} is small, all of it is assigned to the weak receivers and it is solely used to store secret keys. For a slightly larger cache budget M_{tot}, it is assigned to all receivers and secret keys are stored in the cache memories. When the total cache budget M_{tot} exceeds the size required to store the keys for securing the entire communication, the additional space is used to store data at weak receivers. Once the cached data renders the weak receivers equally powerful (in terms of their decoding performance) as the strong receivers, data is also stored in strong receivers' cache memories. When $\delta_z \leq \delta_s$, then our choice is similar, but we start with placing secret keys directly at all the receivers.

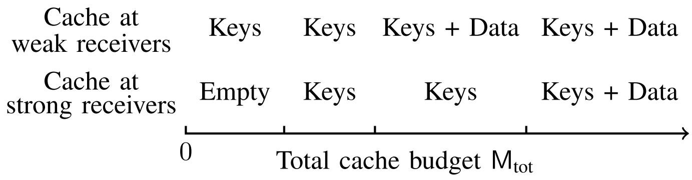


Fig. 7. Cache content at weak and strong receivers in order of increasing total cache budget.

B. Results under Uniform Cache Assignment

For comparison, we also propose a lower bound on $C_{\text{sec,glob}}(\mathsf{M}_{\text{tot}})$ when the cache memory is uniformly assigned over *all receivers*, i.e.

$$\mathsf{M}_w = \mathsf{M}_s = \frac{\mathsf{M}_{\text{tot}}}{\mathsf{K}}. \tag{104}$$

Consider the following $\mathsf{K}+2$ rate-memory pairs:

- $$R_{\text{sym}}^{(0)} := \frac{(\delta_z - \delta_s)^+ \cdot (\delta_z - \delta_w)^+}{\mathsf{K}_w(\delta_z - \delta_s) + \mathsf{K}_s(\delta_z - \delta_w)}, \tag{105a}$$
$$\mathsf{M}_{\text{sym}}^{(0)} := 0; \tag{105b}$$

- $$R_{\text{sym}}^{(1)} := \frac{(1-\delta_w)(1-\delta_s)}{\mathsf{K}_w(1-\delta_s) + \mathsf{K}_s(1-\delta_w)}, \tag{105c}$$
$$\mathsf{M}_{\text{sym}}^{(1)} := \frac{(1-\delta_s)\min\{1-\delta_z, 1-\delta_w\}}{\mathsf{K}_w(1-\delta_s) + \mathsf{K}_s(1-\delta_w)}; \tag{105d}$$

- For $t \in \{1, \ldots, \mathsf{K}_s - 1\}$
$$R_{\text{sym}}^{(t+1)} := \frac{\binom{\mathsf{K}}{t}(1-\delta_w)(1-\delta_s)}{\binom{\mathsf{K}}{t+1}(1-\delta_s) - \binom{\mathsf{K}_s}{t+1}(\delta_w - \delta_s)}, \tag{105e}$$
$$\begin{aligned} \mathsf{M}_{\text{sym}}^{(t+1)} := &\frac{\mathsf{D} \cdot t\binom{\mathsf{K}}{t}(1-\delta_w)(1-\delta_s)}{\mathsf{K}\left[\binom{\mathsf{K}}{t+1}(1-\delta_s) - \binom{\mathsf{K}_s}{t+1}(\delta_w - \delta_s)\right]} \\ &+ \frac{(\mathsf{K}-t)\binom{\mathsf{K}}{t}(1-\delta_s)\min\{1-\delta_z, 1-\delta_w\}}{\mathsf{K}\left[\binom{\mathsf{K}}{t+1}(1-\delta_s) - \binom{\mathsf{K}_s}{t+1}(\delta_w - \delta_s)\right]}; \end{aligned} \tag{105f}$$

- For $t \in \{\mathsf{K}_s, \ldots, \mathsf{K}\}$
$$R_{\text{sym}}^{(t+1)} := \frac{(t+1)(1-\delta_w)}{(\mathsf{K}-t)}, \tag{105g}$$
$$\begin{aligned} \mathsf{M}_{\text{sym}}^{(t+1)} := &\frac{\mathsf{D} \cdot t(t+1)(1-\delta_w)}{\mathsf{K}(\mathsf{K}-t)} \\ &+ \frac{(t+1)}{\mathsf{K}}\min\{1-\delta_z, 1-\delta_w\}. \end{aligned} \tag{105h}$$

Proposition 3 (Lower Bound with Symmetric Caches)**.**

$$\begin{aligned} &C_{\text{sec}}(\mathsf{M}_w = \mathsf{M}_{\text{sym}}, \mathsf{M}_s = \mathsf{M}_{\text{sym}}) \geq \\ &\quad \text{upper hull}\left\{\left(R_{\text{sym}}^{(\ell)}, \mathsf{M}_{\text{sym}}^{(\ell)}\right) : \ \ell \in \{0, \ldots, \mathsf{K}+1\}\right\}. \end{aligned} \tag{106}$$

Proof. It suffices to prove the achievability of the $\mathsf{K}+2$ rate-memory pairs $\{(R_{\text{sym}}^{(\ell)}, \mathsf{M}_{\text{sym}}^{(\ell)}) : \ \ell = 0, \ldots, \mathsf{K}+1\}$. The rate-memory pair $(R_{\text{sym}}^{(0)}, \mathsf{M}_{\text{sym}}^{(0)})$ is achievable by Remark 1. The rate-memory pair $(R_{\text{sym}}^{(1)}, \mathsf{M}_{\text{sym}}^{(1)})$ is achievable because,

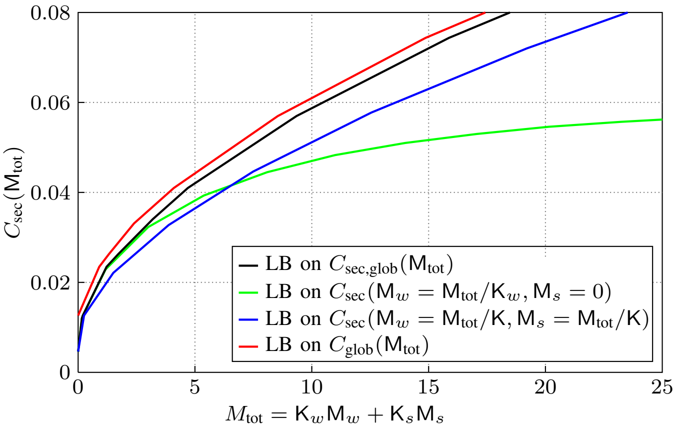


Fig. 8. Lower bounds on $C_{\text{sec,glob}}(\mathsf{M}_{\text{tot}})$ for $\delta_w = 0.7$, $\delta_s = 0.2$, $\delta_z = 0.8$, $D = 50$, $\mathsf{K}_w = 20$, and $\mathsf{K}_s = 10$. Notice that the eavesdropper is weaker than all users.

by Theorem 3, rate $\tilde{R}^{(1)} = R^{(1)}_{\text{sym}}$ is achievable with cache size $\tilde{\mathsf{M}}^{(1)}_w$ at weak receivers and cache size $\tilde{\mathsf{M}}^{(1)}_s$ at strong receivers, and because $\mathsf{M}^{(1)}_{\text{sym}} = \max\left\{\tilde{\mathsf{M}}^{(1)}_w, \tilde{\mathsf{M}}^{(1)}_s\right\}$. For each $t \in \{1, \ldots, \mathsf{K}\}$, the rate-memory pair $(R^{(t+1)}_{\text{sym}}, \mathsf{M}^{(t+1)}_{\text{sym}})$ is achieved by a scheme that combines the Sengupta et al. secure coded caching scheme [6] with a standard BC code. Notice that when $t \geq \mathsf{K}_s$, then each of the secured XORs produced by the Sengupta et al. scheme is intended for at least one weak receiver, and communication is limited by this weak receiver. Each secured XOR, which is of rate $\binom{\mathsf{K}}{t}^{-1} R_{\text{sym}}$, thus requires slightly more than $\binom{\mathsf{K}}{t}^{-1}\frac{nR_{\text{sym}}}{1-\delta_w}$ channel uses to be transmitted reliably. When $t < \mathsf{K}_s$, then $\binom{\mathsf{K}_s}{t+1}$ of the secured XOR are only sent to strong receivers and can thus be sent reliably by using slightly more than $\binom{\mathsf{K}}{t}^{-1}\frac{nR_{\text{sym}}}{1-\delta_s}$ channel uses. The remaining $\binom{\mathsf{K}}{t+1} - \binom{\mathsf{K}_s}{t+1}$ secured XORs are sent to at least one weak receiver, and can thus be sent reliably by using slightly more than $\binom{\mathsf{K}}{t}^{-1}\frac{nR_{\text{sym}}}{1-\delta_w}$ channel uses. □

C. Numerical Comparison

Figure 8 plots the lower bound on $C_{\text{sec,glob}}(\mathsf{M}_{\text{tot}})$ in Corollary 5 (black line) for an example with $\mathsf{K}_w = 20$ weak receivers an $\mathsf{K}_s = 10$ strong receivers. The eavesdropper is degraded with respect to all legitimate receivers. For comparison, the figure also shows our lower bound on the secrecy capacity-memory tradeoff for the scenarios where the available cache memory is uniformly assigned over all weak receivers , i.e., $\mathsf{M}_w = \mathsf{M}_{\text{tot}}/\mathsf{K}_w$ and $\mathsf{M}_s = 0$, and where it is uniformly assigned over *all receivers* (blue line), i.e., $\mathsf{M}_w = \mathsf{M}_s = \mathsf{M}_{\text{tot}}/\mathsf{K}$. Finally, the red line depicts the lower bound on the standard (non-secure) global capacity-memory tradeoff obtained from [25] and [5]. One observes that our lower bound on the global secrecy capacity-memory tradeoff is close to the currently best known lower bound on the non-secure capacity-memory tradeoff. Moreover, similarly to [25], for the global secrecy capacity-memory tradeoff it is suboptimal to assign the cache memories uniformly across users (unless the eavesdropper is stronger than all other receivers). In particular, as Proposition 3 shows, for small cache memories all of it should be assigned uniformly over the weak receivers only.

VII. Coding Schemes when Only Weak Receivers have Cache Memories

This section describes in more detail the coding schemes proposed for the setup when only weak receivers have cache memories. The schemes are also briefly analyzed.

A. Wiretap and Cached Keys

Let Subphase 1 comprise the first βn channel uses and Subphase 2 the last $(1-\beta)n$ channel uses, with β chosen as

$$\beta = \frac{\mathsf{K}_w(\delta_z - \delta_s)}{\mathsf{K}_w(\delta_z - \delta_s) + \mathsf{K}_s(1-\delta_w)}. \tag{107}$$

Fix a small $\epsilon > 0$. Let the secret keys $K_1, \ldots, K_{\mathsf{K}_w}$ be independent of each other and of rate

$$R_{\text{Key}} = \min\left\{\frac{\beta(1-\delta_z)}{\mathsf{K}_w}, R\right\}, \tag{108}$$

where the message rate R is chosen as

$$R = \frac{(\delta_z - \delta_s)(1-\delta_w)}{\mathsf{K}_w(\delta_z - \delta_s) + \mathsf{K}_s(1-\delta_w)} - \epsilon. \tag{109}$$

The scheme requires a cache size at weak receivers equal to

$$\begin{aligned}\mathsf{M}_w = R_{\text{Key}} = \min\Bigg\{&\frac{(\delta_z - \delta_s)(1-\delta_z)}{\mathsf{K}_w(\delta_z - \delta_s) + \mathsf{K}_s(1-\delta_w)}, \\ &\frac{(\delta_z - \delta_s)(1-\delta_w)}{\mathsf{K}_w(\delta_z - \delta_s) + \mathsf{K}_s(1-\delta_w)} - \epsilon\Bigg\}.\end{aligned} \tag{110}$$

We analyze the scheme when averaged over the choice of the random code construction $\mathcal{C}$. By the joint typicality lemma, the average probability of a decoding error at the weak receivers tends to 0 as $n \to \infty$, because

$$R < \frac{\beta(1-\delta_w)}{\mathsf{K}_w}. \tag{111}$$

The probability of error of the wiretap decoding at the strong receivers tends to 0 as $n \to \infty$, because

$$R < \frac{(1-\beta)(\delta_z - \delta_s)}{\mathsf{K}_s}. \tag{112}$$

Notice that by the choice of β in (107), the two constraints (111) and (112) coincide.

We conclude this analysis by verifying the secrecy constraint averaged over the choice of the codebooks $\mathcal{C}$. Since communication in the two phases is independent:

$$\begin{aligned}&I(W_1, \ldots, W_{\mathsf{D}}; Z^n|\mathcal{C}) \\ &= I(W_1, \ldots, W_{\mathsf{D}}; Z_1^{\beta n}|\mathcal{C}) + I(W_1, \ldots, W_{\mathsf{D}}; Z^n_{\beta n+1}|\mathcal{C}) \\ &= I(W_{d_1}, \ldots, W_{d_{\mathsf{K}_w}}; Z_1^{\beta n}|\mathcal{C}) \\ &\quad + I(W_{d_{\mathsf{K}_w+1}}, \ldots, W_{d_{\mathsf{K}}}; Z^n_{\beta n+1}|\mathcal{C}).\end{aligned} \tag{113}$$

The first term in (113) tends to 0 as $n \to \infty$, by Lemma 1 and because the key rate (108) is chosen so that the transmitted codeword is either uniformly distributed over all codewords of a random codebook of rate equal to the message rate or over a subset of codewords that is of rate equal to the eavesdropper's capacity $(1-\delta_z)$. The second term in (113) tends to 0 as $n \to \infty$ because of the properties of a wiretap BC code.

From the analysis we conclude that the average over all choices of the codebooks satisfies the desired properties. There must thus exist at least one good codebook. Letting $\epsilon \to 0$ concludes the achievability proof of the rate-memory pair $R = R^{(1)}$ and $\mathsf{M}_w = \mathsf{M}^{(1)}$ in (62c) and (62d).

B. Cache-Aided Superposition-Jamming

Let

$$\mu := \min\left\{ \frac{\mathsf{K}_w(1-\delta_s)}{\mathsf{K}_s(1-\delta_w)+\mathsf{K}_w(1-\delta_s)}, \frac{\mathsf{K}_w(\delta_z-\delta_s)}{\mathsf{K}_s(1-\delta_w)+\mathsf{K}_w(\delta_w-\delta_s)} \right\}, \tag{114}$$

and set the rate

$$R = \frac{\mu(1-\delta_w)}{\mathsf{K}_w} - \epsilon. \tag{115}$$

Let further $p \in [0, 1/2]$ be so that its binary entropy-function $h_b(p) = 1-\mu$. Choose a key rate

$$R_{\text{Key}} = \min\left\{ \frac{1-\delta_z}{\mathsf{K}_w}, R \right\} \tag{116}$$

and a wiretap binning rate of

$$R_{\text{bin}} = \{(1-\delta_z) - \mu(1-\delta_w)\}^+. \tag{117}$$

In what follows, we will construct a superposition code where:

- The cloud center codewords encode the K_w messages to the weak receivers and their entries are drawn i.i.d. according to $P_U \sim \text{Ber}(1/2)$.
- The satellite codewords encode the K_s messages to the strong receivers together with an additional randomization message of rate R_{bin}, and their entries are drawn conditionally independent according to the conditional distribution $P_{X|U} \sim \text{Ber}(p)$.

Notice that the parameter μ is so that

$$R + \epsilon = \frac{\mu(1-\delta_w)}{\mathsf{K}_w} = \frac{(1-\mu)(1-\delta_s) - R_{\text{bin}}}{\mathsf{K}_s}. \tag{118}$$

Defining $(U, X, Y_1, Y_{\mathsf{K}_w}) \sim P_U P_{X|U} P_{Y_1 Y_{\mathsf{K}_w}|X}$, above equalities are equivalently expressed as

$$R + \epsilon = \frac{I(U; Y_1)}{\mathsf{K}_w} = \frac{I(X; Y_{\mathsf{K}_w+1}|U) - R_{\text{bin}}}{\mathsf{K}_s}, \tag{119}$$

and thus show that our choice of μ makes the decoding constraints at the weak and strong receivers equally stringent.

Placement phase: Generate independent random keys $K_1, \dots, K_w$ of rate R_{Key}. Store each key K_i in the cache memory of Receiver i, for $i \in \mathsf{K}_w$.

Delivery phase: Fix a small $\epsilon > 0$, and generate a cloud center codebook

$$\mathcal{C}_{\text{center}} = \left\{ u^n(1), \dots, u^n(\lfloor 2^{nR} \rfloor) \right\}^{\mathsf{K}_w}, \tag{120}$$

by picking all entries of all codewords i.i.d. according to a Bernoulli-$1/2$ distribution. Then, superposition on each codeword $u^n(w)$, for $w \in \{1, \dots, \lfloor 2^{nR} \rfloor\}^{\mathsf{K}_w}$, a satellite codebook

$$\mathcal{C}_{\text{sat}}(w) = \Big\{ x^n(v, b|w) \colon v \in \{1, \dots, \lfloor 2^{nR} \rfloor\}^{\mathsf{K}_w}, b \in \{1, \dots \lfloor 2^{nR_{\text{bin}}} \rfloor\} \Big\}, \tag{121}$$

by picking the t-th entry of each codeword $x^n(v, b|w)$ equal to the t-th entry of codeword $u^n(w)$ with probability $1-p$ and equal to its binary inverse with probability p.

The transmitter generates the tuples

$$\mathbf{W}_{\text{sec}} := \big(\mathsf{sec}(W_{d_1}, K_1),\ \mathsf{sec}(W_{d_2}, K_2), \mathsf{sec}(W_{d_3}, K_3),\ \dots,\ \mathsf{sec}(W_{d_{\mathsf{K}_w}}, K_{\mathsf{K}_w}) \big), \tag{122}$$

and

$$\mathbf{W}_{\text{sat}} := \big(W_{d_{\mathsf{K}_w+1}}, \dots, W_{d_\mathsf{K}} \big), \tag{123}$$

and draws B uniformly at random over $\{1, \dots, \lfloor 2^{nR_{\text{bin}}} \rfloor\}$. It then sends the codeword

$$x^n\big(\mathbf{W}_{\text{sat}}, B \big| \mathbf{W}_{\text{sec}}\big) \tag{124}$$

over the channel. Each weak receiver $i \in \mathcal{K}_w$ decodes the message $\mathbf{W}_{\text{sec}}$ in the cloud center, and with the key K_i retrieved from its cache memory, it produces a guess of its desired message W_{d_i}. Each strong receiver $j \in \mathcal{K}_s$ decodes both messages $\mathbf{W}_{\text{sec}}, \mathbf{W}_{\text{sat}}$ as well as the randomization message B and extracts the guess of its desired message W_{d_j}.

Analysis: The scheme requires cache memories at the weak receivers of size

$$\mathsf{M}_w = R_{\text{Key}} = \min\left\{ \frac{1-\delta_z}{\mathsf{K}_w}, \frac{(\delta_z-\delta_s)(1-\delta_w)}{\mathsf{K}_s(1-\delta_w)+\mathsf{K}_w(\delta_w-\delta_s)} - \epsilon \right\}. \tag{125}$$

We analyze the scheme on average over the random choice of the codebook $\mathcal{C}$. The average probability of decoding error at a weak receiver tends to 0 as $n \to \infty$, because by (115):

$$\mathsf{K}_w R < \mu(1-\delta_w). \tag{126a}$$

Similarly, the average probability of decoding error at a strong receiver tends to 0 as $n \to \infty$, because

$$R_{\text{bin}} + \mathsf{K}_s R < (1-\mu)(1-\delta_s). \tag{126b}$$

We turn to the secrecy analysis of the scheme. We have:

$$\begin{aligned}
&I(W_1, \dots, W_\mathsf{D}; Z^n|\mathcal{C}) \\
&\quad = I(W_{d_1}, \dots, W_{d_\mathsf{K}}; Z^n|\mathcal{C}) && (127)\\
&\quad = I(W_{d_1}, \dots, W_{d_{\mathsf{K}_w}}; Z^n|\mathcal{C}) \\
&\qquad + I(W_{d_{\mathsf{K}_w+1}}, \dots, W_{d_\mathsf{K}}; Z^n|W_{d_1}, \dots, W_{d_{\mathsf{K}_w}}, \mathcal{C}) && (128)\\
&\quad \le I(W_{d_1}, \dots, W_{d_{\mathsf{K}_w}}; Z^n, W_{d_{\mathsf{K}_w+1}}, \dots, W_{d_\mathsf{K}}, B|\mathcal{C}) \\
&\qquad + I(W_{d_{\mathsf{K}_w+1}}, \dots, W_{d_\mathsf{K}}; Z^n|W_{d_1}, \dots, W_{d_{\mathsf{K}_w}}, \mathcal{C}) && (129)\\
&\quad = I(W_{d_1}, \dots, W_{d_{\mathsf{K}_w}}; Z^n|W_{d_{\mathsf{K}_w+1}}, \dots, W_{d_\mathsf{K}}, B, \mathcal{C}) \\
&\qquad + I(W_{d_{\mathsf{K}_w+1}}, \dots, W_{d_\mathsf{K}}; Z^n|W_{d_1}, \dots, W_{d_{\mathsf{K}_w}}, \mathcal{C}), && (130)
\end{aligned}$$

where the last equality holds by the independence of the messages $W_1, \dots, W_{d_\mathsf{K}}$, the randomization message B, and the codebook $\mathcal{C}$.

Both terms on the right-hand side of (130) tend to 0 as $n \to \infty$, by Lemma 1 and the choice of the key-rate R_{Key} in (116) and the binning rate R_{bin} in (117). To see more specifically that the second term vanishes asymptotically, notice that for any fixed message tuple $W_{d_1} = w_1, W_{d_2} = w_2, \ldots, W_{d_{\mathsf{K}_w}} = w_{\mathsf{K}_w}$, the following two statements apply.

1) The entries of the set (codebook) of satellite codewords $\mathcal{C}_{\text{sat}}(w_1, \ldots, w_{\mathsf{K}_w})$ defined in (131) on top of the next page are i.i.d. Bernoulli-1/2.
2) The codeword used to encode a given message tuple $W_{d_{\mathsf{K}_w+1}} = w_{\mathsf{K}_w+1}, \ldots, W_{d_\mathsf{K}} = w_\mathsf{K}$ is chosen uniformly at random (depending on the realizations of the secret keys $K_1, \ldots, K_{\mathsf{K}_w}$ and the randomization message B) over the subcodebook $\mathcal{S}(w_{\mathsf{K}_w+1}, \ldots, w_\mathsf{K})$ (defined on top of the next page), is of rate
$$\begin{aligned}\frac{1}{n}\log_2\left(\mathcal{S}(w_{\mathsf{K}_w+1}, \ldots, w_\mathsf{K})\right) &\geq R_{\text{bin}} + \mathsf{K}_w R_{\text{Key}} \\ &= (1-\delta_z) \\ &= I(X;Z), \end{aligned} \tag{133}$$
for $X \sim$ Bernoulli-$1/2$.

The sets $\mathcal{S}(w_{\mathsf{K}_w+1}, \ldots, w_\mathsf{K})$ thus all satisfy Condition (34) of Lemma 1, from which follows that for each tuple $W_{d_1} = w_1, W_{d_2} = w_2, \ldots, W_{d_{\mathsf{K}_w}} = w_{\mathsf{K}_w}$, the mutual information
$$\begin{aligned}I(W_{d_{\mathsf{K}_w+1}}, \ldots, W_{d_\mathsf{K}}; Z^n | W_{d_1} = w_1, \ldots, W_{d_{\mathsf{K}_w}} = w_{\mathsf{K}_w}, \mathcal{C}) \\ \to 0 \qquad \text{as} \qquad n \to \infty. \end{aligned} \tag{134}$$
Averaging this last statement over all realizations of the messages $(W_1, \ldots, W_{\mathsf{K}_w})$ proves that the second term in (130) vanishes as $n \to \infty$.

In a similar way, for each realization of the tuple $(W_{d_{\mathsf{K}_w+1}} = w_{\mathsf{K}_w+1}, \ldots, W_{d_\mathsf{K}} = w_\mathsf{K}, B = b)$ the codeword transmitted to convey the tuple $(W_1 = w_1, \ldots, W_{d_{\mathsf{K}_w}} = w_{\mathsf{K}_w})$ is chosen uniformly at random (depending on the secret keys $K_1, \ldots, K_{\mathsf{K}_w}$) over a subset $\mathcal{S}(w_1, \ldots, w_{\mathsf{K}_w})$ of an i.i.d. Bernoulli-1/2 codebook. If $R_{\text{Key}} = R$, then this subset $\mathcal{S}(w_1, \ldots, w_{\mathsf{K}_w})$ is identical for all realizations $(W_1 = w_1, \ldots, W_{d_{\mathsf{K}_w}} = w_{\mathsf{K}_w})$ and thus satisfies Condition (33) of Lemma 1. Otherwise, $\mathsf{K}_w R_{\text{Key}} = (1-\delta_z)$ and the subset $\mathcal{S}(w_1, \ldots, w_{\mathsf{K}_w})$ is of rate equal to $I(X;Z) = (1-\delta_z)$. In this case it satisfies Condition (34) of Lemma 1. We conclude that in any case by Lemma 1, the mutual information
$$\begin{aligned}I(W_{d_1}, \ldots, W_{d_{\mathsf{K}_w}}; Z^n | W_{d_{\mathsf{K}_w+1}} = w_{\mathsf{K}_w+1}, \ldots, W_{d_\mathsf{K}} = w_\mathsf{K}, B, \mathcal{C}) \\ \to 0 \qquad \text{as} \qquad n \to \infty, \end{aligned} \tag{135}$$
and hence also the first term in (130) vanishes as $n \to \infty$.

By standard arguments, one can then conclude that there must exist at least one deterministic codebook such that the probability of decoding error vanishes for $n \to \infty$ and the secrecy constraint (24) constraint holds.

Since $R \to R^{(2)}$ and $\mathsf{M}_w \to \mathsf{M}^{(2)}$ as $\epsilon \to 0$, this establishes achievability of the rate-memory pair $(R^{(2)}, \mathsf{M}_w^{(2)})$ in (62e) and (62f).

C. Secure Cache-Aided Piggyback Coding I

Delivery transmission will be partitioned into three subphases, whose lengths are determined by the parameters $\beta_1, \beta_2, \beta_3$ given on top of the next page. Notice that $\beta_1 + \beta_2 + \beta_3 = 1$.

Message splitting: Fix a small $\epsilon > 0$. Divide each message into two independent submessages
$$W_d = \left[W_d^{(A)}, W_d^{(B)}\right], \quad d \in \mathcal{D}, \tag{136}$$
that are of rates $R^{(A)}$ and $R^{(B)}$ defined in (138) and (139) on top of the next page.

Denote the $\binom{\mathsf{K}_w}{t-1}$ subsets of $\{1, \ldots, \mathsf{K}_w\}$ of size $t-1$ by $G_1^{(t-1)}, \ldots, G_{\binom{\mathsf{K}_w}{t-1}}^{(t-1)}$; the $\binom{\mathsf{K}_w}{t}$ subsets of $\{1, \ldots, \mathsf{K}_w\}$ of size t by $G_1^{(t)}, \ldots, G_{\binom{\mathsf{K}_w}{t}}^{(t)}$; and the $\binom{\mathsf{K}_w}{t+1}$ subsets of $\{1, \ldots, \mathsf{K}_w\}$ of size $t+1$ by $G_1^{(t+1)}, \ldots, G_{\binom{\mathsf{K}_w}{t+1}}^{(t+1)}$.

Divide each message $W_d^{(A)}$ into $\binom{\mathsf{K}_w}{t-1}$ submessages
$$W_d^{(A)} = \left\{W_{d, G_\ell^{(t-1)}}^{(A)} : \quad \ell \in \left\{1, \ldots, \binom{\mathsf{K}_w}{t-1}\right\}\right\}, \tag{140}$$
of rate
$$r^{(A)} = R^{(A)} \binom{\mathsf{K}_w}{t-1}^{-1}, \tag{141}$$
and divide each message $W_d^{(B)}$ into $\binom{\mathsf{K}_w}{t}$ submessages
$$W_d^{(B)} = \left\{W_{d, G_\ell^{(t)}}^{(B)} : \quad \ell \in \left\{1, \ldots, \binom{\mathsf{K}_w}{t}\right\}\right\}, \tag{142}$$
of rate
$$r^{(B)} = R^{(B)} \binom{\mathsf{K}_w}{t}^{-1}. \tag{143}$$

Key generation:

- For each $\ell \in \left\{1, \ldots, \binom{\mathsf{K}_w}{t+1}\right\}$, generate an independent random key $K_{G_\ell^{(t+1)}}$ of rate
$$R_{\text{Key},1} = \binom{\mathsf{K}_w}{t+1}^{-1} \cdot \beta_1 \cdot \min\left\{1-\delta_z, 1-\delta_w\right\}. \tag{144}$$
- For each $\ell \in \left\{1, \ldots, \binom{\mathsf{K}_w}{t}\right\}$, generate an independent random key $K_{G_\ell^{(t)}}$ of rate
$$R_{\text{Key},2} = \binom{\mathsf{K}_w}{t}^{-1} \cdot \beta_2 \cdot \min\left\{1-\delta_z, 1-\delta_w\right\}. \tag{145}$$

Define the binning rate
$$R_{\text{bin}} = \binom{\mathsf{K}_w}{t}^{-1} \cdot \beta_2 \cdot \min\left\{\max\left\{0, \delta_w - \delta_z\right\}, \delta_w - \delta_s\right\}. \tag{146}$$

Placement phase: Placement is as indicated in the following table.

Cache at weak receiver i
$\left\{\left\{W_{d, G_\ell^{(t-1)}}^{(A)}\right\}_{\ell:\, i \in G_\ell^{(t-1)}}, \left\{W_{d, G_\ell^{(t)}}^{(B)}\right\}_{\ell:\, i \in G_\ell^{(t)}}\right\}_{d=1}^{\mathsf{D}}$
$\left\{K_{G_\ell^{(t+1)}}\right\}_{\ell:\, i \in G_\ell^{(t+1)}}, \left\{K_{G_\ell^{(t)}}\right\}_{\ell:\, i \in G_\ell^{(t)}}$

Delivery phase: Is divided into three subphases of lengths $\beta_1 n$, $\beta_2 n$, and $\beta_3 n$.

$$\mathcal{C}_{\text{sat}}(w_1,\ldots,w_{\mathsf{K}_w}) := \Big\{x^n\Big(w_{\mathsf{K}_w+1},\ldots,w_{\mathsf{K}},b\big|\mathsf{sec}\big(w_1,k_1\big),\ldots,\mathsf{sec}\big(w_{\mathsf{K}_w},k_{\mathsf{K}_w}\big)\Big): (w_{\mathsf{K}_w+1},\ldots,w_{\mathsf{K}})\in\big\{1,\ldots,\big\lfloor 2^{nR}\big\rfloor\big\}^{\mathsf{K}_w},\ b\in\big\{1,\ldots,\big\lfloor 2^{nR_{\text{bin}}}\big\rfloor\big\},\ (k_1,\ldots,k_{\mathsf{K}_w})\in\big\{1,\ldots,\big\lfloor 2^{nR_{\text{Key}}}\big\rfloor\big\}^{\mathsf{K}_w}\Big\} \tag{131}$$

$$\mathcal{S}(w_{\mathsf{K}_w+1},\ldots,w_{\mathsf{K}}) := \Big\{x^n\big(w_{\mathsf{K}_w+1},\ldots,w_{\mathsf{K}},b\big|\mathsf{sec}(w_1,k_1),\ldots,\mathsf{sec}(w_{\mathsf{K}_w},k_{\mathsf{K}_w})\big):\ b\in\big\{1,\ldots,\big\lfloor 2^{nR_{\text{bin}}}\big\rfloor\big\},(k_1,\ldots,k_{\mathsf{K}_w})\in\big\{1,\ldots,\big\lfloor 2^{nR_{\text{Key}}}\big\rfloor\big\}^{\mathsf{K}_w}\Big\} \tag{132}$$

$$\beta_1 = \frac{(\mathsf{K}_w-t)(\mathsf{K}_w-t+1)(\delta_z-\delta_s)\min\{\delta_w-\delta_s,\delta_z-\delta_s\}}{(\mathsf{K}_w-t+1)(\delta_z-\delta_s)\big[\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)\min\{\delta_w-\delta_s,\delta_z-\delta_s\}\big]+\mathsf{K}_s^2t(t+1)(1-\delta_w)^2}, \tag{137a}$$

$$\beta_2 = \frac{\mathsf{K}_s(\mathsf{K}_w-t+1)(t+1)(1-\delta_w)(\delta_z-\delta_s)}{(\mathsf{K}_w-t+1)(\delta_z-\delta_s)\big[\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)\min\{\delta_w-\delta_s,\delta_z-\delta_s\}\big]+\mathsf{K}_s^2t(t+1)(1-\delta_w)^2}, \tag{137b}$$

$$\beta_3 = \frac{\mathsf{K}_s^2t(t+1)(1-\delta_w)^2}{(\mathsf{K}_w-t+1)(1-\delta_s)\big[(\mathsf{K}_w-t)(\delta_w-\delta_s)+\mathsf{K}_s(t+1)(1-\delta_w)\big]+\mathsf{K}_s^2t(t+1)(1-\delta_w)^2}. \tag{137c}$$

$$R^{(A)} = \frac{\mathsf{K}_st(t+1)(1-\delta_w)^2(\delta_z-\delta_s)}{(\mathsf{K}_w-t+1)(\delta_z-\delta_s)\big[\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)\min\{\delta_w-\delta_s,\delta_z-\delta_s\}\big]+\mathsf{K}_s^2t(t+1)(1-\delta_w)^2}-\epsilon/2, \tag{138}$$

$$R^{(B)} = \frac{(\mathsf{K}_w-t+1)(t+1)(1-\delta_w)(\delta_z-\delta_s)\min\{\delta_w-\delta_s,\delta_z-\delta_s\}}{(\mathsf{K}_w-t+1)(\delta_z-\delta_s)\big[\mathsf{K}_s(t+1)(1-\delta_w)+(\mathsf{K}_w-t)\min\{\delta_w-\delta_s,\delta_z-\delta_s\}\big]+\mathsf{K}_s^2t(t+1)(1-\delta_w)^2}-\epsilon/2. \tag{139}$$

Subphase 1: This subphase is dedicated to the transmission of the parts of $W_{d_1}^{(B)},\ldots,W_{d_{\mathsf{K}_w}}^{(B)}$ that are not stored in the cache memories of the respective weak receivers. For each $\ell\in\{1,\ldots,\binom{\mathsf{K}_w}{t+1}\}$, the transmitter first calculates the XOR-message

$$W_{\text{XOR},G_\ell^{(t+1)}}^{(B)} := \bigoplus_{i\in G_\ell^{(t+1)}} W_{d_i,G_\ell^{(t+1)}\setminus\{i\}}^{(B)}, \tag{147}$$

and its secured version

$$W_{\text{sec},G_\ell^{(t+1)}}^{(B)} = \mathsf{sec}\big(W_{\text{XOR},G_\ell^{(t+1)}}^{(B)},\ K_{G_\ell^{(t+1)}}\big). \tag{148}$$

It then sends the secured message tuple

$$\mathbf{W}_{\text{sec},w}^{(B)} := \big(W_{\text{sec},G_1^{(t+1)}}^{(B)},\ldots,W_{\text{sec},G_{\binom{\mathsf{K}_w}{t+1}}^{(t+1)}}^{(B)}\big) \tag{149}$$

using a capacity achieving point-to-point code to all weak receivers. After decoding the message tuple (149), each weak receiver i retrieves the secret key $K_{G_\ell^{(t+1)}}$ from its cache memory, for $\ell\in\{1,\ldots,\binom{\mathsf{K}_w}{t+1}\}$ so that $i\in G_\ell^{(t+1)}$, and produces $\hat{W}_{\text{XOR},G_\ell^{(t+1)}}^{(B)}$. Then, it also retrieves the submessages

$$\Big\{W_{d_k,G_\ell^{(t+1)}\setminus\{k\}}^{(B)}\Big\}_{k\in G_\ell^{(t+1)}\setminus\{i\}} \tag{150}$$

from its cache memory, and guesses the desired submessage as:

$$\hat{W}_{d_k,G_\ell^{(t+1)}\setminus\{i\}}^{(B)} = \hat{W}_{\text{XOR},G_\ell^{(t+1)}}^{(B)} \bigoplus_{k\in G_\ell^{(t+1)}\setminus\{i\}} W_{d_k,G_\ell^{(t+1)}\setminus\{k\}}^{(B)}. \tag{151}$$

At the end of Subphase 1, each weak receiver $i\in\mathsf{K}_w$ assembles the guesses produced in (151) with the parts of $W_{d_i}^{(B)}$ it has stored in its cache memory to form a guess $\hat{W}_{d_i}^{(B)}$.

Subphase 2: This subphase is dedicated to the transmission of the parts of $W_{d_i}^{(A)}$ that are not stored in weak receiver i's cache memory, for $i\in\mathcal{K}_w$, and messages $W_{d_j}^{(B)}$, for $j\in\mathcal{K}_s$. Time-sharing is applied over $\binom{\mathsf{K}_w}{t}$ periods, each of length

$$n_2 = \binom{\mathsf{K}_w}{t}^{-1}\cdot\beta_2 n.$$

The periods are labeled $G_1^{(t)},\ldots,G_{\binom{\mathsf{K}_w}{t}}^{(t)}$. For Period $G_\ell^{(t)}$, $\ell\in\{1,\ldots,\binom{\mathsf{K}_w}{t}\}$, the transmitter calculates the XOR-message

$$W_{\text{XOR},G_\ell^{(t)}}^{(A)} := \bigoplus_{i\in G_\ell^{(t)}} W_{d_i,G_\ell^{(t)}\setminus\{i\}}^{(A)} \tag{152a}$$

and its secured version

$$W_{\text{sec},G_\ell^{(t)}}^{(A)} := \mathsf{sec}\big(W_{\text{XOR},G_\ell^{(t)}}^{(A)},K_{G_\ell^{(t)}}\big) \tag{152b}$$

It also forms the tuple of non-secured messages

$$\mathbf{W}^{(B)}_{s,G^{(t)}_\ell} := \left(W^{(B)}_{d_{\mathsf{K}_w+1},G^{(t)}_\ell}, \ldots, W^{(B)}_{d_{\mathsf{K}},G^{(t)}_\ell}\right) \tag{152c}$$

which it sends to the strong receivers. To this end, generate for each period $G^{(t)}_\ell$ a secure piggyback codebook [23], see Figure 3,

$$\begin{aligned}&\mathcal{C}_{\text{spg},G^{(t)}_\ell}\\ &= \Big\{x^{n_2}_{G^{(t)}_\ell}(\ell_{\text{row}};\ \ell_{\text{col}};\ b):\ \ell_{\text{row}} \in \big\{1,\ldots,\lfloor 2^{nr^{(A)}}\rfloor\big\},\\ &\qquad \ell_{\text{col}} \in \big\{1,\ldots,\lfloor 2^{nr^{(B)}}\rfloor\big\},\ b \in \big\{1,\ldots,\lfloor 2^{nR_{\text{bin}}}\rfloor\big\}\Big\}, \end{aligned} \tag{153}$$

by drawing each entry of each codeword i.i.d. according to a Bernoulli-$1/2$ distribution. The rates $r^{(A)}$, $r^{(B)}$, and R_{bin} are defined in (141), (143), and (146).

The transmitter draws an index B uniformly at random over $\{1,\ldots,\lfloor 2^{nR_{\text{bin}}}\rfloor\}$, and sends the codeword

$$x^{n_2}_{G^{(t)}_\ell}\big(W^{(A)}_{\text{sec},G^{(t)}_\ell};\ \mathbf{W}^{(B)}_{s,G^{(t)}_\ell};\ B\big) \tag{154}$$

over the channel.

We now describe the decoding, starting with the decoding at the weak receivers. For each $i \in \mathsf{K}_w$ and each $\ell \in \big\{1,\ldots,\binom{\mathsf{K}_w}{t}\big\}$ so that $i \in G^{(t)}_\ell$, Receiver i decodes message $W^{(A)}_{d_i,G^{(t)}_\ell\backslash\{i\}}$ sent in Period $G^{(t)}_\ell$ by performing the following steps:

1) It retrieves the secret key $K_{G^{(t)}_\ell}$ and the messages $W^{(B)}_{d_{\mathsf{K}_w+1},G^{(t)}_\ell}, \ldots, W^{(B)}_{d_{\mathsf{K}},G^{(t)}_\ell}$ from its cache memory.
2) It extracts the subcodebook

$$\begin{aligned}&\tilde{\mathcal{C}}_{\text{spg},G^{(t)}_\ell}\left(\mathbf{W}^{(B)}_{s,G^{(t)}_\ell}\right) :=\\ &\Big\{x^{n_2}_{G^{(t)}_\ell}\Big(\ell_{\text{row}};\ \mathbf{W}^{(B)}_{s,G^{(t)}_\ell};\ b\Big): \ell_{\text{row}} \in \big\{1,\ldots,\lfloor 2^{n_2 r^{(A)}}\rfloor\big\},\\ &\qquad b \in \big\{1,\ldots,\lfloor 2^{nR_{\text{bin}}}\rfloor\big\}\Big\}.\end{aligned} \tag{155}$$

3) Based on the reduced codebook $\tilde{\mathcal{C}}_{\text{spg},G^{(t)}_\ell}\left(\mathbf{W}^{(B)}_{s,G^{(t)}_\ell}\right)$, it decodes the secured message $W^{(A)}_{\text{sec},G^{(t)}_\ell}$ from its channel outputs in this period $G^{(t)}_\ell$.
4) It applies the inverse mapping $\mathsf{sec}^{-1}_{K_{G^{(t)}_\ell}}$ to the message decoded in step 3) to obtain the guess $\hat{W}^{(A)}_{\text{XOR},G^{(t)}_\ell}$.
5) Finally, it produces the guess

$$\hat{W}^{(A)}_{d_i,G^{(t)}_\ell\backslash\{i\}} = \hat{W}^{(A)}_{\text{XOR},G^{(t)}_\ell} \bigoplus_{k\in G^{(t)}_\ell\backslash\{i\}} W^{(A)}_{d_k,G^{(t)}_\ell\backslash\{k\}}. \tag{156}$$

Strong receivers treat the secure piggyback codebook as a simple wiretap codebook and decode all transmitted messages as well as the bin indices.

At the end of Subphase 2, each weak receiver $i \in \mathcal{K}_w$ assembles the parts of submessage $W^{(A)}_{d_i}$ that it decoded or that are stored in its cache memory to form the guess $\hat{W}^{(A)}_{d_i}$. Similarly, each strong receiver $j \in \mathcal{K}_s$ assembles the decoded parts of $W^{(B)}_{d_j}$ to form the guess $\hat{W}^{(B)}_{d_j}$.

Subphase 3: A wiretap code is used to send the message tuple

$$W^{(A)}_{d_{\mathsf{K}_w+1}}, \ldots, W^{(A)}_{d_{\mathsf{K}}} \tag{157}$$

to all the strong receivers $\mathsf{K}_w+1,\ldots,\mathsf{K}$.

After the last subphase, each Receiver $k \in \mathcal{K}$ assembles its guesses produced for the two submessages $W^{(A)}_{d_k}$ and $W^{(B)}_{d_k}$ to the final guess $\hat{W}_k$.

Analysis: Analysis is performed averaged over the random choice of the codebooks. We first verify that in each of the three subphases the probability of decoding error tends to 0 as $n\to\infty$. Only weak receivers decode during the first subphase. Probability of decoding error vanishes asymptotically as $n\to\infty$, because

$$\frac{\binom{\mathsf{K}_w}{t+1}\binom{\mathsf{K}_w}{t}^{-1}R^{(B)}}{(1-\delta_w)} = \frac{\frac{\mathsf{K}_w-t}{t+1}R^{(B)}}{(1-\delta_w)} < \beta_1. \tag{158}$$

Only strong receivers decode during the third subphase. Probability of decoding error in Subphase 3 vanishes asymptotically, because

$$\frac{\mathsf{K}_s R^{(A)}}{(\delta_z-\delta_s)} < \beta_3. \tag{159}$$

Probability of decoding error at the weak receivers in Subphase 2 vanishes asymptotically, because

$$\frac{\frac{\mathsf{K}_w-t+1}{t}R^{(A)}}{(1-\delta_w)} < \beta_2. \tag{160}$$

Probability of decoding error at the strong receivers in Subphase 2 vanishes asymptotically, because

$$\frac{\frac{\mathsf{K}_w-t+1}{t}R^{(A)} + \mathsf{K}_s R^{(B)} + \binom{\mathsf{K}_w}{t}R_{\text{bin}}}{(1-\delta_s)} < \beta_2. \tag{161}$$

Whenever the decodings in all subphases are successful, all receivers correctly guess their desired messages. Since the decoding error for each subphase vanishes asymptotically, we conclude that also the average overall probability of error vanishes.

We verify the secrecy constraint averaged over the choice of the code construction. For fixed blocklength n:

$$\begin{aligned}&I\left(W_1,\ldots,W_{\mathsf{D}};Z^n|\mathcal{C}\right)\\ &= I\big(W^{(B)}_{d_1},\ldots,W^{(B)}_{d_{\mathsf{K}_w}};Z_1^{\beta_1 n}|\mathcal{C}\big)\\ &\quad + I\big(W^{(A)}_{d_1},\ldots,W^{(A)}_{d_{\mathsf{K}_w}},W^{(B)}_{d_{\mathsf{K}_w+1}},\ldots,W^{(B)}_{d_{\mathsf{K}}};Z^{(\beta_1+\beta_2)n}_{\beta_1 n+1}|\mathcal{C}\big)\\ &\quad + I\big(W^{(A)}_{d_{\mathsf{K}_w+1}},\ldots,W^{(A)}_{d_{\mathsf{K}}};Z^n_{(\beta_1+\beta_2)n+1}|\mathcal{C}\big),\end{aligned} \tag{162}$$

because of the independence of the communications in the three subphases. By Lemma 1, all of the three summands can be bounded by $\epsilon/3$ for sufficiently large n. This holds in particular for the second summand because from the eavesdropper's point of view, the structure of the piggyback codebook is meaningless and for each message tuple $\big(W^{(A)}_{\text{sec},G^{(t)}_\ell};\ \mathbf{W}^{(B)}_{s,G^{(t)}_\ell}\big)$, the transmitted codeword is chosen uniformly at random (depending on B and $K_{G^{(t)}_\ell}$) over a set of $2^{n_2 r_s}$ i.i.d. random codewords where $r_s = R_{\text{Key},2} + R_{\text{bin}} = \min\{1-\delta_z, 1-\delta_s\}$.

Since averaged over the random codebooks, the probability of decoding error and the average mutual information $I(W_1,\dots,W_{\mathsf{D}};Z^n|\mathcal{C})$ both vanish as $n\to\infty$, there must exist at least one choice of the codebooks so that for this choice $\mathsf{P}_e^{\text{Worst}}$ and $I(W_1,\dots,W_{\mathsf{D}};Z^n)$ both vanish asymptotically.

Notice that for each $t\in\{1,\dots,\mathsf{K}_w-1\}$, the rate of communication satisfies

$$R=R^{(A)}+R^{(B)}=R^{(t+2)}-\epsilon, \tag{163}$$

and weak receivers require a cache memory of

$$\begin{aligned}\mathsf{M}_w&=\mathsf{D}\frac{(t-1)}{\mathsf{K}_w}R^{(A)}+\mathsf{D}\frac{t}{\mathsf{K}_w}R^{(B)}+\binom{\mathsf{K}_w-1}{t}R_{\text{Key},1}\\&\quad+\binom{\mathsf{K}_w-1}{t-1}R_{\text{Key},2}\\&=\mathsf{M}^{(t+2)}-\mathsf{D}\cdot\frac{t-1/2}{\mathsf{K}_w}\epsilon.\end{aligned} \tag{164}$$

Letting $\epsilon\to0$ thus concludes the achievability proof of the rate-memory pairs $(R^{(t+2)},\mathsf{M}_w^{(t+2)})$, for $t=1,\dots,\mathsf{K}_w-1$, as defined in (62k)–(62l).

D. Secure Cache-Aided Piggyback Coding II

Let

$$\beta=\frac{\mathsf{K}_w(\delta_z-\delta_s)}{\mathsf{K}_s\min\{1-\delta_z,1-\delta_w\}+\mathsf{K}_w(\delta_z-\delta_s)} \tag{165}$$

Subphase 1 is of length βn and Subphase 2 of length $(1-\beta)n$.

Submessages $\{W_d^{(A)}\}$ are of rate

$$R^{(A)}=\frac{(\delta_z-\delta_s)\min\{1-\delta_z,1-\delta_w\}}{\mathsf{K}_s\min\{1-\delta_z,1-\delta_w\}+\mathsf{K}_w(\delta_z-\delta_s)}-\epsilon/2, \tag{166}$$

and submessages $\{W_d^{(B)}\}$ are of rate

$$R^{(B)}=\frac{\mathsf{K}_w(\delta_z-\delta_s)^2}{\mathsf{K}_s[\mathsf{K}_s\min\{1-\delta_z,1-\delta_w\}+\mathsf{K}_w(\delta_z-\delta_s)]}-\epsilon/2. \tag{167}$$

The key rate is chosen as

$$R_{\text{Key}}=R^{(A)}. \tag{168}$$

We analyze the scheme averaged over the random choice of the codebooks $\mathcal{C}$. Probability of decoding error at the weak receivers in Subphase 1 vanishes asymptotically as $n\to\infty$, because

$$\frac{\mathsf{K}_wR^{(A)}}{(1-\delta_w)}<\beta. \tag{169}$$

Probability of decoding error at the strong receivers in Subphase 1 vanishes asymptotically, because

$$\frac{\mathsf{K}_wR^{(A)}+\mathsf{K}_sR^{(B)}}{(1-\delta_s)}<\beta. \tag{170}$$

Probability of decoding error in Subphase 2 vanishes asymptotically, because

$$\frac{\mathsf{K}_sR^{(A)}}{(\delta_z-\delta_s)}<(1-\beta). \tag{171}$$

As a consequence, also the overall probability of error (averaged over the random choice of the codebooks $\mathcal{C}$) vanishes as $n\to\infty$. Following a similar secrecy analysis as in the previous Subsection VII-C, it can be shown that also the averaged mutual information $I(W_1,\dots,W_{\mathsf{D}};Z^n|\mathcal{C})$ vanishes as $n\to\infty$. This implies that there must exist at least one choice of the codebooks so that for this choice $\mathsf{P}_e^{\text{Worst}}\to0$ and $I(W_1,\dots,W_{\mathsf{D}};Z^n)\to0$ as $n\to\infty$.

Notice now that

$$R=R^{(A)}+R^{(B)}=\frac{\delta_z-\delta_s}{\mathsf{K}_s}-\epsilon. \tag{172}$$

Moreover, the required cache size at each weak receiver is

$$\mathsf{M}_w=\mathsf{D}R^{(B)}+R_{\text{Key}}=\mathsf{M}_w^{(\mathsf{K}_w+2)}-\mathsf{D}\epsilon/2. \tag{173}$$

Taking $\epsilon\to0$, this concludes the proof of achievability of the rate-memory pair $R^{(\mathsf{K}_w+2)},\mathsf{M}_w^{(\mathsf{K}_w+2)}$ in (62g) and (62h).

VIII. Coding Schemes when All Receivers have Cache Memories

A. Cached Keys

Fix a small $\epsilon>0$ and let

$$\beta=\frac{\mathsf{K}_w(1-\delta_s)}{\mathsf{K}_w(1-\delta_s)+\mathsf{K}_s(1-\delta_w)}, \tag{174}$$

and the message rate be

$$R=\frac{\beta(1-\delta_w)}{\mathsf{K}_w}-\epsilon. \tag{175}$$

Notice that by the choice of the parameter β, also

$$R=\frac{(1-\beta)(1-\delta_s)}{\mathsf{K}_s}-\epsilon. \tag{176}$$

Choose key rates

$$R_{\text{Key},w}=\frac{\beta}{\mathsf{K}_w}\cdot\min\{1-\delta_z,1-\delta_w\}, \tag{177}$$

$$R_{\text{Key},s}=\frac{1-\beta}{\mathsf{K}_s}\cdot\min\{1-\delta_z,1-\delta_s\}, \tag{178}$$

and let secret keys $K_1,\dots,K_{\mathsf{K}_w}$ be of rate $R_{\text{Key},w}$ and secret keys $K_{\mathsf{K}_w+1},\dots,K_{\mathsf{K}}$ be of rate $R_{\text{Key},s}$.

The cache requirement at weak receivers is

$$\mathsf{M}_w=R_{\text{Key},w} \tag{179}$$

$$\quad=\frac{(1-\delta_s)}{\mathsf{K}_w(1-\delta_s)+\mathsf{K}_s(1-\delta_w)}\min\{1-\delta_z,1-\delta_w\} \tag{180}$$

and the cache requirement at strong receivers is

$$\mathsf{M}_s=R_{\text{Key},s} \tag{181}$$

$$\quad=\frac{(1-\delta_w)}{\mathsf{K}_w(1-\delta_s)+\mathsf{K}_s(1-\delta_w)}\min\{1-\delta_z,1-\delta_s\} \tag{182}$$

Analysis of the probability of error and of the secrecy constraint (24) are standard and omitted. Letting $\epsilon\to0$ proves achievability of the rate-memory par $(\tilde{R}^{(1)},\tilde{\mathsf{M}}_w^{(1)},\tilde{\mathsf{M}}_s^{(1)})$.

B. Secure Cache-Aided Piggyback Coding with Keys at All Receivers

The scheme follows closely the secure cache-aided piggyback coding I in Subsection VII-C, but for a different choice of rates and time-sharings and with additional secret keys for Subphases 2 and 3, which render wiretap binning useless.

The time-sharing parameters $\beta_1, \beta_2, \beta_3$ and the rates $R^{(A)}$ and $R^{(B)}$ used here are defined on top of the next page.

The following keys are generated:

- For each $\ell \in \left\{1, \ldots, \binom{\mathsf{K}_w}{t+1}\right\}$, generate an independent secret key $K_{G_\ell^{(t+1)}}$ of rate
$$R_{\text{Key},1} = \binom{\mathsf{K}_w}{t+1}^{-1} \cdot \beta_1 \cdot \min\left\{1-\delta_z, 1-\delta_w\right\}. \quad (186)$$
- For each $\ell \in \left\{1, \ldots, \binom{\mathsf{K}_w}{t}\right\}$, generate an independent secret key $K_{G_\ell^{(t)}}$ of rate
$$R_{\text{Key},2} = \binom{\mathsf{K}_w}{t}^{-1} \cdot \beta_2 \cdot \min\left\{1-\delta_z, 1-\delta_w\right\} \quad (187)$$
and K_s independent secret keys $K_{\mathsf{K}_w+1, G_\ell^{(t)}}, \ldots, K_{\mathsf{K}, G_\ell^{(t)}}$ of rate
$$R_{\text{Key},3} = \binom{\mathsf{K}_w}{t}^{-1} \cdot \frac{\beta_2}{\mathsf{K}_s} \cdot \min\left\{\max\left\{0, \delta_w - \delta_z\right\}, \delta_w - \delta_s\right\}. \quad (188)$$
- For $j \in \mathcal{K}_s$, generate an independent secret key K_j of rate
$$R_{\text{Key},4} = \frac{\beta_3}{\mathsf{K}_s} \cdot \min\left\{1-\delta_z, 1-\delta_s\right\}. \quad (189)$$

Placement phase: For each weak receiver $i \in \mathcal{K}_w$, cache

$$\begin{aligned} V_i = &\left\{W^{(A)}_{d, G_\ell^{(t-1)}} : d \in \mathcal{D} \text{ and } i \in G_\ell^{(t-1)}\right\} \\ &\bigcup \left\{W^{(B)}_{d, G_\ell^{(t)}} : d \in \mathcal{D} \text{ and } i \in G_\ell^{(t)}\right\} \\ &\bigcup \left\{K_{G_\ell^{(t+1)}} : i \in G_\ell^{(t+1)}\right\} \bigcup \left\{K_{G_\ell^{(t)}} : i \in G_\ell^{(t)}\right\} \\ &\bigcup \left\{K_{j, G_\ell^{(t)}} : i \in G_\ell^{(t)} \text{ and } j \in \mathcal{K}_s\right\}. \end{aligned} \quad (190)$$

For each strong receiver $j \in \mathcal{K}_s$, cache

$$V_j = K_j \bigcup \left\{K_{j, G_\ell^{(t)}} : \ell \in \left\{1, \ldots, \binom{\mathsf{K}_w}{t}\right\}\right\}. \quad (191)$$

Delivery phase: Is divided into three subphases of lengths $n_1 = \beta_1 n$, $n_2 = \beta_2 n$, and $n_3 = \beta_3 n$, where β_1, β_2, and β_3 are defined in (183) on the next page.

Subphase 1: Is as described in Subsection VII-C.

Subphase 2: Similar to Subsection VII-C, but with extra keys. As in Subsection VII-C, this subphase is split into $\binom{\mathsf{K}_w}{t}$ equally-long periods, which we label by all t-user subsets of $\mathcal{K}_w$ i.e., by $G_1^{(t)}, \ldots, G^{(t)}_{\binom{\mathsf{K}_w}{t}}$. For the transmission in Period $G_\ell^{(t)}$, for $\ell \in \{1, \ldots, \binom{\mathsf{K}_w}{t}\}$, a standard piggyback codebook

$$\begin{aligned} \mathcal{C}_{\text{pg}, G_\ell^{(t)}} = \Big\{ x^{n_2}_{G_\ell^{(t)}}\big(\ell_{\text{row}};\ \ell_{\text{col}}\big) :\ &\ell_{\text{row}} \in \left\{1, \ldots, \left\lfloor 2^{nr^{(A)}} \right\rfloor\right\}, \\ &\ell_{\text{col}} \in \left\{1, \ldots, \left\lfloor 2^{nr^{(B)}} \right\rfloor\right\}\Big\}, \end{aligned} \quad (192)$$

is generated by drawing all entries i.i.d. Bernoulli-1/2. The codebooks are revealed to the transmitter and all receivers, and the rates $r^{(A)}$ and $r^{(B)}$ are chosen as

$$r^{(A)} = R^{(A)} \binom{\mathsf{K}_w}{t-1}^{-1} \quad (193)$$

$$r^{(B)} = R^{(B)} \binom{\mathsf{K}_w}{t}^{-1}. \quad (194)$$

At the beginning of Period $G_\ell^{(t)}$, the transmitter computes

$$W^{(A)}_{\text{sec}, w, G_\ell^{(t)}} = \mathsf{sec}\Big(\bigoplus_{i \in G_\ell^{(t)}} W^{(A)}_{d_i, G_\ell^{(t)} \setminus \{i\}},\ K_{G_\ell^{(t)}}\Big). \quad (195)$$

and

$$\begin{aligned} \mathbf{W}^{(B)}_{\text{sec}, s, G_\ell^{(t)}} := \Big(&\mathsf{sec}\big(W^{(B)}_{d_{\mathsf{K}_w+1}, G_\ell^{(t)}},\ K_{\mathsf{K}_w+1, G_\ell^{(t)}}\big), \\ &\ldots,\ \mathsf{sec}\big(W^{(B)}_{d_{\mathsf{K}}, G_\ell^{(t)}},\ K_{\mathsf{K}, G_\ell^{(t)}}\big)\Big), \end{aligned} \quad (196)$$

and sends the codeword

$$x^{n_2}_{G_\ell^{(t)}}\big(W^{(A)}_{\text{sec}, w, G_\ell^{(t)}};\ \mathbf{W}^{(B)}_{\text{sec}, s, G_\ell^{(t)}}\ \big) \quad (197)$$

over the channel. Decoding of Period $G_\ell^{(t)}$ is performed as follows. All weak receivers $i \in G_\ell^{(t)}$ compute the message tuple to the strong receivers $\mathbf{W}^{(B)}_{\text{sec}, s, G_\ell^{(t)}}$ and perform their decoding steps as described in Subsection VII-C. Strong receivers decode both secured message tuples transmitted in this period. Any given strong receiver $j \in \mathcal{K}_s$ can then recover its desired message

$$W^{(B)}_{d_j, G_\ell^{(t)}} \quad (198)$$

using the secret key $K_{j, G_\ell^{(t)}}$ stored in its cache memory.

At the end of the subphase, each weak receiver $i \in \mathcal{K}_w$ assembles the parts of submessage $W^{(A)}_{d_i}$ that it decoded or it has stored in its cache memory, and forms the guess $\hat{W}^{(A)}_{d_i}$. Similarly, each strong receiver $j \in \mathcal{K}_s$ assembles the decoded parts and forms $\hat{W}^{(B)}_{d_j}$.

Subphase 3: A capacity-achieving code is used to send the secured message tuple

$$\mathsf{sec}\big(W^{(A)}_{d_{\mathsf{K}_w+1}}, K_{\mathsf{K}_w+1}\big), \ldots, \mathsf{sec}\big(W^{(A)}_{d_{\mathsf{K}}}, K_{\mathsf{K}}\big) \quad (199)$$

to the strong receivers $\mathsf{K}_w + 1, \ldots, \mathsf{K}$. Each strong receiver $j \in \mathcal{K}_s$ obtains its desired guess $\hat{W}^{(A)}_{d_j}$ with the help of the secret key K_j stored in its cache memory.

At the end of the entire delivery phase, each Receiver $k \in \mathcal{K}$ assembles its guesses of $W^{(A)}_{d_k}$ and $W^{(B)}_{d_k}$ to produce the final guess $\hat{W}_k$.

Analysis: The scheme is analyzed by averaging over the random code constructions $\mathcal{C}$. We first verify that in each of the three subphases the average probability of decoding error tends to 0 as $n \to \infty$. Only weak receivers decode during the first subphase. Probability of decoding error in Subphase 1 vanishes asymptotically as $n \to \infty$, because

$$\frac{\binom{\mathsf{K}_w}{t+1}\binom{\mathsf{K}_w}{t}^{-1} R^{(B)}}{(1-\delta_w)} = \frac{\frac{\mathsf{K}_w - t}{t+1} R^{(B)}}{(1-\delta_w)} < \beta_1. \quad (200)$$

$$\beta_1 = \frac{(\mathsf{K}_w - t+1)(\mathsf{K}_w - t)(1-\delta_s)(\delta_w-\delta_s)}{(\mathsf{K}_w-t+1)(1-\delta_s)\big[(\mathsf{K}_w-t)(\delta_w-\delta_s)+\mathsf{K}_s(t+1)(1-\delta_w)\big]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2}, \tag{183a}$$

$$\beta_2 = \frac{\mathsf{K}_s(\mathsf{K}_w - t+1)(t+1)(1-\delta_w)(1-\delta_s)}{(\mathsf{K}_w-t+1)(1-\delta_s)\big[(\mathsf{K}_w-t)(\delta_w-\delta_s)+\mathsf{K}_s(t+1)(1-\delta_w)\big]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2}, \tag{183b}$$

$$\beta_3 = \frac{\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2}{(\mathsf{K}_w-t+1)(1-\delta_s)\big[(\mathsf{K}_w-t)(\delta_w-\delta_s)+\mathsf{K}_s(t+1)(1-\delta_w)\big]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2}. \tag{183c}$$

$$R^{(A)} = \frac{\mathsf{K}_s t(t+1)(1-\delta_w)^2(1-\delta_s)}{(\mathsf{K}_w-t+1)(1-\delta_s)\big[(\mathsf{K}_w-t)(\delta_w-\delta_s)+\mathsf{K}_s(t+1)(1-\delta_w)\big]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2} - \epsilon/2, \tag{184}$$

$$R^{(B)} = \frac{(\mathsf{K}_w-t+1)(t+1)(1-\delta_w)(1-\delta_s)(\delta_w-\delta_s)}{(\mathsf{K}_w-t+1)(1-\delta_s)\big[(\mathsf{K}_w-t)(\delta_w-\delta_s)+\mathsf{K}_s(t+1)(1-\delta_w)\big]+\mathsf{K}_s^2 t(t+1)(1-\delta_w)^2} - \epsilon/2 \tag{185}$$

Only strong receivers decode during the third subphase. Probability of decoding error in Subphase 3 vanishes asymptotically, because

$$\frac{\mathsf{K}_s R^{(A)}}{(1-\delta_s)} < \beta_3. \tag{201}$$

In Subphase 2, probability of decoding error at weak receivers vanishes asymptotically, because

$$\frac{\frac{\mathsf{K}_w - t+1}{t} R^{(A)}}{(1-\delta_w)} < \beta_2. \tag{202}$$

Probability of decoding error at strong receivers in Subphase 2 vanishes asymptotically, because

$$\frac{\frac{\mathsf{K}_w - t+1}{t} R^{(A)} + \mathsf{K}_s R^{(B)}}{(1-\delta_s)} < \beta_2. \tag{203}$$

Whenever the decodings in all subphases are successful, all receivers guess their desired messages correctly. As a consequence, also the average of the overall probability of error vanishes as $n \to \infty$.

We analyze the secrecy constraint. For fixed blocklength n:

$$\begin{aligned} & I\big(W_1,\ldots,W_{\mathsf{D}}; Z^n \big| \mathcal{C}\big) \\ &= I\big(W_{d_1}^{(B)},\ldots,W_{d_{\mathsf{K}_w}}^{(B)}; Z_1^{\beta_1 n}\big|\mathcal{C}\big) \\ &\quad + I\big(W_{d_1}^{(A)},\ldots,W_{d_{\mathsf{K}_w}}^{(A)}, W_{d_{\mathsf{K}_w+1}}^{(B)},\ldots,W_{d_{\mathsf{K}}}^{(B)}; Z_{\beta_1 n+1}^{(\beta_1+\beta_2)n}\big|\mathcal{C}\big) \\ &\quad + I\big(W_{d_{\mathsf{K}_w+1}}^{(A)},\ldots,W_{d_{\mathsf{K}}}^{(A)}; Z_{(\beta_1+\beta_2)n+1}^{n}\big|\mathcal{C}\big), \end{aligned} \tag{204}$$

because of the independence of the communications in the three subphases. The sizes of the secret keys have been chosen so that the codewords to be sent in each of the three subphases are chosen uniformly at random over a subset of the random codebooks that are equal to the minimum between $(1-\delta_z)$ and the rates of communication. By Lemma 1 this proves that $I\big(W_1,\ldots,W_{\mathsf{D}}; Z^n\big|\mathcal{C}\big)$ tends to 0 as $n \to \infty$.

We can conclude that there must exist at least one choice of the codebooks so that for this choice both $\mathrm{P}_e^{\text{Worst}} \to 0$ and $I\big(W_1,\ldots,W_{\mathsf{D}}; Z^n\big)$ vanish asymptotically.

For each choice of the parameter $t \in \{1,\ldots,\mathsf{K}_w-1\}$:

$$R = R^{(A)} + R^{(B)} = \tilde{R}^{(t+1)} - \epsilon. \tag{205}$$

Moreover, each weak receiver requires a cache size of

$$\begin{aligned} \mathsf{M}_w &= \frac{\mathsf{D}\big[tR^{(A)} + (t-1)R^{(B)}\big]}{\mathsf{K}_w} + \binom{\mathsf{K}_w - 1}{t} R_{\text{Key},1} \\ &\quad + \binom{\mathsf{K}_w-1}{t-1} R_{\text{Key},2} + \binom{\mathsf{K}_w-1}{t-1}\mathsf{K}_s R_{\text{Key},3} \\ &= \tilde{\mathsf{M}}_w^{(t+1)} - \frac{\mathsf{D}(t-1/2)}{\mathsf{K}_w}\epsilon, \end{aligned} \tag{206}$$

and each strong receiver a cache size of

$$\mathsf{M}_s = R_{\text{Key},4} + \binom{\mathsf{K}_w}{t} R_{\text{Key},3} = \tilde{\mathsf{M}}_s^{(t+1)}. \tag{207}$$

Taking $\epsilon \to 0$ thus proves achievability of the rate-memory triples $(\tilde{R}^{(t+1)}, \tilde{\mathsf{M}}_w^{(t+1)}, \tilde{\mathsf{M}}_s^{(t+1)})$, for $t \in \{1,\ldots,\mathsf{K}_w-1\}$, in (88d)–(88f).

C. Symmetric Secure Cache-Aided Piggyback Coding

Fix $\epsilon > 0$ and define the time-sharing parameters as on top of the next page. Notice that $\beta_1 + \beta_2 + \beta_3 = 1$.

Choose two positive integers $t_w \in \{1,\ldots,\mathsf{K}_w\}$ and $t_s \in \{1,\ldots,\mathsf{K}_s\}$.

Message splitting: Divide each message into two submessages,

$$W_d = \big[W_d^{(A)}, W_d^{(B)}\big], \qquad d \in \mathcal{D}, \tag{208}$$

so that the submessages are of rates defined in (210) and (211) on top of the next page.

Denote the $\binom{\mathsf{K}_w}{t_w}$ subsets of $\{1,\ldots,\mathsf{K}_w\}$ of size t_w by $G_1^{(t_w)},\ldots,G_{\binom{\mathsf{K}_w}{t_w}}^{(t_w)}$ and the $\binom{\mathsf{K}_s}{t_s}$ subsets of $\{1,\ldots,\mathsf{K}_s\}$ of size t_s by $G_1^{(t_s)},\ldots,G_{\binom{\mathsf{K}_s}{t_s}}^{(t_s)}$. Divide every message $W_d^{(A)}$ into $\binom{\mathsf{K}_w}{t_w}$ submessages and every message $W_d^{(B)}$ into $\binom{\mathsf{K}_s}{t_s}$ submessages:

$$W_d^{(A)} = \left\{ W_{d,G_\ell^{(t_w)}}^{(A)} : \quad \ell \in \left\{1,\ldots,\binom{\mathsf{K}_w}{t_w}\right\}\right\}, \tag{212a}$$

$$W_d^{(B)} = \left\{ W_{d,G_\ell^{(t_s)}}^{(B)} : \quad \ell \in \left\{1,\ldots,\binom{\mathsf{K}_s}{t_s}\right\}\right\}. \tag{212b}$$

Key generation:

$$\beta_1 := \frac{\mathsf{K}_w(\mathsf{K}_w - t_w)(t_s+1)(1-\delta_s)^2}{\mathsf{K}_w(\mathsf{K}_w - t_w)(t_s+1)(1-\delta_s)^2 + \mathsf{K}_s(t_w+1)(1-\delta_w)\big[(\mathsf{K}_s - t_s)(1-\delta_w) + \mathsf{K}_w(t_s+1)(1-\delta_s)\big]}, \tag{209a}$$

$$\beta_2 := \frac{\mathsf{K}_w\mathsf{K}_s(t_w+1)(t_s+1)(1-\delta_w)(1-\delta_s)}{\mathsf{K}_w(\mathsf{K}_w - t_w)(t_s+1)(1-\delta_s)^2 + \mathsf{K}_s(t_w+1)(1-\delta_w)\big[(\mathsf{K}_s - t_s)(1-\delta_w) + \mathsf{K}_w(t_s+1)(1-\delta_s)\big]}, \tag{209b}$$

$$\beta_3 := \frac{\mathsf{K}_s(\mathsf{K}_s - t_s)(t_w+1)(1-\delta_w)^2}{\mathsf{K}_w(\mathsf{K}_w - t_w)(t_s+1)(1-\delta_s)^2 + \mathsf{K}_s(t_w+1)(1-\delta_w)\big[(\mathsf{K}_s - t_s)(1-\delta_w) + \mathsf{K}_w(t_s+1)(1-\delta_s)\big]} \tag{209c}$$

$$R^{(A)} = \frac{\mathsf{K}_w(t_w+1)(t_s+1)(1-\delta_w)(1-\delta_s)^2}{\mathsf{K}_w(\mathsf{K}_w - t_w)(t_s+1)(1-\delta_s)^2 + \mathsf{K}_s(t_w+1)(1-\delta_w)\big[(\mathsf{K}_s - t_s)(1-\delta_w) + \mathsf{K}_w(t_s+1)(1-\delta_s)\big]} - \epsilon/2, \tag{210}$$

$$R^{(B)} = \frac{\mathsf{K}_s(t_w+1)(t_s+1)(1-\delta_w)^2(1-\delta_s)}{\mathsf{K}_w(\mathsf{K}_w - t_w)(t_s+1)(1-\delta_s)^2 + \mathsf{K}_s(t_w+1)(1-\delta_w)\big[(\mathsf{K}_s - t_s)(1-\delta_w) + \mathsf{K}_w(t_s+1)(1-\delta_s)\big]} - \epsilon/2 \tag{211}$$

- For each $\ell \in \Big\{1,\ldots,\binom{\mathsf{K}_w}{t_w+1}\Big\}$, generate an independent random key $K_{G_\ell^{(t_w+1)}}$ of rate
$$R_{\text{Key},1} = \binom{\mathsf{K}_w}{t_w+1}^{-1} \cdot \beta_1 \min\{1-\delta_z, 1-\delta_w\}.$$
- For each $\ell \in \Big\{1,\ldots,\binom{\mathsf{K}_s}{t_s+1}\Big\}$, generate an independent random key $K_{G_\ell^{(t_s+1)}}$ of rate
$$R_{\text{Key},2} = \binom{\mathsf{K}_s}{t_s+1}^{-1} \cdot \beta_3 \min\{1-\delta_z, 1-\delta_s\}.$$
- For each $i \in \mathcal{K}_w$ and $j \in \mathcal{K}_s$, generate an independent random key $K_{w,\{i,j\}}$ of rate
$$R_{\text{Key},3} = \frac{\beta_2}{\mathsf{K}_w\mathsf{K}_s} \min\{1-\delta_z, 1-\delta_w\},$$
and an independent random key $K_{s,\{i,j\}}$ of rate
$$R_{\text{Key},4} = \frac{\beta_2}{\mathsf{K}_w\mathsf{K}_s} \min\big\{(\delta_w-\delta_z)^+, 1-\delta_s\big\}.$$

Placement phase: Place the cache contents as shown in the following table.

Cache at weak receiver i	Cache at strong receiver j
$\left\{\left\{W^{(A)}_{d,G_\ell^{(t_w)}}\right\}_{i\in G_\ell^{(t_w)}}\right\}_{d=1}^{\mathsf{D}}$	$\left\{\left\{W^{(B)}_{d,G_\ell^{(t_s)}}\right\}_{j\in G_\ell^{(t_s)}}\right\}_{d=1}^{\mathsf{D}}$
$\left\{K_{G_\ell^{(t_w+1)}}\right\}_{i\in G_\ell^{(t_w+1)}}$	$\left\{K_{G_\ell^{(t_s+1)}}\right\}_{j\in G_\ell^{(t_s+1)}}$
$K_{w,\{i,\mathsf{K}_w+1\}},\ldots K_{w,\{i,\mathsf{K}\}}$	$K_{w,\{1,j\}},\ldots K_{w,\{\mathsf{K}_w,j\}}$
$K_{s,\{i,\mathsf{K}_w+1\}},\ldots K_{s,\{i,\mathsf{K}\}}$	$K_{s,\{1,j\}},\ldots K_{s,\{\mathsf{K}_w,j\}}$

Delivery phase: The delivery phase is divided into three subphases of lengths $\beta_1 n$, $\beta_2 n$ and $\beta_3 n$.

Subphase 1: This phase conveys to each weak receiver $i \in \mathcal{K}_w$, the parts of $W^{(A)}_{d_i}$ that are not stored in its cache memory.

Time-sharing is applied over $\binom{\mathsf{K}_w}{t_w+1}$ equally-long periods of length $n_1 = \beta_1 n\binom{\mathsf{K}_w}{t_w+1}^{-1}$. Label the periods $G_1^{(t+1)},\ldots,G^{(t+1)}_{\binom{\mathsf{K}_w}{t+1}}$.

In Period $G_\ell^{(t+1)}$, for $\ell \in \Big\{1,\ldots,\binom{\mathsf{K}_w}{t_w+1}\Big\}$, the secured XOR-message

$$\mathsf{sec}\Bigg(\bigoplus_{i\in G_\ell^{(t_w+1)}} W^{(A)}_{d_i,G_\ell^{(t_w+1)}\backslash\{i\}},\ K_{G_\ell^{(t_w+1)}}\Bigg) \tag{213}$$

is sent to the subset of receivers $G_\ell^{(t_w+1)}$. Each receiver $i \in G_\ell^{(t_w+1)}$ retrieves the content

$$K_{G_\ell^{(t_w+1)}} \text{ and } \Big\{W^{(A)}_{d_k,G_\ell^{(t_w+1)}\backslash\{k\}}\Big\}_{k\in G_\ell^{(t_w+1)}\backslash\{i\}} \tag{214}$$

from its cache memory. It then decodes the secured message in (213), and with the retrieved cache content (214) it recovers the desired message $W^{(A)}_{d_i,G_\ell^{(t_w+1)}\backslash\{i\}}$.

Subphase 2: Submessages $W^{(B)}_{d_1},\ldots,W^{(B)}_{d_{\mathsf{K}_w}}$ are sent to weak receivers $1,\ldots,\mathsf{K}_w$ and submessages $W^{(A)}_{d_{\mathsf{K}_w+1}},\ldots,W^{(A)}_{d_{\mathsf{K}}}$ to strong receivers $\mathsf{K}_w+1,\ldots,\mathsf{K}$.

Time-sharing is applied over $\mathsf{K}_w\cdot\mathsf{K}_s$ periods, each of length $n_2 = \frac{\beta_2 n}{\mathsf{K}_w\mathsf{K}_s}$. The periods are labeled $\{i,j\}$, for $i \in \mathcal{K}_w$ and $j \in \mathcal{K}_s$. Divide each submessage $W_d^{(B)}$ into K_s parts

$$W_d^{(B)} = \big(W^{(B)}_{d,\mathsf{K}_w+1},\ldots,W^{(B)}_{d,\mathsf{K}}\big) \tag{215a}$$

so that each part is of equal rate $r^{(B)} = R^{(B)}/\mathsf{K}_s$ and for all $j \in \mathcal{K}_s$ and $d \in \mathcal{D}$, part $W^{(B)}_{d,j}$ is stored in strong receiver j's cache memory. Similarly, divide each submessage $W_d^{(A)}$ into K_w parts

$$W_d^{(A)} = \big(W^{(A)}_{d,1},\ldots,W^{(A)}_{d,\mathsf{K}_w}\big), \tag{215b}$$

of equal rate $r^{(A)} = R^{(A)}/\mathsf{K}_w$ so that for each $i \in \mathcal{K}_w$ and $d \in \mathcal{D}$, part $W^{(A)}_{d,i}$ is stored in weak receiver i's cache memory.

We describe the encoding and decoding operations in a period $\{i,j\}$, for $i \in \mathsf{K}_w$ and $j \in \mathsf{K}_s$. In this period, the transmitter sends the codeword

$$x^{n_2}_{\{i,j\}}\Big(\mathsf{sec}\big(W^{(B)}_{d_i,j}, K_{w,\{i,j\}}\big);\ \mathsf{sec}\big(W^{(A)}_{d_j,i}, K_{s,\{i,j\}}\big)\Big) \tag{216}$$

over the channel. At the end of the period, Receiver i first retrieves the cache content

$$K_{s,\{i,j\}},\ W_{d_j,i}^{(A)}, \tag{217}$$

and forms the subcodebook

$$\begin{aligned}\mathcal{C}^{\text{row}}_{\text{pg},\{i,j\}}&\big(\mathsf{sec}\big(W_{d_j,i}^{(A)},K_{s,\{i,j\}}\big)\big)\\ &=\Big\{x^{n_2}_{\{i,j\}}\big(\ell_{\text{row}};\ \mathsf{sec}\big(W_{d_j,i}^{(A)},K_{s,\{i,j\}}\big)\big):\\ &\qquad\qquad \ell_{\text{row}}\in\big\{1,\ldots,\lfloor 2^{nr^{(B)}}\rfloor\big\}\Big\}.\end{aligned} \tag{218}$$

Based on this subcodebook, it then decodes the secured message $\mathsf{sec}\big(W_{d_i,j}^{(B)},K_{w,\{i,j\}}\big)$, and recovers its desired message $W_{d_i,j}^{(B)}$ with the secret key $K_{w,\{i,j\}}$ stored in its cache memory. Receiver j proceeds analogously, except that it retrieves the cache content

$$K_{w,\{i,j\}},\ W_{d_i,j}^{(B)}, \tag{219}$$

and forms the subcodebook

$$\begin{aligned}\mathcal{C}^{\text{col}}_{\text{pg},\{i,j\}}&\big(\mathsf{sec}\big(W_{d_i,j}^{(B)},K_{w,\{i,j\}}\big)\big)\\ &=\Big\{x^{n_2}_{\{i,j\}}\big(\mathsf{sec}\big(W_{d_i,j}^{(B)},K_{w,\{i,j\}}\big);\ \ell_{\text{col}}\big):\\ &\qquad\qquad \ell_{\text{col}}\in\big\{1,\ldots,\lfloor 2^{nr^{(A)}}\rfloor\big\}\Big\}.\end{aligned} \tag{220}$$

It then decodes the secured message $\mathsf{sec}\big(W_{d_j,i}^{(A)},K_{s,\{i,j\}}\big)$, and recovers $W_{d_j,i}^{(A)}$ using the secret key $K_{s,\{i,j\}}$ from its cache memory.

Subphase 3: This subphase conveys to each strong receiver $j\in\mathcal{K}_s$, the parts of submessage $W_{d_j}^{(B)}$ that are not stored in its cache memory. Encoding/decoding operations are obtained from the encoding/decoding operations for Subphase 1 if in the latter the subscript w is replaced by the subscript s and the superscript (A) is replaced by the superscript (B).

Analysis: We analyze the average (over the random choice of the codebooks) probability of decoding error and the average leakage $I(W_1,\ldots,W_{\mathsf{D}};Z^n|\mathcal{C})$.

We start by showing that the probability of decoding error vanishes in each of the three subphases. Only weak receivers perform decoding operations in the first subphase. Probability of error in this subphase thus tends to 0 as $n\to\infty$, because

$$\frac{\frac{\mathsf{K}_w-t_w}{t_w+1}R^{(A)}}{(1-\delta_w)}<\beta_1 \tag{221}$$

In the second subphase, both weak and strong receivers perform decoding operations. The probability of decoding error at weak receivers tends to 0 as $n\to\infty$, because

$$\frac{\mathsf{K}_w R^{(B)}}{(1-\delta_w)}<\beta_2 \tag{222}$$

The probability of decoding error at strong receivers tends to 0 as $n\to\infty$, because

$$\frac{\mathsf{K}_s R^{(A)}}{(1-\delta_s)}<\beta_2. \tag{223}$$

(Notice that by our choice of the subrates $R^{(A)}$ and $R^{(B)}$, the two decoding constraints of Subphase 2, (222) and (223) are equally strong.) Only strong receivers perform decoding operations in the third subphase. The probability of error of these decoding operations tends to 0 as $n\to\infty$, because

$$\frac{\frac{\mathsf{K}_s-t_s}{t_s+1}R^{(B)}}{(1-\delta_s)}<\beta_3 \tag{224}$$

Since all receivers correctly recover their demanded messages when all decoding operations tend to 0, when averaged over the random code construction, the total probability of error tends to 0 as $n\to\infty$.

Communication is secured because the secret keys have been chosen sufficiently long. In fact, following the arguments given in the previous Subsection VIII-B, it can be shown that the average leakage term $I(W_1,\ldots,W_{\mathsf{D}};Z^n|\mathcal{C})$ tends to 0 as $n\to\infty$. By standard arguments, it can then be concluded that there must exist a choice of the codebooks so that for this choice the probability of error $\mathrm{P}_e^{\text{Worst}}$ and the leakage $I(W_1,\ldots,W_{\mathsf{D}};Z^n)$ both vanish asymptotically as $n\to\infty$.

The presented scheme requires weak receivers to have a cache of size

$$\begin{aligned}\mathsf{M}_w&=\frac{\mathsf{D}t_w}{\mathsf{K}_w}R^{(A)}+\frac{(t_w+1)\beta_1\min\{1-\delta_z,1-\delta_w\}}{\mathsf{K}_w}\\ &\quad+\frac{\beta_2\min\{1-\delta_z,2-\delta_w-\delta_s\}}{\mathsf{K}_w}\\ &=\tilde{\mathsf{M}}_w^{(\mathsf{K}_w+t_w(\mathsf{K}_s+1)+t_s)}-\frac{\mathsf{D}t_w}{2\mathsf{K}_w}\epsilon,\end{aligned} \tag{225}$$

and strong receivers a cache of size

$$\begin{aligned}\mathsf{M}_s&=\frac{\mathsf{D}t_s}{\mathsf{K}_s}R^{(B)}+\frac{(t_s+1)\beta_3\min\{1-\delta_z,1-\delta_s\}}{\mathsf{K}_s}\\ &\quad+\frac{\beta_2\min\{1-\delta_z,2-\delta_w-\delta_s\}}{\mathsf{K}_s}\\ &=\tilde{\mathsf{M}}_s^{(\mathsf{K}_w+t_w(\mathsf{K}_s+1)+t_s)}-\frac{\mathsf{D}t_s}{2\mathsf{K}_s}\epsilon.\end{aligned} \tag{226}$$

The rate of the messages is

$$R=R^{(A)}+R^{(B)}=\tilde{R}^{(\mathsf{K}_w+t_w(\mathsf{K}_s+1)+t_s)}-\epsilon. \tag{227}$$

Thus, letting $\epsilon\to 0$ establishes achievability of the desired rate-memory triples in (88g)–(88i).

IX. Summary

We have studied secrecy of cache-aided wiretap erasure BCs with K_w weak receivers, K_s strong receivers and one eavesdropper. We have provided a general upper bound on the secrecy capacity-memory tradeoff for the case when receivers have arbitrary erasure probabilities and arbitrary cache sizes. We have also proposed lower bounds on the secrecy capacity-memory tradeoff for different cache sizes. For some cache arrangements, e.g., for zero cache sizes at strong receivers $\mathsf{M}_s=0$, our upper and lower bounds coincide for small cache sizes. For $\mathsf{M}_s=0$, they also match for large cache sizes. These bounds show that the secrecy constraint can induce a significant loss in capacity compared to the standard non-secure system, especially when $\mathsf{M}_s=0$. They also exhibit that in a secure system, the caching gain with small cache memories is much more important than its non-secure counterpart. This is due to the fact that secret keys can be stored in the caches, which are more useful than cached data. For larger cache sizes, data has to be stored as well and the caching gains of the secure system are similar to the gains in a standard system. We also present a lower bound on the capacity of a scenario where the cache assignment across receivers can be optimized subject to a total cache budget M_{tot}. The lower bound is exact for small cache budgets M_{tot}.

Acknowledgment

We thank the anonymous reviewers and the AE for their valuable comments. We are particular indebted to the AE M. Bloch for telling us that Lemma 1, and thus all our achievability results, hold also under the strong secrecy constraint. Interesting discussions with J. Kliewer are also acknowledged.

Appendix A
Proof of Lemma 2

For each blocklength n, we fix caching, encoding and decoding functions as in (16), (20) and (22) so that both the probability of worst-case error and the secrecy leakage satisfy: [3]

$$\mathrm{P}_e^{\text{Worst}} \xrightarrow[n\to\infty]{} 0, \tag{228a}$$

$$I(W_1,\ldots,W_D;Z^n) \xrightarrow[n\to\infty]{} 0. \tag{228b}$$

We only prove the lemma for the set $\mathcal{S}=\{1,\ldots,k\}$. For other sets $\mathcal{S}\subseteq\{1,\ldots,\mathsf{K}\}$ the proof is similar.

By Fano's inequality and because conditioning can only reduce entropy, there exists a sequence of real numbers $\{\epsilon_n\}_{n=1}^{\infty}$ with

$$\frac{\epsilon_n}{n} \xrightarrow[n\to\infty]{} 0,$$

such that

$$\begin{cases} H(W_{d_1}|Y_1^n,V_1) \leq \dfrac{\epsilon_n}{2\mathsf{K}}, \\ H(W_{d_2}|Y_2^n,V_1,V_2,W_{d_1}) \leq \dfrac{\epsilon_n}{2\mathsf{K}}, \\ \quad\vdots \\ H(W_{d_k}|Y_k^n,V_1,\ldots,V_k,W_{d_1},\ldots,W_{d_{k-1}}) \leq \dfrac{\epsilon_n}{2\mathsf{K}}. \end{cases}$$

Thus,

$$\begin{aligned}
nR &= H(W_{d_1}) \\
&= H(W_{d_1}|Z^n) + I(W_{d_1};Z^n) \\
&\leq H(W_{d_1}|Z^n) + \frac{\epsilon_n}{2} \\
&\leq I(W_{d_1};Y_1^n,V_1) - I(W_{d_1};Z^n) + H(W_{d_1}|Y_1^n,V_1) + \frac{\epsilon_n}{2} \\
&\leq I(W_{d_1};Y_1^n,V_1) - I(W_{d_1};Z^n) + \epsilon_n \\
&\leq I(W_{d_1};Y_1^n|V_1) - I(W_{d_1};Z^n|V_1) + I(W_{d_1};V_1|Z^n) + \epsilon_n \\
&\overset{(a)}{=} \sum_{i=1}^n \Big[I(W_{d_1};Y_{1,i}|V_1,Y_1^{i-1}) - I(W_{d_1};Z_i|V_1,Z_{i+1}^n)\Big] \\
&\quad + n\mathsf{M}_1 + \epsilon_n \\
&\overset{(b)}{=} \sum_{i=1}^n \Big[I(W_{d_1};Y_{1,i}|V_1,Y_1^{i-1}) - I(W_{d_1};Z_i|V_1,Z_{i+1}^n)\Big] \\
&\quad + n\mathsf{M}_1 + \epsilon_n \\
&\quad + \sum_{i=1}^n \Big[I(Z_{i+1}^n;Y_{1,i}|Y_1^{i-1},V_1,W_{d_1}) \\
&\qquad - I(Y_1^{i-1};Z_i|Z_{i+1}^n,V_1,W_{d_1})\Big] \\
&= \sum_{i=1}^n \Big[I(W_{d_1},Z_{i+1}^n;Y_{1,i}|V_1,Y_1^{i-1}) \\
&\qquad - I(W_{d_1},Y_1^{i-1};Z_i|V_1,Z_{i+1}^n)\Big] + n\mathsf{M}_1 + \epsilon_n \\
&\overset{(c)}{=} \sum_{i=1}^n \Big[I(W_{d_1},Z_{i+1}^n;Y_{1,i}|V_1,Y_1^{i-1}) \\
&\qquad - I(W_{d_1},Y_1^{i-1};Z_i|V_1,Z_{i+1}^n)\Big] + n\mathsf{M}_1 + \epsilon_n \\
&\quad - \sum_{i=1}^n \Big[I(Z_{i+1}^n;Y_{1,i}|Y_1^{i-1},V_1) \\
&\qquad - I(Y_1^{i-1};Z_i|Z_{i+1}^n,V_1)\Big] \\
&= \sum_{i=1}^n \Big[I(W_{d_1};Y_{1,i}|V_1,Y_1^{i-1},Z_{i+1}^n) \\
&\qquad - I(W_{d_1};Z_i|V_1,Y_1^{i-1},Z_{i+1}^n)\Big] + n\mathsf{M}_1 + \epsilon_n \\
&\overset{(d)}{\leq} \sum_{i=1}^n \Big[I(W_{d_1};Y_{1,i}|V_1,Y_1^{i-1},Z_{i+1}^n) \\
&\qquad - I(W_{d_1};Z_i|V_1,Y_1^{i-1},Z_{i+1}^n)\Big]^+ + n\mathsf{M}_1 + \epsilon_n \\
&\quad + \sum_{i=1}^n \Big[I(V_1,Y_1^{i-1},Z_{i+1}^n;Y_{1,i}) \\
&\qquad - I(V_1,Y_1^{i-1},Z_{i+1}^n;Z_i)\Big]^+ \\
&\overset{(e)}{=} \sum_{i=1}^n \Big[I(W_{d_1},V_1,Y_1^{i-1},Z_{i+1}^n;Y_{1,i}) \\
&\qquad - I(W_{d_1},V_1,Y_1^{i-1},Z_{i+1}^n;Z_i)\Big]^+ \\
&\quad + n\mathsf{M}_1 + \epsilon_n \\
&\overset{(f)}{\leq} \sum_{i=1}^n \Big[I(W_{d_1},V_1,Y_1^{i-1},Z_{i+1}^n;Y_{1,i}) \\
&\qquad - I(W_{d_1},V_1,Y_1^{i-1},Z_{i+1}^n;Z_i)\Big]^+ \\
&\quad + n\mathsf{M}_1 + \epsilon_n \\
&\quad + \sum_{i=1}^n \Big[I(Y_2^{i-1};Y_{1,i}|W_{d_1},V_1,Y_1^{i-1},Z_{i+1}^n) \\
&\qquad - I(Y_2^{i-1};Z_i|W_{d_1},V_1,Y_1^{i-1},Z_{i+1}^n)\Big]^+ \\
&\overset{(g)}{=} \sum_{i=1}^n \Big[I(W_{d_1},V_1,Y_1^{i-1},Y_2^{i-1},Z_{i+1}^n;Y_{1,i}) \\
&\qquad - I(W_{d_1},V_1,Y_1^{i-1},Y_2^{i-1},Z_{i+1}^n;Z_i)\Big]^+ \\
&\quad + n\mathsf{M}_1 + \epsilon_n
\end{aligned}$$

[3]This convere proof remains valid if the strong secrecy constraint in (228b) is replaced by the weaker constraint $\frac{1}{n}I(W_1,\ldots,W_D;Z^n) \xrightarrow[n\to\infty]{} 0$.

$$\stackrel{(h)}{=} \sum_{i=1}^{n} \Big[I(W_{d_1}, V_1, Y_2^{i-1}, Z_{i+1}^n; Y_{1,i}) - I(W_{d_1}, V_1, Y_2^{i-1}, Z_{i+1}^n; Z_i) \Big]^+ + n\mathsf{M}_1 + \epsilon_n, \tag{229}$$

where (a) holds because $I(W_{d_1}; V_1|Z^n)$ is limited by the entropy of V_1, which cannot exceed $n\mathsf{M}_1$; (b) and (c) follow by the chain rule of mutual information and by applying Csiszar's sum-identity [27, pp. 25]; (d) holds because for all values of x we have: $x \le x^+$ and $0 \le x^+$; (e) and (g) hold because if the eavesdropper is degraded with respect to Receiver 1, then all $[\cdot]^+$ terms are positive and if Receiver 1 is degraded with respect to the eavesdropper then all these terms are zero; (f) holds because $0 \le x^+$ for all values of x; and (g) holds because Receiver 1 is degraded with respect to Receiver 2 and thus the following Markov chain holds:

$$(V_1, W_{d_1}, Z_{i+1}^n, Y_{1,i}, Z_i) \to Y_2^{i-1} \to Y_1^{i-1}.$$

Let Q be a random variable uniform over $\{1, \ldots, n\}$ and independent of all the previously defined random variables. We define the following random variables:

$$U_1 := \Big(W_{d_1}, V_1, Y_2^{Q-1}, Z_{Q+1}^n \Big), \tag{230}$$

$$Y_1 := Y_{1,Q}, \tag{231}$$

$$Z := Z_Q. \tag{232}$$

Dividing by n, we can now rewrite (229) as

$$R \le \big[I(U_1; Y_1|Q) - I(U_1; Z|Q) \big]^+ + \mathsf{M}_1 + \frac{\epsilon_n}{n}. \tag{233}$$

We now derive a similar bound as before, but involving Receivers $1, \ldots, k$. Consider

$$\begin{aligned}
kR &\le \frac{1}{n} H(W_{d_1}, \ldots, W_{d_k}) \\
&= \frac{1}{n} H(W_{d_1}, \ldots, W_{d_k}|Z^n) + \frac{1}{n} I(W_{d_1}, \ldots, W_{d_k}; Z^n) \\
&\le \frac{1}{n} H(W_{d_1}, \ldots, W_{d_k}|Z^n) + \frac{\epsilon_n}{2n} \\
&\stackrel{(a)}{\le} \frac{1}{n} \big[H(W_{d_1}) + H(W_{d_2}|W_{d_1}) + \ldots + H(W_{d_k}|W_{d_{k-1}}, \ldots, W_{d_1}) - I(W_{d_1}, \ldots, W_{d_k}; Z^n) \big] \\
&\quad + \frac{\epsilon_n}{2n} \\
&\stackrel{(b)}{\le} \frac{1}{n} \Big[I(W_{d_1}; Y_1^n, V_1) + I(W_{d_2}; Y_2^n, V_1, V_2|W_{d_1}) + \ldots + I(W_{d_k}; Y_k^n, V_1, \ldots, V_k|W_{d_1}, \ldots, W_{d_{k-1}}) - I(W_{d_1}, \ldots, W_{d_k}; Z^n) \Big] \\
&\quad + \frac{\epsilon_n}{n} \\
&\stackrel{(c)}{=} \frac{1}{n} \Big[I(W_{d_1}; Y_1^n, V_1) - I(W_{d_1}; Z^n) \Big] \\
&\quad + \frac{1}{n} \sum_{\ell=2}^{k} \Big[I(W_{d_\ell}; Y_\ell^n, V_1, \ldots, V_\ell|W_{d_1}, \ldots, W_{d_{\ell-1}}) - I(W_{d_\ell}; Z^n|W_{d_1}, \ldots, W_{d_{\ell-1}}) \Big] \\
&\quad + \frac{\epsilon_n}{n} \\
&= \frac{1}{n} \Big[I(W_{d_1}; Y_1^n|V_1) - I(W_{d_1}; Z^n|V_1) + I(W_{d_1}; V_1|Z^n) \Big] \\
&\quad + \frac{1}{n} \sum_{\ell=2}^{k} \Big[I(W_{d_\ell}; Y_\ell^n|V_1, \ldots, V_\ell, W_{d_1}, \ldots, W_{d_{\ell-1}}) - I(W_{d_\ell}; Z^n|V_1, \ldots, V_\ell, W_{d_1}, \ldots, W_{d_{\ell-1}}) \\
&\qquad + I(W_{d_\ell}; V_1, \ldots, V_\ell|W_{d_1}, \ldots, W_{d_{\ell-1}}, Z^n) \Big] \\
&\quad + \frac{\epsilon_n}{n},
\end{aligned} \tag{234}$$

where (a) follows from the chain rule of mutual information; (b) follows from Fano's inequality; and (c) follows from the chain rule of mutual information.

In a similar way to (233), we can prove that

$$\frac{1}{n} \Big[I(W_{d_1}; Y_1^n|V_1) - I(W_{d_1}; Z^n|V_1) \Big] \le \big[I(U_1; Y_1|Q) - I(U_1; Z|Q) \big]^+. \tag{235}$$

Then, we prove that for each $\ell \in \{2, \ldots, k\}$, the following set of inequalities holds:

$$\begin{aligned}
& I(W_{d_\ell}; Y_\ell^n|V_1, \ldots, V_\ell, W_{d_1}, \ldots, W_{d_{\ell-1}}) - I(W_{d_\ell}; Z^n|V_1, \ldots, V_\ell, W_{d_1}, \ldots, W_{\ell-1}) \\
&\stackrel{(a)}{=} \sum_{i=1}^{n} \big[I(W_{d_\ell}; Y_{\ell,i}|V_1, \ldots, V_\ell, W_{d_1}, \ldots, W_{d_{\ell-1}}, Y_\ell^{i-1}, Z_{i+1}^n) - I(W_{d_\ell}; Z_i|V_1, \ldots, V_\ell, W_{d_1}, \ldots, W_{d_{\ell-1}}, Y_\ell^{i-1}, Z_{i+1}^n) \big] \\
&\stackrel{(b)}{=} \sum_{i=1}^{n} \big[I(W_{d_\ell}; Y_{\ell,i}|V_1, \ldots, V_\ell, W_{d_1}, \ldots, W_{d_{\ell-1}}, Y_2^{i-1}, \ldots, Y_\ell^{i-1}, Z_{i+1}^n) \\
&\qquad - I(W_{d_\ell}; Z_i|V_1, \ldots, V_\ell, W_{d_1}, \ldots, W_{d_{\ell-1}}, Y_2^{i-1}, \ldots, Y_\ell^{i-1}, Z_{i+1}^n) \big] \\
&\stackrel{(c)}{\le} \sum_{i=1}^{n} \big[I(W_{d_\ell}; Y_{\ell,i}|V_1, \ldots, V_\ell, W_{d_1}, \ldots, W_{d_{\ell-1}}, Y_2^{i-1}, \ldots, Y_\ell^{i-1}, Z_{i+1}^n) \\
&\qquad - I(W_{d_\ell}; Z_i|V_1, \ldots, V_\ell, W_{d_1}, \ldots, W_{d_{\ell-1}}, Y_2^{i-1}, \ldots, Y_\ell^{i-1}, Z_{i+1}^n) \big]^+ \\
&\quad + \sum_{i=1}^{n} \big[I(V_\ell; Y_{\ell,i}|V_1, \ldots, V_{\ell-1}, W_{d_1}, \ldots, W_{d_{\ell-1}}, Y_2^{i-1}, \ldots, Y_\ell^{i-1}, Z_{i+1}^n) \\
&\qquad - I(V_\ell; Z_i|V_1, \ldots, V_{\ell-1}, W_{d_1}, \ldots, W_{d_{\ell-1}}, Y_2^{i-1}, \ldots, Y_\ell^{i-1}, Z_{i+1}^n) \big]^+ \\
&\stackrel{(d)}{\le} \sum_{i=1}^{n} \big[I(W_{d_\ell}, V_\ell; Y_{\ell,i}|V_1, \ldots, V_{\ell-1}, W_{d_1}, \ldots, W_{d_{\ell-1}}, Y_2^{i-1}, \ldots, Y_\ell^{i-1}, Z_{i+1}^n) \\
&\qquad - I(W_{d_\ell}, V_\ell; Z_i|V_1, \ldots, V_{\ell-1}, W_{d_1}, \ldots, W_{d_{\ell-1}}, Y_2^{i-1}, \ldots, Y_\ell^{i-1}, Z_{i+1}^n) \big]^+ \\
&\quad + \sum_{i=1}^{n} \big[I(Y_{\ell+1}^{i-1}; Y_{\ell,i}|V_1, \ldots, V_\ell, W_{d_1}, \ldots, W_{d_\ell},
\end{aligned}$$

$$
\begin{aligned}
&\qquad\qquad Y_2^{i-1},\dots,Y_\ell^{i-1},Z_{i+1}^n)\\
&\quad - I(Y_{\ell+1}^{i-1};Z_i|V_1,\dots,V_\ell,W_{d_1},\dots,W_{d_\ell},\\
&\qquad\qquad Y_2^{i-1},\dots,Y_\ell^{i-1},Z_{i+1}^n)\big]^+\\
&=\sum_{i=1}^n \big[I(W_{d_\ell},V_\ell,Y_{\ell+1}^{i-1};Y_{\ell,i}|V_1,\dots,V_{\ell-1},W_{d_1},\dots,W_{d_{\ell-1}},\\
&\qquad\qquad Y_2^{i-1},\dots,Y_\ell^{i-1},Z_{i+1}^n)\\
&\quad - I(W_{d_\ell},V_\ell,Y_{\ell+1}^{i-1};Z_i|V_1,\dots,V_{\ell-1},W_{d_1},\dots,\\
&\qquad\qquad W_{d_{\ell-1}},Y_2^{i-1},\dots,Y_\ell^{i-1},Z_{i+1}^n)\big]^+,
\end{aligned}
\tag{236}
$$

where (a) follows from the chain rule of mutual information and by applying Csiszar's sum-identity; (b) because Receivers $1,\dots,\ell-1$ are degraded with respect to Receiver ℓ, and so the following Markov chain holds:

$$
\begin{aligned}
(W_{d_\ell},Y_{\ell,i},V_1,\dots,V_k,W_{d_1},\dots,W_{d_{\ell-1}},Z_{i+1}^n)\to Y_\ell^{i-1}\\
\to (Y_1^{i-1},\dots,Y_{\ell-1}^{i-1});
\end{aligned}
\tag{237}
$$

(c) holds because for all values of x we have: $x \le x^+$ and $0 \le x^+$; (d) holds because if the eavesdropper is degraded with respect to Receiver ℓ, then all $[\cdot]^+$ terms are positive and if Receiver ℓ is degraded with respect to the eavesdropper than all these terms are zero.

We define for each $k \in \{2,\dots,\mathsf{K}\}$ the random variables

$$
Y_k := Y_{k,Q} \tag{238}
$$

$$
U_k := (W_{d_k},V_k,Y_{k+1}^{Q-1},U_{k-1}). \tag{239}
$$

Dividing by n, we can rewrite constraint (236) as

$$
\begin{aligned}
&\frac{1}{n}\Big[I(W_\ell;Y_\ell^n|V_1,\dots,V_\ell,W_{d_1},\dots,W_{d_{\ell-1}})\\
&\qquad - I(W_\ell;Z^n|V_1,\dots,V_\ell,W_{d_1},\dots,W_{d_{\ell-1}})\Big]\\
&\qquad \le \sum_{\ell=1}^k \big[I(U_\ell;Y_\ell|U_{\ell-1},Q)-I(U_\ell;Z|U_{\ell-1},Q)\big]^+.
\end{aligned}
\tag{240}
$$

Finally, we bound the following sum:

$$
\begin{aligned}
&I(W_{d_1};V_1|Z^n)+\sum_{\ell=2}^k I(W_{d_\ell};V_1,\dots,V_\ell|W_{d_1},\dots,W_{d_\ell-1},Z^n)\\
&\quad \le I(W_{d_1};V_1\dots,V_k|Z^n)\\
&\qquad + \sum_{\ell=2}^k I(W_{d_\ell};V_1,\dots,V_k|W_{d_1},\dots,W_{d_\ell-1},Z^n)\\
&\quad = I(W_{d_1},\dots,W_{d_k};V_1,\dots,V_k|Z^n)\\
&\quad \le n\sum_{\ell=1}^k \mathsf{M}_\ell.
\end{aligned}
\tag{241}
$$

Taking into consideration constraints (235), (240) and (241), we can rewrite constraint (234) as:

$$
\begin{aligned}
kR &\le \sum_{\ell=1}^k \Big[I(U_\ell;Y_\ell|U_{\ell-1},Q)-I(U_\ell;Z|U_{\ell-1},Q)\Big]^+\\
&\quad + \sum_{\ell=1}^k \mathsf{M}_\ell + \frac{\epsilon_n}{n},
\end{aligned}
\tag{242}
$$

where U_0 is a constant.

Letting $n \to \infty$, from constraints (233) and (242), we conclude that Lemma 2 holds.

References

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

[2] R. Timo and M. Wigger, "Joint cache-channel coding over erasure broadcast channels," *Proc. of IEEE Intern. Symp. on Wireless Comm. Systems (ISWCS),* Bruxelles, Belgium, Aug. 2015, pp. 201–205.

[3] S. Saeedi 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.

[4] A. S. Cacciapuoti, M. Caleffi, M. Ji, J. Llorca, and A. M. Tulino,"Speeding up future video distribution via channel-aware caching-aided coded multicast," *IEEE JSAC in Comm.*, vol. 34, no. 8, pp. 2207–2218, Aug. 2016.

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

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

[7] V. Ravindrakumar, P. Panda, N. Karamchandani, and V. Prabhakarany, "Fundamental limits of secretive coded caching," *Proc. of IEEE Intern. Symp. on Inf. Theory (ISIT)*, Barcelona, Spain, Jul. 2016, pp. 425–429.

[8] H. H. Suthan C, I. Chugh, and P. Krishnan, "An improved secretive coded caching scheme exploiting common demands," *IEEE Trans. Veh. Tech.* vol. 66 , no. 11 , pp. 10249–10258, May 2017.

[9] A. A. Zewail and A. Yener, "Coded caching for resolvable networks with security requirements," *Proc. of IEEE Conference on Communications and Network Security (CNS)*, Philadelphia, PA USA, Oct. 2016, pp. 621–625.

[10] Z. H. Awan and A. Sezgin, "Fundamental limits of caching in D2D networks with secure delivery," *Proc. of IEEE International Conference on Communication Workshop (ICCW)*, London, UK, Jun. 2015, pp. 464–469.

[11] F. Gabry, V. Bioglio and I. Land, "On edge caching with secrecy constraints," *Proc. of IEEE International Conference on Communications (ICC)*, Kuala Lumpur, Malaysia, May 2016.

[12] M. K. Kiskaniy and H. R. Sadjadpoury, "A secure approach for caching contents in wireless ad hoc networks," *ArXiv:1709.00132*, Sep. 2017.

[13] L. Xiang, D. W. Kwan Ng, R. Schober, and V. Wong, "Cache-enabled physical layer security for video streaming in backhaul-limited cellular networks," *IEEE Trans. Wireless Comm.*, vol. 17, no. 2, pp. 736–751, Feb. 2018.

[14] F. Engelmann and P. Elia, "A content-delivery protocol, exploiting the privacy benefits of coded caching," *Proc. of 15th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Paris, France, May 2017.

[15] S. Kamel, M. Sarkiss and M. Wigger, "Secure joint cache-channel coding over erasure broadcast channels," *Proc. of IEEE Wireless Communications and Networking Conf. (WCNC)*, San Francisco, CA, Mar. 2017.

[16] S. Kamel, M. Sarkiss, and M. Wigger, "Achieving joint secrecy with cache-channel coding over erasure broadcast channels," *Proc. of IEEE International Conference on Communications (ICC)*, Paris, France, May 2017.

[17] E. Ekrem and S. Ulukus, "Multi-receiver wiretap channel with public and confidential messages," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2165–2177, Apr. 2013.

[18] C. E. Shannon. "Communication theory of secrecy systems," *Bell system technical journal,* vol. 28, no. 4, pp. 656–715, 1949.

[19] H. Yamamoto, "Rate-distortion theory for the Shannon cipher system," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 827–835, May 1997.

[20] W. Kang, and N. Liu. "Wiretap channel with shared key," *Proc. of 2010 ITW*, Dublin, UK, 30. Aug.–3. Sep. 2010, pp. 1-5.

[21] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1562–1575, Apr. 2006.

[22] Y.-K. Chia and A. El Gamal, "3-receiver broadcast channels with common and confidential messages," *IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, Korea, pp. 1849–1853, Jun 2009.

[23] A. Mansour, R. Schaefer, and H. Boche, “On the capacity of broadcast channels with degraded message sets and message cognition under different secrecy constraints.” *ArXiv:1501.04490*, Jan. 2015.
[24] R. F. Schaefer, A. Khisti, and H. V. Poor, “Secure broadcasting using independent secret keys,” *IEEE Trans. Comm.*, vol. 66, no. 2, pp. 644-661, Feb. 2018.
[25] S. Saeedi Bidokhti, M. Wigger and A. Yener, “Benefits of cache assignment on degraded broadcast channels.” *ArXiv:1702.08044*, Feb. 2017.
[26] E. Tuncel, “Slepian-Wolf coding over broadcast channels,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1469–1482, Apr. 2006.
[27] A. El Gamal and Y. H. Kim, Network Information Theory, 2011, Cambridge Univ. Press.

Sarah Kamel (S’17, M’18) obtained the engineering diploma in Communications and Computer Science from the Lebanese University, Faculty of Engineering, in 2013. She obtained the M.Sc. in Digital Communication Systems in 2013, and the Ph.D. degree in Communications and Electronics in 2017, both from Telecom ParisTech. In 2013-2017, she was a Ph.D. candidate at CEA LIST Communicating Systems Lab, working on cryptographic and information-theoretic aspects of security in wireless communications. In 2017-2018, she was a PostDoctoral researcher in the Digital Communications group at Telecom ParisTech. Her research interests include coded caching, physical layer security, polar and lattice codes, and lattice-based cryptography.

Mireille Sarkiss (S05M09) received her degree in electrical engineering from Lebanese University, Faculty of Engineering, Lebanon, in 2003, and the M.S. and Ph.D. degrees in communications and electronics from Telecom ParisTech, France, in 2004 and 2009, respectively. From February 2009 to June 2010, she pursued postdoctoral research at the Department of Communications and Electronics, Telecom ParisTech. From October 2010 to November 2018, she was with the Communicating Systems Laboratory, CEA, LIST, France, as a Research Engineer. In December 2018, she joined Telecom SudParis, France, as an assistant professor. Her research interests include wireless communications, particularly, cooperative communications and energy-efficient communications, lattice coding and decoding, distributed coding, resource allocation and computation offloading in Mobile Edge Computing and secrecy capacity in cache-aided networks.

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

Ghaya Rekaya-Ben Othman was born in Tunisia, in 1977. She received the degree in electrical engineering from ENIT, Tunisia, in 2000, and the Ph.D. degree from Telecom ParisTech (ex-ENST), France, in 2004. In 2005, she joined, the Department of Communications and Electronics, at Telecom-ParisTech as an Assistant Professor. Since 2012, she has been a Full Professor at Telecom-ParisTech. She is the recipient of the City of Paris Award of the Best Woman Scientist in 2007. Her research work concerns : Coding for MIMO systems, Decoding for MIMO systems, Cooperative communications and Optical fiber communications.