Decentralized joint cache-channel coding over erasure broadcast channels

Sarah Kamel, Mireille Sarkiss, Michèle Wigger

Decentralized Joint Cache-Channel Coding over Erasure Broadcast Channels

Sarah Kamel
LTCI, Télécom ParisTech
Université Paris-Saclay
Paris, France
sarah.kamel@telecom-paristech.fr

Mireille Sarkiss
Communicating Systems Laboratory
CEA, LIST
Gif-sur-Yvette, France
mireille.sarkiss@cea.fr

Michèle Wigger
LTCI, Télécom ParisTech
Université Paris-Saclay
Paris, France
michele.wigger@telecom-paristech.fr

***Abstract*—We derive upper bounds on the rate-memory tradeoff of cache-aided erasure broadcast channels with K_w weak receivers and K_s strong receivers. We follow a decentralized placement scenario, where coordination is not needed prior to the delivery phase. We study two setups: a standard scenario without eavesdropper and a wiretap scenario with an external eavesdropper. For both scenarios, we propose joint cache-channel coding schemes that efficiently exploit the cache contents and take into consideration the users' channel characteristics at the same time. We show that the decentralized placement strategy causes only a small increase in delivery rate compared to centralized strategy. Similarly, when cache sizes are moderate, the rate is increased only slightly by securing the communication against external eavesdroppers. This is not the case when cache memories are small and large.**

***Index Terms*—Coded caching, receiver caching, joint cache-channel coding, erasure broadcast channels, weak secrecy.**

I. Introduction

Coded caching was introduced by Maddah-Ali and Nieson in [1], [2] to reduce network congestion during periods of peak-traffic by prestoring fragments of popular contents at users' caches during off-peak hours. Designing carefully the placement phase allowed them to set a coded-multicasting delivery phase where the transmitter serves simultaneously multiple users while benefiting from their cached contents. In these works, the delivery communication takes place over an error-free broadcast channel (BC) and all users have equal cache sizes. In addition, centralized and decentralized placement strategies were studied showing significant gains over traditional caching with uncoded unicast transmission. In the centralized setup [1], the central server is aware of the active users in the network. Yet without any knowledge on the users' demands, it coordinates the placement of fragments of all possibly requested files in the receivers' cache memories. Decentralized algorithms lack this coordination [2] because they cannot depend on the set of active users. Indeed, in practice, the identities or the number of active users may not be known to the server ahead delivery. This might be due to users' mobility and their connection to another server between placement and delivery phases. Therefore, in the decentralized placement scenario, the users fill their caches randomly and independently from each other with an equal number of bits. Then, prior to the delivery phase, the server is informed about the active users in the network, the caching contents and the requests of each user. Various extensions of the decentralized coded caching scheme have since then been proposed, for example in [3]–[5]. In particular, Sengupta et al. [3] combined decentralized coded caching with a secure delivery scheme that XORs the delivery messages with prestored random keys to secure the communication from an external eavesdropper. All these works assumed as [2] that the delivery communication occurs over a common noise-free bit-pipe to all receivers.

In this paper, we investigate decentralized coded caching with and without secrecy constraint when the delivery phase takes place over an erasure BC with a set of weak receivers that have cache memories and a set of strong receivers without cache memories. The centralized caching counterparts of these setups were studied in [6], [7] and in our previous works [8]–[10]. As in [8], in this paper communication needs to be kept secret from an external eavesdropper that does not have access to the cache memories but to a degraded version of the channel outputs at the weak receivers. The eavesdropper is not allowed to learn any information about each of the messages *individually*. A stronger secrecy constraint where the eavesdropper is not allowed to learn any information about all the possibly requested messages was considered in [9], [10].

We present coding schemes and upper bounds on the delivery rate-memory tradeoff, i.e., the smallest rate of transmission for given cache sizes, for the setups with and without secrecy constraint. Our coding schemes build on a combination of various piggyback codes [6], [7] in the standard setup and a version thereof involving secure piggyback coding [8] and one-time pads with non-requested message bits in the secure setup. Piggyback coding is a joint cache-channel coding scheme where the encoding and decoding operations exploit simultaneously the cache contents and the channel statistics, improving thus further the caching gains compared to separate cache-channel coding schemes [2].

Unlike the secure caching schemes of [3], we resign from placing secret keys in the cache memories. In fact, since we require only that each file is individually kept secure, such secret keys are not very helpful. They can simply be replaced by bits of files that are not requested because XORs of bits belonging to different messages are secure by definition. From a practical point of view, caching secret keys is not

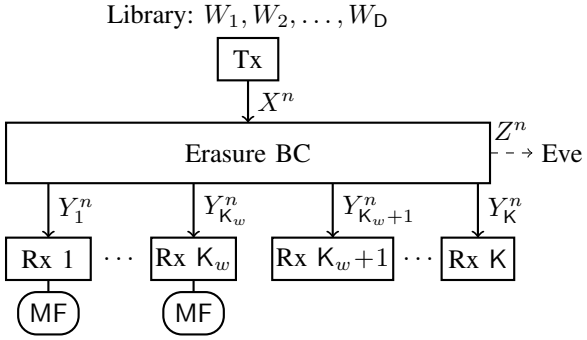

Fig. 1. Erasure BC with K_w weak receivers with cache memories of size MF bits and $\mathsf{K}_s = \mathsf{K} - \mathsf{K}_w$ strong receivers without caches. We study the scenarios with and without eavesdropper, depicted with a dashed line.

desirable because it requires certain coordination between the transmitter and some receivers before the delivery phase [3], which otherwise is not required in decentralized caching.

II. Problem Definition

We first describe the standard scenario without eavesdropper and then the scenario where communication needs to be kept secret from an external eavesdropper.

A. Standard Scenario without Eavesdropper

We consider an erasure BC with a single transmitter and K receivers as shown in Figure 1 without the eavesdropper. The input alphabet of the BC is $\mathcal{X} := \{0,1\}$ and all receivers have the same output alphabet $\mathcal{Y} := \mathcal{X} \cup \Delta$, where Δ indicates the loss of a bit at a receiver. The K receivers are partitioned into two sets $\mathcal{K}_w := \{1, \ldots, \mathsf{K}_w\}$ and $\mathcal{K}_s := \{\mathsf{K}_w + 1, \ldots, \mathsf{K}\}$, and the receivers in each of the two sets have same channel statistics. Specifically, the K_w receivers in set $\mathcal{K}_w$ are weak and have the same erasure probability $\delta_w > 0$ whereas the $\mathsf{K}_s = \mathsf{K} - \mathsf{K}_w$ receivers in $\mathcal{K}_s$ are strong and have the same erasure probability $\delta_s > 0$ such that

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

Each weak receiver has access to a local cache memory of size MF bits, while strong receivers have no cache memories.

The transmitter can access a library of $\mathsf{D} > \mathsf{K}$ independent files (messages) $W_1, \ldots, W_{\mathsf{D}}$, each consisting of F i.i.d. random bits. We denote the b-th bit of file W_d by $W_{d,b}$. Every receiver $k \in \mathcal{K} := \{1, \ldots, \mathsf{K}\}$ demands exactly one file W_{d_k} from the library. So, $d_k \in \mathcal{D}$ describes the demand of Receiver k, and $\mathbf{d} := (d_1, \ldots, d_{\mathsf{K}}) \in \mathcal{D}^{\mathsf{K}}$ the *demand vector* of all the receivers.

Communication takes place in two phases: a *decentralized placement phase* where each weak receiver fills its cache memory with randomly chosen bits from the library and a *centralized delivery phase* where the demanded files W_{d_k}, for $k \in \mathcal{K}$, are conveyed to the receivers. During the placement phase, the demand vector $\mathbf{d}$ is unknown to the transmitter and the receivers. As is standard for decentralized caching, the cache placement at a given receiver cannot depend on the number of receivers K in the system. That means each weak receiver $i \in \mathcal{K}_w$ computes its cache content V_i by means of a randomized placement function $g_i : \{1, \ldots, 2^{\mathsf{F}}\}^{\mathsf{D}} \to \{1, \ldots, 2^{\mathsf{MF}}\}$ that does not depend on K:

$$V_i := g_i(W_1, \ldots, W_{\mathsf{D}}). \tag{2}$$

Prior to the delivery phase, the demand vector $\mathbf{d}$ as well as the realization of all randomized placement functions $g_1, \ldots, g_{\mathsf{K}_w}$ are learned by the transmitter and all the legitimate receivers. The delivery phase is of length n. That means, for a given demand vector $\mathbf{d}$, the transmitter sends inputs

$$X^n = f_{\mathbf{d}}(W_1, \ldots, W_{\mathsf{D}}), \tag{3}$$

for some choice of the encoding function $f_{\mathbf{d}} : \{1, \ldots, 2^{\mathsf{F}}\}^{\mathsf{D}} \to \mathcal{X}^n$ that can depend on the demand vector $\mathbf{d}$ as well as on the realizations of the placement functions $g_1, \ldots, g_{\mathsf{K}_w}$.

Each weak receiver $i \in \mathcal{K}_w$ decodes its demanded message W_{d_i} based on the observed outputs $Y_i^n := (Y_{i,1}, \ldots, Y_{i,n})$ and its cache content V_i:

$$\hat{W}_i := \varphi_i(Y_i^n, V_i), \qquad i \in \mathcal{K}_w, \tag{4}$$

for some function $\varphi_i : \mathcal{Y}^n \times \mathcal{V} \to \{1, \ldots, 2^{\mathsf{F}}\}$. Each strong receiver $j \in \mathcal{K}_s$ decodes its demanded message based only on the observed outputs Y_j^n:

$$\hat{W}_j := \varphi_j(Y_j^n), \qquad j \in \mathcal{K}_s, \tag{5}$$

for some function $\varphi_j : \mathcal{Y}^n \to \{1, \ldots, 2^{\mathsf{F}}\}$. Notice that all functions $\varphi_1^{(n)}, \ldots, \varphi_{\mathsf{K}}^{(n)}$ can depend on the demand vector $\mathbf{d}$ and the realizations of the placement functions $g_1, \ldots, g_{\mathsf{K}_w}$.

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

$$\mathrm{P}_e^{\text{Worst}} := \max_{\mathbf{d} \in \mathcal{D}^{\mathsf{K}}} \mathrm{P}\left[\bigcup_{k=1}^{\mathsf{K}} \left\{\hat{W}_k \neq W_{d_k}\right\}\right]. \tag{6}$$

Definition 1. *A rate-memory pair (R, M) is* achievable *for the described setup, if for every $\epsilon > 0$ and sufficiently large blocklength n, there exist caching, encoding, and decoding functions so that*

$$\mathrm{P}_e^{\text{Worst}} \leq \epsilon \tag{7}$$

Definition 2. *For cache memory size MF, the* rate-memory tradeoff *$R^\star(\mathsf{M})$ is the smallest rate R so that the pair (R, M) is achievable:*

$$R^\star(\mathsf{M}) := \inf\left\{R\colon\ (R, \mathsf{M}) \text{ achievable}\right\}. \tag{8}$$

A first main goal of this paper is to provide a good upper bound on $R^\star(\mathsf{M})$.

B. Scenario with an Eavesdropper

We also consider the scenario with an external eavesdropper Eve in Figure 1. The eavesdropper has no access to the cache memories. During the delivery phase, it observes the channel outputs $Z^n := (Z_1, \ldots, Z_n)$, where each Z_t is the output of a binary erasure channel from the input X_t. For simplicity,

the eavesdropper is assumed to be weaker than all legitimate receivers. So its erasure probability δ_z satisfies

$$0 < \delta_s \leq \delta_w \leq \delta_z \leq 1. \tag{9}$$

Communication needs to be kept secret from the eavesdropper in the sense that Z^n should provide no information about any of the messages $W_1, \ldots, W_{\mathsf{D}}$ *individually*. To avoid confusion, and emphasize its dependence on the secrecy constraint, we call the length of the delivery phase in this setup n_{sec} and the corresponding delivery rate $R_{\text{sec}} := \frac{n_{\text{sec}}}{\mathsf{F}}$.

Definition 3. *A rate-memory pair* $(R_{\text{sec}}, \mathsf{M})$ *is* securely achievable *if for every* $\epsilon > 0$ *and sufficiently large blocklength* n_{sec}*, there exist caching, encoding, and decoding functions so that the worst-case error probability defined in* (6) *satisfies:*

$$\mathrm{P}_e^{\text{Worst}} \leq \epsilon, \tag{10a}$$

and the described individual secrecy constraint holds:

$$\frac{1}{n} I(W_d; Z^n) < \epsilon, \qquad \forall d \in \mathcal{D}. \tag{10b}$$

Definition 4. *For cache memory size* MF*, the* secrecy rate-memory tradeoff $R^\star_{\text{sec}}(\mathsf{M})$ *is the smallest rate* R_{sec} *so that the pair* $(R_{\text{sec}}, \mathsf{M})$ *is securely achievable:*

$$R^\star_{\text{sec}}(\mathsf{M}) := \inf\big\{R_{\text{sec}}\colon\ (R_{\text{sec}}, \mathsf{M}) \text{ securely achievable}\big\}. \tag{11}$$

A second main goal of this paper is to provide a good upper bound on $R^\star_{\text{sec}}(\mathsf{M})$.

III. Non-Secure Decentralized Caching

Consider the standard scenario without secrecy constraint. Define for each $\ell \in \{0, 1, \ldots, \mathsf{K}_w\}$:

$$\gamma^{(\ell)} := \left(\frac{\mathsf{M}}{\mathsf{D}}\right)^{\ell} \left(1 - \frac{\mathsf{M}}{\mathsf{D}}\right)^{\mathsf{K}_w - \ell}. \tag{12}$$

Theorem 1. *The rate-memory tradeoff is upper bounded as:*

$$R^\star(\mathsf{M}) \leq \frac{\sum_{\ell=1}^{\mathsf{K}_w} \binom{\mathsf{K}_w}{\ell} \gamma^{(\ell-1)}}{1-\delta_w} + \frac{\mathsf{K}_s \gamma^{(0)}}{1-\delta_s} + \frac{\sum_{\ell=1}^{\mathsf{K}_w} \binom{\mathsf{K}_w}{\ell} \left[\mathsf{K}_s(1-\delta_w)\gamma^{(\ell)} - (\delta_w - \delta_s)\gamma^{(\ell-1)}\right]^+}{(1-\delta_w)(1-\delta_s)}. \tag{13}$$

Proof. Based on the schemes in Sections III-A and III-B. □

Figure 2 illustrates the upper bound in Theorem 1 at hand of an example. The figure also shows the performance of a separation-based scheme that combines the Maddah-Ali & Niesen decentralized coded caching scheme [2] for the weak receivers with a stand-alone capacity-achieving code for erasure BCs. This code delivers to the weak receivers the XORs produced by the coded caching scheme, and to the strong receivers their requested files. The lower-most line shows the performance of the centralized scheme in [7].

Figure 2 shows that joint cache-channel coding reduces the delivery rate significantly compared to separate coding.

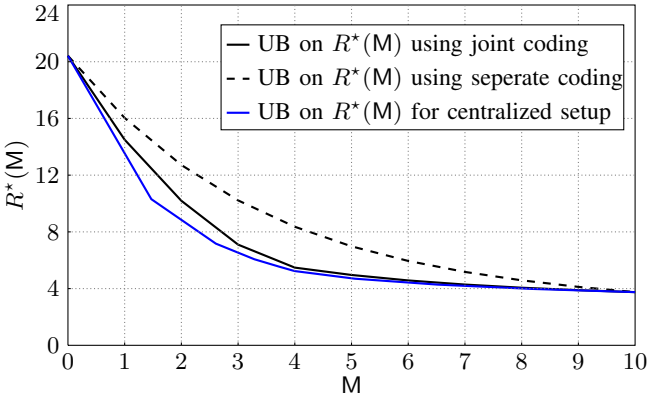

Fig. 2. Upper bounds on $R^\star(\mathsf{M})$ for $\delta_w = 0.7$, $\delta_s = 0.2$, $\mathsf{K}_w = 5$, $\mathsf{K}_s = 3$, and $\mathsf{D} = 10$.

It also shows that the increase in the delivery rate caused by decentralized caching is small.

A. Scheme for $\mathsf{K}_w = 3$ and $\mathsf{K}_s = 1$

We first describe the scheme for $\mathsf{K}_w = 3$ and $\mathsf{K}_s = 1$. Fix a file size F and a small positive $\epsilon > 0$. Let

$$\mathsf{M}_\epsilon := \mathsf{M} - \epsilon. \tag{14}$$

Decentralized placement phase: For each $d \in \mathcal{D}$ and each $b \in \{1, \ldots, \mathsf{F}\}$, every weak receiver $i \in \mathcal{K}_w$ caches the b-th bit of file W_d, i.e., $W_{d,b}$, with probability $\frac{\mathsf{M}_\epsilon}{\mathsf{D}}$, independently of all other bits and of all other receivers. Notice that by the weak law of large numbers and by (14), the probability that any given weak receiver has stored no more than MF bits tends to 1 as $\mathsf{F} \to \infty$. So, the described cache placement satisfies the constraint on the cache sizes with very high probability.

For any given subset of receivers $G \subseteq \mathcal{K}_w$, define for each file W_d the bits stored exclusively at receivers in G:

$$W_{d,G} := \big\{W_{d,b}\colon W_{d,b} \text{ cached exclusively at receivers in } G\big\}.$$

The cache content at the three weak receivers is:

$$V_1 = \big\{W_{d,\{1\}}, W_{d,\{1,2\}}, W_{d,\{1,3\}}, W_{d,\{1,2,3\}}\big\}_{d=1}^{\mathsf{D}}, \tag{15a}$$

$$V_2 = \big\{W_{d,\{2\}}, W_{d,\{1,2\}}, W_{d,\{2,3\}}, W_{d,\{1,2,3\}}\big\}_{d=1}^{\mathsf{D}}, \tag{15b}$$

$$V_3 = \big\{W_{d,\{3\}}, W_{d,\{1,3\}}, W_{d,\{2,3\}}, W_{d,\{1,2,3\}}\big\}_{d=1}^{\mathsf{D}}. \tag{15c}$$

Notice that by the weak law of large numbers, for each $G \subseteq \mathcal{K}_w$ and each $d \in \mathcal{D}$:

$$\frac{|W_{d,G}|}{\mathsf{F}} \to \left(\frac{\mathsf{M}_\epsilon}{\mathsf{D}}\right)^{|G|} \left(1 - \frac{\mathsf{M}_\epsilon}{\mathsf{D}}\right)^{\mathsf{K}_w - |G|} \quad \text{as } \mathsf{F} \to \infty. \tag{16}$$

Thus, for large file sizes F and for any file W_d and $\ell \in \{1, 2, 3\}$, approximately the same number of bits is exclusively cached at any subset of ℓ weak receivers. We therefore define $\mathsf{F}^{(0)}, \mathsf{F}^{(1)}, \mathsf{F}^{(2)}, \mathsf{F}^{(3)}$ as the expected number of bits that are not cached at all, that are cached at a single weak receiver, at pairs of weak receivers, and at all three weak receivers, respectively. We have for $\ell \in \{0, \ldots, \mathsf{K}_w\}$:

$$\mathsf{F}^{(\ell)} := \gamma_\epsilon^{(\ell)} \mathsf{F} \quad \text{where} \quad \gamma_\epsilon^{(\ell)} := \left(\frac{\mathsf{M}_\epsilon}{\mathsf{D}}\right)^{\ell} \left(1 - \frac{\mathsf{M}_\epsilon}{\mathsf{D}}\right)^{\mathsf{K}_w - \ell}. \tag{17}$$

Notice that $\gamma_\epsilon^{(\ell)} \to \gamma^{(\ell)}$ as $\epsilon \to 0$.

Centralized delivery phase: Each receiver $k \in \{1,\dots,4\}$ demands message W_{d_k}. Before transmission starts, the transmitter divides each of the submessages $W_{d_4,\{1\}}$, $W_{d_4,\{2\}}, W_{d_4,\{3\}}, W_{d_4,\{1,2\}}, W_{d_4,\{2,3\}}, W_{d_4,\{1,3\}}, W_{d_4,\{1,2,3\}}$ intended to Receiver 4 into two parts

$$W_{d_4,\{i\}} = \left(W^{(1)}_{d_4,\{i\}}, W^{(2)}_{d_4,\{i\}}\right), \qquad i \in \{1,2,3\}, \tag{18}$$

$$W_{d_4,\{i,i'\}} = \left(W^{(1)}_{d_4,\{i,i'\}}, W^{(2)}_{d_4,\{i,i'\}}\right), \qquad i,i' \in \{1,2,3\} \text{ and } i \neq i', \tag{19}$$

$$W_{d_4,\{1,2,3\}} = \left(W^{(1)}_{d_4,\{1,2,3\}}, W^{(2)}_{d_4,\{1,2,3\}}\right). \tag{20}$$

The two parts in (18) are of sizes $\mathsf{F}^{(1,1)}$ and $\mathsf{F}^{(1,2)}$, the two parts in (19) of sizes $\mathsf{F}^{(2,1)}$ and $\mathsf{F}^{(2,2)}$, and the two parts in (20) of sizes $\mathsf{F}^{(3,1)}$ and $\mathsf{F}^{(3,2)}$, such that for $\ell = 1,2,3$,

$$\mathsf{F}^{(\ell,1)} = \min\left\{\mathsf{F}^{(\ell)}, \mathsf{F}^{(\ell-1)}\frac{\delta_w - \delta_s}{1-\delta_w}\right\} \tag{21}$$

$$\mathsf{F}^{(\ell,2)} = \max\left\{0, \mathsf{F}^{(\ell)} - \mathsf{F}^{(\ell-1)}\frac{\delta_w - \delta_s}{1-\delta_w}\right\}; \tag{22}$$

The delivery phase applies time-sharing over 4 subphases of lengths n_1, n_2, n_3 and n_4 bits, that sum up to n bits in total. Subphases 1 and 2 are further divided into three equally-long periods of $n_1/3$ and $n_2/3$ bits, respectively.

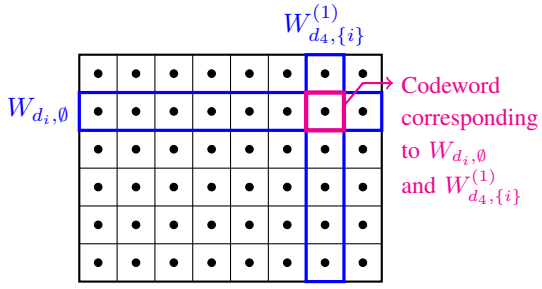

Fig. 3. Structure of standard piggyback codebook with codewords arranged in an array. Here the rows encode $W_{d_i,\emptyset}$ and the columns $W^{(1)}_{d_4,\{i\}}$.

Consider Subphase 1. For each of the three periods $i \in \{1,2,3\}$, the transmitter creates a standard piggyback [6] codebook

$$\mathcal{C} := \left\{\mathbf{x}_1(w_r, w_c): \quad w_r \in \left\{1,\dots,2^{\mathsf{F}^{(0)}}\right\}, \; w_c \in \left\{1,\dots,2^{\mathsf{F}^{(1,1)}}\right\}\right\}, \tag{23}$$

with entries drawn i.i.d. according to a Bernoulli-$1/2$ distribution. The codewords of such a codebook are arranged in an array with rows encoding the message w_r and columns encoding the message w_c, see Figure 3. The transmitter then uses such a piggyback codebook in period $i \in \{1,2,3\}$ to send $W_{d_i,\emptyset}$ to Receiver i and $W^{(1)}_{d_4,\{i\}}$ to Receiver 4. That means, it sends the codeword

$$\mathbf{x}_1\big(W_{d_i,\emptyset}, W^{(1)}_{d_4,\{i\}}\big). \tag{24}$$

Weak receiver i has stored $W^{(1)}_{d_4,\{i\}}$ in its cache and can decode based on the restricted codebook $\mathcal{C}_{\text{sub},i}\big(W^{(1)}_{d_4,\{i\}}\big)$ consisting only of the codewords in the column indicated by $W^{(1)}_{d_4,\{i\}}$:

$$\mathcal{C}_{\text{sub},i}\big(W^{(1)}_{d_4,\{i\}}\big) := \left\{\mathbf{x}_1\big(w_r, W^{(1)}_{d_4,\{i\}}\big): \quad w_r \in \left\{1,\dots,2^{\mathsf{F}^{(0)}}\right\}\right\}.$$

Its decoding performance is thus the same as if this message $W^{(1)}_{d_4,\{i\}}$ had not been sent at all. Strong receiver 4 has no cache memory and decodes both messages $W_{d_i,\emptyset}$ and $W^{(1)}_{d_4,\{i\}}$ based on the entire codebook $\mathcal{C}$.

Decoding in this subphase is reliable, when

$$n_1 = \max\left\{\frac{3\mathsf{F}^{(0)}}{1-\delta_w}, \frac{3\mathsf{F}^{(0)} + 3\mathsf{F}^{(1,1)}}{1-\delta_s}\right\} + \epsilon\mathsf{F} = \frac{3\mathsf{F}^{(0)}}{1-\delta_w} + \epsilon\mathsf{F}. \tag{25}$$

Here, the second equality holds by the choice of $\mathsf{F}^{(1,1)}$ in (21).

Consider now Subphase 2. For the transmission in Period 1, the transmitter uses a standard piggyback codebook to transmit

$$W_{w,\{1,2\}} := W_{d_1,\{2\}} \oplus W_{d_2,\{1\}} \tag{26}$$

to the weak receivers 1 and 2 and the message $W^{(1)}_{d_4,\{1,2\}}$ to the strong receiver 4. That means, it applies a codebook as in Figure 3, but where message $W_{d_i,\emptyset}$ is replaced by $W_{w,\{1,2\}}$ and message $W^{(1)}_{d_4,\{i\}}$ by $W^{(1)}_{d_4,\{1,2\}}$. The transmitter then sends the codeword in row $W_{w,\{1,2\}}$ and column $W^{(1)}_{d_4,\{1,2\}}$ of this codebook over the channel. Receivers 1 and 2 can retrieve message $W^{(1)}_{d_4,\{1,2\}}$ from their cache memories, and thus decode message $W_{w,\{1,2\}}$ based only on the column of the codebook that corresponds to $W^{(1)}_{d_4,\{1,2\}}$. Receiver 1 then retrieves message $W_{d_2,\{1\}}$ from its cache memory and recovers its desired message part $W_{d_1,\{2\}} = W_{w,\{1,2\}} \oplus W_{d_2,\{1\}}$. Receiver 2 proceeds analogously to recover $W_{d_2,\{1\}}$. Receiver 4 decodes both messages $W_{w,\{1,2\}}$ and $W^{(1)}_{d_4,\{1,2\}}$ based on the entire codebook.

Using similar steps, in Period 2, messages $W_{d_1,\{3\}}$, $W_{d_3,\{1\}}$ and $W^{(1)}_{d_4,\{1,3\}}$ are sent to Receivers 1, 3 and 4. In Period 3, messages $W_{d_2,\{3\}}$, $W_{d_3,\{2\}}$ and $W^{(1)}_{d_4,\{2,3\}}$ are sent to Receivers 2, 3 and 4. Decoding in this subphase is reliable, when

$$n_2 = \max\left\{\frac{3\mathsf{F}^{(1)}}{1-\delta_w}, \frac{3\mathsf{F}^{(1)} + 3\mathsf{F}^{(2,1)}}{1-\delta_s}\right\} + \epsilon\mathsf{F} = \frac{3\mathsf{F}^{(1)}}{1-\delta_w} + \epsilon\mathsf{F}, \tag{27}$$

where the second equality holds by (21).

Consider Subphase 3. The transmitter uses a standard piggyback codebook to transmit the message

$$W_{w,\{1,2,3\}} = W_{d_1,\{2,3\}} \oplus W_{d_2,\{1,3\}} \oplus W_{d_3,\{1,2\}} \tag{28}$$

to Receivers 1, 2 and 3 and the message $W^{(1)}_{d_4,\{1,2,3\}}$ to Receiver 4. Decoding is done in a similar way as in Subphase 2. Decoding in this subphase is reliable, when

$$n_3 = \max\left\{\frac{\mathsf{F}^{(2)}}{1-\delta_w}, \frac{\mathsf{F}^{(2)} + \mathsf{F}^{(3,1)}}{1-\delta_s}\right\} + \epsilon\mathsf{F} = \frac{\mathsf{F}^{(2)}}{1-\delta_w} + \epsilon\mathsf{F}, \tag{29}$$

where the second equality holds by (21).

In Subphase 4, the transmitter uses a capacity-achieving

point-to-point code to transmit to Receiver 4 the missing parts of its message:

$$\Big(W_{d_4,\emptyset}, W^{(2)}_{d_4,\{1\}}, W^{(2)}_{d_4,\{2\}}, W^{(2)}_{d_4,\{3\}}, W^{(2)}_{d_4,\{1,2\}}, W^{(2)}_{d_4,\{1,3\}}, W^{(2)}_{d_4,\{2,3\}}, W^{(2)}_{d_4,\{1,2,3\}}\Big). \quad (30)$$

Decoding in this subphase is reliable, when

$$n_4 = \frac{\mathsf{F}^{(0)} + 3\mathsf{F}^{(1,2)} + 3\mathsf{F}^{(2,2)} + \mathsf{F}^{(3,2)}}{1-\delta_s} + \epsilon\mathsf{F}. \quad (31)$$

As $\epsilon \to 0$, the described coding scheme achieves the upper bound on $R^\star(\mathsf{M})$ in Theorem 1 for $\mathsf{K}_w = 3$ and $\mathsf{K}_s = 1$.

B. General Scheme

Fix a file size F and a small $\epsilon > 0$. Let $\mathsf{M}_\epsilon = \mathsf{M} - \epsilon$.

Decentralized placement phase: The placement is described in Section III-A. The cache content at a given weak receiver i can then be written as:

$$V_i = \{W_{d,G} \colon \ G \text{ so that } i \in G\}_{d=1}^{\mathsf{D}}. \quad (32)$$

For sufficiently large file sizes F and for any file W_d and all $\ell \in \{1,\dots,\mathsf{K}_w\}$, approximately the same number of bits is exclusively cached at any of the subsets of ℓ weak receivers. As before, let $\mathsf{F}^{(\ell)}$ be the expected number of bits cached at a given size-ℓ subset of receivers and $\mathsf{F}^{(0)}$ be the number of bits cached at no receiver. See (17) for the value of each $\mathsf{F}^{(\ell)}$.

Centralized delivery phase: The delivery phase is divided into $\mathsf{K}_w + 1$ subphases. Denote by n_ℓ the number of bits sent in Subphase ℓ and by n the total number of sent bits. Each subphase $\ell \in \{1,\dots,\mathsf{K}_w\}$ is further divided into $\binom{\mathsf{K}_w}{\ell}$ periods, each intended to ℓ weak receivers and all the strong receivers. Subphase $\mathsf{K}_w + 1$ is intended only to the strong receivers.

Before transmission starts, the transmitter divides each of the strong receivers' submessages into 2 parts:

$$W_{d_j,G} = \left(W^{(1)}_{d_j,G}, W^{(2)}_{d_j,G}\right), \quad \forall j \in \mathcal{K}_s, \forall G \subseteq \mathcal{K}_w. \quad (33)$$

If G is of size $|G| = \ell$, then the two parts in (33) are of sizes

$$\mathsf{F}^{(\ell,1)} = \min\left\{\mathsf{F}^{(\ell)}, \mathsf{F}^{(\ell-1)}\frac{\delta_w - \delta_s}{\mathsf{K}_s(1-\delta_w)}\right\}, \quad (34)$$

$$\mathsf{F}^{(\ell,2)} = \max\left\{0, \mathsf{F}^{(\ell)} - \mathsf{F}^{(\ell-1)}\frac{\delta_w - \delta_s}{\mathsf{K}_s(1-\delta_w)}\right\}. \quad (35)$$

Consider Subphase $\ell \in \{1,\dots \mathsf{K}_w\}$. Let $G^{(\ell)}_1, \dots, G^{(\ell)}_{\binom{\mathsf{K}_w}{\ell}}$ denote the $\binom{\mathsf{K}_w}{\ell}$ subsets of $\{1,\dots,\mathsf{K}_w\}$ of size ℓ. In each period $p \in \{1,\dots,\binom{\mathsf{K}_w}{\ell}\}$ of Subphase ℓ, the transmitter uses the standard piggyback codebook in Figure 3 to transmit the message

$$W_{w,G^{(\ell)}_p} := \bigoplus_{i \in G^{(\ell)}_p} W_{d_i, G^{(\ell)}_p \setminus \{i\}} \quad (36)$$

to all weak receivers in $G^{(\ell)}_p$, and the message

$$\mathbf{W}^{(1)}_{s,G^{(\ell)}_p} := \left(W^{(1)}_{d_{\mathsf{K}_w+1},G^{(\ell)}_p}, \dots, W^{(1)}_{d_{\mathsf{K}},G^{(\ell)}_p}\right) \quad (37)$$

to all the K_s strong receivers.

The following choice of the length of Subphase ℓ ensures that the probability of decoding error vanishes at all the receivers as $\mathsf{F} \to \infty$:

$$\begin{aligned} n_\ell &= \max\left\{\frac{\binom{\mathsf{K}_w}{\ell}\mathsf{F}^{(\ell-1)}}{1-\delta_w}, \frac{\binom{\mathsf{K}_w}{\ell}\left(\mathsf{F}^{(\ell-1)} + \mathsf{K}_s\mathsf{F}^{(\ell,1)}\right)}{1-\delta_s}\right\} + \epsilon\mathsf{F} \\ &= \frac{\binom{\mathsf{K}_w}{\ell}\mathsf{F}^{(\ell-1)}}{1-\delta_w} + \epsilon\mathsf{F}, \end{aligned} \quad (38)$$

where the second equality holds by (34).

In subphase $\mathsf{K}_w + 1$, the transmitter uses a standard BC code to send the missing parts of their messages to the strong receivers. The probability of decoding error tends to 0 as $\mathsf{F} \to \infty$ in this last subphase, if

$$n_{\mathsf{K}_w+1} = \frac{\mathsf{K}_s\mathsf{F}^{(0)} + \mathsf{K}_s\sum_{\ell=1}^{\mathsf{K}_w}\binom{\mathsf{K}_w}{\ell}\mathsf{F}^{(\ell,2)}}{1-\delta_s} + \epsilon\mathsf{F}. \quad (39)$$

The total number of sent bits thus satisfies

$$\begin{aligned} \frac{n}{\mathsf{F}} = \frac{\sum_{\ell=1}^{\mathsf{K}_w}\binom{\mathsf{K}_w}{\ell}\gamma_\epsilon^{(\ell-1)}}{1-\delta_w} + \frac{\mathsf{K}_s\gamma_\epsilon^{(0)} + \mathsf{K}_s\sum_{\ell=1}^{\mathsf{K}_w}\binom{\mathsf{K}_w}{\ell}\frac{\mathsf{F}^{(\ell,2)}}{\mathsf{F}}}{1-\delta_s} \\ +(\mathsf{K}_w+1)\epsilon. \end{aligned} \quad (40)$$

Combining (35) and (40), and letting $\epsilon \to 0$ and thus $\gamma_\epsilon^{(\ell)} \to \gamma^{(\ell)}$, establishes the rate-memory tradeoff in (13).

IV. Secure Decentralized Caching

Consider the standard scenario without secrecy constraint. Let $\gamma^{(\ell)}$, for $\ell \in \{0,1,\dots,\mathsf{K}_w\}$, be as defined in (12).

Theorem 2. *The secrecy rate-memory tradeoff is upper bounded as:*

$$\begin{aligned} R^\star_{\text{sec}}(\mathsf{M}) \le{}& \frac{\sum_{\ell=1}^{\mathsf{K}_w}\binom{\mathsf{K}_w}{\ell}\gamma^{(\ell-1)}}{1-\delta_w} + \frac{\mathsf{K}_w\gamma_{\text{bin}}}{1-\delta_w} + \frac{\mathsf{K}_s\gamma^{(0)}}{\delta_z-\delta_s} \\ &+ \frac{\sum_{\ell=2}^{\mathsf{K}_w}\binom{\mathsf{K}_w}{\ell}\left[\mathsf{K}_s(1-\delta_w)\gamma^{(\ell)} - (\delta_w-\delta_s)\gamma^{(\ell-1)}\right]^+}{(1-\delta_w)(\delta_z-\delta_s)} \\ &+ \frac{\mathsf{K}_w\left[\mathsf{K}_s(1-\delta_w)\gamma^{(1)} - (\delta_w-\delta_s)\left(\gamma^{(0)}+\gamma_{\text{bin}}\right)\right]^+}{(1-\delta_w)(\delta_z-\delta_s)}, \end{aligned} \quad (41)$$

where, if $\gamma^{(1)} < \left(\gamma^{(0)} + \gamma_{\text{bin}}\right)\frac{\delta_w-\delta_s}{\mathsf{K}_s(1-\delta_w)}$, *then*

$$\gamma_{\text{bin}} := \left[\frac{\left[(\mathsf{D}-\mathsf{K})(1-\delta_w) + \mathsf{K}_w(1-\delta_z)\right]\gamma^{(0)}}{\mathsf{K}_w(\delta_z-\delta_w)} - \frac{\mathsf{K}_s(1-\delta_w)\gamma^{(1)}}{(\delta_z-\delta_w)} - \frac{(\mathsf{D}-\mathsf{K})(1-\delta_w)}{\mathsf{K}_w(\delta_z-\delta_w)}\right]^+, \quad (42a)$$

and otherwise,

$$\gamma_{\text{bin}} := \left[\frac{\left[(\mathsf{D}-\mathsf{K}_s)(1-\delta_w) - \mathsf{K}_w(\delta_z-\delta_s)\right]\gamma^{(0)}}{\mathsf{K}_w(\delta_z-\delta_s)} - \frac{(\mathsf{D}-\mathsf{K})(1-\delta_w)}{\mathsf{K}_w(\delta_z-\delta_s)}\right]^+. \quad (42b)$$

The upper bound in Theorem 2 is illustrated in Figure 4 at hand of an example. The figure also shows an upper bound attained with separate cache-channel coding that follows the scheme in [2] and secures the messages by XORing them with other messages and by using wiretap erasure BC codes. For comparison, the figure also shows the bounds on $R^\star(\mathsf{M})$ obtained in the previous section.

Figure 4 shows that the secrecy constraint increases the delivery rate when cache memories are small or large. However, for moderate cache memories, the rates are very close to the non-secure rates.

Outline of secure coding scheme: To achieve the individual secrecy constraint in (10b), we follow the non-secure scheme explained in Section III-B, but where we secure the non-XORed messages either by XORing them with messages from the library or by adding random binning. We detail out the changes.

In our scheme of Section III-B, communications in Subphases $\ell = 2, \ldots, \mathsf{K}_w$ are already secure, i.e., they satisfy (10b). The reason is two-fold. On one hand, we only send XORs of demanded files to the weak receivers that do not provide any information about each of the files individually. On the other hand, the XORs sent to the weak receivers are sufficiently long to act as wiretap binning for the messages sent to the strong receivers. No changes are thus required for these subphases, and we choose

$$n_{\text{sec},\ell} = n_\ell = \frac{\binom{\mathsf{K}_w}{\ell}\mathsf{F}^{(\ell-1)}}{1-\delta_w} + \epsilon\mathsf{F}, \qquad \ell \in \{2, \ldots, \mathsf{K}_w\}. \quad (43)$$

To render Subphase 1 secure, the standard piggyback code is replaced by a secure piggyback code [8] with binning of bin size $\mathsf{F}_{\text{bin}} := \gamma_{\text{bin}}\mathsf{F} + \epsilon\mathsf{F}$, where γ_{bin} is defined in (42). Moreover, each message $W_{d_i,\emptyset}$ is divided into two parts $W^{(1)}_{d_i,\emptyset}, W^{(2)}_{d_i,\emptyset}$ of sizes $\tilde{\mathsf{F}}^{(0,1)} = \min\left\{\mathsf{F}^{(0)}, \frac{\mathsf{D}-\mathsf{K}}{\mathsf{K}_w}\left(\mathsf{F}-\mathsf{F}^{(0)}\right)\right\}$ and $\tilde{\mathsf{F}}^{(0,2)} = \mathsf{F}^{(0)} - \tilde{\mathsf{F}}^{(0,1)}$, and prior to encoding it with the secure piggyback codebook, $W^{(1)}_{d_i,\emptyset}$ is XORed with some of the bits stored in the cache memory of weak receiver i that do not belong to files requested by any user. Decoding in this subphase is reliable for large F, if the subphase is of length

$$n_{\text{sec},1} = \frac{\mathsf{K}_w\left(\mathsf{F}^{(0)} + \mathsf{F}_{\text{bin}}\right)}{1-\delta_w} + \epsilon\mathsf{F}. \quad (44)$$

Communication in Subphase $\mathsf{K}_w + 1$ is rendered secure by replacing the standard BC code with a wiretap BC code. This subphase is decoded reliably for large F, if

$$n_{\text{sec},\mathsf{K}_w+1} = \frac{\mathsf{K}_s\mathsf{F}^{(0)} + \mathsf{K}_s\sum_{\ell=1}^{\mathsf{K}_w}\binom{\mathsf{K}_w}{\ell}\tilde{\mathsf{F}}^{(\ell,2)}}{\delta_z - \delta_s} + \epsilon\mathsf{F}, \quad (45)$$

where for $\ell = 2, \ldots, \mathsf{K}_w$, $\tilde{\mathsf{F}}^{(\ell,2)} := \mathsf{F}^{(\ell,2)}$ as defined in (35), and

$$\tilde{\mathsf{F}}^{(1,2)} := \max\left\{0, \mathsf{F}^{(1)} - \left(\mathsf{F}^{(0)} + \mathsf{F}_{\text{bin}}\right)\frac{\delta_w - \delta_s}{\mathsf{K}_s(1-\delta_w)}\right\}. \quad (46)$$

Combining (43), (44) and (45), and letting $\epsilon \to 0$, yields the secrecy rate-memory tradeoff in (41).

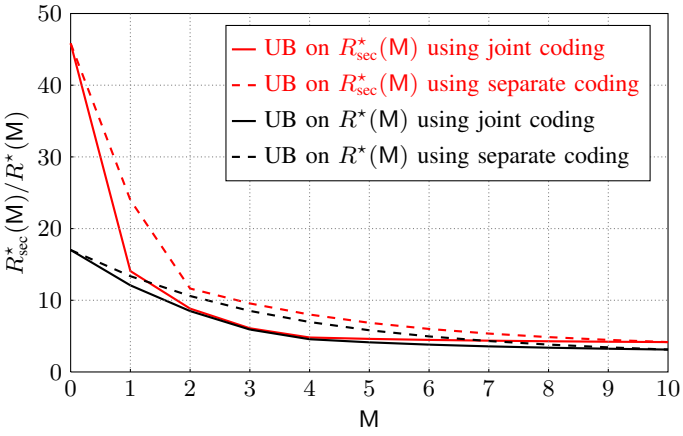

Fig. 4. Upper bounds on $R^\star_{\text{sec}}(\mathsf{M})$ and $R^\star(\mathsf{M})$ for $\delta_w = 0.7$, $\delta_s = 0.2$, $\delta_z = 0.8$, $\mathsf{K}_w = 5$, $\mathsf{K}_s = 3$, and $\mathsf{D} = 10$.

V. Summary

We have derived upper bounds on the decentralized rate-memory tradeoff of a K-receiver erasure BC, where K_w receivers are weak and have cache memories and K_s receivers are strong and have no caches. We have studied two scenarios for the delivery phase: a standard scenario without eavesdropper and a wiretap scenario with an external eavesdropper. Our upper bounds are achieved by joint cache-channel coding schemes and attain delivery rates only slightly higher than in the case of a centralized placement. In the wiretap scenario, for small and large cache memories, the required delivery rate is significantly increased compared to the standard scenario. However, when cache sizes are moderate, the rates in both cases are close.

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

[3] 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.

[4] M. M. Amiri, Q. Yang, and D. Gündüz, "Decentralized caching and coded delivery with distinct cache capacities," *IEEE Trans. on Communications*, vol. 65, no. 11, pp. 4657–4669, Nov. 2017.

[5] S. Wang, W. Li, X. Tian, and Hui Liu, "Coded caching with heterogenous cache sizes," Online: arXiv:1504.01123v3, 2015.

[6] S. Saeedi Bidokhti, R. Timo, and M. Wigger, "Noisy broadcast networks with receiver caching." Online: stanford.edu/saeedi/jrnlcache.pdf, 2016.

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

[8] 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.

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

[10] S. Kamel, M. Wigger, and M. Sarkiss, "Coded caching for wiretap broadcast channels," *Proc. of IEEE Inf. Theory Workshop (ITW)*, Kaohsiung, Taiwan, Nov. 2017.