On Coding for Real-Time Streaming under Packet Erasures

Derek Leong*†, Asma Qureshi*, and Tracey Ho*
*Department of Electrical Engineering, California Institute of Technology, Pasadena, California 91125, USA
†Institute for Infocomm Research, Singapore 138632, Singapore
dleong@i2r.a-star.edu.sg, aqureshi@caltech.edu, tho@caltech.edu

***Abstract*—We consider a real-time streaming system where messages created at regular time intervals at a source are encoded for transmission to a receiver over a packet erasure link; the receiver must subsequently decode each message within a given delay from its creation time. We study a bursty erasure model in which all erasure patterns containing erasure bursts of a limited length are admissible. For certain classes of parameter values, we provide code constructions that asymptotically achieve the maximum message size among all codes that allow decoding under all admissible erasure patterns. We also study an i.i.d. erasure model in which each transmitted packet is erased independently with the same probability; the objective is to maximize the decoding probability for a given message size. We derive an upper bound on the decoding probability for any time-invariant code, and show that the gap between this bound and the performance of a family of time-invariant intrasession codes is small in the high reliability regime.**

***Index Terms*—Erasure correction, real-time streaming.**

I. Introduction

We consider a real-time streaming system where messages created at regular time intervals at a source are encoded for transmission to a receiver over a packet erasure link; the receiver must subsequently decode each message within a given delay from its creation time.

Two erasure models are studied in this paper. The first is a bursty erasure model in which all erasure patterns containing erasure bursts of a limited length are admissible. The objective is to find a code that achieves the maximum message size, among all codes that allow all messages to be decoded by their respective decoding deadlines under all admissible erasure patterns. The second is an i.i.d. erasure model in which each transmitted packet is erased independently with the same probability. The objective here is to maximize the decoding probability for a given message size.

In our previous work [1], we showed that a time-invariant intrasession code is asymptotically optimal over all codes (time-varying and time-invariant, intersession and intrasession) as the number of messages goes to infinity, for a sliding window erasure model, and for the bursty erasure model when the maximum erasure burst length is sufficiently short or long. Intrasession coding is attractive due to its relative simplicity (it allows coding within the same message but not across different messages), but it is not known in general when intrasession coding is sufficient or when intersession coding is necessary.

This work was supported in part by the Air Force Office of Scientific Research under Grant FA9550-10-1-0166.

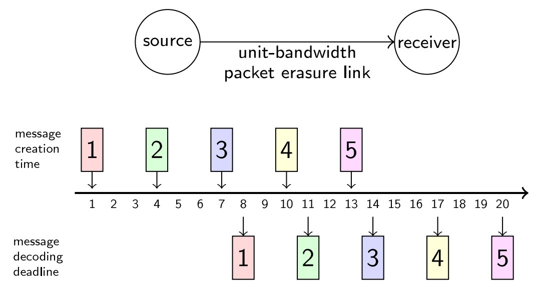


Fig. 1. Real-time streaming system for $(c, d) = (3, 8)$. Each of the messages $\{1, \ldots, 5\}$ is assigned a unique color. Messages are created at regular intervals of c time steps at the source, and must be decoded within a delay of d time steps from their respective creation times at the receiver. At each time step t, the source transmits a single data packet of normalized unit size over the packet erasure link.

Our Contribution: For the bursty erasure model, we show that diagonally interleaved codes derived from specific systematic block codes are asymptotically optimal over all codes for certain classes of parameter values. For the i.i.d. erasure model, we derive an upper bound on the decoding probability for any time-invariant code, and show that the gap between this bound and the performance of a family of time-invariant intrasession codes is small in the high reliability regime.

Related Work: Martinian *et al.* [2], [3] and Badr *et al.* [4] provide constructions of streaming codes that minimize the decoding delay for certain types of bursty erasure models. Tree codes or anytime codes, for which the decoding failure probability decays exponentially with delay, are considered in [5]–[7].

We begin with a formal definition of the problem in Section II. In Sections III and IV, we present our results for the bursty erasure model and the i.i.d. erasure model, respectively. Detailed results, figures, examples, and proofs can be found in the extended paper [8].

II. Problem Definition

Consider a discrete-time data streaming system comprising a source and a receiver, with a directed unit-bandwidth packet erasure link from the source to the receiver. Independent messages of uniform size $s > 0$ are created at regular intervals of $c \in \mathbb{Z}^+$ time steps at the source. At each time step $t \in \mathbb{Z}^+$, the source transmits a single data packet of normalized unit size over the packet erasure link; either the entire packet is received instantaneously by the receiver at time step t, or the entire packet is erased and never received. The receiver must subsequently decode each message within a delay of $d \in \mathbb{Z}^+$ time steps from its creation time. Fig. 1 depicts this real-time streaming system for an instance of (c, d).

More precisely, let random variable M_k denote message k; the random variables $\{M_k\}$ are independent, and $H(M_k) = s$ for each $k \in \mathbb{Z}^+$. To simplify our definition of the encoding functions, we shall further assume that $M_1, M_2, \ldots$ are identically distributed, and nonpositive messages $M_0, M_{-1}, \ldots$ are zeros.

Each message $k \in \mathbb{Z}^+$ is created at time step $(k-1)c + 1$, and is to be decoded by time step $(k-1)c + d$. Let W_k be the *coding window* for message k, which we define as the interval of d time steps between its creation time and decoding deadline, i.e., $W_k \triangleq \{(k-1)c + 1, \ldots, (k-1)c + d\}$. We shall assume that $d > c$ so as to avoid the degenerate case of nonoverlapping coding windows for which it is sufficient to code individual messages separately.

The unit-size packet transmitted at each time step $t \in \mathbb{Z}^+$ must be a function of messages created at time step t or earlier. Let random variable X_t denote the packet transmitted at time step t; we have $H(X_t) \leq 1$ for each $t \in \mathbb{Z}^+$. For brevity, we define $X[A] \triangleq (X_t)_{t \in A}$.

Because we are dealing with hard message decoding deadlines and fixed-size messages and packets, we consider a given message k to be decodable from the packets received at time steps $t \in A$ if and only if $H\big(M_k \,\big|\, X[A]\big) = 0$.

Consider the first n messages $\{1, \ldots, n\}$, and the union of their (overlapping) coding windows T_n given by

$$T_n \triangleq W_1 \cup \cdots \cup W_n = \{1, \ldots, (n-1)c + d\}.$$

An *erasure pattern* $E \subseteq T_n$ specifies a set of erased packet transmissions over the link; the packets transmitted at time steps $t \in E$ are erased, while those transmitted at time steps $t \in T_n \backslash E$ are received. An *erasure model* essentially describes a distribution of erasure patterns.

For a given pair of positive integers a and b, we define the *offset quotient* $q_{a,b}$ and *offset remainder* $r_{a,b}$ to be the unique integers satisfying $a = q_{a,b}\, b + r_{a,b}$, $q_{a,b} \in \mathbb{Z}_0^+$, and $r_{a,b} \in \{1, \ldots, b\}$, where $\mathbb{Z}_0^+$ denotes the set of nonnegative integers, i.e., $\mathbb{Z}^+ \cup \{0\}$. Note that this definition departs from the usual definition of quotient and remainder in that $r_{a,b}$ can be equal to b but not zero.

III. Bursty Erasure Model

In this section, we look at erasure patterns that contain erasure bursts of a limited length. Consider the first n messages $\{1, \ldots, n\}$, and the union of their (overlapping) coding windows T_n. Let $\mathcal{E}_n^{\mathsf{B}}$ be the set of erasure patterns in which each erasure burst is an interval of at most z erased time steps, and consecutive erasure bursts are separated by a guard interval or gap of at least $d - z$ unerased time steps, i.e.,

$$\begin{aligned}\mathcal{E}_n^{\mathsf{B}} \triangleq \Big\{ & E \subseteq T_n : \\ & (t \notin E \ \wedge\ t{+}1 \in E) \Rightarrow |E \cap \{t{+}1, \ldots, t{+}z{+}1\}| \leq z, \\ & (t \in E \ \wedge\ t{+}1 \notin E) \Rightarrow |E \cap \{t{+}1, \ldots, t{+}d{-}z\}| = 0 \Big\}.\end{aligned}$$

The objective is to construct a code that allows all n messages $\{1, \ldots, n\}$ to be decoded by their respective decoding deadlines under any erasure pattern $E \in \mathcal{E}_n^{\mathsf{B}}$. Let s_n^{B} be the maximum message size that can be achieved by such a code, for a given choice of (n, c, d, z).

This model can be seen as an instance of a more general class of bursty erasure models where the maximum erasure burst length and the minimum guard interval length can be arbitrarily specified. In a similar bursty erasure model considered by Martinian *et al.* [2], [3] and Badr *et al.* [4], the maximum erasure burst length (given by B) is z, while the minimum guard interval length (given by T) is $d - 1$. For the same choice of (d, z), our model captures a larger set of erasure patterns and is therefore stricter (the respective cut-set bounds reflect this comparison).

In [1], we showed that a time-invariant intrasession code is asymptotically optimal when d is a multiple of c, or when the maximum erasure burst length z is sufficiently short or long. Here, we present a family of time-invariant intersession codes that are asymptotically optimal in several other cases.

A. Diagonally Interleaved Codes

Consider a systematic block code $\mathcal{C}$ that encodes a given vector of $d - \alpha$ information symbols $\mathbf{a} = (a[1], \ldots, a[d{-}\alpha])$ as a codeword vector of d symbols $(a[1], \ldots, a[d{-}\alpha], b[1], \ldots, b[\alpha])$, where each symbol has a normalized size of $\frac{1}{d}$. For each $i \in \{1, \ldots, \alpha\}$, we define an encoding function g_i so that the parity symbol $b[i]$ is given by $b[i] = g_i(\mathbf{a})$.

For a given choice of (c, d, α), we can derive a time-invariant *diagonally interleaved code* for a message size of $s = \frac{d-\alpha}{d} c$ by interleaving codeword symbols produced by the component systematic block code $\mathcal{C}$ in a diagonal pattern.

First, to facilitate code construction, we represent the derived code by a table of symbols, with each cell in the table assigned one symbol of size $\frac{1}{d}$. Fig. 2 illustrates our construction for an instance of (c, d, α). Let $x_t[i]$ denote the symbol in column $t \in \mathbb{Z}$ and row $i \in \{1, \ldots, d\}$. The unit-size packet transmitted at each time step t is composed of the d symbols $x_t[1], \ldots, x_t[d]$ in column t of the table. Rows $\{1, \ldots, d - \alpha\}$ of the table are populated by information symbols, while rows $\{d - \alpha + 1, \ldots, d\}$ are populated by parity symbols.

Next, we divide each message k into $(d - \alpha)c$ submessages or information symbols denoted by $M_k[1], \ldots, M_k[(d - \alpha)c]$, with each symbol having a size of $\frac{s}{(d-\alpha)c} = \frac{1}{d}$. The information symbols corresponding to each message k are assigned evenly to the columns representing the first c time steps in

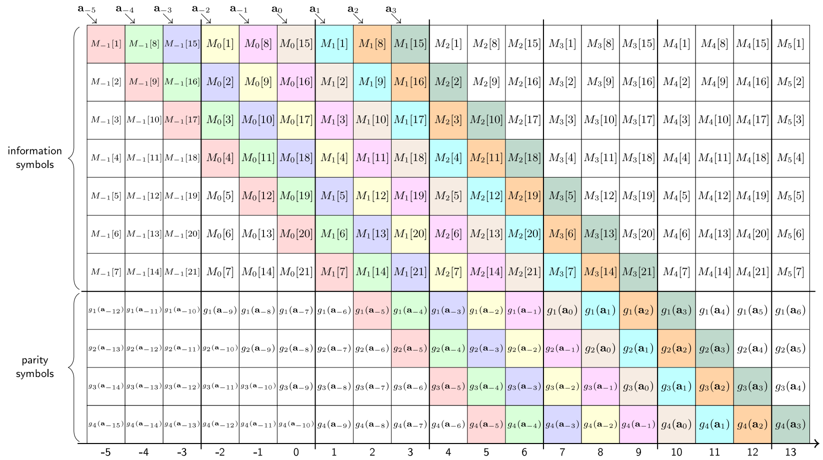


Fig. 2. Construction of the diagonally interleaved code, for $(c, d, \alpha) = (3, 11, 4)$. Rows $\{1, \ldots, d-\alpha\}$ of the table are populated by information symbols, while rows $\{d-\alpha+1, \ldots, d\}$ are populated by parity symbols. The d symbols on each diagonal spanning across d consecutive time steps constitute one codeword produced by the component systematic block code $\mathcal{C}$.

coding window W_k, so that

$$x_t[i] = M_{q_{t,c}+1}[(r_{t,c}-1)(d-\alpha)+i]$$

for each $i \in \{1, \ldots, d-\alpha\}$. To obtain the parity symbols for column t, we apply the component systematic block code $\mathcal{C}$ to the information symbols on each diagonal, so that

$$x_t[d-\alpha+i] = g_i\left(\left(x_{t-i-(d-\alpha)+\ell}[\ell]\right)_{\ell=1}^{d-\alpha}\right)$$

for each $i \in \{1, \ldots, \alpha\}$. Thus, the d symbols on each diagonal spanning across d consecutive time steps in the derived code constitute one codeword produced by $\mathcal{C}$. Note that the information symbols for nonexistent messages (i.e., nonpositive messages and messages after the actual final message) are assumed to be zeros so that all codeword symbols are well defined.

B. Optimality of Diagonally Interleaved Codes

Diagonally interleaved codes that are derived from systematic block codes $\mathcal{C}$ with certain properties turn out to be asymptotically optimal in certain cases. These sufficient code properties are given by the following lemma:

Lemma 1. *Consider the diagonally interleaved code for a given choice of $(c, d, \alpha{=}z)$ satisfying $c \le z \le d-c$. Suppose that the d symbols of the codeword vector $(a[1], \ldots, a[d-z], b[1], \ldots, b[z])$ produced by the component systematic block code $\mathcal{C}$ are transmitted sequentially across an erasure link, one symbol per time step, over the time interval $L \triangleq \{1, \ldots, d\}$. For each $j \in \{1, \ldots, d\}$, let $E_j^{\mathsf{z}} \subseteq L$ be the erasure pattern that contains a single wrap-around erasure burst of exactly z erased time steps (which may wrap around the last and first time steps in the interval) with the jth time step in the interval as the "leading" erasure, i.e.,*

$$E_j^{\mathsf{z}} \triangleq \left\{r_{j+\ell,d} : \ell \in \{0, \ldots, z-1\}\right\}.$$

Let $\mathcal{E}^{\mathsf{z}}$ be the set of all such erasure patterns, i.e., $\mathcal{E}^{\mathsf{z}} \triangleq \left\{E_1^{\mathsf{z}}, \ldots, E_d^{\mathsf{z}}\right\}$. If the systematic block code $\mathcal{C}$ satisfies both of the following symbol decoding requirements, then the diagonally interleaved code derived from $\mathcal{C}$ achieves a message size of $\frac{d-z}{d}c$ for the bursty erasure model:

D1) *For each $i \in \{1, \ldots, c\}$, the information symbol $a[i]$ is decodable by the $(d-c+i)$th time step in interval L under any erasure pattern $E^{\mathsf{z}} \in \mathcal{E}^{\mathsf{z}}$.*

D2) *The information symbols $a[c{+}1], \ldots, a[d{-}z]$ are decodable by the last time step in interval L under any erasure pattern $E^{\mathsf{z}} \in \mathcal{E}^{\mathsf{z}}$.*

The condition $c \le z \le d-c$ is actually implied by the symbol decoding requirements: the first information symbol $a[1]$ would otherwise be undecodable by its decoding deadline under erasure pattern E_1^{z} because by that time step, no parity symbols would have been transmitted if $c > z$, and no symbols would have been received if $z > d-c$. Note that the use of a systematic MDS code as the component systematic block code $\mathcal{C}$ may not be sufficient here because of the additional decoding deadlines imposed on individual symbols.

The following theorem shows that a degenerate diagonally interleaved code that uses only intrasession coding is asymptotically optimal over all codes for the specified parameter conditions:

Theorem 1. *Consider the bursty erasure model for a given choice of (c, d, z) satisfying* all *of the following three conditions: 1) d is not a multiple of c; 2) $c \le z \le d-c$; and 3) d is a multiple of $d-z$. Let $\mathcal{C}$ be a systematic block code that encodes a given vector of $d-z$ information symbols $\mathbf{a} = (a[1], \ldots, a[d{-}z])$ as a codeword vector of d symbols $(a[1], \ldots, a[d{-}z], b[1], \ldots, b[z])$, where each symbol has a normalized size of $\frac{1}{d}$, and the parity symbol $b[i]$ is given by*

$$b[i] = g_i(\mathbf{a}) \triangleq a[r_{i,d-z}]$$

for each $i \in \{1, \ldots, z\}$. The diagonally interleaved code derived from $\mathcal{C}$ is asymptotically optimal over all codes in the following sense: it achieves a message size of $\frac{d-z}{d}c$, which *is equal to the asymptotic maximum achievable message size* $\lim_{n\to\infty} s_n^{\mathsf{B}}$.

The following two theorems describe diagonally interleaved codes that are asymptotically optimal over all codes for the specified parameter conditions:

Theorem 2. *Consider the bursty erasure model for a given choice of* (c,d,z) *satisfying* all *of the following five conditions: 1)* d *is not a multiple of* c*; 2)* $c \le z \le d-c$*; 3)* d *is not a multiple of* $d-z$*; 4)* $z < d-z$*; and 5)* z *is a multiple of* r'*, where* $r' \triangleq r_{d-z,z} \in \{1,\ldots,z\}$*. Let* $\mathcal{C}$ *be a systematic block code that encodes a given vector of* $d-z$ *information symbols* $\mathbf{a} = (a[1],\ldots,a[d{-}z])$ *as a codeword vector of* d *symbols* $(a[1],\ldots,a[d{-}z],b[1],\ldots,b[z])$*, where each symbol has a normalized size of* $\frac{1}{d}$*, and the parity symbol* $b[i]$ *is given by*

$$b[i] = g_i(\mathbf{a}) \triangleq \left(\bigoplus_{k=1}^{\frac{d-z-r'}{z}} a[(k{-}1)z{+}i] \right) \oplus a[d{-}z{-}r'{+}r_{i,r'}]$$

for each $i \in \{1,\ldots,z\}$*. The diagonally interleaved code derived from* $\mathcal{C}$ *is asymptotically optimal over all codes in the following sense: it achieves a message size of* $\frac{d-z}{d}c$*, which is equal to the asymptotic maximum achievable message size* $\lim_{n\to\infty} s_n^{\mathsf{B}}$.

Theorem 3. *Consider the bursty erasure model for a given choice of* (c,d,z) *satisfying* all *of the following five conditions: 1)* d *is not a multiple of* c*; 2)* $c \le z \le d-c$*; 3)* d *is not a multiple of* $d-z$*; 4)* $z > d-z$*; and 5)* z' *is a multiple of* r'*, where* $z' \triangleq r_{z,d-z} \in \{1,\ldots,d{-}z{-}1\}$ *and* $r' \triangleq r_{d-z,z'} \in \{1,\ldots,z'\}$*. Let* $\mathcal{C}$ *be a systematic block code that encodes a given vector of* $d-z$ *information symbols* $\mathbf{a} = (a[1],\ldots,a[d{-}z])$ *as a codeword vector of* d *symbols* $(a[1],\ldots,a[d{-}z],b[1],\ldots,b[z])$*, where each symbol has a normalized size of* $\frac{1}{d}$*, and the parity symbol* $b[i]$ *is given by*

$$b[i] = g_i(\mathbf{a}) \triangleq \begin{cases} \left(\bigoplus_{k=1}^{\frac{d-z-r'}{z'}} a[(k{-}1)z'{+}i] \right) \oplus a[d{-}z{-}r'{+}r_{i,r'}] & \text{if } i \in \{1,\ldots,z'\}, \\ a[r_{i-z',d-z}] & \text{if } i \in \{z'+1,\ldots,z\} \end{cases}$$

for each $i \in \{1,\ldots,z\}$*. The diagonally interleaved code derived from* $\mathcal{C}$ *is asymptotically optimal over all codes in the following sense: it achieves a message size of* $\frac{d-z}{d}c$*, which is equal to the asymptotic maximum achievable message size* $\lim_{n\to\infty} s_n^{\mathsf{B}}$.

IV. IID Erasure Model

In this section, we consider a random erasure model in which each packet transmitted over the link is erased independently with the same probability p_{e}. For brevity, let S_k denote the success event "message k is decodable by its decoding deadline, i.e., time step $(k-1)c+d$", and let $\overline{S_k}$ denote the complementary failure event. We restrict our attention to time-invariant codes here in the interest of practicality.

Definition (Time-Invariant Code)**.** A code is time-invariant if there exist causal and deterministic encoding functions $f_1,\ldots,f_c$ and a finite encoder memory size $m_{\mathsf{E}} \in \mathbb{Z}^+$ such that the packet transmitted at each time step $(k-1)c+i$, where $k \in \mathbb{Z}^+$, $i \in \{1,\ldots,c\}$, is given by the function f_i applied to the m_{E} most recent messages, i.e.,

$$X_{(k-1)c+i} = f_i\big(\underbrace{M_k, M_{k-1},\ldots,M_{k-m_{\mathsf{E}}+1}}_{m_{\mathsf{E}}\text{ most recent messages}}\big).$$

Consider the i.i.d. erasure model for a given choice of (c,d,p_{e},s). We shall adopt the *decoding probability* $\mathbb{P}[S_k]$, i.e., the probability that a given message k is decodable by its decoding deadline, as the primary performance metric. The decoder memory size is assumed to be unbounded so that the decoder has access to all received packets. Let the random subset $U_k \subseteq T_k$ be the unerased time steps that are no later than the decoding deadline for message k; the received packets that can be used by the decoder for decoding message k are therefore given by $X[U_k]$. Consequently, the decoding probability $\mathbb{P}[S_k]$, where $k \in \mathbb{Z}^+$, can be expressed in terms of U_k as follows:

$$\begin{aligned} &\mathbb{P}[S_k] = \mathbb{P}\left[H\big(M_k \,\big|\, X[U_k]\big) = 0\right] \\ &= \sum_{U_k \subseteq T_k} \mathbb{1}\left[H\big(M_k \,\big|\, X[U_k]\big) = 0\right] \cdot (1-p_{\mathsf{e}})^{|U_k|}(p_{\mathsf{e}})^{|T_k|-|U_k|}. \end{aligned}$$

By combining the proof techniques of [1] and [9, Lemma 1], we can derive an upper bound on the decoding probability $\mathbb{P}[S_k]$ for any time-invariant code:

Theorem 4. *Consider the i.i.d. erasure model for a given choice of* (c,d,p_{e},s)*. For any time-invariant code with encoder memory size* m_{E}*, the probability that a given message* $k \ge m_{\mathsf{E}}$ *is decodable by its decoding deadline is upper-bounded as follows:*

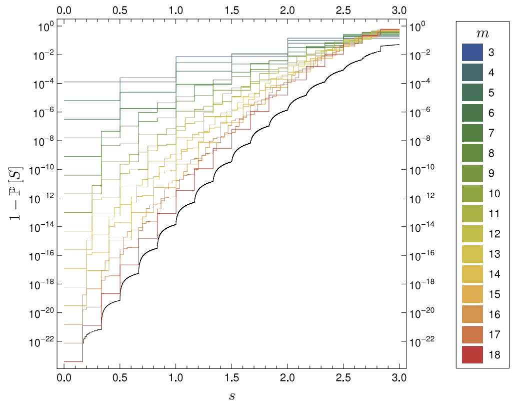
$$\mathbb{P}[S_k] \le \sum_{z=0}^{d} \left\lfloor \min\left(\frac{(d-z)c}{d\,s}, 1\right) \binom{d}{z} \right\rfloor (1-p_{\mathsf{e}})^{d-z}(p_{\mathsf{e}})^{z}.$$

Note that the decoding probability $\mathbb{P}[S_k]$ for the early messages $k < m_{\mathsf{E}}$ can potentially be higher than that for the subsequent messages $k \ge m_{\mathsf{E}}$ because the decoder already knows the nonpositive messages (which are assumed to be zeros).

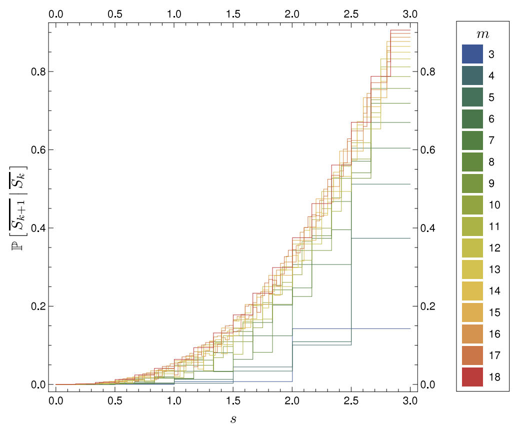
For real-time streaming applications that are sensitive to bursts of decoding failures, it may be useful to adopt the *burstiness of undecodable messages* as a secondary performance metric. One way of measuring this burstiness is to compute the conditional probability $\mathbb{P}\left[\overline{S_{k+1}} \,\middle|\, \overline{S_k}\right]$, i.e., the conditional probability that the next message is undecodable by its decoding deadline given that the current message is undecodable by its decoding deadline. The higher this conditional probability is, the more likely it is to remain "stuck" in a burst of undecodable messages.

A. Performance of Symmetric Codes

In [1], we introduced a time-invariant intrasession code that divides each packet evenly among all *active* messages



(a)



(b)

Fig. 3. Plots of (a) the decoding failure probability $1 - \mathbb{P}[S]$, and (b) the burstiness of undecodable messages as measured by the conditional probability $\mathbb{P}\left[\overline{S_{k+1}} \,\middle|\, \overline{S_k}\right]$, where $k \in \mathbb{Z}^+$, against message size s, for the family of symmetric time-invariant intrasession codes, for $(c, d, m_{\mathsf{E}}, p_{\mathsf{e}}) = (3, 18, 6, 0.05)$. Spreading parameter $m \in \{c, \ldots, d'\}$, where $d' = \min(d, m_{\mathsf{E}}c) = 18$, gives the size of the effective coding window for each code. The black curve in (a) describes a lower bound on the decoding failure probability for any time-invariant code, as given by Theorem 4.

(a message k is active at time step t if and only if t falls within its coding window, i.e., $t \in W_k$). An appropriate code (e.g., random linear coding, MDS code) is then applied to the allocation of packet space so that each message can be decoded whenever the total amount of received data that encodes that message is at least the message size s. Here, we generalize this code construction to obtain a family of *symmetric time-invariant intrasession codes*. For each symmetric code, we define a *spreading parameter* $m \in \{c, \ldots, d'\}$, where $d' \triangleq \min(d, m_{\mathsf{E}}c)$, so that the *effective coding window* for message k is given by

$$W'_k \triangleq \{(k-1)c + 1, \ldots, (k-1)c + m\}.$$

We subsequently redefine *active* messages in terms of the *effective* coding window instead of the coding window (which corresponds to $m = d$).

1) Decoding Probability: Fig. 3a shows how the family of symmetric codes perform in terms of the decoding probability $\mathbb{P}[S]$, for an instance of (c, d, m_{E}). These plots and other empirical observations suggest that 1) maximal spreading (i.e., $m = d'$) performs well, i.e., achieves a relatively high $\mathbb{P}[S]$, when the message size s and the packet erasure probability p_{e} are small, while minimal spreading (i.e., $m = c$) performs well when s and p_{e} are large (this echoes the analytical findings of [9]); and 2) although this family of codes may not always contain an optimal time-invariant intrasession code, we can find good codes (with decoding probabilities close to the upper bound of Theorem 4) among them when s and p_{e} are small.

2) Burstiness of Undecodable Messages: Fig. 3b shows how the family of symmetric codes perform in terms of the burstiness of undecodable messages as measured by the conditional probability $\mathbb{P}\left[\overline{S_{k+1}} \,\middle|\, \overline{S_k}\right]$, for an instance of (c, d, m_{E}). These plots and other empirical observations suggest that over a wide range of message sizes s and packet erasure probabilities p_{e}, minimal spreading (i.e., $m = c$) performs well, i.e., achieves a relatively low $\mathbb{P}\left[\overline{S_{k+1}} \,\middle|\, \overline{S_k}\right]$, while maximal spreading (i.e., $m = d'$) performs poorly. This agrees with the intuition that for a pair of consecutive messages, a greater overlap in their effective coding windows would tend to increase the correlation between their decodabilities. In the case of minimal spreading (i.e., $m = c$), the decodability of a message is independent of the decodability of other messages because the effective coding windows do not overlap at all.

3) Trade-off between Performance Metrics: Our results show that for the family of symmetric codes, a trade-off exists between the decoding probability $\mathbb{P}[S]$ and the burstiness of undecodable messages as measured by the conditional probability $\mathbb{P}\left[\overline{S_{k+1}} \,\middle|\, \overline{S_k}\right]$ when the message size s and packet erasure probability p_{e} are small (this is a regime of interest because it supports a high decoding probability). Specifically, although maximal spreading (i.e., $m = d'$) achieves a high decoding probability, it also exhibits a higher burstiness of undecodable messages. Thus, a symmetric code with a suboptimal decoding probability but lower burstiness may be preferred for an application that is sensitive to bursty undecodable messages.

References

[1] D. Leong and T. Ho, "Erasure coding for real-time streaming," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, Massachusetts, USA, Jul. 2012.

[2] E. Martinian and C.-E. W. Sundberg, "Low delay burst erasure correction codes," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2002, pp. 1736–1740.

[3] E. Martinian and M. Trott, "Delay-optimal burst erasure code construction," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2007, pp. 1006–1010.

[4] A. Badr, A. Khisti, W.-T. Tan, and J. Apostolopoulos, "Streaming codes for channels with burst and isolated erasures," *arXiv:1208.0072*, Aug. 2012. [Online]. Available: http://arxiv.org/abs/1208.0072

[5] L. J. Schulman, "Coding for interactive communication," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 1745–1756, Nov. 1996.

[6] A. Sahai, "Anytime information theory," Ph.D. dissertation, Massachusetts Institute of Technology, 2001.

[7] R. T. Sukhavasi, "Distributed control and computing: Optimal estimation, error correcting codes, and interactive protocols," Ph.D. dissertation, California Institute of Technology, 2012.

[8] D. Leong, A. Qureshi, and T. Ho, "On coding for real-time streaming under packet erasures." [Online]. Available: http://purl.org/net/2013.01

[9] D. Leong, A. G. Dimakis, and T. Ho, "Distributed storage allocations," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4733–4752, Jul. 2012.