Correcting Erasure Bursts with Minimum Decoding Delay

Zhi Li
Department of
Electrical Engineering
Stanford University
Stanford, CA 94305, USA
Email: leeoz@stanford.edu

Ashish Khisti
Department of Electrical
& Computer Engineering
University of Toronto
Toronto, ON M5S 3G4, Canada
Email: akhisti@comm.utoronto.ca

Bernd Girod
Department of
Electrical Engineering
Stanford University
Stanford, CA 94305, USA
Email: bgirod@stanford.edu

***Abstract*—Erasure correcting codes are widely used in upper layers of packet-switched networks, where the packet erasures often exhibit bursty patterns. The conventional wisdom to deal with bursty erasures is to apply block interleaving to break down the bursty patterns prior to error correcting coding, or use long-block Reed-Solomon codes. We show that they unnecessarily lead to sub-optimal decoding delay. In this work, with the problem model of multiple erasure bursts present in a coding block, we study the fundamental tradeoff among the rate, decoding delay and burst correction performance of erasure correcting codes. Focusing on a class of codes achieving the Singleton bound, we show that the lowest delay to recover any individual symbol not only depends on how many bursts are present in a coding block, but also on whether the source symbols are encoded causally or non-causally. We also describe a few practical linear code constructions that achieve the performance limit discussed.**

I. Introduction

Erasure correcting codes are widely used in upper layers of packet-switched networks for many delay-sensitive applications, such as Internet Protocol Television (IPTV), Voice over IP (VoIP), video conferencing and distance learning.

In many scenarios, packet erasures often exhibit bursty patterns. In IPTV systems, one main source of error is the impulse noise over digital subscriber lines (DSL) [1], which typically lasts up to dozens of milliseconds. In wireless communications, a prominent feature of wireless media is time-varying multipath fading, often causing channel transition between good and bad channel state [2]. In packet-switched networks, when a router is overwhelmed by traffic, it tends to drop packets in bursts.

Erasure correcting coding inevitably introduces decoding delay. By intuition, the characteristics of erasure patterns should affect the amount of decoding delay needed to recover all the erasures. For example, Figure 1 shows three channel realizations with the same number of erasures but different erasure patterns. In (a), as the erasures are spread apart in time, we can apply a $(3, 2)$ Reed-Solomon code to encode every 2 source packets into 3 channel packets, resulting in a decoding delay of 2 packets. In (b), the erasures are lumped in bursts. One way is to apply interleaving to spread out the erasures, then treat the resulting pattern the same way as for (a). However, intuitively, there is some structure in the lumped bursts that one can exploit, and interleaving would simply destroy them. Another way is to apply a long $(12, 8)$ Reed-Solomon code and recover all the erasures at the end of the block, resulting in a delay of 11 packets. But there is another way: as we will show in this work, for (b) it is possible to construct a $(12, 8)$ burst erasure code that exploit the erasure burstiness, such that *each packet* is guaranteed to be decoded with a delay of 8 packets. In (c), erasures that are both speaded apart and lumped together are present. When the code needs to deal with erasures of random patterns, one should not expect to do better than applying a $(12, 8)$ Reed-Solomon code with a decoding delay of 11 packets.

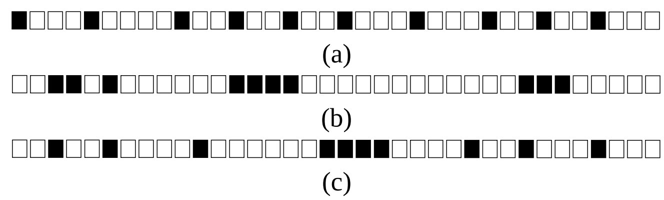

Fig. 1. Channels with the same number of erasures but different erasure patterns. A similar discussion of these erasure patterns appeared in Chapter 1 of [3], but the focus was on a single erasure burst and i.i.d. erasures. In this work, we consider the problem model of multiple erasure bursts within a block.

In this work, we consider a problem model where Z number of erasure bursts, each of length B, present in a coding block. This model encompasses the single burst case ($Z = 1$) and the random pattern case ($B = 1$) as special cases. Based on this model, we study the fundamental tradeoff among the rate, decoding delay and burst correction performance of erasure correcting codes. Focusing on a class of codes achieving the Singleton bound, our main finding is that the lowest delay to recover any individual symbol not only depends on how many bursts are present in a coding block, but also on whether the source symbols are encoded causally or non-causally.

To develop this result, we first discusses related work in Section II and formulate the problem in Section III. In Section IV, we state a theorem that characterizes the rate-delay-burst tradeoff of Singleton-achieving non-causal codes. In Section V, we show the surprising result that for a subclass of the causal codes, the performance bound can be tightened further.

We also show a converse result for the systematic codes, which can be regarded as the block code counterpart of the periodic erasure channel argument for the streaming codes [4]. In Section VI, we briefly discuss the case when the codes are non-Singleton-achieving.

II. Related Work

Our study is inspired by [4], [5], [3], where Martinian et al. studied systematic delay-optimal streaming codes for a single erasure burst. They first design a family of block codes, then use diagonal interleaving [5] to convert the block codes into streaming codes.

For a systematic block code that correct one B-burst with delay T, it is sufficient to use a $(T, T-B)$ cyclic code over finite field Q, of generator matrix

$$G = \begin{pmatrix} I_{T-B} & P_{(T-B)\times B} \end{pmatrix} \tag{1}$$

where $I_M \in Q^{M\times M}$ denotes identity matrix of dimension $M \times M$ and $P_{(T-B)\times B}$ is the parity sub-matrix. Martinian et al. propose a construction of systematic $(T+B, T)$ block code with the same delay and burst correction performance, but of better rate, having generator matrix:

$$G = \begin{pmatrix} I_T & \begin{matrix} I_B \\ P_{(T-B)\times B} \end{matrix} \end{pmatrix} \tag{2}$$

where $P_{(T-B)\times B}$ is the same as in (1). An intuition behind this construction is that since it does not correct wrap-around bursts, the rate can be improved. A converse result based on a periodic erasure channel argument shows that the constructed streaming code achieves the rate-delay-burst bound [4]:

$$T \geq \max\left(\frac{R}{1-R}\, ,\, 1\right)\cdot B. \tag{3}$$

Delay-optimal erasure codes for broadcast channel is studied by Khisti et al. in [6].

This work differs from [4], [5], [3], [6] as follows. i) Block codes are considered instead of streaming codes. The conversion from block codes to streaming codes can be achieved trivially [5]. ii) A broader class of causal/non-causal codes are considered instead of systematic codes. iii) A more general class of bursty erasure patterns consisting of multiple bursts within each coding block are considered instead of a single burst.

III. Problem Setup

Throughout the paper, random variables are denoted by sans-serif letters (e.g., s, x, y). The entropy of random variable x is denoted by $H(\mathsf{x})$, and the conditional entropy of x given y is denoted by $H(\mathsf{x}\,|\,\mathsf{y})$. A random variable vector $(\mathsf{x}[a] \ldots \mathsf{x}[b])$ is sometimes denoted by $\mathsf{x}[a, ..., b]$.

Consider a sequence of packets transmitted over a communication channel, where each packet is modeled as a channel symbol $\mathsf{x}[i] \in Q$ over a finite field Q injected at some time i. The channel symbol can either be erased (denoted by symbol $\star$) or passed to the receiver intact. For a block of N channel symbols, denote by $\mathcal{B} \subseteq \{1, ..., N\}$ an *erasure pattern*, i.e., the set of time indices when erasures occur. At the receiver, the received channel symbols are

$$\mathsf{y}[i] = \begin{cases} \star, & i \in \mathcal{B} \\ \mathsf{x}[i], & \text{otherwise}, \end{cases} \qquad i = 1, ..., N.$$

An *erasure pattern collection* is a set of erasure patterns $\{\mathcal{B}\} \subseteq 2^{\{1,...,N\}}$. For example, let a B*-burst* be B consecutive symbol erasures. The set of all patterns consisting of Z number of B-bursts is an erasure pattern collection.

A block code $C = (E, D)$ for an erasure channel consists of an encoding function E and a decoding function D. The operation of the code is described as follows.

Encoding. The encoding function $E : Q^K \mapsto Q^N$ takes in K source symbols $\mathsf{S} := (\mathsf{s}[1] \ldots \mathsf{s}[K]) \in Q^K$ and maps them into N channel symbols $\mathsf{X} := (\mathsf{x}[1] \ldots \mathsf{x}[N]) \in Q^N$. The rate of the code is

$$R = \frac{K}{N}. \tag{4}$$

E can be written as $E = \{E_i\}_{i=1}^N$, where E_i is the encoding function for channel symbol $\mathsf{x}[i]$. Two special cases are:

- The code is *causal*, if in the encoding function the current channel symbol is a function of the current and previous source symbols, i.e., $\mathsf{x}[i] = E_i(\mathsf{s}[1, ..., i])$, $i = 1, ..., N$. Causal codes are a special case of the more general class of *non-causal* codes.
- The code is *systematic* (which is a special case of causal codes), if in the encoding function the first K channel symbols is a replica of the K source symbols, i.e., $\mathsf{x}[i] = \mathsf{s}[i]$, $i = 1, ..., N$.

If the code is linear, the encoding function can be represented in matrix multiplication as

$$\mathsf{X} = E(\mathsf{S}) = \mathsf{S} \cdot G$$

where $\mathsf{S} \in Q^{1\times K}$, $\mathsf{X} \in Q^{1\times N}$ and $G \in Q^{K\times N}$ is the generator matrix. For systematic code, $G = [I\ P]$, where $I \in Q^{K\times K}$ is an identity matrix and $P \in Q^{K\times(N-K)}$ is a parity check matrix. For causal code, $G = [U\ P]$, where $U \in Q^{K\times K}$ is an upper-trianglar matrix.

Decoding. The decoding function $D : \{Q, \star\}^N \mapsto Q^K$ takes in the received channel symbols $\mathsf{Y} := (\mathsf{y}[1] \ldots \mathsf{y}[N]) \in \{Q, \star\}^N$ and maps them to the reconstructed source symbol $\widehat{\mathsf{S}} := (\widehat{\mathsf{s}}[1] \ldots \widehat{\mathsf{s}}[K]) \in Q^K$. D can be written as $D = \{D_i\}_{i=1}^K$, where D_i is the decoding function for $\widehat{\mathsf{s}}[i]$. If $\widehat{\mathsf{S}} = \mathsf{S}$ for every $\mathcal{B} \in \{\mathcal{B}\}$, we say code C is *feasible* for the erasure pattern collection $\{\mathcal{B}\}$. Furthermore, if the decoding of each source symbol is subject to a decoding delay T, i.e., $\widehat{\mathsf{s}}[i] = D_i(\mathsf{y}[1, ..., i+T])$, $i = 1, ..., K$, we say the code is *delay-T*.

In this work, we focus on a class of code that achieves the Singleton bound. The canonical *Singleton bound* is a simple converse result that governs the performance of any error correction codes. It states that for a code C with encoding function $E : Q^K \mapsto Q^N$ and minimum distance d, $N \geq K + d - 1$. For an erasure code correcting Z B-bursts, we must have the minimum distance larger than the number of erasures, therefore,

$$N \geq K + ZB. \tag{5}$$

If $N = K + ZB$, we refer to the code as *Singleton-achieving*.

IV. Non-Causal Codes

In this section, we state a theorem that characterizes the rate-delay-burst tradeoff of Singleton-achieving non-causal codes. The main theorem is stated as follows.

Theorem 1. *It is possible to construct a delay-T rate-R Singleton-achieving non-causal block erasure code feasible for any erasure patterns of Z B-bursts, if*

$$T \geq T^* := \max\left(\frac{Z}{1-R} - 1\ ,\ Z\right) \cdot B. \tag{6}$$

Conversely, if $T < T^*$, no feasible code can be constructed.

Proof: The achievability of the code is proven by construction in Section IV-A . The converse part is proven in Section IV-B. ■

A. Theorem 1: Achievability

The main idea of this code construction is to group the source and channel symbols into independently decodable sub-blocks, and interleave the inter-sub-block symbols to keep the intra-sub-block symbols well separated, such that a B-burst can only corrupt at most one symbol in each sub-block. Furthermore, the positions of the symbols in each sub-block need to be carefully selected in order to meet a per-sub-block decoding deadline.

Proposition 2. *The encoding matrix G of a rate-R delay-T^* (as in (6) of Theorem 1) non-causal block erasure code feasible for any erasure patterns of Z B-bursts can be constructed as follows.*

- *Compute T^*according to (6). Determine the matrix dimension $K \times N$ as $K = T^* - ZB + B$ and $N = T^* + B$.*
- *Assign the K source symbols into B sub-blocks evenly. Denoted by λ the remainder of K divided by B. The first $B - \lambda$ sub-blocks each have $\lfloor \frac{K}{B} \rfloor$ source symbols; the remaining λ each have $\lfloor \frac{K}{B} \rfloor + 1$ source symbols. Let the number of source and channel symbols in sub-block i be K_i and N_i, respectively, with $N_i = K_i + Z$, $i = 1, ..., B$.*
- *For sub-block i, $i = 1, ..., B$, create an (N_i, K_i) maximum distance separable (MDS) code generator matrix $G_i = (g_i(1) \dots g_i(N_i))$, where $g_i(j) \in Q^{K_i \times 1}$ is the j-th column of matrix G_i.*
- *Fill in the matrix G as follows. Initialize G as a $K \times N$ all-zero matrix. Each sub-block i corresponds to K_i rows of G, $i = 1, ..., B$, from top to bottom. For sub-block $i = 1, ..., B - \lambda$, put $g_i(1)$, $g_i(2)$,...,$g_i(N_i)$ at column $i + \lambda$, $i + \lambda + B$,...,$i + \lambda + (N_i - 1)B$, respectively. For sub-block $i = B - \lambda + 1, ..., B$, put $g_i(1)$, $g_i(2)$,...,$g_i(N_i)$ at column $i + \lambda - B$, $i + \lambda$,...,$i + \lambda + (N_i - 2)B$, respectively.*

Example. For $R = 4/13$, $B = 3$, $Z = 3$, we compute $T^* = 10$, $K = 4$, $N = 13$. The constructed generator matrix over GF(3) is

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 2 \end{pmatrix} \tag{7}$$

The decoding is straightforward. One simply need to decode each sub-block MDS code independently. Since for each sub-block, Z B-burst can only corrupt at most Z symbols, each sub-block can always be decoded successfully.

We can verify that this code meets $T = T^*$. If $R \geq \frac{1}{Z+1}$, the required decoding delay simplifies to $T = N - B$. Thus, the first $B - 1$ source symbols are urgent symbols with a deadline $N - B + i$, $i = 1, ..., B - 1$; the rest soruce symbols are non-urgent and can be decoded at the end of the block. In the proposed construction, each sub-block i is arranged in a way that its last column $g_i(N_i)$ is always put at the $(N - B + i)$-th column of G. This ensures that each sub-block i has at least one source symbol decoded with delay $N - B$. If $R < \frac{1}{Z+1}$, to meet the required decoding delay, it is sufficient to use a $(ZB + B, B)$ code.

In the special case that $\frac{T}{B}$ is an integer, the constructed generator matrix degenerates to a form of $G = G_1 \otimes I_B$, where G_1 is the MDS code generator matrix for a single sub-block, I_B is an $B \times B$ identity matrix and $\otimes$ denotes Kronecker product. This corresponds to interleaving by length B followed by encoding with generator matrix G_1.

B. Theorem 1: The Converse

In this section, we show the converse result that if $T < T^*$, no feasible code can be constructed. The key to prove this is to identify a set of conditions (i.e., erasure patterns) and use an entropy argument to show that the code cannot possibly meet them all at the same time. We first introduce a useful entropy lemma:

Lemma 3. *Let $\mathsf{X} = (\mathsf{x}_1, ..., \mathsf{x}_n)$, and $\mathsf{X}_{\Omega \backslash i} = (\mathsf{x}_1, ...\mathsf{x}_{i-1}, \mathsf{x}_{i+1}, ..., \mathsf{x}_n)$. If $H(\mathsf{s}) > 0$ and*

$$H(\mathsf{s} \,|\, \mathsf{X}_{\Omega \backslash i}) = 0, \quad i = 1, ..., n, \tag{8}$$

then

$$H(\mathsf{X}) < \sum_{i=1}^{n} H(\mathsf{x}_i).$$

Proof: See Appendix A. ■

Consider that the source symbol block $\mathsf{s}[1], ..., \mathsf{s}[K]$ is encoded into the channel symbol block $\mathsf{x}[1], ..., \mathsf{x}[N]$, where $N = K + ZB$. Let $\mathsf{x}[1, ..., K]$ be divided into non-overlapping segments of length B except for the last segment, i.e., let the index sets be $\mathcal{I}_1 := \{1, ..., B\}$, $\mathcal{I}_2 := \{B + 1, ..., 2B\}$, ..., $\mathcal{I}_{\lceil K/B \rceil} := \{(\lceil \frac{K}{B} \rceil - 1)B + 1, ..., K\}$, then $\mathsf{x}[1, ..., K] = (\mathsf{x}[\mathcal{I}_1], ..., \mathsf{x}[\mathcal{I}_{\lceil K/B \rceil}])$.

TABLE I
ERASURE PATTERNS CONSIDERED IN THE PROOF OF THE CONVERSE PART OF THEOREM 1. IN THIS EXAMPLE, $N = 11$, $K = 5$, $B = 2$ AND $Z = 3$.

Time	1	2	3	4	5	6	7	8	9	10	11
$\mathcal{B}_1$	$\star$	$\star$				$\star$	$\star$	$\star$	$\star$		
$\mathcal{B}_2$			$\star$	$\star$		$\star$	$\star$	$\star$	$\star$		
$\mathcal{B}_3$					$\star$	$\star$	$\star$	$\star$	$\star$		
$\mathcal{B}_0$						$\star$	$\star$	$\star$	$\star$	$\star$	$\star$

Since the coding scheme must work on source of any distribution, we can arbitrarily assume that the source symbols $\mathsf{s}[i]$, $i = 1, ..., K$ are i.i.d. ~ uniform over Q. Thus $H(\mathsf{s}[i]) = H(\mathsf{s}) := \log Q$, $i = 1, ..., K$, and

$$H(\mathsf{s}[1], ..., \mathsf{s}[K]) = K \cdot H(\mathsf{s}). \tag{9}$$

Furthermore, since each $\mathsf{x}[i]$ is over Q, we have

$$H(\mathsf{x}[i]) \leq H(\mathsf{s}), \quad i = 1, ..., N. \tag{10}$$

We would like to show

$$T \geq K + (Z-1)B. \tag{11}$$

Assume the opposite is true, for example, $T = K+(Z-1)B-1$. Then $\mathsf{s}[1]$ is recoverd at time $K + (Z-1)B$. Consider a set of erasure patterns (refer to Table I for an example):

$$\mathcal{B}_i = \mathcal{I}_i \bigcup \{K+1, ..., K+(Z-1)B\}, \quad i = 1, ..., \left\lceil \frac{K}{B} \right\rceil.$$

That $\mathsf{s}[1]$ is recoverd at time $K + (Z-1)B$ implies

$$H(\mathsf{s}[1] \,|\, \mathsf{x}[1, ..., K] \setminus \mathsf{x}[\mathcal{I}_i]) = 0, \quad i = 1, ..., \left\lceil \frac{K}{B} \right\rceil.$$

Applying Lemma 3 with $\mathsf{s} = \mathsf{s}[1]$ and $\mathsf{x}_i = \mathsf{x}[\mathcal{I}_i]$, $i = 1, ..., \lceil \frac{K}{B} \rceil$, we have

$$\begin{aligned} H(\mathsf{x}[1, ..., K]) &< \sum_{i=1}^{\lceil K/B \rceil} H(\mathsf{x}[\mathcal{I}_i]) \\ &\leq \sum_{i=1}^{K} H(\mathsf{x}[i]) \leq K \cdot H(\mathsf{s}). \end{aligned} \tag{12}$$

Now, consider the erasure pattern $\mathcal{B}_0 = \{K+1, ..., N\}$ (refer to Table I for an example). The feasibility of the code implies that $\mathsf{s}[1], ..., \mathsf{s}[K]$ must be recovered from the remaining symbols $\mathsf{x}[1], ..., \mathsf{x}[K]$. This implies $H(\mathsf{s}[1, ..., K] \,|\, \mathsf{x}[1, ..., K]) = 0$. But

$$\begin{aligned} 0 &= H(\mathsf{s}[1, ..., K] \,|\, \mathsf{x}[1, ..., K]) \\ &= H(\mathsf{s}[1, ..., K]) - H(\mathsf{x}[1, ..., K]) \qquad (13) \\ &> K \cdot H(\mathsf{s}) - K \cdot H(\mathsf{s}) = 0 \qquad (14) \end{aligned}$$

where in (13) we use the fact that $\mathsf{x}[1, ..., K]$ is a function of $\mathsf{s}[1, ..., K]$, implying $H(\mathsf{x}[1, ..., K] \,|\, \mathsf{s}[1, ..., K]) = 0$. (14) follows (9) and (12). The above equations lead to contradiction; thus, the assumption is false and (11) is true.

Next, consider next the erasure pattern $\mathcal{B}_1 = \{1, ..., ZB\}$. Clearly, The earliest time to decode $\mathsf{s}[1]$ is at $ZB + 1$. We have

$$T \geq ZB. \tag{15}$$

Combining (4), (5), (11) and (15), we conclude (6).

V. CAUSAL CODES

In this section, we discuss the rate-delay-burst performance limit of Singleton-achieving causal codes.

First of all, since the causal codes are a special class of non-causal codes, Theorem 1 applies for causal codes as well. For the single burst case $Z = 1$, (6) degenerates to (3). In this case, the code construction (2) achieves the performance limit. It is tempting to think that one can generalize the construction (2) to the $Z \geq 2$ case and find codes that meet (6). The surprising result we find is that no such code can be constructed in general. As we show in the following proposition, when $Z \geq 2$ and $K = B+1$, the best code constructable has decoding delay of one more symbol than (6).

Proposition 4. *It is possible to construct a delay-T Singleton-achieving causal block erasure code with the number of source symbols $K = B+1$ and rate $R = \frac{B+1}{ZB+B+1}$, feasible for any erasure patterns of $Z \geq 2$ B-bursts, if*

$$T \geq T^{**} := \left(\frac{Z}{1-R} - 1 \right) B + 1. \tag{16}$$

Conversely, if $T < T^{**}$, no feasible code can be constructed.

Proof: The achievability of the code is proven by construction in Section V-A . The converse part is proven in Section V-B. ■

In Section V-C, we also show a converse result that bounds the rate-burst-delay performance of any systematic codes (either Singleton-achieving or not).

A. Proposition 4: Achievability

The key idea of the construction to achieve $T = T^{**}$ is the following. A naive way of extending (2) to a multiple-burst code is to simply repeat the parity part of (2) Z times after the systematic part. One can easily verify that this approach fails under the erasure pattern of the first ZB channel symbols being corrupted. However, a remedy is to delay the recovery by one symbol to "save enough space" to remove the non-urgent source symbols before recovering the urgent source symbols.

Proposition 5. *The encoding matrix G of a delay-T^{**} (as in (16) of Proposition 4) Singleton-achieving causal block erasure code with the number of source symbols $K = B+1$ and rate $R = \frac{B+1}{ZB+B+1}$, feasible for any erasure patterns of $Z \geq 2$ B-bursts, can be constructed as*

$$\left(\begin{array}{ccc} I_{B+1} & \underbrace{\begin{array}{ccc} I_B & \cdots & I_B \\ P_{1\times B} & & P_{1\times B} \end{array}}_{Z-1 \text{ times}} & \begin{array}{cc} 0_{B\times 1} & I_{B-1} \\ \multicolumn{2}{c}{W_{2\times B}} \end{array} \end{array} \right),$$

where $P_{1\times B}$ is the parity matrix of a $(B+1, 1)$ systematic MDS code (e.g., a repitition code) and $W_{2\times B}$ is the parity matrix of a $(B+2, 2)$ systematic MDS code. Furthermore,

$$\left(\begin{array}{cc} \begin{array}{c} 1 \\ P_{1\times B}[B] \end{array} & W_{2\times B}[1] \end{array} \right) \tag{17}$$

must be full-rank, where $A[i]$ denotes the i-th column of matrix A.

Example. For $R = 5/13$, $B = 4$, $Z = 2$, the constructed generator matrix over $\mathrm{GF}(2^2)$ is

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 2 & 1 & 3 & 3 \end{pmatrix}. \quad (18)$$

We can verify that this code meets $T = T^{**}$. The required decoding delay can be written as $T = N - B + 1$, and there are $B - 2$ urgent symbols that need to be recovered by the end of the block. If one of the bursts completely corrupts the last B symbols, then the remaining symbols can be recovered in a similar way as the decoding of (2) (see [5]). If the last B symbols are intact but the first ZB symbols are erased, the urgent symbols can be recovered in the following steps. First, for column $ZB+1$ and $ZB+2$, decode the non-urgent symbols by inverting (18). Second, remove the non-urgent symbols from column $ZB+3, ..., N$ to recover the urgent symbols. One can also verify that if the last B symbols are corrupted partially, the urgent symbols can be recovered on time as well.

B. *Proposition 4: The Converse*

In this section, we show the converse result that if $T < T^{**}$, no feasible code can be constructed.

Consider that the source symbol block $\mathsf{s}[1], ..., \mathsf{s}[K]$ is encoded into the channel symbol block $\mathsf{x}[1], ..., \mathsf{x}[N]$, where $N = K + ZB$ and $K = B + 1$.

Similar to Section IV-B, since the coding scheme must work on source of any distribution, we can arbitrarily assume that the source symbols $\mathsf{s}[i]$, $i = 1, ..., K$ are i.i.d. ~ uniform over Q, thus $H(\mathsf{s}[i]) = H(\mathsf{s}) := \log Q$, $i = 1, ..., K$. Based on this assumption, we establish a set of necessary conditions that the code must meet: The feasibility of the code implies that for any erasure pattern $\mathcal{B} \in \{\mathcal{B} : \mathcal{B}$ consists of ZB-bursts$\}$,

$$H(\mathsf{s}[1], ..., \mathsf{s}[K] \,|\, \mathsf{x}[\{1, ..., N\} \setminus \mathcal{B}]) = 0. \quad (19)$$

From (19) and the Singleton-achieving constraint, it is clear that for any $i \neq j$, $i, j \in \{1, ..., N\} \setminus \mathcal{B}$, $\mathcal{B} \in \{\mathcal{B} : \mathcal{B}$ consists of ZB-bursts$\}$:

$$\mathsf{x}[i] \text{ and } \mathsf{x}[j] \text{ are statistically independent.} \quad (20)$$

Furthermore, by (19) and the Singleton-achieving constraint,

$$H(\mathsf{x}[i]) = H(\mathsf{s}), \quad i = 1, ..., N. \quad (21)$$

From (21), if $H(\mathsf{x}[i] \,|\, \mathsf{s}[j]) = 0$ for some i, j, then

$$\begin{aligned} H(\mathsf{s}[j] \,|\, \mathsf{x}[i]) &= H(\mathsf{x}[i] \,|\, \mathsf{s}[j]) + H(\mathsf{s}[j]) - H(\mathsf{x}[i]) \\ &= 0 + H(\mathsf{s}) - H(\mathsf{s}) = 0. \quad (22) \end{aligned}$$

The Singleton-achieving and causality constraints together imply:

$$H(\mathsf{x}[1] \,|\, \mathsf{s}[1]) = 0. \quad (23)$$

TABLE II

ERASURE PATTERNS CONSIDERED IN THE PROOF OF THEOREM **??**. IN THIS EXAMPLE, $N = 13$, $K = 4$, $B = 3$ AND $Z = 3$.

Time	1	2	3	4	5	6	7	8	9	10	11	12	13
$\mathcal{B}_1$	⋆	⋆	⋆	⋆	⋆	⋆	⋆	⋆	⋆				
$\mathcal{B}_2$	⋆	⋆	⋆	⋆	⋆	⋆		⋆	⋆	⋆			
$\mathcal{B}_3$	⋆	⋆	⋆	⋆	⋆	⋆			⋆	⋆	⋆		
$\mathcal{B}_0$		⋆	⋆	⋆	⋆	⋆	⋆				⋆	⋆	⋆

$$\begin{aligned} H(\mathsf{s}[1] \,|\, \mathsf{x}[ZB+1], \mathsf{x}[ZB+2]) &= 0, \quad (26) \\ H(\mathsf{s}[1] \,|\, \mathsf{x}[ZB+2], \mathsf{x}[ZB-B+1]) &= 0, \quad (27) \\ H(\mathsf{s}[1] \,|\, \mathsf{x}[ZB-B+1], \mathsf{x}[ZB-B+2]) &= 0. \quad (28) \end{aligned}$$

$$\begin{aligned} H(\mathsf{x}[ZB+2] \,|\, \mathsf{s}[1], \mathsf{x}[ZB+1]) &= 0, \quad (29) \\ H(\mathsf{x}[ZB-B+1] \,|\, \mathsf{s}[1], \mathsf{x}[ZB+2]) &= 0, \quad (30) \\ H(\mathsf{x}[ZB-B+2] \,|\, \mathsf{s}[1], \mathsf{x}[ZB-B+1]) &= 0. \quad (31) \end{aligned}$$

By (22), we must have $H(\mathsf{s}[1] \,|\, \mathsf{x}[1]) = 0$. This implies that $\mathsf{s}[1]$ and $\mathsf{x}[1]$ must be one-one.

We would like to show that

$$T \geq K + ZB - B + 1. \quad (24)$$

Assume the opposite is true, for example, let

$$T = K + ZB - B. \quad (25)$$

Then $\mathsf{s}[1]$ must be recoverd at time $T + 1$. We will show that this leads to contradiction. The intuition behind the proof-by-contradiction is as follows. There are three types of requirements that the code must meet: i) feasibility, i.e., (19), ii) decoding delay constraint T (for this proof, only the delay constraint of $\mathsf{s}[1]$ is necessary) and iii) causality, i.e., (23). The causality suggests that $\mathsf{s}[1]$ and $\mathsf{x}[1]$ are one-one, thus $\mathsf{s}[2], ..., \mathsf{s}[K]$ can only be conveyed by $\mathsf{x}[2], ..., \mathsf{x}[N]$. Some symbols from $\mathsf{x}[2], ..., \mathsf{x}[T+1]$ are also used for early recovery of $\mathsf{s}[1]$. It turns out that too many channel symbols are consumed by $\mathsf{s}[1]$, and not enough are left for $\mathsf{s}[2], ..., \mathsf{s}[K]$. The formal argument is as follows.

Step 1: decoding delay constraint. Let $\mathcal{J}_i := \{ZB - B + i, ..., ZB - 1 + i\}$, $i = 1, 2, 3$. Consider a set of 3 erasure patterns (refer to Table II for an example):

$$\mathcal{B}_i = \{1, ..., ZB - B\} \bigcup \mathcal{J}_i, \quad i = 1, 2, 3.$$

The fact that $\mathsf{s}[1]$ must be recovered by $T + 1$ suggests that $H(\mathsf{s}[1] \,|\, \mathsf{x}[\{ZB - B + 1, ..., T + 1\} \setminus \mathcal{J}_i]) = 0$, $i = 1, 2, 3$. Write this out, we have (26) ~ (28).

Now we show (26) implies (29). First, by (20), if we let $\mathcal{B}_0 = \{2, ..., ZB - B + 1\} \bigcup \{N - B + 1, N\}$, we can deduce that $\{\mathsf{x}[1], \mathsf{x}[ZB + 1]\}$ are mutually independent. Since $\mathsf{s}[1]$ and $\mathsf{x}[1]$ are one-one,

$$H(\mathsf{s}[1] \,|\, \mathsf{x}[ZB+1]) = H(\mathsf{s}[1]) = H(\mathsf{s}). \quad (32)$$

We have:

$$\begin{aligned}
& H(\mathsf{x}[ZB+2]\,|\,\mathsf{s}[1],\mathsf{x}[ZB+1]) \\
= \;& H(\mathsf{x}[ZB+2]\,|\,\mathsf{x}[ZB+1]) \\
& -H(\mathsf{s}[1]\,|\,\mathsf{x}[ZB+1]) \qquad (33)\\
\leq \;& H(\mathsf{s})-H(\mathsf{s})=0, \qquad (34)
\end{aligned}$$

where (33) follows (26); (34) follows (21) and (32). Thus, we have proven that (26) implies (29). Following the same procedure, we can prove (27) implies (30) and (28) implies (31). (29)~(31) together imply:

$$H(\mathsf{x}[ZB-B+2]\,|\,\mathsf{s}[1],\mathsf{x}[ZB+1])=0. \qquad (35)$$

Step 2: feasibility. Consider the erasure pattern $\mathcal{B}_0 = \{2,...,ZB-B+1\}\bigcup\{N-B+1,...,N\}$ (refer to Table II for an example). Using the fact that the code is feasible, we have:

$$\begin{aligned}
0 \;=\;& H(\mathsf{s}[1,...,B+1]\,|\,\mathsf{x}[1], \\
& \mathsf{x}[ZB-B+2,...,ZB+1]) \\
=\;& H(\mathsf{s}[2,...,B+1]\,|\,\mathsf{s}[1], \\
& \mathsf{x}[ZB-B+2,...,ZB+1]) \qquad (36)
\end{aligned}$$

where (36) follows that $\mathsf{s}[1]$ and $\mathsf{x}[1]$ are one-one.

Step 3: combining. From (35) and (36),

$$\begin{aligned}
0 \;=\;& H(\mathsf{s}[2,...,B+1]\,|\,\mathsf{s}[1],\mathsf{x}[ZB-B+2,...,ZB+1]) \\
=\;& H(\mathsf{s}[2,...,B+1]\,|\,\mathsf{s}[1],\mathsf{x}[ZB-B+2], \\
& \mathsf{x}[ZB-B+3,...,ZB],\mathsf{x}[ZB+1]) \\
=\;& H(\mathsf{s}[2,...,B+1]\,|\,\mathsf{s}[1], \\
& \mathsf{x}[ZB-B+3,...,ZB+1]) \qquad (37)
\end{aligned}$$

where (37) follows the fact that by (35), $\mathsf{x}[ZB-B+2]$ is completely determined by $\mathsf{s}[1]$ and $\mathsf{x}[ZB+1]$. In (37), as $\mathsf{s}[1]$ is independent of $\mathsf{s}[2,...,B+1]$, we must recover $\mathsf{s}[2,...,B+1]$ from $\mathsf{x}[ZB-B+3,...,ZB+1]$. But this is impossible, since the number of channel symbols in $\mathsf{x}[ZB-B+3,...,ZB+1]$ is $B-1$, which is not enough to recover B source symbols in $\mathsf{s}[2,...,B+1]$. Thus, the assumption (25) is untrue and we have (24).

Finally, combining (4), (5) and (24), we conclude (16).

C. The Converse for Systematic Codes

In this section, we show a converse result that bounds the rate-delay-burst tradeoff of any systematic codes. This result can be regarded as a block-code counterpart of the periodic erasure channel argument for the streaming codes [4].

Theorem 6. *For a delay-T rate-R systematic block erasure code feasible for any erasure patterns of Z B-bursts, (6) applies.*

Proof: Consider the channel symbol block $\mathsf{x}[1],...,\mathsf{x}[N]$, where $\mathsf{x}[i]=\mathsf{s}[i]$, $i=1,...,K$. We would like to show (11). Assume the opposite is true, for example, let $T=K+(Z-1)B-1$. Then $\mathsf{s}[1]$ is recoverd at time $K+(Z-1)B$. Consider the erasure pattern

$$\mathcal{B}_0=\{1,...,B\}\bigcup\{K+1,...,K+(Z-1)B\}.$$

That $\mathsf{s}[1]$ is recovered at time $K+(Z-1)B$ suggests it must be decoded from the remaining symbols $\mathsf{s}[B+1],...\mathsf{s}[K]$. This implies

$$H(\mathsf{s}[1]\,|\,\mathsf{s}[B+1],...,\mathsf{s}[K])=0. \qquad (38)$$

But since the coding scheme must work on source of any distribution, if we assume $\mathsf{s}[i]$, $i=1,...,K$ are mutually independent source symbols, (38) is false. Therefore, (11) must be true.

Finally, similar to Theorem 1, we have (15). Combining (4), (5), (11) and (15), we conclude (6). ■

Remark. Theorem 6 applies to both Singleton-achieving and non-Singleton-achieving codes.

VI. Beyond the Singleton-achieving Codes

In Theorem 1 and Proposition 4, we have imposed the constraint that the codes must be Singleton-achieving, i.e., $N=K+ZB$. In the proof of the converse for Theorem 1, the Singleton-achieving constraint is necessary to lead a contradiction among the erasure patterns $\mathcal{B}_0$ and $\mathcal{B}_1$,...,$\mathcal{B}_{\lceil K/B\rceil}$. In the proof of the converse for Proposition 4, this constraint is necessary to enforce independence among channel symbols in (20), as well as to lead to a contradiction among the erasure patterns.

The natural question to raise is, do the derived bounds still hold if we remove this constraint by letting $N>K+ZB$? In other words, let $N=K+ZB+\Delta N$, $\Delta N>0$, can we improve the rate-delay-burst tradeoff beyond (6) or (16)? In fact, for Theorem 1 and non-causal codes, one can show that when this constraint is removed, the delay can be reduced to $T<K+(Z-1)B$. Consider the example where $K=5$, $Z=1$, $B=2$ and $N=8$, we can design a non-causal code of delay $T=4$ with the following generator matrix:

$$\begin{pmatrix}
1&0&0&0&1&0&0&0\\
0&0&1&0&1&0&0&0\\
0&1&0&0&0&0&1&0\\
0&0&0&1&0&0&0&1\\
0&0&0&0&0&1&0&1
\end{pmatrix}.$$

However, one can verify that its rate-delay-burst tradeoff is not improved beyond (6), since although the delay is reduced, its rate is reduced as well, as we augment the block length. After some attempts of constructing better codes, our general feel is that it may not be justifiable to reduce the rate in order to improve the rate-delay-burst tradeoff. This conjecture remains an open question for future research.

VII. Conclusion

In this work, we have studied the decoding delay effect of erasure correction codes when multiple erasure bursts are present in a coding block. Through a set of converse results and code constructions, our main finding is that the lowest delay to recover any individual symbol not only depends on how many bursts are present in a coding block, but also on whether the source symbols are encoded causally or non-causally.

As this work is on-going, many questions remain open. First, a rate-delay-burst bound for the general Singleton-achieving causal codes as well as code constructions that achieves this bound are unknown. Second, it is still unknown whether the rate-delay-burst bound can be improved if the Singleton-achieving constraint is removed. Lastly, it will be interesting to consider the decoding delay problem on other types of channels.

ACKNOWLEDGMENT

The authors had useful conversations with Emin Martinian, Yuxin Chen and Jinhua Jiang in developing the ideas.

APPENDIX A: PROOF OF LEMMA 3

Proof: Consider

$$\begin{aligned}
H\left(\mathsf{x}_1 \,|\, \mathsf{x}_2^n\right) &\leq H\left(\mathsf{s}, \mathsf{x}_1 \,|\, \mathsf{x}_2^n\right) \\
&= H\left(\mathsf{s} \,|\, \mathsf{x}_2^n\right) + H\left(\mathsf{x}_1 \,|\, \mathsf{s}, \mathsf{x}_2^n\right) \\
&= H\left(\mathsf{x}_1 \,|\, \mathsf{s}, \mathsf{x}_2^n\right) \qquad (39) \\
&\leq H\left(\mathsf{x}_1 \,|\, \mathsf{s}, \mathsf{x}_3^n\right) \\
&= H\left(\mathsf{s} \,|\, \mathsf{x}_1, \mathsf{x}_3^n\right) + H\left(\mathsf{x}_1 \,|\, \mathsf{x}_3^n\right) - H\left(\mathsf{s} \,|\, \mathsf{x}_3^n\right) \\
&= H\left(\mathsf{x}_1 \,|\, \mathsf{x}_3^n\right) - H\left(\mathsf{s} \,|\, \mathsf{x}_3^n\right) \qquad (40) \\
&\leq H\left(\mathsf{x}_1\right) - H\left(\mathsf{s} \,|\, \mathsf{x}_3^n\right) \qquad (41)
\end{aligned}$$

where x_i^j denotes $(\mathsf{x}_i, ..., \mathsf{x}_j)$. In (39) and (40) we apply (8) for $i = 1$ and $i = 2$, respectively. Apply (41), we have:

$$\begin{aligned}
H(\mathsf{X}) &= H\left(\mathsf{x}_1 \,|\, \mathsf{x}_2^n\right) + H\left(\mathsf{x}_2^n\right) \\
&\leq H\left(\mathsf{x}_1\right) - H\left(\mathsf{s} \,|\, \mathsf{x}_3^n\right) + H\left(\mathsf{x}_2^n\right) \\
&\leq \sum_{i=1}^{n} H(\mathsf{x}_i) - H\left(\mathsf{s} \,|\, \mathsf{x}_3^n\right).
\end{aligned}$$

If $H\left(\mathsf{s} \,|\, \mathsf{x}_3^n\right) > 0$, the conclusion follows; otherwise we have

$$H\left(\mathsf{s} \,|\, \mathsf{x}_3^n\right) = 0. \qquad (42)$$

Similarly, apply (42), together with (8) for $i = 3$ to upper bound $H\left(\mathsf{x}_1^2 \,|\, \mathsf{x}_3^n\right)$, we end up with

$$H\left(\mathsf{x}_1^2 \,|\, \mathsf{x}_3^n\right) \leq H\left(\mathsf{x}_1^2\right) - H\left(\mathsf{s} \,|\, \mathsf{x}_4^n\right).$$

If $H\left(\mathsf{s} \,|\, \mathsf{x}_4^n\right) > 0$, the conclusion follows; otherwise we have $H\left(\mathsf{s} \,|\, \mathsf{x}_4^n\right) = 0$. Repeat the same argument until we exhaust all the cases. ■

REFERENCES

[1] A. Begen, "Error control for IPTV over xDSL networks," in *Proc. Consumer Communications and Networking Conference, 2008. (CCNC 2008)*.

[2] V. Subramanian and V. G. Subramanian, "Broadband fading channels: Signal burstiness and capacity," *IEEE Transactions on Information Theory*, vol. 48, pp. 809–827, 1999.

[3] E. Martinian, "Dynamic information and constraints in source and channel coding," *Ph.D. Thesis, Massachusetts Inst. of Technology*, September 2004.

[4] C.-E. W. S. E. Martinian, "Burst erasure correction codes with low decoding delay," *IEEE Trans. on Information Theory*, vol. 50, pp. 2494–2502, October 2004.

[5] E. Martinian and M. Trott, "Delay-optimal burst erasure code construction," in *Proc. International Symposium on Information Theory*, July 2007.

[6] A. Badr, A. Khisti, and E. Martinian, "Diversity embedded streaming codes: Construction and optimality," in *Proc. Globecom*, 2010.