Burst-Erasure Correcting Codes with Optimal Average Delay

Nitzan Adler and Yuval Cassuto

Electronics
Computers
Communications

Burst-Erasure Correcting Codes with Optimal Average Delay

Nitzan Adler and Yuval Cassuto, *Senior Member, IEEE*

Abstract

The objective of low-delay codes is to protect communication streams from erasure bursts by minimizing the time between the packet erasure and its reconstruction. Previous work has concentrated on the constant-delay scenario, where all erased packets need to exhibit the same decoding delay. We consider the case of heterogeneous delay, where the objective is to minimize the average delay across the erased packets in a burst. We derive delay lower bounds for the average case, and show that they match the constant-delay bounds only at a single rate point 0.5. We then construct codes with optimal average delays for the entire range of code rates. The construction for rates under 0.5 achieves optimality for every erasure instance, while the construction for rates above 0.5 is optimal for an infinite number, but not all, of the erasure instances. The paper also studies the benefits of delay heterogeneity within the application of sensor communications. It is shown that a carefully designed code can significantly improve the temporal precision at the receiving node following erasure-burst events.

Keywords

Low-delay codes, erasure codes, burst erasures, average delay, codes for sensor communications

I. Introduction

There are many practical scenarios where a communication system needs to reconstruct corrupt or lost data with minimal delay. These scenarios are common in communication devices with small buffers, and in systems interacting with the physical world. When delay becomes a major concern, one needs to explicitly introduce it to the coding model. A very elegant coding model involving decoding delay has been introduced by Martinian [1], which in particular showed that MDS codes, a common panacea for erasures, are *not* optimal when burst-erasure correction is needed with low delay. The paradigm developed by Martinian – its constructions and bounds – was the basis for several follow-up works promoting different scenarios of low-delay communication: [2] (flexible construction), [3] (multiple bursts), and [4] (multi-user). Additional works addressing low-delay coded communications have appeared in [5], [6], [7], [8], [9], [10], [11], [12], [13].

The prior work has concentrated on the case where every packet in the stream needs to exhibit the same delay. There are many practical scenarios where this restriction is not necessary. For example, in many control networks (e.g. automotive networks), the nodes not only forward data, but also perform computations on it. In such networks it is preferable to obtain part of the data very early, and start the computation while additional packets are being reconstructed. For such scenarios we are considering in this paper *heterogeneous delay*, and seek to minimize the *average delay* of reconstruction, calculated over the packets erased in a burst-erasure instance.

As it turns out, there is a big gap between the achievable delays in the average and constant regimes. It is possible to reduce the delay considerably if one lifts the constant requirement. In particular, in Section III we derive bounds for the average case, and show that they only match the constant case at a single rate point $R = 0.5$. The bounds mark the fundamental limits to the *average* decoding delay given the code rate R and erasure burst length B. The bound is divided to three rate regions, compared to only two regions in the bound for constant delay [1]. Analytical and constructive reasonings about heterogeneous delay are simplified by introducing a new definition of delay we call *recovery delay*. We then move in Section IV to construct codes that match the average-delay bounds for the entire range of code rates. One construction for rates $R \geq 0.5$ achieves optimality for an infinite number of burst instances characterized by their phase with respect to the construction. Another construction for rates $R < 0.5$ achieves optimality for every burst instance. The constructions are given as infinite families of codes with fairly flexible parameters. We first characterize the parameter families that allow average-delay optimality, and focus on them in the constructions. Rates that allow optimality are of the form $m/(m+1)$ for $R \geq 0.5$, and $1/(m+1)$ for $R < 0.5$ (m an integer). For $R \geq 0.5$ we construct codes with optimal average delay for any choice of m, and for any burst length B. We choose to present this as two constructions: one when m and B are co-prime, and another for general m and B. The special case of m and B co-prime enjoys a more regular structure and simpler proofs, thus winning the main attention in the paper.

The new proposal of low average-delay codes is principally motivated by sensor communications in delay-sensitive control applications. In such applications, measurement data packets are transmitted to a remote node over a lossy channel, and the

Nitzan Adler and Yuval Cassuto are with the Department of Electrical Engineering, Technion – Israel Institute of Technology, Haifa Israel (emails: {nitzansh@tx,ycassuto@ee}.technion.ac.il).

receiving node performs time-critical computations on these packets upon their reconstruction. The proposed heterogeneous-delay codes are ideal for these scenarios because they maximize the number of reconstructed packets at any time following the erasure burst. Geared toward such applications, in Section V we illustrate the use of heterogeneous-delay codes in sensor communications where packets carry measurement data with a *significance hierarchy*. In this case the code needs to prioritize high-significance packets to be decoded with a lower delay than low-significance ones. These delay priorities can be viewed as an extension of other priority mechanisms in known packet transmission schemes [14], [15]. The main result of this section is a code construction showing excellent precision performance following an erasure burst, and in particular significantly better than a known code with optimal constant delay. The discussion of Section V also reveals the important aspect of designing the code with an appropriate reconstruction order for the packets.

II. Low Delay Codes

A. Previous work – constant delay

The low-delay coding paradigm was founded in the work of Martinian [1]. We now briefly review its model and main results. A low-delay erasure coding scheme facilitates packet transmission subject to bursts of erasures. Each packet comprises a number of symbols taken from some alphabet. The packets are encoded by a causal encoder, i.e., the coded symbols in a packet depend only on information symbols from past and present packets. The objective of the low-delay code is to reconstruct the symbols of an erased packet *at a minimum temporal delay* from the time of the packet's original transmission. Hence low-delay codes are characterized by a tradeoff between three parameters: the erasure-burst length B, the reconstruction delay T – both B and T are measured in units of packets – and the code rate R defined as the ratio between the information content of the packet and its total size. These three parameters are shown in [1] to satisfy the following inequality relation

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

The implication of (1) is a lower bound on the reconstruction delay given the burst length and the code rate. Another important result of [1] is that this bound is achievable, through an explicit family of codes whose parameters satisfy (1) with equality.
The above model and the codes constructed for it have the property of *constant delay*, i.e, each erased packet in a burst is reconstructed with the same delay from the time of its transmission. Our objective in this paper is to lift the constant-delay property, and examine codes that provide different delays to different packets in the burst.

B. Heterogeneous delay

To deal with heterogeneous delay, we define T_i to be the delay of the i-th erased packet in a given erasure burst. That is, T_i packets were transmitted between the transmission of packet i and its eventual reconstruction, i.e., until its last missing information symbol is reconstructed. T_1 is the delay of the first erased packet in the burst, i.e, the oldest one, and T_B is the last, most recently erased packet. This is shown in Figure 1. A natural performance measure for heterogeneous delay is the average delay across the erasure burst.

Definition 1 *The average delay over all B erased packets is defined as*

$$\bar{T} = \frac{\sum_{i=1}^{B} T_i}{B} \,.$$

Figure 1: The non-constant delay T_i.
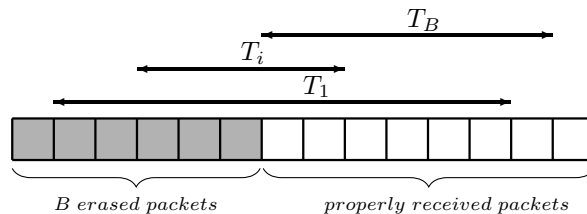
When using a constant-delay code, then clearly $\bar{T} = T$, so the average delay must obey the lower bound (1). However, for the heterogeneous-delay case, (1) in general is *not* a lower bound on $\bar{T}$. It is therefore interesting to investigate the possibility to improve the average delay beyond the best-achievable constant-delay codes. This investigation will include both lower bounds

on the average delay, and code constructions with lower average delays. Toward that objective the following definition will be useful.

Definition 2 *The recovery delay of the i-th packet, denoted κ_i, is the number of packets that were received from the end of the erasure burst until the i-th packet is fully recovered.*

Hence the recovery delay differs from the standard delay of [1] (see sub-section II-A) in that it does not include the packets erased in the burst. An example illustrating this is given in Fig 2.
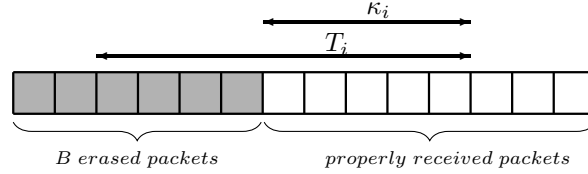

Figure 2: The recovery delay κ_i vs. the standard delay T_i.

It is a simple but helpful observation that the *average* recovery delay is equal to the average standard delay plus some constant depending only on B.

Proposition 1 *A code that can correct a burst erasure with length B and has an average delay $\bar{T}$ and an average recovery delay $\bar{\kappa}$ must satisfy*

$$\bar{T} = \bar{\kappa} + \frac{B-1}{2} \ . \tag{2}$$

Proof: According to the definition of κ_i we can write the following relation between κ_i and T_i

$$\kappa_i = T_i - (B - i) \tag{3}$$

and by calculating the average $\bar{\kappa}$ over all B erased packets we obtain the relation (2). ∎

We will now see some simple examples for the profile of the recovery delay $\{\kappa_i\}_{i=1}^{B}$ in comparison to the profile of the standard delay $\{T_i\}_{i=1}^{B}$.

Example 1 *Constant delay: In this case the delay has a constant value, i.e., $\forall i \quad T_i = T$. Therefore the recovery delay of the first erased packet κ_1 is the shortest one, and according to (3) the recovery delay of the next erased packet κ_2 must be larger than κ_1 exactly by 1 in order to satisfy condition $T_1 = T_2 = T$. In general we can say that $\kappa_i = \kappa_{i-1} + 1$, so the packet with the longest recovery delay has $\kappa_B = T$. The resulting recovery delay profile $\{\kappa_i\}_{i=1}^{B}$ now follows, and illustrated pictorially in Figure 3.*

$$\{\kappa_i\}_{i=1}^{B} = \{\kappa_1 = T - (B-1), \ \kappa_2 = T - (B-2), \ \dots, \kappa_{B-1} = T - 1, \ \kappa_B = T\} \ .$$

The average recovery delay is clearly $\bar{\kappa} = T - \frac{B-1}{2}$.
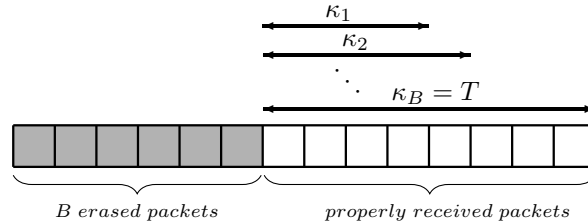
Figure 3: The profile of the recovery delay $\{\kappa_i\}_{i=1}^{B}$ for a constant delay.

Example 2 *MDS codes: When low-delay codes are constructed from $[n, k]$ MDS block codes as suggested in [1] (using techniques from [16]), packets are reconstructed after the arrival of k properly received packets[1] following the erasure burst of length $n - k$.*

---

[1] To simplify the discussion we ignore cases where a packet reconstruction is split between before and after the burst.

Thus we have for this case that the recovery delay is constant $\kappa_i = \kappa = k \quad \forall i$. According to (3) we can find out that the profile of the standard delay is as follows, and illustrated in Figure 4.

$$\{T_i\}_{i=1}^{B} = \{T_1 = k + (B-1), \; T_2 = k + (B-2), \; \ldots \; T_{B-1} = k+1, \; T_B = k\} \;,$$

The average standard delay is
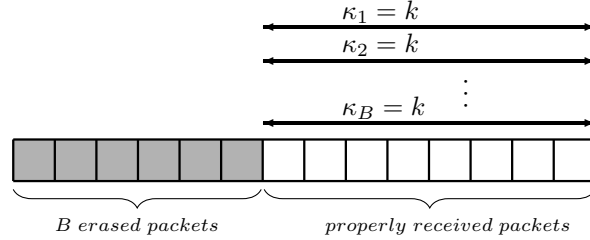
$$\bar{T} = k + \frac{B-1}{2} \; .$$

Figure 4: The profile of the recovery delay $\{\kappa_i\}_{i=1}^{B}$ for MDS codes.

III. BOUNDS ON THE AVERAGE DELAY

The known delay lower bounds like the one in (1) only apply to constant delay, and it is not clear what limits exist for the average delay in the heterogeneous case. Hence in this section we seek such bounds on the average delay. Similarly to the previous work, our discussion will be based on causal encoders.

Theorem 2 *Let $\{T_i\}_{i=1}^{B}$ be the delay profile obtained for a given decoding instance following an erasure burst of length B. Then for a code with rate R the average delay $\bar{T}$ must satisfy*

$$\bar{T} \geq \begin{cases} \frac{R}{1-R} \cdot \frac{B+1}{2} + \frac{B-1}{2} & ; \quad R \geq \frac{1}{2} \\[2mm] \frac{B}{2} \cdot \left(\frac{R}{1-R} + 1 \right) & ; \quad \frac{1}{1+B} \leq R < \frac{1}{2} \\[2mm] \frac{B+1}{2} & ; \quad R < \frac{1}{1+B} \end{cases} \tag{4}$$

It can be seen in (4) that the lower bound is split to three rate intervals, applying to the rate intervals $R < \frac{1}{1+B}$, $\frac{1}{1+B} \leq R < \frac{1}{2}$ and $R \geq \frac{1}{2}$, respectively. Our proof will be divided into two parts: $R \geq \frac{1}{2}$ and $R < \frac{1}{2}$.
We first give the following useful lemma.

Lemma 3 *A rate-R code carrying packets with S information symbols and P parity symbols each has a recovery-delay profile $\{\kappa_i\}_{i=1}^{B}$ that satisfies*

$$\kappa_{\max} \geq \frac{BS}{P} = B \cdot \frac{R}{1-R} \;,$$

where $\kappa_{\max} \triangleq \max_{1 \leq i \leq B} \{\kappa_i\}$.

Proof: Because the encoder is causal, reconstructing all B erased packets requires at least BS/P packets following the burst. There is at least one packet i not yet fully decoded after $BS/P - 1$ packets following the burst, and this packet thus must have $\kappa_i \geq BS/P$. ∎

For convenience, Lemma 3 is stated for systematic codes, but a similar result follows for non-systematic codes. Note that $\kappa_{\max}$ is the *guard space*, which is the minimal number of unerased packets required for full recovery of all B erased packets. We are now ready to prove Theorem 2 for $R \geq \frac{1}{2}$.

Proof for $R \geq \frac{1}{2}$: In this proof we will use an inductive version of the proof of Lemma 3. From the lemma we already know that $\kappa_{\max}$ is at least $\lceil B \cdot \frac{S}{P} \rceil$, since every κ_i must be an integer. Now looking at the other $B-1$ packets, the maximal recovery delay among them is at least $\lceil (B-1) \frac{S}{P} \rceil$. Continuing this argument inductively we get the following lower bound for $\bar{\kappa}$

$$\bar{\kappa}_{R \geq \frac{1}{2}} \geq \frac{\lceil \frac{S}{P} \rceil + \cdots + \lceil (B-1) \cdot \frac{S}{P} \rceil + \lceil B \cdot \frac{S}{P} \rceil}{B} \geq \frac{\frac{S}{P} + \cdots + (B-1) \cdot \frac{S}{P} + B \cdot \frac{S}{P}}{B} = \frac{R}{1-R} \cdot \frac{B+1}{2} \; . \tag{5}$$

Note that we did not need the $R \geq \frac{1}{2}$ assumption to get the bound, but the assumption (equiv. $S/P \geq 1$) is necessary to make the second inequality tight by setting $\frac{S}{P} \in \mathbb{N}$. Now with the help of (2) we can write a lower bound for the average delay

$$\bar{T} = \bar{\kappa} + \frac{B-1}{2} \geq \frac{R}{1-R} \cdot \frac{B+1}{2} + \frac{B-1}{2} . \tag{6}$$

∎

Since the lower bound can be achieved only if $\frac{S}{P} \in \mathbb{N}$, we must choose the parameters S and P so they will satisfy $S = mP$, $m \in \mathbb{N}$, therefore in order to attain the lower bound for $R \geq \frac{1}{2}$ the coding rate must be of the form $R = \frac{m}{m+1}$. This condition can be satisfied only if $R \geq \frac{1}{2}$. When $R < \frac{1}{2}$ the lower bound (6) is still correct but certainly unachievable, therefore we would like to find a better lower bound for $R < \frac{1}{2}$.

Proof for $R < \frac{1}{2}$: To make the proof simpler we will prove a special case where $\frac{P}{S} \in \mathbb{N}$ and $\frac{BS}{P} \in \mathbb{N}$. In Appendix A we show that for $R < \frac{1}{2}$, if at least one of these conditions is not satisfied, the average delay becomes strictly higher, i.e., the lower bound still applies to the general case $\frac{P}{S} \in \mathbb{Q}$, and $\frac{BS}{P} \in \mathbb{Q}$ and it may be attained only when $\frac{P}{S} \in \mathbb{N}$ and $\frac{BS}{P} \in \mathbb{N}$. For this special case we can calculate the lower bound as follows

$$\bar{\kappa}_{R<\frac{1}{2}} \geq \frac{\left\lceil \frac{S}{P} \right\rceil + \left\lceil 2\frac{S}{P} \right\rceil + \cdots + \left\lceil (B-1) \cdot \frac{S}{P} \right\rceil + \left\lceil B \cdot \frac{S}{P} \right\rceil}{B} = \frac{\overbrace{1+1+\ldots+1}^{\frac{P}{S} \ times} + \overbrace{2+2+\ldots+2}^{\frac{P}{S} \ times} + \ldots + \overbrace{\frac{BS}{P} + \ldots + \frac{BS}{P}}^{\frac{P}{S} \ times}}{B}$$

$$= \frac{\frac{P}{S} \cdot \sum_{i=1}^{B \cdot \frac{S}{P}} i}{B} = \frac{P}{S} \cdot \frac{B \cdot \frac{S}{P} \left(B \cdot \frac{S}{P} + 1 \right)}{2B} = \frac{B}{2} \cdot \frac{S}{P} + \frac{1}{2} = \frac{B}{2} \cdot \frac{R}{1-R} + \frac{1}{2} .$$

The first inequality follows from the same argument as for $R \geq \frac{1}{2}$ as in (5). The first equality follows from the assumptions $\frac{P}{S} \in \mathbb{N}$ and $\frac{BS}{P} \in \mathbb{N}$. In this case, by each received packet after the burst we can reconstruct at most $\frac{P}{S}$ erased packets.

For the sub-interval $R < \frac{1}{1+B}$ the bound can be further tightened to $\bar{\kappa} \geq 1$, which follows trivially from the fact that each κ_i is at least 1. So altogether we obtain

$$\bar{\kappa}_{R<\frac{1}{2}} \geq \begin{cases} \frac{B}{2} \cdot \frac{R}{1-R} + \frac{1}{2} & ; \quad \frac{1}{1+B} \leq R < \frac{1}{2} \\ \\ 1 & ; \quad R < \frac{1}{1+B} \end{cases}$$

and with the help of (2) we get exactly the bottom two arguments of the lower bound (4).

∎

Figure 5 shows the lower bound for the average delay (solid) in comparison to the known lower bound for constant delay (dashed). It can be seen that there is a significant gap between the two bounds for all rates, except for the singular point $R = \frac{1}{2}$ where the two bounds coincide. The equality between these two bounds at that single point has an intuitive explanation. When $R = \frac{1}{2}$ then $S = P$, so obtaining the optimal average delay must be done by reconstructing exactly one different erased packet at each time unit. By choosing a specific order of reconstruction (whereby the order of reconstruction is identical to the transmission order), we get a constant-delay profile that is also optimal in average delay.

IV. CONSTRUCTIONS WITH OPTIMAL AVERAGE DELAY

After deriving lower bounds on the average delay, our objective is to find code constructions that attain these bounds. Similar to the bound derivation in Section III, the constructions of this section will split to codes with $R \geq \frac{1}{2}$ and codes with $R < \frac{1}{2}$.

A. Notations and definitions

A code construction is specified below through its encoder. The code is systematic, so each packet contains a systematic part with information symbols and a redundancy part with parity symbols. Each packet has an integer time index representing its order in the packet sequence. The encoder is causal, so parity symbols of a packet at time i are computed only from packets with time indices smaller than i. We now list some notations that will be used in the sequel.

- $s_j[i]$ – the j-th information symbol at time i, $s_j[i] \in F$ where F is a finite field. We refer to j as the *serial index* and to i as the *time index*.
- $s_j^l[i]$ – a set of information symbols at time i containing $\{ s_j[i], s_{j+1}[i], \ldots, s_l[i] \}$.
- $P\{ s_{j_1}[i_1], s_{j_2}[i_2], \ldots, s_{j_k}[i_k] \}$ – is a generalized parity-check function taking k information symbols, and returning one symbol from the same field F. P needs to satisfy the property that any of its arguments can be reconstructed from
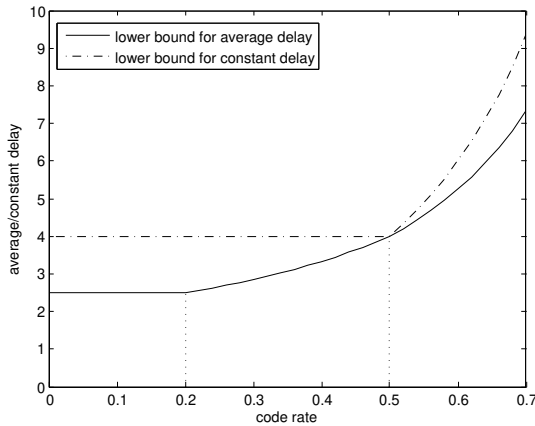
Figure 5: Lower bounds on decoding delay: average delay vs. constant delay, for $B = 4$.

the output and the remaining $k - 1$ arguments. One may think of P as the simple element-wise parity function over a finite field. We refer to P's k arguments and its output as a *parity group* (of size $k + 1$). Since P is a symmetric function in its arguments, we specify the arguments of P as a set.

- $\vec{x}[i]$ – the full packet at time i, including the information symbols and the parity symbol.

B. Construction for rates $R \geq \frac{1}{2}$

In this part we construct codes for $R \geq \frac{1}{2}$, which arguably is the more practically interesting case. Recall from Section III that achieving the lower bound (4) requires $R = \frac{m}{1+m}$ with $m \in \mathbb{N}$, so the construction will consider only rates of this form. In the first construction we use B and m that are co-prime, i.e., $gcd(B, m) = 1$.

Construction 1 *For co-prime B and m we define the following encoder specifying the packet output at time i*

$$\vec{x}[i] = \left(s_0^{m-1}[i] \ , \ x^P[i]\right),$$

where

$$x^P[i] \triangleq P\left\{\left\{s_{\langle i \rangle_m}[i - \langle i \rangle_m - 1 - jB]\right\}_{\forall j \in J}\right\}$$

is the parity symbol, and

$$J \triangleq \left\{j \in \mathbb{Z} \ : \ -\left\lfloor \frac{\langle i \rangle_m}{B} \right\rfloor \leq j \leq -\left\lceil \frac{\langle i \rangle_m + 2}{B} \right\rceil + (m+1)\right\} . \tag{7}$$

We use the notation $\langle i \rangle_m \triangleq i \bmod m$.

The encoder adds a parity symbol to every m information symbols, yielding a code with rate $R = \frac{m}{1+m}$. It can be seen that the encoder of Construction 1 is systematic and causal. The range of j in (7) follows from the inequality

$$i - mB - (B - 1) \leq i - \langle i \rangle_m - 1 - jB \leq i - 1 \ ; \quad j \in \mathbb{Z} \ ,$$

where the upper bound is from causality, and the lower bound is from the fact that the reconstruction procedure takes mB time units after a burst of B erased packets. In Table I we can see an example of the encoding for $B = 3$ and $m = 2$.

If $B > m$ we get a simpler interval of j values, in which case (7) becomes $J = \{0, 1, 2, \ldots, m\}$. Therefore if $B > m$ we can write the encoding of the parity symbol as

$$x^P[i] = P\left\{\left\{s_{\langle i \rangle_m}[i - \langle i \rangle_m - 1 - jB]\right\}_{j=0}^{m}\right\} . \tag{8}$$

The decoder *The packet $\vec{x}[i]$, with m information symbols and an encoded parity symbol, is received by the decoder at time i. As long as no erasure occurred, the decoder uses only the information symbols. When the decoder identifies an erasure with burst-length B, it performs a reconstruction procedure during the next mB time units. Let us assume that the first packet received after the burst is at time index $\tilde{i}$, i.e., the erased packets are at time units $\tilde{i} - 1, \tilde{i} - 2, \ldots, \tilde{i} - B$. For each received packet at times $\tilde{i} \leq i \leq \tilde{i} + mB - 1$ the decoder finds the missing argument in $x^P[i]$ by finding $j \in J$ that satisfies the condition*

$$i - \langle i \rangle_m - 1 - jB \in \left\{\tilde{i} - 1, \tilde{i} - 2, \ldots, \tilde{i} - B\right\} . \tag{9}$$

Now the decoder uses $x^P[i]$ and the known other m arguments of $x^P[i]$ to reconstruct the following symbol that was erased in the burst

$$s_{\langle i \rangle_m}\left[i - \langle i \rangle_m - 1 - jB\right] . \tag{10}$$

Example 3 *For $B = 3$ and $m = 2$ the encoder is specified in Table I. Suppose packets $\vec{x}[-3]$, $\vec{x}[-2]$ and $\vec{x}[-1]$ were erased, meaning the first received packet after the burst is $\vec{x}[0]$. Then, according to the specification of the decoder, the parity of $\vec{x}[0]$ will reconstruct $s_0[-1]$ from $s_0[-4]$ and $s_0[-7]$. Then the parity of $\vec{x}[1]$ will reconstruct $s_1[-1]$ from $s_1[-4]$ and $s_1[-7]$, and so forth until the parity of $\vec{x}[5]$ will reconstruct $s_1[-3]$ from $s_1[0]$ and $s_1[3]$. The resulting recovery delays of $\vec{x}[-3]$, $\vec{x}[-2]$ and $\vec{x}[-1]$ are 6, 4 and 2, respectively.*

$$
\begin{aligned}
\vec{x}[-4] &= \left(s_0^1[-4] \ , \ P\{s_0[-5] \ , \ s_0[-8] \ , \ s_0[-11]\}\right) \\
\vec{x}[-3] &= \left(s_0^1[-3] \ , \ P\{s_1[-5] \ , \ s_1[-8] \ , \ s_1[-11]\}\right) \\
\vec{x}[-2] &= \left(s_0^1[-2] \ , \ P\{s_0[-3] \ , \ s_0[-6] \ , \ s_0[-9]\}\right) \\
\vec{x}[-1] &= \left(s_0^1[-1] \ , \ P\{s_1[-3] \ , \ s_1[-6] \ , \ s_1[-9]\}\right) \\
\vec{x}[0] &= \left(s_0^1[0] \ , \ P\{s_0[-1] \ , \ s_0[-4] \ , \ s_0[-7]\}\right) \\
\vec{x}[1] &= \left(s_0^1[1] \ , \ P\{s_1[-1] \ , \ s_1[-4] \ , \ s_1[-7]\}\right) \\
\vec{x}[2] &= \left(s_0^1[2] \ , \ P\{s_0[1] \ , \ s_0[-2] \ , \ s_0[-5]\}\right) \\
\vec{x}[3] &= \left(s_0^1[3] \ , \ P\{s_1[1] \ , \ s_1[-2] \ , \ s_1[-5]\}\right) \\
\vec{x}[4] &= \left(s_0^1[4] \ , \ P\{s_0[3] \ , \ s_0[0] \ , \ s_0[-3]\}\right) \\
\vec{x}[5] &= \left(s_0^1[5] \ , \ P\{s_1[3] \ , \ s_1[0] \ , \ s_1[-3]\}\right)
\end{aligned}
$$

TABLE I: Example for the encoder of Construction 1 when $B = 3$ and $m = 2$.

One may observe that decoding in Example 3 yields average delay of 5 (average recovery delay of 4), which satisfies the bound (4) with equality. We now turn to show that this optimality applies in more generality. To prove the precise optimality statement we first give the following definition.

Definition 3 *The burst's phase shift ϕ is defined as $\langle \tilde{i} \rangle_m$, where $\tilde{i}$ is the time unit of the first received packet after the length B burst.*

Theorem 4 *For a code specified by Construction 1, any burst with $\phi = 0$ has an average delay of*

$$\bar{T} = \frac{mB + m + B - 1}{2} .$$

Proof: In order to simplify the proof we will discuss only the case when $B > m$, therefore $J = \{0, 1, 2, \ldots, m\}$. The proof for the complement case $B < m$ can be found in Appendix B. Before arguing about delays, we prove the correctness of the decoder, i.e., that any length B burst with any phase shift can be successfully reconstructed during mB time units after the burst. Examining (8), we see that all the time arguments are distant at least B time units from one another, so there can be up to one missing argument in every parity function. The correctness proof will work in two parts: in part 1 we show that each of the mB packets received after the burst can reconstruct an information symbol in an erased packet; in part 2 we show that the received packets reconstruct *distinct* erased symbols. <u>Part 1</u>: we now show that every packet with time index $\tilde{i} + \alpha$, $0 \le \alpha \le mB - 1$, contains in its parity group a former information symbol with time index $\tilde{i} - \beta$, $1 \le \beta \le B$. That is, every received packet at the interval $\{\tilde{i}, \ldots, \tilde{i} + mB - 1\}$ contains in its parity group an information symbol from an erased packet, and as said above it is the only argument that is missing. According to (8), the time arguments of the parity function at time $\tilde{i} + \alpha$ are

$$\left[\tilde{i} + \alpha - \langle \tilde{i} + \alpha \rangle_m - 1 - jB\right]_{j=0}^m . \tag{11}$$

The decoder needs to find the index $j \in \{0, 1, \ldots, m\}$ that points from the received packet to the erased symbol to reconstruct. For the packet received at time $\tilde{i} + \alpha$, let us examine the expression $\langle \alpha \rangle_B - \langle \tilde{i} + \alpha \rangle_m$. Its range is

$$1 - m \ \le \ \langle \alpha \rangle_B - \langle \tilde{i} + \alpha \rangle_m \ \le \ B - 1 ,$$

since $0 \le \langle \alpha \rangle_B \le B - 1$ and $0 \le \langle \tilde{i} + \alpha \rangle_m \le m - 1$. When $1 - m \le \langle \alpha \rangle_B - \langle \tilde{i} + \alpha \rangle_m \le 0$ we choose $j = \lfloor \frac{\alpha}{B} \rfloor$, and when $1 \le \langle \alpha \rangle_B - \langle \tilde{i} + \alpha \rangle_m \le B - 1$ we choose $j = \lfloor \frac{\alpha}{B} \rfloor + 1$. In both cases $j \in \{0, 1, \ldots, m\}$, since $0 \le \lfloor \frac{\alpha}{B} \rfloor \le m - 1$ according to the definition of α. In the first case (11) becomes

$$\tilde{i} + \alpha - \langle \tilde{i} + \alpha \rangle_m - 1 - \left\lfloor \frac{\alpha}{B} \right\rfloor B \ = \ \tilde{i} + \langle \alpha \rangle_B - \langle \tilde{i} + \alpha \rangle_m - 1,$$

which is bounded by

$$\tilde{i} - m \ \leq \ \tilde{i} + \langle\alpha\rangle_B - \langle\tilde{i}+\alpha\rangle_m - 1 \ \leq \ \tilde{i} - 1 \ .$$

Since $B > m$ it has the required form $\tilde{i} - \beta$ where $1 \leq \beta < B$. In the second case

$$\tilde{i} + \alpha - \langle\tilde{i}+\alpha\rangle_m - 1 - \left(\left\lfloor\frac{\alpha}{B}\right\rfloor + 1\right)B \ = \ \tilde{i} + \langle\alpha\rangle_B - \langle\tilde{i}+\alpha\rangle_m - 1 - B,$$

which is bounded by

$$\tilde{i} - B \ \leq \ \tilde{i} + \langle\alpha\rangle_B - \langle\tilde{i}+\alpha\rangle_m - 1 - B \ \leq \ \tilde{i} - 2 \ ,$$

also having the required form. Altogether we proved that every received packet with time index $\tilde{i} + \alpha, \ 0 \leq \alpha \leq mB - 1$, can reconstruct an erased information symbol. <u>Part 2</u>: we now show that each information symbol appears in not more than one parity group with a time index of the form $\tilde{i} + \alpha, \ 0 \leq \alpha \leq mB - 1$. We assume in way of contradiction that there exist two distinct integers α' and α'' where $0 \leq \alpha', \alpha'' \leq mB - 1$, such as the two packets $\tilde{i} + \alpha'$ and $\tilde{i} + \alpha''$ contain the same former information symbol in their parity group, therefore

$$s_{\langle\tilde{i}+\alpha'\rangle_m}\left[\tilde{i} + \alpha' - \langle\tilde{i}+\alpha'\rangle_m - 1 - j'B\right] = s_{\langle\tilde{i}+\alpha''\rangle_m}\left[\tilde{i} + \alpha'' - \langle\tilde{i}+\alpha''\rangle_m - 1 - j''B\right] \ .$$

We assume w.l.o.g. that $a' < a''$. Both of the indices (the time index and the serial index) should be equal. Out of the equality of the serial indices we can conclude that

$$\langle\tilde{i}+a'\rangle_m = \langle\tilde{i}+\alpha''\rangle_m \ \Longleftrightarrow \ m|\alpha'' - \alpha' \ . \tag{12}$$

Out of the equality of the time indices and (12)

$$\alpha' - j'B = \alpha'' - j''B \ \Longleftrightarrow \ \alpha'' - \alpha' = (j'' - j')B \ .$$

Since $gcd(B, m) = 1$ and $m|\alpha'' - \alpha'$, we know that $m|j'' - j'$, therefore $mB|\alpha'' - \alpha'$, and since $0 \leq \alpha', \alpha'' \leq mB - 1$ we can conclude that $\alpha'' - \alpha' = 0 \Longleftrightarrow \alpha' = \alpha''$, a contradiction. So each received packet with time index $\tilde{i} + \alpha, \ 0 \leq \alpha \leq mB - 1$, contains in its parity group different former information symbols.

Combining parts 1 and 2, we find that by receiving the mB packets with time indices $\tilde{i}, \ldots, \tilde{i} + mB - 1$ we reconstruct all the mB information symbols erased in the length B burst. This fact is true regardless of the burst phase shift.

After proving the correctness of the decoder, we analyze its average delay based on finding which erased information symbol is reconstructed at each time index. According to (8), we can see that by every m sequential packets, starting from $\langle i\rangle_m = 0$, at the time interval $\tilde{i} \leq i \leq \tilde{i} + mB - 1$, we obtain the same reconstructed time index, since they have the same $i - \langle i\rangle_m$. On the other hand, the serial index of these m reconstructed symbols, which is defined by $\langle i\rangle_m$, increases by 1 with every received packet. Therefore, by these m sequential packets, we reconstruct all m information symbols of the same erased packet. For the case $\phi = 0$, the serial index of the first reconstructed symbol is $\langle i\rangle_m = 0$, so the reconstruction procedure starts exactly at the beginning of a series of m received packets reconstructing the same erased packet. Hence the first reconstructed packet is fully recovered after m time units. Afterwards, the time index jumps to a different erased value every m received packets, and according to the correctness proof we know that all erased time indices will be successfully reconstructed. Therefore, the recovery delay profile for $\phi = 0$ is

$$\{\kappa_i\}_{i=1}^B = \{m, 2m, 3m, \ldots, Bm\} \ , \tag{13}$$

and the average recovery delay and standard delay are

$$\bar{\kappa} = \frac{mB + m}{2} \ \Longleftrightarrow \ \bar{T} = \frac{mB + m + B - 1}{2} \ .$$

∎

Example 4 $\underline{B = m + 1}$: *We will now demonstrate the order of reconstruction when $\phi = 0$, for the special case $B = m+1$. We can see that $\overline{gcd(B, m)} = 1$ and $B > m$. Table II shows which erased symbol is being reconstructed at each time unit. We can see that every m sequential received packets are used for the reconstruction of m different symbols of the same packet, and after mB received packets all B erased packets were successfully recovered, therefore the recovery delay profile is as (13). In this special case we start the reconstruction with the last erased packet $\tilde{i} - 1$, and continue with $\tilde{i} - 2$, and so on until the last reconstructed packet is the first to be erased $\tilde{i} - B$. For different relations between m and B we will get different order of reconstruction, but with the same average delay.*

We reiterate that the average-delay lower bound of Theorem 2 applies to *every decoding instance*, and thus Theorem 4 implies the following optimality statement.

Corollary 5 *Construction 1 achieves the optimal average delay for an infinite number of decoding instances.*

time index of received packet	the reconstructed symbol	the value of j in (10)
$\tilde{i}$	$s_0\left[\tilde{i}-1\right]$	0
$\tilde{i}+1$	$s_1\left[\tilde{i}-1\right]$	0
$\tilde{i}+2$	$s_2\left[\tilde{i}-1\right]$	0
$\vdots$	$\vdots$	$\vdots$
$\tilde{i}+m-1$	$s_{m-1}\left[\tilde{i}-1\right]$	0
$\tilde{i}+m$	$s_0\left[\tilde{i}-2\right]$	1
$\tilde{i}+m+1$	$s_1\left[\tilde{i}-2\right]$	1
$\vdots$	$\vdots$	$\vdots$
$\tilde{i}+2m-1$	$s_{m-1}\left[\tilde{i}-2\right]$	1
$\tilde{i}+2m$	$s_0\left[\tilde{i}-3\right]$	2
$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$
$\tilde{i}+mB-1$	$s_{m-1}\left[\tilde{i}-B\right]$	m

TABLE II: Example for the order of reconstruction where $B=m+1$ and $\phi=0$ that shows which erased information symbol is reconstructed at each time unit in the interval $\tilde{i} \le i \le \tilde{i}+mB-1$.

Where infinite decoding instances refers to all instances with $\phi=0$.

In the following we give the average delay of Construction 1 for a general phase shift.

Theorem 6 *For a given burst with a non-zero phase shift $1 \le \phi \le m-1$, a code specified by Construction 1 achieves the average delay*

$$\bar{T} = \frac{mB+m+B-1}{2} + \frac{(B-1)(m-\phi)}{B} . \tag{14}$$

Proof: The correctness of the decoder for all phase shifts was established in Theorem 4. The difference between the two cases ($\phi=0$ and $\phi \neq 0$) is that for $\phi \neq 0$ the serial index of the first reconstructed symbols is nonzero, so the reconstruction procedure starts somewhere in the middle of an erased packet, according to ϕ. Therefore, by the first $m-\phi$ received packets we reconstruct $m-\phi$ different information symbols of the same erased packet, but afterwards the reconstruction of a different erased packet starts, before full recovery of the first one. So the first fully recovered packet has a recovery delay of $2m-\phi$, and the next one $3m-\phi$ and so on, until $B-1$ erased packets are fully recovered. Only then we return to the packet whose reconstruction we had started with, so its recovery delay is mB. In Table III we see an example for the order of reconstruction when $B=m+1$ and $\phi=1$. Finally, the recovery delay profile when $\phi \neq 0$ is

$$\{\kappa_i\}_{i=1}^{B} = \{2m-\phi,\, 3m-\phi,\, \ldots,\, Bm-\phi,\, Bm\} \ ,$$

and the average delay is as in (14). ∎

If we define $\phi=0$ as $\phi=m$, the average-delay expression in (14) will be correct for all phase shifts $\phi \in \{1, 2, \ldots, m\}$. Moreover, according to (14) we can calculate the expected average delay of Construction 1, assuming burst positions are drawn uniformly

$$E_\phi\{\bar{T}\} = \frac{mB+m+B-1}{2} + \frac{(B-1)(m-1)}{2B} . \tag{15}$$

To evaluate the delay performance of Construction 1, we plot in Figure 6 the min ($\phi=0$), max ($\phi=1$), and expected (15) average delay. We see that for every $B>2$ Construction 1 gives superior delay compared to optimal constant-delay codes, including with respect to the maximum average delay among the phase shifts. The gap between the average delay of Construction 1 and the optimal constant delay follows from the fact that in Construction 1 the majority of reconstructed packets have significantly shorter delays than the optimal constant delay, and only a small number of packets have longer or equal delays.

The requirement for Construction 1 to have $gcd(m,B)=1$ is not fundamental, and optimal average delay can also be obtained for m and B that are not co-prime. This is demonstrated in the following Construction 2.

time index of received packet	the reconstructed symbol	the value of j in (10)
$\tilde{i}$	$s_1\left[\tilde{i}-2\right]$	0
$\tilde{i}+1$	$s_2\left[\tilde{i}-2\right]$	0
$\vdots$	$\vdots$	$\vdots$
$\tilde{i}+m-2$	$s_{m-1}\left[\tilde{i}-2\right]$	0
$\tilde{i}+m-1$	$s_0\left[\tilde{i}-3\right]$	1
$\tilde{i}+m$	$s_1\left[\tilde{i}-3\right]$	1
$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$
$\tilde{i}+(B-1)\,m-2$	$s_{m-1}\left[\tilde{i}-B\right]$	$m-1$
$\tilde{i}+(B-1)\,m-1$	$s_0\left[\tilde{i}-1\right]$	$m-1$
$\vdots$	$\vdots$	$\vdots$
$\tilde{i}+Bm-2$	$s_{m-1}\left[\tilde{i}-1\right]$	$m-1$
$\tilde{i}+Bm-1$	$s_0\left[\tilde{i}-2\right]$	m

TABLE III: Example for the order of reconstruction where $B=m+1$ and $\phi=1$ that shows which erased information symbol is reconstructed at each time unit in the interval $\tilde{i}\leq i\leq \tilde{i}+mB-1$.
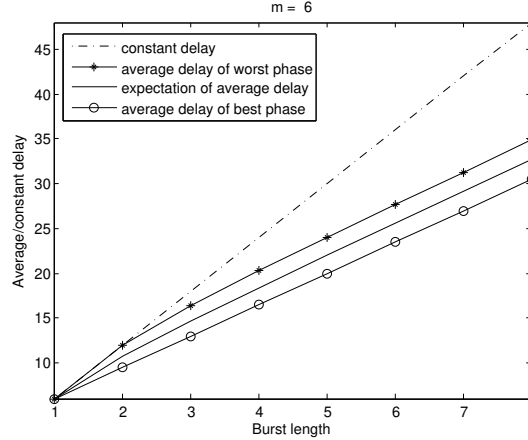

Figure 6: The average delay of Construction 1 compared to optimal constant delay. For Construction 1 three curves are plotted: the min, max, and expected among the m burst phase shifts.

Construction 2 *For any B and m we define the following encoder specifying the packet output at time i*

$$\vec{x}\left[i\right]=\left(s_0^{m-1}\left[i\right]\,,\;x^P\left[i\right]\right)$$

where

$$x^P\left[i\right]\triangleq P\left\{\left\{s_{\langle i\rangle_m}\left[i-\langle i\rangle_{mB}-1-\lambda(i)-jB\right]\right\}_{\forall j\,\in\,J}\right\}$$

and

$$\lambda(i)\triangleq\left\lfloor\frac{\langle i\rangle_{mB}}{m}\right\rfloor$$

and

$$J\triangleq\left\{j\in\mathbb{Z}\;:\;-\left\lfloor\frac{\langle i\rangle_{mB}+\lambda(i)}{B}\right\rfloor\leq j\leq-\left\lceil\frac{\langle i\rangle_{mB}+\lambda(i)+2}{B}\right\rceil+(m+1)\right\}\;.$$

It can be seen that similarly to Construction 1, Construction 2 is also systematic and causal, and its coding rate is $R=\frac{m}{m+1}$. The difference is the new definitions of $x^P\left[i\right]$ and J. The decoder for Construction 2 is very similar to the decoder given for

Construction 1 in that every m consecutive received packets reconstruct one full erased packet, where the burst phase shift ϕ determines at which serial index the reconstruction begins. Therefore, straightforward extensions of Theorems 4 and 6 can prove that Construction 2 has the same optimal average delay for $\phi = 0$ as Construction 1, and also the same average delay for a general ϕ. The main caveat in moving to B and m not co-prime is that the construction repeats itself every mB time units and not every m as Construction 1. This means that the reconstruction order and delay profile (when each packet in the burst is fully reconstructed) depends not only on ϕ, but more generally on $\langle \tilde{i} \rangle_{mB}$. In Section V we show an application of low-delay codes that requires highly regular reconstruction orders and delay profiles, in which case co-prime m and B are required.

C. Coding rate $R < \frac{1}{2}$

In this part we present codes with rate $\frac{1}{1+B} \leq R < \frac{1}{2}$ that achieve the optimal average delay. In this case, according to the proof of Theorem 2, the lower bound for $\frac{1}{1+B} \leq R < \frac{1}{2}$ may be attained only if the rate has the form $R = \frac{1}{1+m}$ and also $\frac{B}{m} \in \mathbb{N}$. In the following construction, beside the requirement $\frac{B}{m} \in \mathbb{N}$, we also demand that $\frac{B}{m}$ and m are co-prime and also $\frac{B}{m} > m$. Note that in this construction we extend the parity function P to act on vectors with size m, where the vector action of P is the element-wise parity function.

Construction 3 *For a given B and m that satisfy the conditions: $\frac{B}{m} \in \mathbb{N}$, $gcd\left(\frac{B}{m}, m\right) = 1$ and $\frac{B}{m} > m$, we define the following encoder specifying the packet output at time i*

$$\vec{x}[i] = \left(s_0^{m-1}[i] \; , \; \vec{x}^P[i]\right),$$

where

$$\vec{x}^P[i] = \left(P\left\{\left\{s_0^{m-1}\left[i-1-j\cdot\frac{B}{m}\right]\right\}_{j \in J}\right\}, P\left\{\left\{s_0^{m-1}\left[i-2-j\cdot\frac{B}{m}\right]\right\}_{j \in J}\right\}, \ldots, P\left\{\left\{s_0^{m-1}\left[i-m-j\cdot\frac{B}{m}\right]\right\}_{j \in J}\right\}\right).$$

The input to the encoder is a vector of m information symbols, and the parity component added by the encoder are m^2 parity symbols divided as m parity vectors of size m each. The index set J is now

$$J = \left\{ \begin{array}{ll} \{0, m\} & if \quad \langle i \rangle_m = 0 \\ m - t(i) & if \quad \langle i \rangle_m \neq 0 \end{array} \right. \tag{16}$$

where $t(i)$ is the unique integer t in the range $[1, m-1]$ that satisfies

$$\left\langle t \cdot \frac{B}{m} \right\rangle_m = m - \langle i \rangle_m \quad . \tag{17}$$

Since m and $\frac{B}{m}$ are co-prime, all the members in the set $\left\{\langle 1 \cdot \frac{B}{m}\rangle_m, \langle 2 \cdot \frac{B}{m}\rangle_m, \ldots, \langle(m-1) \cdot \frac{B}{m}\rangle_m\right\}$ are different from one another, therefore t in (17) must be unique. The encoder adds m^2 parity symbols to every m information symbols, so the code rate is $R = \frac{1}{1+m}$. Similarly to the case $R \geq \frac{1}{2}$, it can be seen that the code is systematic and causal. In this construction, the number of arguments in each parity function is not uniform: when $\langle i \rangle_m \neq 0$ there is only one argument in every parity function, i.e., this is simple repetition, and when $\langle i \rangle_m = 0$ we have two arguments in every parity function. We may also notice that the parity functions contain only full sets of m information symbols of a single packet (s_0^{m-1}), and not fractions of packets as we had in the constructions for $R \geq \frac{1}{2}$.

Before defining the decoder for Construction 3, we would like to simplify the presentation of $\vec{x}^P$ by defining the parameter $l \in \{1, 2, \ldots, m\}$, so now we can present $\vec{x}^P$ as

$$\vec{x}^P[i] = \left\{P\left\{\left\{s_0^{m-1}\left[i-l-j\cdot\frac{B}{m}\right]\right\}_{j \in J}\right\}\right\}_{l=1}^m . \tag{18}$$

The decoder *The packet $\vec{x}[i]$ with m information symbols and m^2 encoded parity symbols is received by the decoder at time i. As long as no erasure occurs, the decoder uses only the information symbols. When the decoder identifies an erasure with burst-length B, it performs a reconstruction procedure during the next $\frac{B}{m}$ time units. Let us assume that the first packet received after the burst is at time index $\tilde{i}$, i.e., the erased packets are at time units $\tilde{i}-1, \tilde{i}-2, \ldots, \tilde{i}-B$. By each received packet at times $\tilde{i} \leq i \leq \tilde{i}+\frac{B}{m}-1$ the decoder reconstructs m full erased packets, i.e., m^2 erased symbols, by finding for each $l \in \{1, 2, \ldots, m\}$ the parameter $j(l) \in J$ that satisfies the condition*

$$i - l - j(l) \cdot \frac{B}{m} \in \left\{\tilde{i}-1, \tilde{i}-2, \ldots, \tilde{i}-B\right\} .$$

Now the decoder uses $\vec{x}^P[i]$, and if $\langle i \rangle_m = 0$ also the other known argument of each parity function in $\vec{x}^P[i]$, to reconstruct the following m full erased packets

$$\left\{ s_0^{m-1}\left[i - l - j(l) \cdot \frac{B}{m} \right] \right\}_{l=1}^{m} .$$

Theorem 7 *For a code specified by Construction 3, any burst with any phase shift has an average delay*

$$\bar{T} = \frac{B}{2}\left(\frac{1}{m} + 1 \right) . \tag{19}$$

The proof of Theorem 7, similarly to the case $R \geq \frac{1}{2}$ in Theorem 4, will also consist of two steps: first the decoder's correctness, and then the calculation of the average delay. The following Lemma establishes the decoder's correctness (proof in Appendix C).

Lemma 8 *For a code specified by Construction 3, for any decoding instance with burst length B, all erased packets are successfully reconstructed during $\frac{B}{m}$ time units after the burst, where in each time unit in the interval $\tilde{i} \leq i \leq \tilde{i} + \frac{B}{m} - 1$, different m erased packet are reconstructed.*

By using Lemma 8 it is quite simple to calculate the average delay for all phase shifts and by that to prove Theorem 7.

Proof: According to Lemma 8, we know that for all phase shifts the recovery delay of the first m reconstructed packets is 1, and the recovery delay of the next m reconstructed packets is 2, and so on, until finally the recovery delay of the last reconstructed packets is $\frac{B}{m}$. So the full recovery delay profile for any phase shift is

$$\{\kappa_i\}_{i=1}^{B} = \{ \overbrace{1, \ldots, 1}^{m \text{ times}}, \overbrace{2, \ldots, 2}^{m \text{ times}}, \ldots, \overbrace{\frac{B}{m}, \ldots, \frac{B}{m}}^{m \text{ times}} \} ,$$

and the average recovery delay is

$$\bar{\kappa} = \frac{1}{2} \cdot \left(\frac{B}{m} + 1 \right) ,$$

and together with (2) we obtain (19). ∎

It can be seen that (19) is exactly the lower bound (4) for $R = \frac{1}{1+m}$, so we can write the following Corollary.

Corollary 9 *Construction 3 has optimal average delay for $\frac{1}{1+B} \leq R < \frac{1}{2}$ for all decoding instances (independent on the phase shift).*

In Construction 3 the phase shift determines the packets' order of reconstruction, but it does *not* affect the average delay, because unlike in the high-rate case, only full packets are reconstructed for all phase shifts.

Example 5 *For $B = 6$ and $m = 2$ the encoder of Construction 3 is specified in Table IV. Suppose packets $\vec{x}[-6], \vec{x}[-5], \ldots, \vec{x}[-1]$ were erased, so the first received packet is $\vec{x}[0]$, meaning $\phi = 0$. According to the specification of the decoder, the parities of $\vec{x}[0]$ will reconstruct all information symbols of packets $\vec{x}[-1]$ and $\vec{x}[-2]$ using $\vec{x}[-7]$ and $\vec{x}[-8]$, so their recovery delay is 1. Then the parities of $\vec{x}[1]$ will reconstruct packets $\vec{x}[-3]$ and $\vec{x}[-4]$, with recovery delay of 2. Finally, the parities of $\vec{x}[2]$ will reconstruct packets $\vec{x}[-5]$ and $\vec{x}[-6]$ using $\vec{x}[1]$ and $\vec{x}[0]$, with recovery delay of 3. So $\bar{\kappa} = 2 \iff \bar{T} = 4.5$, which is exactly the average delay in (19) for $B = 6$ and $m = 2$.*

We will now choose a different burst erasure with a different phase shift. Suppose packets $\vec{x}[-5], \vec{x}[-4], \ldots, \vec{x}[0]$ were erased, so the first received packet is $\vec{x}[1]$, meaning $\phi = 1$. In this case by $\vec{x}[1]$ we reconstruct packets $\vec{x}[-3]$ and $\vec{x}[-4]$ with recovery delay of 1. Then by $\vec{x}[2]$ we reconstruct packets $\vec{x}[-5]$ and $\vec{x}[0]$ using $\vec{x}[-6]$ and $\vec{x}[1]$, with recovery delay of 2, and finally we reconstruct packets $\vec{x}[-1]$ and $\vec{x}[-2]$ with recovery delay 3. Therefore, also in this case $\bar{\kappa} = 2 \iff \bar{T} = 4.5$. So for both phase shifts we attained the optimal average delay.

Construction 3 is an example for a construction that achieves the lower bound for $R < \frac{1}{2}$, for all phase shifts. But in fact it can be extended to a larger family of constructions with the same property, where the only necessary condition for optimality is $\frac{B}{m} \in \mathbb{N}$. It can be seen that Construction 3 repeats itself every m time units, where the extended family of constructions includes also codes with longer periods, such as B.

$$\vec{x}\,[-6] = \left(s_0^1\,[-6]\ ,\ P\left\{s_0^1\,[-7]\ ,\ s_0^1\,[-13]\right\}\ ,\ P\left\{s_0^1\,[-8]\ ,\ s_0^1\,[-14]\right\}\right)$$
$$\vec{x}\,[-5] = \left(s_0^1\,[-5]\ ,\ s_0^1\,[-9]\ ,\ s_0^1\,[-10]\right)$$
$$\vec{x}\,[-4] = \left(s_0^1\,[-4]\ ,\ P\left\{s_0^1\,[-5]\ ,\ s_0^1\,[-11]\right\}\ ,\ P\left\{s_0^1\,[-6]\ ,\ s_0^1\,[-12]\right\}\right)$$
$$\vec{x}\,[-3] = \left(s_0^1\,[-3]\ ,\ s_0^1\,[-7]\ ,\ s_0^1\,[-8]\right)$$
$$\vec{x}\,[-2] = \left(s_0^1\,[-2]\ ,\ P\left\{s_0^1\,[-3]\ ,\ s_0^1\,[-9]\right\}\ ,\ P\left\{s_0^1\,[-4]\ ,\ s_0^1\,[-10]\right\}\right)$$
$$\vec{x}\,[-1] = \left(s_0^1\,[-1]\ ,\ s_0^1\,[-5]\ ,\ s_0^1\,[-6]\right)$$
$$\vec{x}\,[0] = \left(s_0^1\,[0]\ ,\ P\left\{s_0^1\,[-1]\ ,\ s_0^1\,[-7]\right\}\ ,\ P\left\{s_0^1\,[-2]\ ,\ s_0^1\,[-8]\right\}\right)$$
$$\vec{x}\,[1] = \left(s_0^1\,[1]\ ,\ s_0^1\,[-3]\ ,\ s_0^1\,[-4]\right)$$
$$\vec{x}\,[2] = \left(s_0^1\,[2]\ ,\ P\left\{s_0^1\,[1]\ ,\ s_0^1\,[-5]\right\}\ ,\ P\left\{s_0^1\,[0]\ ,\ s_0^1\,[-6]\right\}\right)$$
$$\vec{x}\,[3] = \left(s_0^1\,[3]\ ,\ s_0^1\,[-1]\ ,\ s_0^1\,[-2]\right)$$

TABLE IV: Example for the encoder of Construction 3 when $B = 6$ and $m = 2$.

V. APPLICATIONS

In this section we examine low average-delay codes from an applications perspective. Our objective is twofold: first is to motivate our constructions by showing how realistic applications can benefit from low *average* delay; second is to explicitly deal with the important issue of *packet reconstruction order*, which was not material for proving the average-delay performance, but is critical for real applications. The application we consider here is *incremental-resolution sensing*, in which transmitted packets carry sensor measurement data in different resolutions. It is important to note that this chosen application is merely an example, and low average-delay codes may be useful well beyond this example.

A. Communication model

We model the sensor as taking a measurement at some frequency, and representing the measured value as symbols of M significance levels. The significance level of a symbol represents the uncertainty about the measurement in the absence of the symbol. Hence higher-significance symbols are "more important" than lower-significance ones. In our communication model, every measurement is translated to a sequence of *information packets*, where each packet is associated with a significance level between SL_{M-1} (the most significant) down to SL_0 (the least significant). All packets are assumed to have the same number m of information symbols. The information packets are encoded and transmitted over a channel introducing, as usual, burst erasures of up to B packets. When no erasures occur, we clearly prefer receiving the packets in decreasing order of significance level: SL_{M-1} first, then SL_{M-2}, and so on. This implies that the transmission order must also be from SL_{M-1} downto SL_0, as depicted in Figure 7 (transmission order shown from right to left). By receiving the high significance packets first, we allow the receiver to start computing on the received value before the entire packet sequence is received, and with good precision that improves gracefully as the sequence continues to arrive.

When an erasure burst does occur, the code-design objective for this model is that the delays of the erased packets will be according to the significance hierarchy, i.e., erased packets with a higher significance level will have a shorter delay than erased packets with a lower significance level.
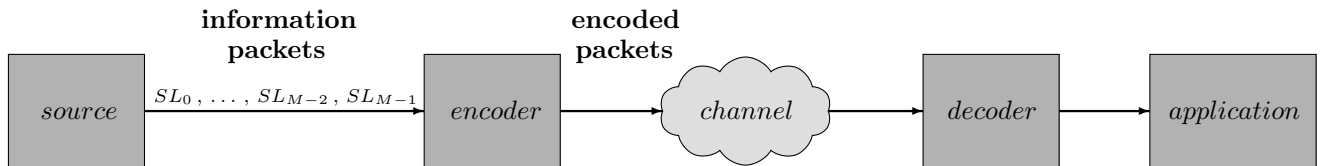

Figure 7: Low delay coding with a significance hierarchy.

To align the incremental-resolution sensing model with low-delay coding, we assume that each packet carries m information symbols of the same significance level, and for further simplicity that each such symbol is a single bit. Such scenario is suitable in the case of m sensors, each providing an M-bit measurement vector split between M consecutive packets. This scenario can be extended to mb sensors, $b \in \mathbb{N}$, or to the case where a sensor contributes multiple information bits to each packet. It follows that the unit of transmission is a *set* of M consecutive packets carrying a single measurement from each of m sensors. The transmission of each packet <u>set</u> starts at time $\hat{i}$ that satisfies $\langle \hat{i} \rangle_M = 0$. Since the transmission order is in decreasing significance levels, at time $\hat{i}$ the packet has significance level SL_{M-1}, at time $\hat{i}+1$ the packet has significance level SL_{M-2}, and so on

until time $\hat{i} + M - 1$ when an SL_0 packet is transmitted. The schedule of significance levels continues periodically, so that in general, at time i the packet carries symbols with significance SL_r, where

$$r = M - \langle i \rangle_M - 1 \ . \tag{20}$$

The system performance lies upon the precision of the sensor measurements as available at the receiver side. To analyze this precision, we first set the policy by which the receiver reproduces a remote measurement. When all M bits of a measurement $(\rho_0, \dots, \rho_{M-1})$ are available to the receiver, the (normalized) measurement value $U \in [0, 1)$ is trivially reproduced as

$$U = \sum_{r=0}^{M-1} \rho_r \cdot 2^{r-M} \ .$$

When a bit ρ_r is missing, the receiver replaces it with the estimate $\tilde{\rho}_r = 0.5$ when calculating the approximate value $\tilde{U}$. As a result, the total *magnitude* of estimation error due to missing bits $A \subseteq \{0, \dots, M-1\}$ equals

$$E = \sum_{r \in A} 2^{r-M-1} \ .$$

Note that E is not greater than 0.5, which is the worst possible error when estimating a value $U \in [0, 1)$. To evaluate the precision performance of a low-delay code, we define $A[i]$ as the multi-set of significance levels of all the bits missing at time i due to erased packets not yet fully recovered. Then we obtain the *total temporal error* of a low-delay code as

$$E_{tot}[i] = \sum_{r \in A[i]} 2^{r-M-1} \ .$$

Later in the section we derive the function $E_{tot}[i]$ exhibited by low-delay codes in the occurrence of burst erasures.

B. Code construction and error analysis

It is clear that lowering the total temporal error $E_{tot}[i]$ requires that the receiver reconstructs the higher-significance bits as early as possible after they are erased. In the coding terminology we need a code that not only has low average delay, but also guarantees that high-significance packets will have short delays. A low-delay code construction does not in general have this property, even if it is known to have optimal average delay.

1) Code construction: To that end, we now show a family of codes that have both optimal average delay and delays that improve with the significance level. This code family is obtained by restricting the parameters of Construction 1 to m and B satisfying $B = am - 1$, a an integer. Note that Construction 1 requires a weaker relation of m and B to be co-prime. As previously in the paper, we will discuss only the worst case where the burst length is exactly B.

The following Theorem describes the delays of the erased packets when using the suggested code family with a significance hierarchy of packets with $M = m$. It gives delay bounds per each significance level, whereby the delay gets shorter as the packet significance level gets higher.

Theorem 10 *Given a code specified by Construction 1, where in addition $B = am - 1$ for some integer a. For packets transmitted with cyclically ordered significance levels $SL_{m-1}, \dots, SL_r, \dots, SL_0, SL_{m-1}, \dots$ starting at time $i = 0$, the delay of a packet with SL_r after a B-burst erasure is upper bounded by*

$$T(r) \leq (m - r) B + m = am^2 - r(am - 1) \ . \tag{21}$$

Proof: In this proof we show that the delay of $B-1$ erased packets must satisfy (21) with equality, i.e, $T(r) = (m - r) B + m$, while the delay of one singular packet with lower delay than the upper bound (21) is discussed in Appendix D.

First note that since $gcd(am - 1, m) = 1$ for every a, $B = am - 1$ implies the condition B and m co-prime required for Construction 1. As before define $\tilde{i}$ as the time unit of the first received packet after the length B burst. Recall from the proof of Theorem 4 that by any m sequential packets, starting from time index $\langle i \rangle_m = 0$ in the interval $\tilde{i} \leq i \leq \tilde{i} + mB - 1$, we reconstruct all m information symbols of a single erased packet. Thus excluding the singular erased packet whose reconstruction starts at time $\tilde{i}$ (immediately following the burst), each of the other $B - 1$ packets is reconstructed fully in an interval of m sequential time units of the form $\left[\tilde{i} + cm - \phi \ , \ \tilde{i} + (c+1) m - \phi - 1 \right]$, where $1 \leq c \leq B - 1$ and ϕ is the burst phase shift. Referring to one of these $B - 1$ erased packets, assume its time index is $\tilde{i} - \beta$, with $1 \leq \beta \leq B$, its significance level is r, and it is being reconstructed at the time interval $\left[\tilde{i} + cm - \phi \ , \ \tilde{i} + (c+1) m - \phi - 1 \right]$, with $1 \leq c \leq B - 1$. To find the significance level of this packet we use (20) and the fact $M = m$ to get

$$\langle \tilde{i} - \beta \rangle_m = m - r - 1 \ . \tag{22}$$

From the specification of the code construction, specifically from (10), we tie between the packet reconstruction start time and transmission time

$$\tilde{i} + cm - \phi - 1 - j'B = \tilde{i} - \beta, \tag{23}$$

where j' is the j that satisfies (9). To get the delay T of this packet, we subtract from the reconstruction end time $\tilde{i} + (c+1)m - \phi - 1$ the left-hand side of (23), and obtain

$$T(j') = j'B + m \ . \tag{24}$$

Now we would like to express the delay as a function of the significance level r, for which we find the relation between j' and r

$$m - r - 1 = \langle \tilde{i} - \beta \rangle_m = \langle \underbrace{\tilde{i} + cm - \phi}_{=0 \bmod m} - 1 - j'B \rangle_m = \langle -j'B - 1 \rangle_m = \langle -j'(am - 1) - 1 \rangle_m = \langle j' - 1 \rangle_m \ .$$

The first equality is (22), the second is taking (23) modulo m, the fourth is from the code parameters $B = am - 1$, and the rest are simple reorganizations. Since $\beta \geq 1$, we know from (23) that $\tilde{i} + cm - \phi - 1 - j'B \leq \tilde{i} - 1 \Rightarrow cm - \phi \leq j'B$, and since $c \geq 1$, $cm - \phi$ is positive, therefore $j' > 0$. Moreover, from (7) we also know that $j' \leq m$, so $1 \leq j' \leq m \Rightarrow \langle j' - 1 \rangle_m = j' - 1$, hence

$$j' - 1 = m - r - 1 \Rightarrow j' = m - r \ . \tag{25}$$

By substituting (25) in (24) we obtain

$$T(r) = B(m - r) + m = am^2 - r(am - 1) \ . \tag{26}$$

$\blacksquare$

Examining (21), indeed we see that as r increases T decreases, thus appropriately favoring high-significance packets in the decoding schedule. Another property of this construction is that for the $B - 1$ packets that satisfy (21) with equality, erased packets sharing the same significance level have the same delay, which implies that within significance levels packets are reconstructed according to their transmission order. The following example illustrates these properties with a concrete transmission instance.

Example 6 *Table V demonstrates the packet reconstruction order of Construction 1 with $B = am - 1$, taking $m = 4$, $a = 3$ (viz. $B = 11$), and showing a decoding instance with $\phi = 2$. The first reconstructed symbol has time index $\tilde{i} - 3$, but the packet that contains it will not be fully recovered until the very end of the reconstruction procedure. The first three fully reconstructed packets are the most significant packets, with significance level SL_3, and all of them have the same delay of 15. Afterwards, all the erased SL_2 packets are reconstructed with delay of 26, and then the SL_1 packets with delay of 37. Finally, the erased SL_0 packets are reconstructed, where most of them have delay of 48, except for the last reconstructed one, which is the singular SL_0 packet that was already partially reconstructed at the beginning, with a lower delay of 46. A similar reconstruction order will be obtained for any m, a and $\phi \neq 0$, while when $\phi = 0$ the SL_0 packet with time index $\tilde{i} - 1$ will be the first one to be fully recovered with delay of m.*

time interval of received packets	the value of j in (10)	time index of reconstructed packet	time index of reconstructed packet $\bmod m$	significance level	recovery delay & standard delay
$[\tilde{i} \ , \ \tilde{i} + 1]$	0	$\tilde{i} - 3$	$\langle \tilde{i} - 3 \rangle_4 = 3$	SL_0	not reconstructed yet
$[\tilde{i} + 2 \ , \ \tilde{i} + 5]$	1	$\tilde{i} - 10$	$\langle \tilde{i} - 10 \rangle_4 = 0$	SL_3	$\kappa = 6 \ , \ T = 15$
$[\tilde{i} + 6 \ , \ \tilde{i} + 9]$	1	$\tilde{i} - 6$	$\langle \tilde{i} - 6 \rangle_4 = 0$	SL_3	$\kappa = 10 \ , \ T = 15$
$[\tilde{i} + 10 \ , \ \tilde{i} + 13]$	1	$\tilde{i} - 2$	$\langle \tilde{i} - 2 \rangle_4 = 0$	SL_3	$\kappa = 14 \ , \ T = 15$
$[\tilde{i} + 14 \ , \ \tilde{i} + 17]$	2	$\tilde{i} - 9$	$\langle \tilde{i} - 9 \rangle_4 = 1$	SL_2	$\kappa = 18 \ , \ T = 26$
$[\tilde{i} + 18 \ , \ \tilde{i} + 21]$	2	$\tilde{i} - 5$	$\langle \tilde{i} - 5 \rangle_4 = 1$	SL_2	$\kappa = 22 \ , \ T = 26$
$[\tilde{i} + 22 \ , \ \tilde{i} + 25]$	2	$\tilde{i} - 1$	$\langle \tilde{i} - 1 \rangle_4 = 1$	SL_2	$\kappa = 26 \ , \ T = 26$
$[\tilde{i} + 26 \ , \ \tilde{i} + 29]$	3	$\tilde{i} - 8$	$\langle \tilde{i} - 8 \rangle_4 = 2$	SL_1	$\kappa = 30 \ , \ T = 37$
$[\tilde{i} + 30 \ , \ \tilde{i} + 33]$	3	$\tilde{i} - 4$	$\langle \tilde{i} - 4 \rangle_4 = 2$	SL_1	$\kappa = 34 \ , \ T = 37$
$[\tilde{i} + 34 \ , \ \tilde{i} + 37]$	4	$\tilde{i} - 11$	$\langle \tilde{i} - 11 \rangle_4 = 3$	SL_0	$\kappa = 38 \ , \ T = 48$
$[\tilde{i} + 38 \ , \ \tilde{i} + 41]$	4	$\tilde{i} - 7$	$\langle \tilde{i} - 7 \rangle_4 = 3$	SL_0	$\kappa = 42 \ , \ T = 48$
$[\tilde{i} + 42 \ , \ \tilde{i} + 43]$	4	$\tilde{i} - 3$	$\langle \tilde{i} - 3 \rangle_4 = 3$	SL_0	$\kappa = 44 \ , \ T = 46$

TABLE V: The packet reconstruction order of Construction 1 with $B = am - 1$ for $m = 4$, $a = 3$, and $\phi = 2$.

2) Error analysis: Now that we know the delays offered to different significance levels by the code construction, we would like to analyze the resulting measurement-estimation error at the receiver in the model we defined in Section V-A. Specifically, our main interest will be in evaluating the total temporal error $E_{tot}[i]$ at each time i between the start of the length-B burst erasure, and until all erased packets are fully reconstructed. Recall that $E_{tot}[i]$ sums the error magnitudes of all missing bits in erased packets not yet reconstructed at time i. It is immediate to calculate $E_{tot}[i]$ exactly for all i when using Construction 1 with $B = am-1$. Knowing ϕ of the burst gives the times in which each significance level is erased, and the reconstruction times of the packets are calculated from (26). For comparison, $E_{tot}[i]$ can also be calculated exactly for optimal constant-delay codes with the same parameters. In Figure 8 we plot the total temporal error functions for $m = 5$ and $a = 3$ (viz. $B = 14$), for two different phase shifts: (a) $\phi = 0$ and (b) $\phi = 3$. In the same plots we show the same function when using a previously known optimal constant-delay code.

The plots of Figure 8 immediately reveal the benefits of this scheme: the drops in the error function are steep following the end of the erasure burst. As a result, the receiver can process the measurements with good precision relatively early after they are erased (compare to constant-delay codes that are extremely late to lower the error function). The difference between the two ϕ values is expressed in the maximal value of the error function (higher in (b) than in (a)), which is attributed to the different multiplicities of significance levels erased in the burst. In general for $B = am-1$ the maximal value of $E_{tot}[i]$ increases with ϕ. In addition, the small initial drop in (a) shows the unique case of $\phi = 0$, wherein prior to reconstructing the SL_{m-1} packets a singular SL_0 packet is fully reconstructed.
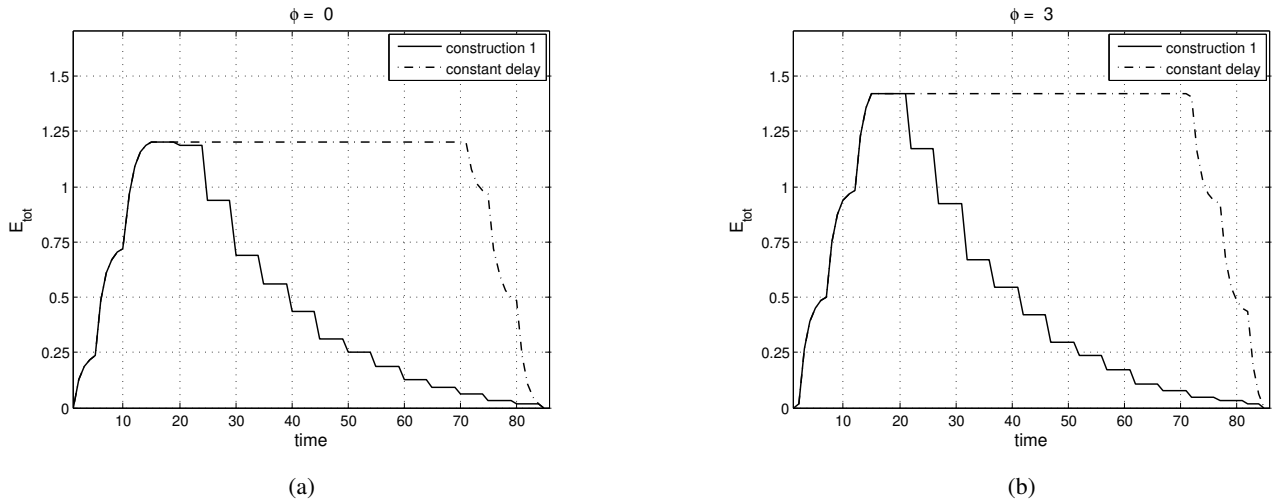

Figure 8: $E_{tot}[i]$ vs. the time i, for a single burst erasure with length $B = am-1$, where $m = 5$ and $a = 3$. Construction 1 (solid line), and optimal constant-delay construction (dashed line). (a) $\phi = 0$, and (b) $\phi = 3$.

It will be convenient to give the performance of low-delay codes using the single-letter measure of total receiver error due to a full B-burst erasure. We define the *full-burst total error* E_{acc} by summing $E_{tot}[i]$ over all i affected by a single B-burst erasure. This means summing in the time interval $\tilde{i} - B \le i \le \tilde{i} + mB - 1$, hence

$$E_{acc} = \sum_{i=\tilde{i}-B}^{\tilde{i}+mB-1} E_{tot}[i] \ .$$

Note that E_{acc} is represented by the area under the graph in Figure 8. A closed form expression for E_{acc} can be derived as a function of m, a and ϕ when using Construction 1. Here we give an expression for $\bar{E}_{acc}$, the full-burst total error averaged over all ϕ values

$$\bar{E}_{acc} = \frac{1}{m} - 2a + a^2 m + \frac{am-1}{2} + \left(2a - \frac{1}{m} - a^2 m - \frac{a^2 m^2}{2} + \frac{3-m}{4} \right) \cdot 2^{-m} \ , \tag{27}$$

with an exception when $a = 1$, where we should subtract $\frac{2^{-m-1}}{m}$ from (27). In comparison, for an optimal constant-delay construction, we obtain

$$\bar{E}_{acc}^{(\text{const})} = \frac{\left(1 - 2^{-m}\right)\left(am - 1\right)^2}{2} \ .$$

In Figure 9 we plot $\bar{E}_{acc}$ and $\bar{E}_{acc}^{(\text{const})}$ as a function of the burst length, when m (and the code rate) is fixed. The precision gap between the two coding alternatives widens as the burst length grows.
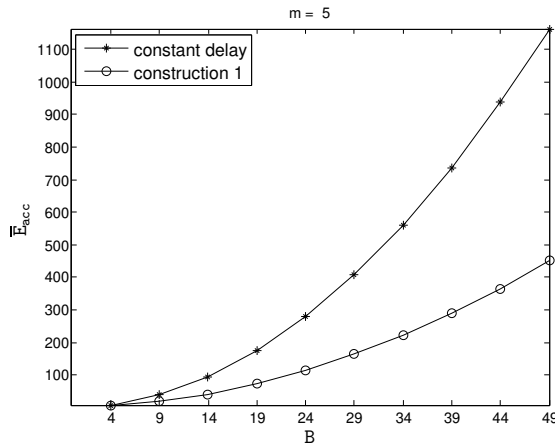
Figure 9: $\bar{E}_{acc}$ vs. $B = am - 1$, $m = 5$: showing the total error due to a single burst erasure. For Construction 1 (circle markers) and for an optimal constant delay construction (star markers).

C. Order of reconstruction

The good precision performance shown in the previous sub-section comes from a code construction implementing a suitable *reconstruction order* for the erased packets. This motivates the following discussion on packet reconstruction order in low-delay codes. The center point of this discussion is that optimal average-delay codes can have a degree of freedom to set the reconstruction order without sacrificing the average-delay optimality. That is the case with Construction 1 of this paper. It can be seen that a more general version of Construction 1 can be given as the following Construction 4, while carrying over all the properties proved in Section IV. The generalization is done by defining the parameter τ that shifts the time arguments of the parity symbols.

Construction 4 *For co-prime B and m we define the following encoder specifying the packet output at time i*

$$\vec{x}[i] = \left(s_0^{m-1}[i] \ , \ x^P[i] \right)$$

where

$$x^P[i] = P\left\{ \left\{ s_{\langle i \rangle_m}[i - \langle i \rangle_m - \tau - jB] \right\}_{\forall j \in J} \right\} \ ; \quad 1 \le \tau \le B$$

is the parity symbol, and

$$J = \left\{ j \in \mathbb{Z} \ : \ -\left\lfloor \frac{\langle i \rangle_m + \tau - 1}{B} \right\rfloor \le j \le -\left\lceil \frac{\langle i \rangle_m + \tau + 1}{B} \right\rceil + (m+1) \right\} \ .$$

Each of the B possible values for τ in Construction 4 gives an optimal average-delay code, with $\tau = 1$ being the special case of Construction 1. Varying the parameter τ imposes a cyclic shift on the packets' order of reconstruction. If one is not careful to match the τ parameter to the application using the code, sub-optimal performance may result. This can be seen in Figure 10 showing the total temporal error of Construction 4 with $\tau = 3$, used to reconstruct packets with a significance hierarchy as specified in Section V-A. The code parameters in this example are the same as in Figure 8 ($m = 5$, $a = 3$), but the resulting performance is much worse.

VI. CONCLUSION

This paper defines a natural variation on the established model of low-delay codes. It shows that average-case optimality spans a much better tradeoff between correctability, rate, and delay. The most interesting open problem from this work is whether optimal average delay can be achieved for $R \ge 0.5$ in all erasure phases, and if not, what limits exist for the expected average delay across phases. A natural extension of this framework is to deal with channels introducing errors in addition to erasures, including random errors not in bursts. Another interesting direction is studying the application of low average-delay codes to sensor communications within the intersection of information theory and control.
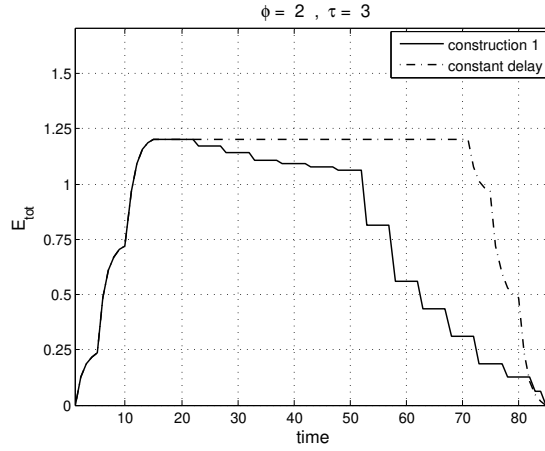
Figure 10: $E_{tot}[i]$ vs. the time i, for a single burst erasure with length $B = am - 1$, where $m = 5$, $a = 3$, and $\phi = 2$. Construction 4 with $\tau = 3$ (solid line), and optimal constant-delay construction (dashed line).

APPENDIX A
PROOF OF THEOREM 2 FOR $\frac{1}{1+B} \leq R < \frac{1}{2}$ IN THE GENERAL CASE.

In order to complete the proof of the lower bound of the average delay when $\frac{1}{1+B} \leq R < \frac{1}{2}$, we should prove that the lower bound shown for the special case $\frac{P}{S} \in \mathbb{N}$ and $\frac{BS}{P} \in \mathbb{N}$ is also correct for the general case $\frac{P}{S} \in \mathbb{Q}$ and $\frac{BS}{P} \in \mathbb{Q}$. Further, we show that the bound may be attained only when $\frac{P}{S} \in \mathbb{N}$ and $\frac{BS}{P} \in \mathbb{N}$.

Proof: Our objective is to find a lower bound on the expression

$$
\frac{\left\lceil \frac{S}{P} \right\rceil + \cdots + \left\lceil (B-1) \cdot \frac{S}{P} \right\rceil + \left\lceil B \cdot \frac{S}{P} \right\rceil}{B} . \tag{28}
$$

When $\frac{1}{1+B} \leq R < \frac{1}{2}$ then $S < P$, so the equation $\left\lceil \frac{S}{P} \right\rceil + \cdots + \left\lceil (B-1) \cdot \frac{S}{P} \right\rceil + \left\lceil B \cdot \frac{S}{P} \right\rceil$ contains $\left\lfloor \frac{P}{S} \right\rfloor$ times $'1'$, $\left\lfloor \frac{2P}{S} \right\rfloor - \left\lfloor \frac{P}{S} \right\rfloor$ times $'2'$, $\left\lfloor \frac{3P}{S} \right\rfloor - \left\lfloor \frac{2P}{S} \right\rfloor$ times $'3'$, and so on until $\left\lfloor \frac{\lfloor \alpha \rfloor P}{S} \right\rfloor - \left\lfloor \frac{(\lfloor \alpha \rfloor - 1)P}{S} \right\rfloor$ times $' \lfloor \alpha \rfloor '$, where $\alpha = B \cdot \frac{S}{P}$, and there is also another part $B - \left\lfloor \frac{\lfloor \alpha \rfloor P}{S} \right\rfloor$ times $' \lfloor \alpha \rfloor + 1'$ (this part is equal to zero when $\frac{BS}{P} \in \mathbb{N}$). So we can write

$$
\frac{\left\lceil \frac{S}{P} \right\rceil + \cdots + \left\lceil (B-1) \cdot \frac{S}{P} \right\rceil + \left\lceil B \cdot \frac{S}{P} \right\rceil}{B} =
$$

$$
\frac{1 \cdot \left\lfloor \frac{P}{S} \right\rfloor + 2 \cdot \left(\left\lfloor \frac{2P}{S} \right\rfloor - \left\lfloor \frac{P}{S} \right\rfloor \right) + 3 \cdot \left(\left\lfloor \frac{3P}{S} \right\rfloor - \left\lfloor \frac{2P}{S} \right\rfloor \right) + \cdots + \lfloor \alpha \rfloor \cdot \left(\left\lfloor \frac{\lfloor \alpha \rfloor P}{S} \right\rfloor - \left\lfloor \frac{(\lfloor \alpha \rfloor - 1)P}{S} \right\rfloor \right) + (\lfloor \alpha \rfloor + 1) \cdot \left(B - \left\lfloor \frac{\lfloor \alpha \rfloor P}{S} \right\rfloor \right)}{B} =
$$

$$
\frac{(\lfloor \alpha \rfloor + 1) \cdot B - \left(\left\lfloor \frac{P}{S} \right\rfloor + \left\lfloor \frac{2P}{S} \right\rfloor + \cdots + \left\lfloor \frac{\lfloor \alpha \rfloor P}{S} \right\rfloor \right)}{B} \geq \frac{(\lfloor \alpha \rfloor + 1) \cdot B - \left(\frac{P}{S} + \frac{2P}{S} + \cdots + \frac{\lfloor \alpha \rfloor P}{S} \right)}{B} .
$$

This inequality becomes equality only when $\frac{P}{S} \in \mathbb{N}$.

$$
= \frac{(\lfloor \alpha \rfloor + 1) \cdot B - \frac{P}{S} \sum_{i=1}^{\lfloor \alpha \rfloor} i}{B} = \lfloor \alpha \rfloor + 1 - \frac{\frac{P}{S} \cdot \lfloor \alpha \rfloor \cdot (\lfloor \alpha \rfloor + 1)}{2B} = (\lfloor \alpha \rfloor + 1) \cdot \left(1 - \frac{\lfloor \alpha \rfloor}{2\alpha} \right) .
$$

Now we use the formula $\lfloor x \rfloor = x - \{x\}$, where $\{x\}$ is the fractional part of x.

$$
= (\alpha - \{\alpha\} + 1) \cdot \left(1 - \frac{\alpha - \{\alpha\}}{2\alpha} \right) = \frac{\alpha + 1}{2} + \frac{\{\alpha\}}{2\alpha} - \frac{\{\alpha\}^2}{2\alpha} .
$$

The quadratic function in $\{\alpha\}$ attains its minimum at $\{\alpha\} = 0$, therefore the expression is lower bounded by

$$
\frac{\alpha + 1}{2} + \frac{\{\alpha\}}{2\alpha} - \frac{\{\alpha\}^2}{2\alpha} \geq \frac{\alpha + 1}{2},
$$

with equality when $\alpha = \frac{BS}{P} \in \mathbb{N}$.

So we proved that expression (28) is lower bounded by

$$\frac{B \cdot \frac{S}{P} + 1}{2} = \frac{B}{2} \cdot \frac{R}{1 - R} + \frac{1}{2}$$

in the general case, which is exactly the lower bound in Theorem 2 when $\frac{1}{1+B} \leq R < \frac{1}{2}$. The proof also shows that the lower bound is tight only when $\frac{P}{S} \in \mathbb{N}$ and $\frac{BS}{P} \in \mathbb{N}$. ∎

APPENDIX B
THE PROOF OF THEOREMS 4 AND 6 FOR $B < m$

We proved Theorem 4 and 6 under the assumption that $B > m$. We will now show that they apply to *every* co-prime B and m.

Proof: We discuss in this proof only the decoder's correctness part, since after proving correctness the recovery-delay profile is obtained in the same way as in the proof for the case $B > m$, for both $\phi = 0$ (Theorem 4) and $\phi \neq 0$ (Theorem 6).

The parity symbol is

$$x^P[i] = P\left\{\{s_{\langle i\rangle_m}[i - \langle i\rangle_m - 1 - jB]\}_{\forall j \in J}\right\} \tag{29}$$

where

$$J \triangleq \left\{j \in \mathbb{Z} \ : \ -\left\lfloor\frac{\langle i\rangle_m}{B}\right\rfloor \leq j \leq -\left\lceil\frac{\langle i\rangle_m + 2}{B}\right\rceil + (m+1)\right\} \ . \tag{30}$$

As was discussed in the proof for $B > m$, by examining (29) we know that all the time arguments are distant at least B time units from one another, so there can be up to one missing argument in every parity function. Similar to the previous case our first step is to show that every packet with time index $\tilde{i} + \alpha$ where $0 \leq \alpha \leq mB - 1$ contains in its parity group a former information symbol with time index $\tilde{i} - \beta$ where $1 \leq \beta \leq B$. According to (29) the time arguments of the parity function at time $\tilde{i} + \alpha$ are

$$\left[\tilde{i} + \alpha - \langle\tilde{i} + \alpha\rangle_m - 1 - jB\right]_{j \in J} \ , \tag{31}$$

and we would like to find $j \in J$ that satisfies the requirement. This time we examine the expression $\alpha - \left\langle\langle\tilde{i} + \alpha\rangle_m\right\rangle_B - 1$ which is bounded by

$$-B \ \leq \ \alpha - \left\langle\langle\tilde{i} + \alpha\rangle_m\right\rangle_B - 1 \ \leq \ mB - 2$$

since $0 \leq \left\langle\langle\tilde{i} + \alpha\rangle_m\right\rangle_B \leq B - 1$ and $0 \leq \alpha \leq mB - 1$. Our chosen j will be $j = \gamma - \left\lfloor\frac{\langle\tilde{i} + \alpha\rangle_m}{B}\right\rfloor$, where γ is the unique integer satisfying

$$(\gamma - 1)B \ \leq \ \alpha - \left\langle\langle\tilde{i} + \alpha\rangle_m\right\rangle_B - 1 \ \leq \ \gamma B - 1 \ ; \quad 0 \leq \gamma \leq m \ . \tag{32}$$

By using the chosen j in (31) we obtain

$$\tilde{i} + \alpha - \langle\tilde{i} + \alpha\rangle_m - 1 - \left(\gamma - \left\lfloor\frac{\langle\tilde{i} + \alpha\rangle_m}{B}\right\rfloor\right)B \ = \ \tilde{i} + \alpha - \left\langle\langle\tilde{i} + \alpha\rangle_m\right\rangle_B - 1 - \gamma B \ ,$$

and according to (32) we conclude that

$$\tilde{i} - B \ \leq \ \tilde{i} + \alpha - \left\langle\langle\tilde{i} + \alpha\rangle_m\right\rangle_B - 1 - \gamma B \ \leq \ \tilde{i} - 1 \ ,$$

which is exactly the requirement that we sought.

What left to do is to make sure that the chosen j satisfies condition (30) for $i = \tilde{i} + \alpha$. It is easy to see that it satisfies the lower bound since when $\gamma = 0$ the chosen j is equal to the lower bound, and for any other value of γ j has a higher value. Regarding the upper bound we should prove that the following inequality is true for every γ

$$\gamma - \left\lfloor\frac{\langle\tilde{i} + \alpha\rangle_m}{B}\right\rfloor \ \leq \ m + 1 - \left\lceil\frac{\langle\tilde{i} + \alpha\rangle_m + 2}{B}\right\rceil \ \Rightarrow \ \left\lceil\frac{\langle\tilde{i} + \alpha\rangle_m + 2}{B}\right\rceil - \left\lfloor\frac{\langle\tilde{i} + \alpha\rangle_m}{B}\right\rfloor \ \leq \ m + 1 - \gamma \ . \tag{33}$$

We can find that

$$L \triangleq \left\lceil\frac{\langle\tilde{i} + \alpha\rangle_m + 2}{B}\right\rceil - \left\lfloor\frac{\langle\tilde{i} + \alpha\rangle_m}{B}\right\rfloor \ = \ \begin{cases} 1 & if \quad 0 \leq \left\langle\langle\tilde{i} + \alpha\rangle_m\right\rangle_B \leq B - 2 \\ 2 & if \quad \left\langle\langle\tilde{i} + \alpha\rangle_m\right\rangle_B = B - 1 \end{cases} \ .$$

When $0 \leq \gamma \leq m-1$ the inequality (33) is always true since $1 \leq L \leq 2$. When $\gamma = m$ (33) becomes $L \leq 1$, and it is true only when $0 \leq \left\langle \left\langle \tilde{i} + \alpha \right\rangle_m \right\rangle_B \leq B - 2$, but if $\gamma = m$ and also $\left\langle \left\langle \tilde{i} + \alpha \right\rangle_m \right\rangle_B = B - 1$ the lower bound (32) becomes

$$(m-1) B \ \leq \ \alpha - (B-1) - 1 \ \Rightarrow \ \alpha \geq mB \ ,$$

and we know that $0 \leq \alpha \leq mB - 1$ so there is a contradiction. Therefore, $\gamma = m$ and $\left\langle \left\langle \tilde{i} + \alpha \right\rangle_m \right\rangle_B = B - 1$ is not a valid option, so (33) is true for all valid cases.

We now proved that every fully received packet with time index $\tilde{i} + \alpha$, $0 \leq \alpha \leq mB - 1$, contains an erased information symbol. Regarding the second part of the proof that shows that each information symbol appears in not more than one parity group, the same proof from the case $B > m$ goes through here too. Finally, by combining the two parts we proved the correctness of the decoder also for $B < m$. ∎

APPENDIX C
THE PROOF OF LEMMA 8

We would like to prove the decoder's correctness of Construction 3, i.e., that for every phase shift, all B erased packets are reconstructed during the first $\frac{B}{m}$ time units after the burst, and that in every time unit in the interval $\tilde{i} \leq i \leq \tilde{i} + \frac{B}{m} - 1$ we reconstruct exactly m different erased packets.

Proof: According to (16) and (18) we can see that there can be up to one missing argument in every parity group, since when $\langle i \rangle_m \neq 0$ there is only one argument in each parity group, and when $\langle i \rangle_m = 0$ the time indices of the two arguments are distant B time units from one another, so they cannot be erased together. Therefore, if an erased argument appears in the parity group of a received packet it is successfully reconstructed

In the first part of the proof we would like to show that every packet with a time index $\tilde{i} + \alpha$ where $0 \leq \alpha \leq \frac{B}{m} - 1$ contains in its parity group m time indices with the form $\tilde{i} - \beta$ where $1 \leq \beta \leq B$, i.e., all the fully received packets at the interval $\{\tilde{i}, \ldots, \tilde{i} + \frac{B}{m} - 1\}$ contain in every one of their m parity groups a packet that was erased. The serial indices are not important in this case since the parity functions contain all m informations symbols of a packet. The m time indices of the arguments in packet $\tilde{i} + \alpha$ are

$$\left\{ \tilde{i} + \alpha - l - j(l) \cdot \frac{B}{m} \right\}_{l=1}^{m} . \tag{34}$$

If $\left\langle \tilde{i} + \alpha \right\rangle_m \neq 0$ then according to (16) $j(l) = m - t(\tilde{i} + \alpha)$, so (34) becomes

$$\left\{ \tilde{i} + \alpha - l - B + t(\tilde{i} + \alpha) \cdot \frac{B}{m} \right\}_{l=1}^{m} ,$$

which can be lower and upper bounded according to the boundaries of α, t and l by

$$\tilde{i} - B \ < \ \tilde{i} - m - B + \frac{B}{m} \ \leq \ \tilde{i} + \alpha - l - B + t(\tilde{i} + \alpha) \cdot \frac{B}{m} \ \leq \ \tilde{i} - 2 \ ,$$

where the left inequality is from the fact that $\frac{B}{m} > m$.

When $\left\langle \tilde{i} + \alpha \right\rangle_m = 0$ we choose the value of $j(l)$ to be 0 or m. If $\alpha < l$ we choose $j(l) = 0$, and then the time indices in (34) are bounded by

$$\tilde{i} - B \ < \ \tilde{i} - m \ \leq \ \tilde{i} + \alpha - l \ \leq \ \tilde{i} - 1$$

according to the boundaries of α and l and the fact that $\alpha \leq l - 1$, when the left inequality is from the fact that $m < \frac{B}{m} \leq B$. If $\alpha \geq l$ we choose $j(l) = m$, and now (34) is bounded by

$$\tilde{i} - B \ \leq \ \tilde{i} + \alpha - l - B \ \leq \ \tilde{i} - B + \frac{B}{m} - 2 \ .$$

The upper bound must satisfy $\tilde{i} - B + \frac{B}{m} - 2 \geq \tilde{i} - B$ since $\frac{B}{m} > m \Rightarrow \frac{B}{m} \geq 2$. By combining all these cases we completed the first part of the proof.

In the second part we want to prove that each time index appears in not more than one parity group with a time index of the form $\tilde{i} + \alpha$ where $0 \leq \alpha \leq \frac{B}{m} - 1$. We will assume by contradiction that there are two distinct integers $0 \leq \alpha', \alpha'' \leq \frac{B}{m} - 1$ such that the packets $\tilde{i} + \alpha'$ and $\tilde{i} + \alpha''$ contain in their parity group the same time index, therefore

$$\tilde{i} + \alpha' - l' - j' \cdot \frac{B}{m} \ = \ \tilde{i} + \alpha'' - l'' - j'' \cdot \frac{B}{m} \ . \tag{35}$$

If $\left\langle \tilde{i} + \alpha' \right\rangle_m \neq 0$ and $\left\langle \tilde{i} + \alpha'' \right\rangle_m \neq 0$, from (16) we obtain the following equation

$$\tilde{i} + \alpha' + t' \cdot \frac{B}{m} - \left(\tilde{i} + \alpha'' + t'' \cdot \frac{B}{m} \right) = l' - l'' \ .$$

According to condition (17) we know that $m | \tilde{i} + \alpha' + t'\frac{B}{m}$ and $m | \tilde{i} + \alpha'' + t''\frac{B}{m}$, therefore $m | l' - l''$. This conclusion is true also if $\left\langle \tilde{i} + \alpha' \right\rangle_m = 0$, or $\left\langle \tilde{i} + \alpha'' \right\rangle_m = 0$ or both, since when $\left\langle \tilde{i} + \alpha' \right\rangle_m = 0$ we know that $m | \tilde{i} + \alpha'$ and also $m | \left\{ j \cdot \frac{B}{m} \right\}_{j \in \{0, m\}}$, and the same for $\left\langle \tilde{i} + \alpha'' \right\rangle_m = 0$. Since $l', l'' \in \{1, \dots, m\} \Rightarrow |l' - l''| \leq m - 1$ and $m | l' - l''$ we conclude that $l' - l'' = 0$.

By using $l' = l''$ in (35) we obtain

$$\alpha' - \alpha'' = \frac{B}{m} (j' - j'') \ ,$$

so $\frac{B}{m} | \alpha' - \alpha''$. We will assume w.l.o.g. that $\alpha' \geq \alpha''$, and since $0 \leq \alpha', \alpha'' \leq \frac{B}{m} - 1$ we know that $0 \leq \alpha' - \alpha'' \leq \frac{B}{m} - 1$, and together with $\frac{B}{m} | \alpha' - \alpha''$ we find out that $\alpha' - \alpha'' = 0 \Rightarrow \alpha' = \alpha''$, which is a contradiction. So we proved that each time index can appear in not more than one parity group in the interval $\tilde{i} \leq i \leq \tilde{i} + \frac{B}{m} - 1$.

By combining the two parts of the proof we conclude that by every received packet at the interval $\tilde{i} \leq i \leq \tilde{i} + \frac{B}{m} - 1$ we reconstruct m erased packets, and they are different, so during $\frac{B}{m}$ time units we reconstruct all B erased packets. ∎

APPENDIX D
THE PROOF OF THEOREM 10 FOR THE SINGULAR PACKET

Theorem 10 was proved for $B - 1$ erased packets that satisfy (21) with equality. We will now discuss the remaining singular packet that contains the first reconstructed symbol.

Proof: The proof will be divided to three scenarios: (a) $\phi = 0$, (b) $\phi \neq 0$ and ($a \neq 1$ or $\phi \neq m - 1$) , (c) $a = 1$ and $\phi = m - 1$.

(a) When $\phi = 0$, as discussed in the proof of Theorem 4, the singular packet is fully recovered by the first m received packets, so it is reconstructed in the time interval $\left[\tilde{i} , \tilde{i} + m - 1 \right]$, and will be fully recovered at $\tilde{i} + m - 1$. The time index of the reconstructed packet is $\tilde{i} - 1$ (obtained by using $j = 0$ in (10)), therefore the delay is m.

(b) When $\phi \neq 0$, the reconstruction of the singular packet is split between the two intervals $\left[\tilde{i} , \tilde{i} + m - \phi - 1 \right]$ and $\left[\tilde{i} + mB - \phi , \tilde{i} + mB - 1 \right]$, therefore it is the last fully reconstructed packet, which is fully recovered at $\tilde{i} + mB - 1$. When $a \neq 1$ or $\phi \neq m - 1$, the time index of the singular packet is $\tilde{i} - \phi - 1$ (again by using $j = 0$), so the delay is $mB + \phi$.

(c) When $a = 1$ and $\phi = m - 1$, the packet is still fully recovered at $\tilde{i} + mB - 1$ as in (b), but now the time index of the first reconstructed symbol is different and equal to $\tilde{i} - 1$ (by using $j = -1$), so the delay is mB.

The delay of the singular packet, for all scenarios, is summarized as follows

$$T_{singular} = \begin{cases} m & \text{if} & \phi = 0 \\ mB & \text{if} & a = 1 \quad \text{and} \quad \phi = m - 1 \\ mB + \phi & \text{if} & \text{otherwise} \end{cases} \ .$$

The time index of the singular packet in the first two cases (a) and (b) is $\tilde{i} - \phi - 1$, and since $\left\langle \tilde{i} - \phi - 1 \right\rangle_m = m - 1$, this is an SL_0 packet. The expression of the upper-bound in (21) when $r = 0$ is $mB + m$, which is greater than the calculated delays m or $mB + \phi$, respectively. In the third scenario (c), the time index of the singular packet is $\tilde{i} - 1$, and since $\phi = m - 1$ we know that $\left\langle \tilde{i} - 1 \right\rangle_m = m - 2$, therefore it is an SL_1 packet. The expression in (21) when $r = 1$ is mB, which is equal to the calculated delay in this scenario. Therefore, the inequality (21) is satisfied for all three scenarios.

It can also be shown that for the third scenario $a = 1$ and $\phi = m - 1$, there is no erased SL_0 packet, therefore the singular packet always has the lowest erased significance level. ∎

REFERENCES

[1] E. Martinian and C.-E. Sundberg, "Burst erasure correction codes with low decoding delay," *Information Theory, IEEE Transactions on*, vol. 50, no. 10, pp. 2494–2502, 2004.

[2] E. Martinian and M. Trott, "Delay-optimal burst erasure code construction," in *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*. IEEE, 2007, pp. 1006–1010.

[3] Z. Li, A. Khisti, and B. Girod, "Correcting erasure bursts with minimum decoding delay," in *Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on*. IEEE, 2011, pp. 33–39.

[4] A. Badr, A. Khisti, and E. Martinian, "Diversity embedded streaming erasure codes (de-sco): Constructions and optimality," *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 5, pp. 1042–1054, 2011.

[5] A. Khisti and J. P. Singh, "On multicasting with streaming burst-erasure codes," in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*. IEEE, 2009, pp. 2887–2891.

[6] A. Badr, D. Lui, and A. Khisti, "Multicast streaming codes (mu-sco) for burst erasure channels," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*.   IEEE, 2010, pp. 383–390.

[7] Z. Li, A. Khisti, and B. Girod, "Forward error protection for low-delay packet video," in *Packet Video Workshop (PV), 2010 18th International*.   IEEE, 2010, pp. 1–8.

[8] A. Badr, D. Lui, and A. Khisti, "Streaming codes for multicast over burst erasure channels," *Information Theory, IEEE Transactions on*, vol. 61, no. 8, pp. 4181–4208, 2015.

[9] A. Badr, A. Khisti, W.-t. Tan, and J. Apostolopoulos, "Streaming codes with partial recovery over channels with burst and isolated erasures," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 9, no. 3, pp. 501–516, 2015.

[10] Z. Li, A. C. Begen, and B. Girod, "Delay-optimal burst erasure codes for parallel links," in *Network Coding (NetCod), 2011 International Symposium on*. IEEE, 2011, pp. 1–6.

[11] A. Badr, A. Khisti, W.-T. Tan, and J. Apostolopoulos, "Robust streaming erasure codes using mds constituent codes," in *Information Theory (CWIT), 2013 13th Canadian Workshop on*.   IEEE, 2013, pp. 158–163.

[12] D. Leong, A. Qureshi, and T. Ho, "On coding for real-time streaming under packet erasures," in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*.   IEEE, 2013, pp. 1012–1016.

[13] D. Leong and T. Ho, "Erasure coding for real-time streaming," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*.   IEEE, 2012, pp. 289–293.

[14] A. Albanese, J. Blömer, J. Edmonds, M. Luby, and M. Sudan, "Priority encoding transmission," *Information Theory, IEEE Transactions on*, vol. 42, no. 6, pp. 1737–1744, 1996.

[15] Y. Wei and T. Ho, "On prioritized coding for real-time streaming under packet erasures," in *Communication, Control, and Computing (Allerton), 2013 51st Annual Allerton Conference on*.   IEEE, 2013, pp. 327–334.

[16] G. Forney Jr, "Burst-correcting codes for the classic bursty channel," *Communication Technology, IEEE Transactions on*, vol. 19, no. 5, pp. 772–781, 1971.