

Streaming Codes with Partial Recovery over Channels with Burst and Isolated Erasures

Ahmed Badr, *Student Member, IEEE*, Ashish Khisti, *Member, IEEE*, Wai-Tian Tan, *Member IEEE* and John Apostolopoulos, *Fellow IEEE*

Abstract—We study forward error correction codes for low-delay, real-time streaming communication over packet erasure channels. Our encoder operates on a stream of source packets in a sequential fashion, and the decoder must output each packet in the source stream within a fixed delay. We consider a class of practical channel models with correlated erasures and introduce new “streaming codes” for efficient error correction over these channels. For our analysis, we propose a simplified class of erasure channels that introduce both burst and isolated erasures within the same decoding window. We demonstrate that the previously proposed streaming codes can lead to significant number of packet losses over such channels. Our proposed constructions involve a layered coding approach, where a burst-erasure code is first constructed, and additional layers of parity-checks are concatenated to recover from the isolated erasure patterns. We also introduce another construction that requires a significantly smaller field-size, but incurs some performance loss. Numerical simulations over the Gilbert-Elliott and Fritchman channel models indicate that by addressing patterns involving both burst and isolated erasures within the same window, our proposed codes achieve significant gains over previously proposed streaming codes.

Index Terms—Low-Delay Codes, Real-Time Streaming Communication, Burst and Isolated Errors, Gilbert-Elliott Channels, Application Layer Error Correction

I. INTRODUCTION

AN increasing number of multimedia applications require low-delay and real-time communication of information. For example, the end-to-end latency in interactive video conferencing must be less than 200 ms. Extensive analysis of video conferencing systems such as Skype (see e.g., [1], [2] and references therein) has revealed a number of interesting paradigms. For example it is observed that the forward error correction (FEC) redundancy ratio of Skype is about 4.5 times of the actual loss rate. Such a significant overhead is most likely driven by the stringent delays constraints. Thus the study of low-delay FEC is a fertile area of research.

Fundamental limits of communication under delay and real-time constraints can be very different from the classical Shannon capacity. It is well known for example that the Shannon

capacity of an erasure channel only depends on the fraction of packets lost over the channel. When delay constraints are imposed, the pattern of packet losses becomes significant. In practical systems, packet losses are often temporally correlated due to effects such as fading, congestion or buffer overflow. Such losses are often modelled using Markov models [3] such as the Gilbert-Elliott channel. When the channel is in the “good” state, it behaves as an i.i.d. erasure channel, whereas in the “bad” state, it behaves as a burst erasure channel. In this paper, we carefully examine the local dynamics of such channels and propose a new class of error correction codes that significantly improve the packet loss rate over earlier approaches.

In [4], [5] the authors introduce a class of streaming codes for the burst erasure channel model. These codes operate on a stream of source packets in a sequential manner, and the decoder also reconstructs the source packets sequentially. The codes in [4], [5] are optimal for the proposed channel model. However, when erasures can happen at arbitrary positions instead of a burst, which we refer to as *isolated erasure* patterns, these codes incur significant packet losses. More recently in [6], we proposed a class of streaming codes that can correct from certain types of burst and isolated erasure patterns. These codes were shown to simultaneously attain a large column distance and column span, which are required for the isolated-erasure and burst-erasure recovery respectively. However, these codes were not designed to handle burst and isolated erasures occurring within the same decoding window. We show in the present paper that such error events can be significant over the Gilbert-Elliott channels, particularly when the decoding delay is moderately large. We propose a new class of codes that guarantee partial recovery for such patterns and observe significant performance gains in the simulations.

Streaming codes for channels that can introduce multiple erasure bursts, with a certain minimum guard interval separating them, have been independently developed in [7], [8]. However, to our knowledge these works do not aim for a robust extension involving both burst and isolated erasures as considered in this paper. As mentioned before, the most closely related work to the present paper is our earlier work in [6]. In the sequel, we explain why these constructions are not suitable when the decoding window introduces both burst and isolated erasures. Our proposed codes involves suitable modifications of the constructions in [6] based on a layered code design. However, the construction and analysis are considerably more challenging due to the complexity of the erasure patterns considered here. In the broader literature, real-time streaming

Ahmed Badr and Ashish Khisti are with the School of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5S 3G4 Canada (e-mail: abadr,akhisti@comm.utoronto.ca)

John Apostolopoulos and Wai-Tian Tan were with Hewlett-Packard Laboratories, Palo Alto, CA 94304 when this work was done. They are now with Cisco Systems, CA.

This work was supported by a discovery grant from National Science and Engineering Research Council of Canada, Hewlett Packard through a HP-IRP Award and an Ontario Early Researcher Award. Preliminary results of this work were presented at the 2013 CWIT Conference, Toronto, Canada and the 2013 ISIT Conference, Istanbul, Turkey.

over i.i.d. erasure channels has been studied in [9] where it is established that the playback delay increases logarithmically with the time elapsed since the start of transmission with or without feedback. Hybrid approaches with FEC and feedback are studied in [10], [11], while tree codes for streaming over i.i.d. erasure channels are studied in [12].

In the rest of this paper, we review some previously proposed streaming codes in Section II and then introduce the class of Partial Recovery Codes in Section III. We propose an alternate construction that requires a smaller field size in Section IV. Simulation results of the Gilbert-Elliott and Fritchman channel models are presented in Section V while Conclusion appear in Section VI.

II. BACKGROUND

We review two different classes of streaming codes, m -MDS and MiDAS codes from earlier works. These codes will constitute an important building block in our proposed constructions. We discuss their erasure correction properties in the streaming setup and motivate our new construction.

A. Problem Setup

At each time instant $i \geq 0$, a source packet $\mathbf{s}[i] \in \mathbb{F}_q^k$ consisting of k symbols over a field \mathbb{F}_q , arrives at the encoder, and a channel packet $\mathbf{x}[i] \in \mathbb{F}_q^n$ consisting of n symbols is transmitted, i.e.,

$$\mathbf{x}[i] = f_i(\mathbf{s}[0], \dots, \mathbf{s}[i]), \quad (1)$$

where $f_i(\cdot)$ denotes the encoding function at time i . The rate of the code is $R = k/n$. The channel is a packet erasure channel. At time i either the entire packet is received perfectly by the receiver i.e., $\mathbf{y}[i] = \mathbf{x}[i]$, or it is completely erased i.e., $\mathbf{y}[i] = \star$. In practice, packet losses occur due to a variety of impairments such as channel fading, buffer overflow, congestion etc. Such packet losses are often modelled (see e.g., [3] and references therein) using statistical models such as the Gilbert-Elliott [13], [14] or the Fritchman [15] model. However, the direct analysis of such models appears rather involved. In the sequel, we will introduce a class of *sliding window erasure channels*. Such models provide useful approximations to the erasure patterns in the underlying statistical models and are simpler to analyze.

The receiver must reconstruct each source packet with a delay of T packets, i.e.,

$$\hat{\mathbf{s}}[i] = \gamma_i(\mathbf{y}[0], \mathbf{y}[1], \dots, \mathbf{y}[i+T]), \quad (2)$$

where γ_i is the decoding function at time i . We assume that $\hat{\mathbf{s}}[i] \in \mathbb{F}_q^k \cup \{\star\}$, where the symbol \star denotes a failure in decoding the packet by its deadline. A code consisting of a causal encoder (1) and delay-constrained decoder (2) will be referred to as a *streaming code*. In this work, we will examine the average loss probability achieved by various streaming codes under different channel models.

B. m -MDS Codes

The m -MDS codes [16]–[18] are a class of (n, k, m) systematic linear convolutional codes that map each source packet $\mathbf{s}[i] \in \mathbb{F}_q^k$ to a channel packet $\mathbf{x}[i] \in \mathbb{F}_q^n$. Since the code

is assumed to be systematic, we can express $\mathbf{x}[i] = (\mathbf{s}[i], \mathbf{p}[i])$, where $\mathbf{p}[i] \in \mathbb{F}_q^{n-k}$, the parity-check packet at time i , is a linear combination of the source packets

$$\mathbf{p}[i] = \left(\sum_{l=0}^m \mathbf{s}^\dagger[i-l] \cdot \mathbf{H}_l \right)^\dagger. \quad (3)$$

Here m denotes the memory of the code¹, while the matrices $\mathbf{H}_l \in \mathbb{F}_q^{k \times (n-k)}$.

The m -MDS codes, correspond to a specific choice of matrices \mathbf{H}_l such that the code has a maximum column distance [16]. Intuitively, such a property guarantees that each symbol of a parity-check packet contributes an independent linear equation within the window of length $m+1$. Thus the decoder can recover the erased source packets, once it has received sufficiently many parity-check packets. While such a property is satisfied with high probability if the matrices \mathbf{H}_l are selected at random, the m -MDS codes provide deterministic constructions with these properties. In our setup, we can express the error correction properties of these codes via the following Lemma, which will be useful in our subsequent analysis.

Lemma 1. *Consider a systematic (n, k, m) m -MDS code of rate $R = \frac{k}{n}$ that maps each source packet $\mathbf{s}[i] \in \mathbb{F}_q^k$ into a channel packet $\mathbf{x}[i] \in \mathbb{F}_q^n$ that is transmitted over the channel at time i . For each $j \in \{0, 1, \dots, m\}$, the following is satisfied:*

- P1. *Suppose that in the window $[0, j]$, the channel introduces no more than $N = \lfloor (1-R)(j+1) \rfloor$ erasures in arbitrary locations, then $\mathbf{s}[0]$ is recovered by time $t = j$.*
- P2. *Suppose an erasure burst of length B spans the interval $[0, B-1]$. All the packets $\mathbf{s}[0], \dots, \mathbf{s}[B-1]$ are simultaneously recovered by time $t = j$ provided that $B \leq \lfloor (1-R)(j+1) \rfloor$.*
- P3. *Suppose that there are two erasure bursts spanning the intervals $[0, B_1-1]$ and $[r, r+B_2-1]$. All the $B_1 + B_2$ packets $\mathbf{s}[0], \dots, \mathbf{s}[B_1-1], \mathbf{s}[r], \dots, \mathbf{s}[r+B_2-1]$ are recovered by time $t = j$ provided that*

$$r \leq \frac{B_1}{1-R} \quad \text{and} \quad B_1 + B_2 \leq \lfloor (1-R)(j+1) \rfloor. \quad (4)$$

Proof: We refer the reader to [6] for the proof of properties P1 and P2. The proof of property P3 is presented in Appendix A. ■

Even though Properties P1–P3 of Lemma 1 pertain to the erasures in the interval $[0, j]$, they can be immediately translated to different types of sliding window erasure channel models. In property P1, the channel can introduce up to N erasures in arbitrary locations in any sliding window of length $W = j+1$. We will denote this channel by $C_1(N, W)$. Property P1 essentially guarantees that the m -MDS code recovers every source packet with a delay of $T = j$, provided that $N \leq \lfloor (1-R)(j+1) \rfloor$.

In property P2, the channel is only allowed to introduce one erasure burst of length no greater than B in any sliding window of length $W = j+1$. Such a channel will be denoted

¹We use \dagger instead of the conventional T to denote the vector/matrix transpose operation since T is used to denote the delay constraint. Throughout the paper, we will treat $\mathbf{s}[i]$ and $\mathbf{x}[i]$ as column vectors. For convenience, we will not use the \dagger notation when the dimensions are clear.

by $\mathcal{C}_{\text{II}}(B, W)$. Property P2 guarantees that using an m -MDS code, the entire burst of length $B \leq \lfloor (1-R)(j+1) \rfloor$ is recovered simultaneously.

Similarly in property P3, the channel can introduce two such erasure bursts.

The performance of m -MDS can be further improved over channels that introduce a mixture of burst and isolated erasures as discussed next.

C. MiDAS Codes

In [6], [19], a class of sliding window erasure channels that introduce both burst and isolated erasures was studied. The channel $\mathcal{C}_{\text{III}}(N, B, W)$ can only introduce up to N erasures in arbitrary locations or no more than B erasures in a single burst in any window consisting of W consecutive symbols. Note that upon setting $N = 1$, the model reduces to the burst erasure channel $\mathcal{C}_{\text{II}}(B, W)$ and upon setting $B = N$, the channel reduces to $\mathcal{C}_{\text{I}}(N, W)$. More generally $N \leq B$ must hold since the burst erasure is a specific type of erasure pattern. A new family of codes that achieve a near optimal rate was proposed for this class of channels.

Theorem 1 (Badr et al. [6], [19]). *For the $\mathcal{C}_{\text{III}}(N, B, W)$ channel, the maximum achievable rate R with a delay of T packets, such that $W \geq T + 1$ and $T \geq B$, satisfies:*

$$\frac{T-N}{T-N+B} \leq R \leq \frac{T-N+1}{T-N+B+1}. \quad (5)$$

We will review the associated code, Maximum Distance And Span (MiDAS) code, in Theorem 1 next. The basic idea behind this construction is to first construct a burst erasure code for the $\mathcal{C}_{\text{II}}(B, W)$ channel with a delay of T . Thereafter, we append a separate layer of parity-checks to make the code robust to N isolated erasures. We present the details of the encoder and decoder, as such a layering approach will also be used in our subsequent constructions.

Encoder:

- 1) Split each source packet $\mathbf{s}[i] \in \mathbb{F}_q^k$ into two groups of sub-packets, $\mathbf{u}[i] \in \mathbb{F}_q^{k^u}$ and $\mathbf{v}[i] \in \mathbb{F}_q^{k^v}$ as follows,

$$\mathbf{s}[i] = \underbrace{(u_0[i], \dots, u_{k^u-1}[i])}_{=\mathbf{u}[i]}, \underbrace{(v_0[i], \dots, v_{k^v-1}[i])}_{=\mathbf{v}[i]}, \quad (6)$$

where $k^u + k^v = k$.

- 2) Apply a $(k^v + k^u, k^v, T)$ m -MDS code to the $\mathbf{v}[\cdot]$ sub-packets of rate $R^v = \frac{k^v}{k^v+k^u}$ generating the k^u parity-check symbols $\mathbf{p}^v[i] = (p_0^v[i], \dots, p_{k^u-1}^v[i])$.
- 3) Combine a shifted version of the $\mathbf{u}[\cdot]$ packets with the $\mathbf{p}^v[\cdot]$ parities, i.e., $\mathbf{q}[i] = \mathbf{p}^v[i] + \mathbf{u}[i-T]$.
- 4) Apply a $(k^u + k^s, k^u, T)$ m -MDS code to the $\mathbf{u}[\cdot]$ sub-packets of rate $R^u = \frac{k^u}{k^u+k^s}$ generating the k^s parity-check symbols $\mathbf{p}^u[i] = (p_0^u[i], \dots, p_{k^s-1}^u[i])$.
- 5) Concatenate the parity-check packets and the source packet to generate the overall channel packets,

$$\mathbf{x}[i] = (\mathbf{u}[i], \mathbf{v}[i], \mathbf{p}^v[i] + \mathbf{u}[i-T], \mathbf{p}^u[i]).$$

Note that at each time the construction takes $k^u + k^v$ source symbols in packet $\mathbf{s}[i]$ and maps them to a channel packet $\mathbf{x}[i]$ consisting of $2k^u + k^v + k^s$ symbols. It follows that the rate of the code is $R = \frac{k^u+k^v}{2k^u+k^v+k^s}$. We select

$$k^u = MB, \quad k^v = M(T-B), \quad k^s = \frac{MNB}{T-N+1},$$

where M is selected such that k^s is also an integer. It can be readily verified that these choice of parameters satisfies the lower bound in Theorem 1. Furthermore the rate R^v of the $(k^v + k^u, k^v, T)$ constituent m -MDS in step (2), and the rate R^u in step (4) are given by:

$$R^v = \frac{T-B}{T}, \quad R^u = \frac{T-N+1}{T+1}, \quad (7)$$

which satisfy $(1-R^v)T = B$ and $(1-R^u)(T+1) = N$.

Decoder: In the analysis of the decoder, we consider the interval $[0, T]$ and show that the decoder can recover $\mathbf{s}[0]$ by time $t = T$ if there is either an erasure burst of length B or smaller, or up to N isolated erasures in this interval. Once $\mathbf{s}[0]$ has been recovered, its effect can be subtracted from all future packets, and then the interval $[1, T+1]$ can be similarly considered to recover $\mathbf{s}[1]$. Thus by exploiting the time-invariance of the code, every source packet can be recovered.

For the case of burst erasures, we suppose that an erasure burst of length B spans the interval $[0, B-1]$. We are then guaranteed that there are no additional erasures in the interval $[B, T]$ since $W \geq T+1$ is assumed. The decoder proceeds in two steps. For each $j \in \{B, \dots, T-1\}$, the decoder recovers the parity-check packets $\mathbf{p}^v[j]$, by subtracting the unerased $\mathbf{u}[j-T]$ from the associated $\mathbf{q}[j] = \mathbf{p}^v[j] + \mathbf{u}[j-T]$ packets. Then property P2 in Lemma 1 can be used to recover the B erased sub-packets $\mathbf{v}[0], \dots, \mathbf{v}[B-1]$ at time $T-1$ since $(1-R^v)T = B$ (cf. (7)). In the second step, the decoder computes the parity-check packets $\mathbf{p}^v[T]$ and subtracts it from $\mathbf{q}[T] = \mathbf{p}^v[T] + \mathbf{u}[0]$ to recover the underlying $\mathbf{u}[0]$ by its deadline. Hence, $\mathbf{s}[0] = (\mathbf{u}[0], \mathbf{v}[0])$ is decoded at time T .

In the case of N isolated erasures in the interval $[0, T]$, the decoder first recovers $\mathbf{v}[0]$ at time $t = T-1$ using the parity-checks $\mathbf{q}[\cdot]$, and $\mathbf{u}[0]$ at time $t = T$ using the parity-checks $\mathbf{p}^u[\cdot]$ constructed in step 4. The recovery of $\mathbf{v}[0]$ by time $T-1$ follows in a fashion similar to that in the case of a burst erasure. However, we use P1 in Lemma 1 instead which shows that the number of isolated erasures under which the recovery of $\mathbf{v}[0]$ is possible is given by $N^v = (1-R^v)T = B$. Since $N \leq B$ holds, the recovery of $\mathbf{v}[0]$ by time $t = T-1$ is guaranteed. For recovering $\mathbf{u}[0]$, note that the associated code $(\mathbf{u}[i], \mathbf{p}^u[i])$ is a m -MDS code with rate R^u (cf. (7)) that satisfies $(1-R^u)(T+1) = N$. The recovery of $\mathbf{u}[0]$, and in turn $\mathbf{s}[0]$, follows from property P1 in Lemma 1.

D. Example - MiDAS Codes

Table I illustrates an example of a MiDAS code for $\mathcal{C}_{\text{III}}(N=2, B=3, W=8)$ and a delay $T=7$. Such a code recovers a burst of maximum length $B=3$, or $N=2$ isolated erasures in any window of length $W=8$. Suppose that an erasure burst spans the interval $[0, 2]$. First $\mathbf{v}[0], \mathbf{v}[1]$, and $\mathbf{v}[2]$ are recovered at time $t=6$, using the parity-check packets $\mathbf{p}^v[3], \dots, \mathbf{p}^v[6]$, which are recovered by cancelling the unerased sub-packets $\mathbf{u}[-4], \dots, \mathbf{u}[-1]$ from the parity-check packets $\mathbf{q}[3], \dots, \mathbf{q}[6]$. Thereafter, $\mathbf{u}[0]$ can be recovered at time $t=7$ by cancelling $\mathbf{p}^u[7]$ etc. In contrast, if there are up to $N=2$ isolated erasures in the interval $[0, 7]$ then the parity-checks $\mathbf{p}^v[t]$ in the interval $t \in [0, 6]$ can be used to

	[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
$k^u = 15$	$\mathbf{u}[0]$	$\mathbf{u}[1]$	$\mathbf{u}[2]$	$\mathbf{u}[3]$	$\mathbf{u}[4]$	$\mathbf{u}[5]$	$\mathbf{u}[6]$	$\mathbf{u}[7]$	$\mathbf{u}[8]$	$\mathbf{u}[9]$
$k^v = 20$	$\mathbf{v}[0]$	$\mathbf{v}[1]$	$\mathbf{v}[2]$	$\mathbf{v}[3]$	$\mathbf{v}[4]$	$\mathbf{v}[5]$	$\mathbf{v}[6]$	$\mathbf{v}[7]$	$\mathbf{v}[8]$	$\mathbf{v}[9]$
$k^u = 15$	$\mathbf{p}^v[0]$	$\mathbf{p}^v[1]$	$\mathbf{p}^v[2]$	$\mathbf{p}^v[3]$	$\mathbf{p}^v[4]$	$\mathbf{p}^v[5]$	$\mathbf{p}^v[6]$	$\mathbf{p}^v[7]$	$\mathbf{p}^v[8]$	$\mathbf{p}^v[9]$
	$+\mathbf{u}[-7]$	$+\mathbf{u}[-6]$	$+\mathbf{u}[-5]$	$+\mathbf{u}[-4]$	$+\mathbf{u}[-3]$	$+\mathbf{u}[-2]$	$+\mathbf{u}[-1]$	$+\mathbf{u}[0]$	$+\mathbf{u}[1]$	$+\mathbf{u}[2]$
$k^s = 6$	$\mathbf{p}_2^y[0]$	$\mathbf{p}_2^y[1]$	$\mathbf{p}_2^y[2]$	$\mathbf{p}_2^y[3]$	$\mathbf{p}_2^y[4]$	$\mathbf{p}_2^y[5]$	$\mathbf{p}_2^y[6]$	$\mathbf{p}_2^y[7]$	$\mathbf{p}_2^y[8]$	$\mathbf{p}_2^y[9]$

TABLE I: A MiDAS code desinged for $\mathcal{C}(K = 0, N = 2, B = 3, W = 8)$ and a delay of $T = 7$. The rate of the code is $R = \frac{35}{56} = \frac{5}{8}$ which matches the lower-bound in Theorem 1.

recover $\mathbf{v}[0]$ while the parity-checks in the last layer, $\mathbf{p}^u[t]$ and in the interval $t \in [0, 7]$ can be used to recover $\mathbf{u}[0]$.

In practice, the values of N and B are dictated by the number of isolated erasures (associated with the good state), and the burst-length (associated with the bad state) of the Gilbert-Elliott channel that must be corrected by a streaming code. However, these are not the only patterns that must be corrected. It turns out that erasure patterns consisting of an erasure burst followed or preceded by an isolated erasure can also cause significant error events. Such patterns arise for example when the channel in the bad state transitions into a good state, where it introduces an isolated erasure. The MiDAS codes can lead to significant losses over these erasure patterns. Consider for example the code in Table I and suppose the channel introduces erasures at times $t = 0, 1, 6$. i.e., a burst of length $B' = 2$ is followed by an isolated erasure. Such a pattern is not included in the channel $\mathcal{C}_{\text{III}}(N = 2, B = 3, W = 8)$. It can be seen that the isolated erasure at time $t = 6$ will interfere with the recovery of $\mathbf{u}[0]$ and $\mathbf{u}[1]$ at time $t = 7$ and $t = 8$ respectively as the parity-check symbols $\mathbf{p}^v[t]$ will involve $\mathbf{v}[6]$ which is erased. Thus we need new code constructions that deal with such erasure patterns.

III. PARTIAL RECOVERY CODES

We introduce a class of streaming codes that can handle erasure patterns consisting of a burst erasure accompanied by an isolated erasure. We first formally define such a channel model and then discuss the code construction.

A. System Model and Main Results

Definition 1 (Channel $\mathcal{C}(K, N, B, W)$). *In any sliding window of length W , the channel $\mathcal{C}(K, N, B, W)$ can introduce only one of the following patterns:*

- A single erasure burst of maximum length B plus no more than K isolated erasures where $B + K < W$ or,
- A maximum of $N \leq B + K$ erasures in arbitrary locations.

Remark 1. *Note that if $B + K \geq W$ the capacity is trivially zero, and so the constraint $B + K < W$ in Definition 1 must be imposed. Also, the condition $N \leq B + K$ follows since a burst plus isolated erasures is included in the isolated erasures pattern.*

When $K = 0$, the model reduces to the $\mathcal{C}_{\text{III}}(N, B, W)$ used in the MiDAS code construction. In this section, we will focus on the case of $K = 1$ which appears to be of

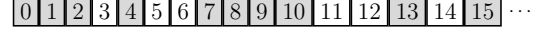


Fig. 1: An example of channel $\mathcal{C}(1, 2, 3, 5)$ in Definition 1: In any sliding window of length $W = 5$ there is either a single erasure burst of length up to $B = 3$ and possibly one isolated erasure, or $N = 2$ isolated erasures.

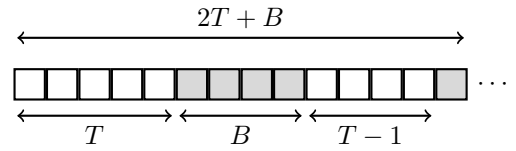


Fig. 2: An isolated erasure associated with a burst in channel $\mathcal{C}_{\text{II}}(N, B, W = 2T + B)$

most practical interest. We will initially consider the case when $N = 1$ i.e., the $\mathcal{C}(K = 1, N = 1, B, W)$. Such a channel allows us to focus on the burst-erasure patterns. We will subsequently show in Section III-D how our construction can be extended to handle the case where $N > 1$, by adding an extra layer of parity-checks. We further focus on the case when $W \geq 2T + B$, where T is the decoding delay. This allows us to consider an isolated erasure happening either T packets before or after an erasure burst.

Definition 2 (Associated Isolated Erasure). *An isolated erasure is defined to be associated with the erasure burst, if it occurs within the T packets before or after this burst.*

Note that every erasure burst has at most one associated isolated erasure (cf. Fig. 2). Conversely every isolated erasure can be associated with no more than one erasure burst. Instead of attempting to recover every source packet for such erasure patterns we introduce a class of partial recovery codes as discussed below. The advantage of this approach is that the overhead required is smaller, which is desirable in practice.

Definition 3 (Partial Recovery Code (PRC)). *A Partial Recovery Code (PRC) for $\mathcal{C}(K = 1, N = 1, B, W \geq 2T + B)$ recovers all but one source packet with a delay of T in each pattern consisting of an erasure burst and its associated isolated erasure.*

Our main result is the following achievable rate for partial recovery codes.

Theorem 2 (Partial Recovery Codes (PRC)). *There exists a partial recovery code satisfying Definition 3, for the $\mathcal{C}(K =$*

$$1, N = 1, B, W \geq 2T + B \text{ channel of rate}$$

$$R = \max_{B < \Delta \leq T} \frac{\Delta(T - \Delta) + (B + 1)}{\Delta(T - \Delta) + (B + 1)(T - \Delta + 2)}. \quad (8)$$

□

The coding scheme associated with Theorem 2 is presented next. The decoding analysis is presented in Appendix B.

B. Partial Recovery Codes - Code Construction

Let k^u, k^v, k^r and Δ be integers that will be specified in the sequel. We outline the main steps in the code construction below.

- 1) **Source Splitting:** Split each source packet $\mathbf{s}[i] \in \mathbb{F}_q^k$ into two groups of sub-packets, $\mathbf{u}[i] \in \mathbb{F}_q^{k^u}$ and $\mathbf{v}[i] \in \mathbb{F}_q^{k^v}$ as in (6).
- 2) **Construction of \mathcal{C}_{12} :** We apply a $(k^v + k^u + k^r, k^v, T)$ m -MDS code, $\mathcal{C}_{12} : (\mathbf{v}[i], \mathbf{p}[i])$, of rate $R_{12} = \frac{k^v}{k^v + k^u + k^r}$ to the $\mathbf{v}[\cdot]$ sub-packets to generate parity-check packets $\mathbf{p}[\cdot] \in \mathbb{F}_q^{k^u + k^r}$,

$$\mathbf{p}[i] = \left(\sum_{l=0}^T \mathbf{v}^\dagger[i-l] \cdot \mathbf{H}_l \right)^\dagger, \quad (9)$$

where $\mathbf{H}_l \in \mathbb{F}_q^{k^v \times (k^u + k^r)}$ for $l = 0, \dots, T$ are the matrices associated with the m -MDS code (3).

- 3) **Parity-Check Splitting:** We split each $\mathbf{p}[i]$ into two groups of sub-packets, $\mathbf{p}^1[i] \in \mathbb{F}_q^{k^u}$ and $\mathbf{p}^2[i] \in \mathbb{F}_q^{k^r}$ by assigning the first k^u symbols in $\mathbf{p}[i]$ to $\mathbf{p}^1[i]$ and the remaining k^r symbols of $\mathbf{p}[i]$ to $\mathbf{p}^2[i]$. We can express:

$$\mathbf{p}^k[i] = \left(\sum_{l=0}^T \mathbf{v}^\dagger[i-l] \cdot \mathbf{H}_l^k \right)^\dagger, \quad k = 1, 2 \quad (10)$$

where the matrices \mathbf{H}_l^k are defined as $\mathbf{H}_l = [\mathbf{H}_l^1 \mid \mathbf{H}_l^2]$ for $l = 0, \dots, T$. It can be shown that both \mathbf{H}_l^1 and \mathbf{H}_l^2 satisfy the m -MDS property [18, Theorem 2.4] and therefore the codes $\mathcal{C}_1 : (\mathbf{v}[i], \mathbf{p}^1[i])$ and $\mathcal{C}_2 : (\mathbf{v}[i], \mathbf{p}^2[i])$ are both m -MDS codes².

- 4) **Repetition Code:** We combine a shifted copy of $\mathbf{u}[\cdot]$ with the $\mathbf{p}^1[\cdot]$ parity-checks to generate $\mathbf{q}[i] = \mathbf{p}^1[i] + \mathbf{u}[i - \Delta]$. Here $\Delta \in \{B+1, \dots, T\}$ denotes the shift applied to the $\mathbf{u}[\cdot]$ stream before embedding it onto the $\mathbf{p}^1[\cdot]$ stream.

- 5) **Channel Packet:** We concatenate the parity-check packets to the source packet to construct the channel packet,
$$\mathbf{x}[i] = (\mathbf{s}[i], \mathbf{q}[i], \mathbf{p}^2[i]). \quad (11)$$

The rate of the code in (11) is clearly $R = \frac{k^u + k^v}{2k^u + k^v + k^r}$. We further select the codes \mathcal{C}_{12} and \mathcal{C}_2 to have the following rates:

$$R_{12} \triangleq \frac{k^v}{k^v + k^u + k^r} = \frac{\Delta - B - 1}{\Delta}, \quad (12)$$

$$R_2 \triangleq \frac{k^r}{k^v + k^r} = \frac{T - \Delta + 1}{T - \Delta + 2}. \quad (13)$$

The rate R_{12} is chosen such that the corresponding code \mathcal{C}_{12} is capable of correcting a burst of length $B+1$ within a delay of $\Delta-1$ symbols. Hence, (12) satisfies $(1-R_{12})(\Delta) = B+1$. Similarly, the rate R_2 is chosen such that the \mathcal{C}_2 code is capable of recovering a single erasure within a delay of $T-\Delta+1$. Thus, (13) satisfies $(1-R_2)(T-\Delta+2) = 1$. It will be justified

²The proof of this property uses the fact that a subset of linearly independent columns is linearly independent and is rather straightforward and is thus omitted.

	[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
$k^u = 6$	$\mathbf{u}[0]$	$\mathbf{u}[1]$	$\mathbf{u}[2]$	$\mathbf{u}[3]$	$\mathbf{u}[4]$	$\mathbf{u}[5]$	$\mathbf{u}[6]$	$\mathbf{u}[7]$	$\mathbf{u}[8]$
$k^v = 4$	$\mathbf{v}[0]$	$\mathbf{v}[1]$	$\mathbf{v}[2]$	$\mathbf{v}[3]$	$\mathbf{v}[4]$	$\mathbf{v}[5]$	$\mathbf{v}[6]$	$\mathbf{v}[7]$	$\mathbf{v}[8]$
$k^u = 6$	$\mathbf{p}^1[0]$	$\mathbf{p}^1[1]$	$\mathbf{p}^1[2]$	$\mathbf{p}^1[3]$	$\mathbf{p}^1[4]$	$\mathbf{p}^1[5]$	$\mathbf{p}^1[6]$	$\mathbf{p}^1[7]$	$\mathbf{p}^1[8]$
$k^r = 2$	$+\mathbf{u}[-6]$	$+\mathbf{u}[-5]$	$+\mathbf{u}[-4]$	$+\mathbf{u}[-3]$	$+\mathbf{u}[-2]$	$+\mathbf{u}[-1]$	$+\mathbf{u}[0]$	$+\mathbf{u}[1]$	$+\mathbf{u}[2]$
	$\mathbf{p}^2[0]$	$\mathbf{p}^2[1]$	$\mathbf{p}^2[2]$	$\mathbf{p}^2[3]$	$\mathbf{p}^2[4]$	$\mathbf{p}^2[5]$	$\mathbf{p}^2[6]$	$\mathbf{p}^2[7]$	$\mathbf{p}^2[8]$

TABLE II: A Partial Recovery Code with $B = 3$ achieving a rate of $\frac{5}{9}$ for a delay of $T = 7$.

that these two properties are sufficient for partial recovery on channel $\mathcal{C}(1, 1, B, W \geq 2T + B)$.

We further use the following values of k^u, k^v and k^r that satisfy (12) and (13),

$$\begin{aligned} k^u &= (B+1)(T-\Delta+1) - (\Delta-B-1) \\ k^v &= (T-\Delta+1)(\Delta-B-1) \\ k^r &= \Delta-B-1, \end{aligned} \quad (14)$$

and the corresponding rate of the PRC code is

$$R = \max_{B < \Delta \leq T} \frac{\Delta(T-\Delta) + (B+1)}{\Delta(T-\Delta) + (B+1)(T-\Delta+2)}, \quad (15)$$

which meets the rate given in Theorem 2. This completes the details of the code construction. The decoding steps are described in detail in Appendix B.

Remark 2. In our construction, the parameter $\Delta \in \{B+1, \dots, T\}$ is a free parameter that can be optimized. Selecting Δ to be small, will reduce the effective delay of the burst-erasure code, and require the size of parity-checks $\mathbf{p}^1[\cdot]$ to be large. However, the duration for recovering the isolated erasure is longer and hence the size of $\mathbf{p}^2[\cdot]$ can be smaller. Conversely selecting a larger value of Δ will require the size of $\mathbf{p}^2[\cdot]$ to be large while the size of $\mathbf{p}^1[\cdot]$ can be smaller. The optimization in (15) reflects such a tradeoff.

Remark 3. If we assume that the source alphabet is sufficiently large such that the integer constraints can be ignored then the optimal shift is given by

$$\Delta^* = \arg \max_{\Delta} R(\Delta) = T + 1 - \sqrt{T - B}, \quad (16)$$

and the corresponding rate is given by

$$R^* = \frac{(T+2)\sqrt{T-B} - 2(T-B)}{(T+B+3)\sqrt{T-B} - 2(T-B)}, \quad (17)$$

which is achieved using

$$k^u = k^{u^*} = (B+2)\sqrt{T-B} - (T-B)$$

$$k^v = k^{v^*} = (T-B)(\sqrt{T-B} - 1)$$

$$k^r = k^{r^*} = (T-B) - \sqrt{T-B}.$$

In particular, when $T \gg 1$ and $B \gg 1$ holds the rate expression (17) reduces to

$$R^* \approx \frac{T - 2\sqrt{T-B}}{T+B - 2\sqrt{T-B}}. \quad (18)$$

Comparing (18) to the capacity of the burst erasure channel $C = \frac{T}{T+B}$ (see [4]), we see effect of an additional isolated erasure reduces the rate by no more than $\frac{2\sqrt{T-B}}{T+B}$.

C. Example - PRC

We illustrate a PRC construction in Table II for the channel with parameters $N = 1, K = 1, B = 3, W = 17$ and with delay $T = 7$. The values of the code parameters are $k^u = 6, k^v = 4, k^r = 2$ and $\Delta = 6$ and the rate is $\frac{5}{9}$.

Encoder

- 1) Each source packet $\mathbf{s}[i] \in \mathbb{F}_q^{10}$ consisting of ten symbols is split it into $\mathbf{u}[i]$ and $\mathbf{v}[i]$, consisting of $k^u = 6$ and $k^v = 4$ symbols respectively.
- 2) Let $k^r = 2$ and apply a $(k^u + k^v + k^r, k^v) = (12, 4)$ m -MDS code \mathcal{C}_{12} to the $\mathbf{v}[i]$ sub-packets to generate parity-checks $\mathbf{p}[i] \in \mathbb{F}_q^8$. Note that the rate of \mathcal{C}_{12} equals $R_{12} = \frac{1}{3}$.
- 3) We split each $\mathbf{p}[i]$ into $\mathbf{p}^1[i]$ and $\mathbf{p}^2[i]$ consisting of $k^u = 6$ and $k^r = 2$ symbols as shown in Table II. Note that the codes $\mathcal{C}_1 : (\mathbf{v}[i], \mathbf{p}^1[i])$ and $\mathcal{C}_2 : (\mathbf{v}[i], \mathbf{p}^2[i])$ are both m -MDS codes of rates $R_1 = 2/5$ and $R_2 = 2/3$ respectively.
- 4) We generate $\mathbf{q}[i] = \mathbf{p}^1[i] + \mathbf{u}[i - \Delta]$ and let the transmitted packet be $\mathbf{x}[i] = (\mathbf{u}[i], \mathbf{v}[i], \mathbf{q}[i], \mathbf{p}^2[i])$, which corresponds to one column in Table II.

The overall rate equals $R = 5/9$.

Decoder

We start by considering the case when the burst erasure precedes the isolated erasure. Without loss of generality suppose that the burst erasure spans the interval $[0, 2]$ and the isolated erasure occurs at time $t > 2$.

- $t \in [3, 4]$: We use code \mathcal{C}_{12} in the interval $[0, \Delta - 1] = [0, 5]$ to recover the erased sub-packets $\mathbf{v}[0], \mathbf{v}[1], \mathbf{v}[2], \mathbf{v}[t]$ simultaneously at time $\tau = 5$. Towards this end we use the parity-checks $\mathbf{p}^1[\cdot]$ obtained by cancelling the $\mathbf{u}[\cdot]$ sub-packet in this interval as well as the parity-checks $\mathbf{p}^2[\cdot]$. Our erasure pattern consists of a burst erasure of length $B_1 = 3$ spanning the interval $[0, 2]$ and an isolated erasure at time either $t = 3$ or $t = 4$. Note that $(1 - R_{12})(\tau + 1) = 4$ and that the conditions associated with property P3 in Lemma 1 are satisfied, and thus the recovery of the $\mathbf{v}[\cdot]$ follows.

For the recovery of $\mathbf{u}[0], \mathbf{u}[1], \mathbf{u}[2]$ and $\mathbf{u}[t]$, we need to subtract associated the $\mathbf{p}^1[\cdot]$ parity-check packet from $\mathbf{q}[6], \mathbf{q}[7], \mathbf{q}[8]$ and $\mathbf{q}[t + \Delta]$. This follows since all the $\mathbf{v}[\cdot]$ sub-packets are recovered already by time $t = 5$.

- $t \geq 5$: We use code \mathcal{C}_{12} in the interval $[0, 4]$ to recover the erased sub-packets $\mathbf{v}[0], \mathbf{v}[1], \mathbf{v}[2]$ at time $\tau = 4$. This follows by applying property P2 in Lemma 1 since $(1 - R_{12})(\tau + 1) \geq 3$.

Once the erasure burst has been recovered, the erased sub-packet $\mathbf{v}[t]$ can be recovered at time $\tau = t + T - \Delta + 1 = t + 2$ using the parity-checks $\mathbf{p}^2[\cdot]$ associated with \mathcal{C}_2 in the interval $[t, t + 2]$. This again follows from Lemma 1 since

$$(1 - R_2)(T - \Delta + 2) = \left(1 - \frac{2}{3}\right) 3 = 1 \quad (19)$$

which suffices to recover the missing $\mathbf{v}[t]$.

For the recovery of $\mathbf{u}[0], \mathbf{u}[1], \mathbf{u}[2]$, we need to use $\mathbf{q}[6], \mathbf{q}[7], \mathbf{q}[8]$ respectively. If $t = 5$ then we recover $\mathbf{v}[5]$ at time $\tau = 7$ and then compute $\mathbf{p}^1[6], \dots, \mathbf{p}^1[8]$ and in turn recover all the missing $\mathbf{u}[\cdot]$ sub-packets. If $t \in [6, 8]$, then we will not be able to recover the associated $\mathbf{u}[t - 6]$ but the remaining two sub-packets can be recovered. Thus we will have one non-recovered packet in this case. If

$t > 8$ then clearly all the three $\mathbf{u}[\cdot]$ sub-packets are recovered and complete recovery is achieved.

For the case when the isolated erasure occurs at time 0 and the burst follows it spanning the interval $[t, t + 2]$, the decoder proceeds as follows.

- $t = 1$: We use code \mathcal{C}_{12} in the interval $[0, 5]$ to recover $\mathbf{v}[0], \dots, \mathbf{v}[3]$ at time $\tau = 5$. The decoder computes $\mathbf{p}^1[t + 1]$ in the interval $[6, 9]$ and subtracts them to recover $\mathbf{u}[0], \dots, \mathbf{u}[3]$. All the erased packets are recovered for this erasure pattern.
- $t \geq 2$: We recover $\mathbf{v}[0]$ at time 1 using the code \mathcal{C}_{12} before the start of the erasure burst. In particular, taking $\tau = 1$ note that

$$(1 - R_{12})(\tau + 1) = \left(1 - \frac{1}{3}\right) 2 > 1, \quad (20)$$

which by Lemma 1 suffices to recover $\mathbf{v}[0]$ at time $\tau = 1$. We next show how $\mathbf{v}[t], \mathbf{v}[t + 1], \mathbf{v}[t + 2]$ can be recovered using the code \mathcal{C}_{12} in the window $[t, t + 5]$. If $t \in \{2, 3\}$ then in addition to $\mathbf{v}[t], \mathbf{v}[t + 1], \mathbf{v}[t + 2]$ we should also account for the erasure at time 6 due to the repetition of $\mathbf{u}[0]$. Thus there are a total of 4 erasures in the window $[t, t + 5]$ with one burst of length three. The recovery from such a pattern follows by invoking P3 in Lemma 1. For $t \in [4, 6]$, the erasure burst overlaps with the repetition of $\mathbf{u}[0]$ and thus there are only three erasures and the recovery follows by the application of Lemma 1. For $t > 6$, the sub-packet $\mathbf{u}[0]$ is recovered at time 6 and then the erasure burst in $[t, t + 2]$ can be recovered separately.

For general values of B and T , similar decoding steps are used and provided in Appendix B.

D. Robust Partial Recovery Codes

While the PRC constructed in Section III-B is for the special case of $N = 1$, we now extend such construction to be able to deal with $N > 1$ isolated erasures, i.e., the $\mathcal{C}(K = 1, N, B, W \geq 2T + B)$ channel. Our development parallels the layered code design approach for MiDAS in Section II-C, where a burst erasure code is first constructed, and then an additional layer of parity-checks is appended to correct isolated erasures. We will refer to this construction as robust PRC. As will be apparent, these codes ensure *partial recovery* (cf. Definition 3) when the pattern consists of a burst and one isolated erasures but *perfect recovery* when the pattern consists of N isolated erasures.

We start with the PRC code for burst plus one isolated erasure constructed in Section III-B. The associated rate is $R = \frac{k^u + k^v}{2k^u + k^v + k^r}$ where we select k^u, k^v and k^r according to (14). We then add an additional layer of parity-checks to protect the $\mathbf{u}[\cdot]$ sub-packets. In particular, we apply a $(k^u + k^s, k^u, T)$ m -MDS code to the $\mathbf{u}[\cdot]$ sub-packets generating k^s parity-check symbols $\mathbf{w}[i] = (w_0[i], \dots, w_{k^s-1}[i])$. We then concatenate the generated parity-checks to the channel packet in (11), i.e., the channel packet of the robust version is given by,

$$\mathbf{x}[i] = (\mathbf{s}[i], \mathbf{q}[i], \mathbf{p}^2[i], \mathbf{w}[i]). \quad (21)$$

It remains to specify the value of k^s in the above construction such that $\mathbf{s}[0] = (\mathbf{u}[0], \mathbf{v}[0])$ is recovered by time

T . As in the case of MiDAS codes, $\mathbf{u}[0]$ and $\mathbf{v}[0]$ will be recovered separately. For $\mathbf{v}[0]$, we consider the interval $[0, \Delta - 1]$. We note that the interfering $\mathbf{u}[\cdot]$ sub-packets in this interval are not erased and can be cancelled out from $\mathbf{q}[\cdot]$ to recover the underlying parity-checks $\mathbf{p}^1[\cdot]$. Hence, we use property P1 in Lemma 1 to recover $\mathbf{v}[0]$ by time $\Delta - 1$ since, $(1 - R_{12})\Delta = B + 1 \geq N$. For $\mathbf{u}[0]$ to be recovered using the $(\mathbf{u}[i], \mathbf{w}[i])$ m -MDS code of rate $\frac{k^u}{k^u + k^s}$ in the interval $[0, T]$, the following needs to be satisfied,

$$\left(1 - \frac{k^u}{k^u + k^s}\right)(T + 1) \geq N \Rightarrow k^s \geq \frac{Nk^u}{T - N + 1}. \quad (22)$$

Thus by selecting k^s as above we complete our construction for any $N > 1$.

IV. PARTIAL RECOVERY CODES USING MDS CODES

In this section, we propose an alternative construction, which replaces m -MDS codes in the PRC construction with a diagonally interleaved MDS block code. While such codes incur some performance loss, they require a significantly lower field size and might be more suitable in certain applications. We will refer to these codes as PRC-MDS codes to differentiate them with the construction in Section III-B.

Theorem 3 (PRC using MDS codes). *There exists a partial recovery code for $\mathcal{C}(K = 1, N = 1, B, W \geq 2T + B)$ of rate,*

$$R = \max_{B < \Delta \leq T} \frac{(T - \Delta + 1)\Delta}{(T - \Delta + 1)(\Delta + B + 1) + (\Delta - B - 1)}, \quad (23)$$

that satisfy Definition 3 and with a field-size that increases as $\mathcal{O}(T^3)$. \square

We assume that each source packet $\mathbf{s}[i]$ consists of $(T - \Delta + 1)\Delta$ symbols, $\mathbf{s}[i] = (s_0[i], \dots, s_{(T-\Delta+1)\Delta-1}[i])$ and generate a total of $(B + 1)(T - \Delta + 1) + \Delta - B - 1$ parity-check symbols, as discussed below.

We divide $\mathbf{s}[i]$ into two groups, $\mathbf{s}[i] = (\mathbf{u}[i], \mathbf{v}[i])$, where

$$\begin{aligned} \mathbf{u}[i] &= (u_0[i], \dots, u_{(T-\Delta+1)(B+1)-1}[i]) \\ &= (s_0[i], \dots, s_{(T-\Delta+1)(B+1)-1}[i]) \end{aligned} \quad (24)$$

$$\begin{aligned} \mathbf{v}[i] &= (v_0[i], \dots, v_{(T-\Delta+1)(\Delta-B-1)-1}[i]) \\ &= (s_{(T-\Delta+1)(B+1)}[i], \dots, s_{(T-\Delta+1)\Delta-1}[i]). \end{aligned} \quad (25)$$

We then apply a systematic $(\Delta, \Delta - B - 1)$ MDS code to the $\mathbf{v}[\cdot]$ sub-packets with interleaving factor of $T - \Delta + 1$, generating $(T - \Delta + 1)(B + 1)$ parity-check packets, $\mathbf{p}^1[\cdot] = (p_0^1[i], \dots, p_{(T-\Delta+1)(B+1)-1}^1[i])$. The resulting codewords starting at $v_j[i]$ are expressed as:

$$\mathbf{c}_j^1[i] = \begin{bmatrix} v_j[i] \\ v_{j+(T-\Delta+1)}[i+1] \\ v_{j+2(T-\Delta+1)}[i+2] \\ \vdots \\ v_{j+(T-\Delta+1)(\Delta-B-2)}[i+\Delta-B-2] \\ p_j^1[i+\Delta-B-1] \\ p_{j+(T-\Delta+1)}^1[i+\Delta-B] \\ \vdots \\ p_{j+(T-\Delta+1)B}^1[i+\Delta-1] \end{bmatrix}, \quad (26)$$

for $j = \{0, 1, \dots, T - \Delta\}$. We combine the $\mathbf{u}[\cdot]$ sub-packets with the parity-check packets $\mathbf{p}^1[\cdot]$ after applying a shift of Δ to the earlier, i.e., $\mathbf{q}[i] = \mathbf{p}^1[i] + \mathbf{u}[i - \Delta]$.

Next we generate a second set of parity-check packets $\mathbf{p}^2[\cdot]$ by applying a $(T - \Delta + 2, T - \Delta + 1)$ MDS code to the $\mathbf{v}[\cdot]$ sub-packets with interleaving factor of $\Delta - B - 1$. Thus

$$\mathbf{c}_j^2[i] = \begin{bmatrix} v_j[i] \\ v_{j+(\Delta-B-1)}[i+1] \\ v_{j+2(\Delta-B-1)}[i+2] \\ \vdots \\ v_{j+(\Delta-B-1)(T-\Delta)}[i+T-\Delta] \\ p_j^2[i+T-\Delta+1] \end{bmatrix}, \quad (27)$$

for $j = \{0, 1, \dots, \Delta - B - 2\}$ are the codewords.

We finally concatenate the parity-check packets, $\mathbf{p}^2[\cdot]$, to the previously generated parity-check packets $\mathbf{q}[\cdot]$. The channel packet is given by,

$$\mathbf{x}[i] = (\mathbf{s}[i], \mathbf{q}[i], \mathbf{p}^2[i]). \quad (28)$$

One can see that the rate of the constructed code in (28) is:

$$R = \frac{(T - \Delta + 1)\Delta}{(T - \Delta + 1)(\Delta + B + 1) + (\Delta - B - 1)}, \quad (29)$$

which matches that in Theorem 3.

The decoding analysis is analogous to PRC in Appendix B. We refer the reader to [20] for the details, but omit it in the paper due to page constraints.

We next discuss the field size required for the PRC-MDS codes. Note that in the construction, the source vector must consist of $q_1 = (T - \Delta + 1)\Delta$ symbols. Secondly we construct block MDS codes with parameters $(\Delta, \Delta - B - 1)$ and $(T - \Delta + 2, T - \Delta + 1)$. For these codes to simultaneously exist, each symbol must belong to a field of size greater than $q_2 = \max(\Delta, T - \Delta + 2)$, see, e.g., [21]. Thus, for a given pair B and a delay T , a field size of $q = q_1 \cdot q_2$, i.e., $\mathcal{O}(T^3)$, is sufficient to construct the corresponding PRC-MDS code, as stated in Theorem 3. In contrast, the field size for the m -MDS codes used in III-B can increase exponentially in T as noted in [18].

The PRC-MDS construction exhibits some loss in the achievable rate (23), when compared to (8). This happens because the PRC construction in III-B is able to use the code \mathcal{C}_{12} , which consists of $(\mathbf{v}[i], \mathbf{p}^1[i], \mathbf{p}^2[i])$ to recover the erased symbols $\mathbf{v}[i]$ from the erasure burst. However, the PRC-MDS code with diagonally interleaved block codes only uses the code \mathcal{C}_1 , which consists of $(\mathbf{v}[i], \mathbf{p}^1[i])$ for recovering $\mathbf{v}[i]$ from the burst. Thus the overhead associated with $\mathbf{p}^1[i]$ is higher for the PRC-MDS construction. Nevertheless, as we will see in numerical evaluations, the rate loss compared is small, since the contribution of the $\mathbf{p}^2[i]$ in \mathcal{C}_{12} is generally small.

A. Numerical Comparisons

Fig. 3 compares the rates achieved by different codes as a function of the delay when the erasure burst length $B = 20$ is fixed. The uppermost curve is the capacity of the burst erasure channel when there are no isolated erasures, but the burst length $B' = 21$. The lowermost black curve shows the rate achieved by the m -MDS code. We note that this rate is significantly smaller than the PRC codes. The rate of the PRC-MDS construction only exhibits a negligible loss compared to the PRC codes.

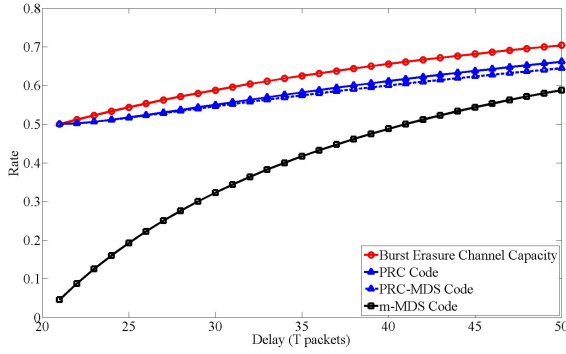


Fig. 3: Achievable rate for PRC and PRC-MDS construction for $B = 20$. We also illustrate the capacity of the burst erasure channel and the rate of the baseline m -MDS code.

	[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
$k^u = 8$	$u_0[0]$	$u_0[1]$	$u_0[2]$	$u_0[3]$	$u_0[4]$	$u_0[5]$	$u_0[6]$	$u_0[7]$	$u_0[8]$
	$u_1[0]$	$u_1[1]$	$u_1[2]$	$u_1[3]$	$u_1[4]$	$u_1[5]$	$u_1[6]$	$u_1[7]$	$u_1[8]$
	$u_2[0]$	$u_2[1]$	$u_2[2]$	$u_2[3]$	$u_2[4]$	$u_2[5]$	$u_2[6]$	$u_2[7]$	$u_2[8]$
	$u_3[0]$	$u_3[1]$	$u_3[2]$	$u_3[3]$	$u_3[4]$	$u_3[5]$	$u_3[6]$	$u_3[7]$	$u_3[8]$
	$u_4[0]$	$u_4[1]$	$u_4[2]$	$u_4[3]$	$u_4[4]$	$u_4[5]$	$u_4[6]$	$u_4[7]$	$u_4[8]$
	$u_5[0]$	$u_5[1]$	$u_5[2]$	$u_5[3]$	$u_5[4]$	$u_5[5]$	$u_5[6]$	$u_5[7]$	$u_5[8]$
	$u_6[0]$	$u_6[1]$	$u_6[2]$	$u_6[3]$	$u_6[4]$	$u_6[5]$	$u_6[6]$	$u_6[7]$	$u_6[8]$
	$u_7[0]$	$u_7[1]$	$u_7[2]$	$u_7[3]$	$u_7[4]$	$u_7[5]$	$u_7[6]$	$u_7[7]$	$u_7[8]$
$k^v = 4$	$v_0[0]$	$v_0[1]$	$v_0[2]$	$v_0[3]$	$v_0[4]$	$v_0[5]$	$v_0[6]$	$v_0[7]$	$v_0[8]$
	$v_1[0]$	$v_1[1]$	$v_1[2]$	$v_1[3]$	$v_1[4]$	$v_1[5]$	$v_1[6]$	$v_1[7]$	$v_1[8]$
	$v_2[0]$	$v_2[1]$	$v_2[2]$	$v_2[3]$	$v_2[4]$	$v_2[5]$	$v_2[6]$	$v_2[7]$	$v_2[8]$
	$v_3[0]$	$v_3[1]$	$v_3[2]$	$v_3[3]$	$v_3[4]$	$v_3[5]$	$v_3[6]$	$v_3[7]$	$v_3[8]$
$k^p = 8$	$p_0^1[0]$	$p_0^1[1]$	$p_0^1[2]$	$p_0^1[3]$	$p_0^1[4]$	$p_0^1[5]$	$p_0^1[6]$	$p_0^1[7]$	$p_0^1[8]$
	$p_1^1[0]$	$p_1^1[1]$	$p_1^1[2]$	$p_1^1[3]$	$p_1^1[4]$	$p_1^1[5]$	$p_1^1[6]$	$p_1^1[7]$	$p_1^1[8]$
	$p_2^1[0]$	$p_2^1[1]$	$p_2^1[2]$	$p_2^1[3]$	$p_2^1[4]$	$p_2^1[5]$	$p_2^1[6]$	$p_2^1[7]$	$p_2^1[8]$
	$p_3^1[0]$	$p_3^1[1]$	$p_3^1[2]$	$p_3^1[3]$	$p_3^1[4]$	$p_3^1[5]$	$p_3^1[6]$	$p_3^1[7]$	$p_3^1[8]$
	$p_4^1[0]$	$p_4^1[1]$	$p_4^1[2]$	$p_4^1[3]$	$p_4^1[4]$	$p_4^1[5]$	$p_4^1[6]$	$p_4^1[7]$	$p_4^1[8]$
	$p_5^1[0]$	$p_5^1[1]$	$p_5^1[2]$	$p_5^1[3]$	$p_5^1[4]$	$p_5^1[5]$	$p_5^1[6]$	$p_5^1[7]$	$p_5^1[8]$
	$p_6^1[0]$	$p_6^1[1]$	$p_6^1[2]$	$p_6^1[3]$	$p_6^1[4]$	$p_6^1[5]$	$p_6^1[6]$	$p_6^1[7]$	$p_6^1[8]$
	$p_7^1[0]$	$p_7^1[1]$	$p_7^1[2]$	$p_7^1[3]$	$p_7^1[4]$	$p_7^1[5]$	$p_7^1[6]$	$p_7^1[7]$	$p_7^1[8]$
$k^s = 2$	$p_0^2[0]$	$p_0^2[1]$	$p_0^2[2]$	$p_0^2[3]$	$p_0^2[4]$	$p_0^2[5]$	$p_0^2[6]$	$p_0^2[7]$	$p_0^2[8]$
	$p_1^2[0]$	$p_1^2[1]$	$p_1^2[2]$	$p_1^2[3]$	$p_1^2[4]$	$p_1^2[5]$	$p_1^2[6]$	$p_1^2[7]$	$p_1^2[8]$
	$p_2^2[0]$	$p_2^2[1]$	$p_2^2[2]$	$p_2^2[3]$	$p_2^2[4]$	$p_2^2[5]$	$p_2^2[6]$	$p_2^2[7]$	$p_2^2[8]$
	$p_3^2[0]$	$p_3^2[1]$	$p_3^2[2]$	$p_3^2[3]$	$p_3^2[4]$	$p_3^2[5]$	$p_3^2[6]$	$p_3^2[7]$	$p_3^2[8]$

TABLE III: A PRC-MDS code construction with $B = 3$ achieving a rate of $R = \frac{12}{22}$ for a delay of $T = 7$. The double and single boxes denote the symbols of the codewords $c_0^1[0]$ and $c_1^2[6]$ respectively while shaded columns denote erased channel packets. We note that each of the parity-check symbols $p_j^1[t]$ is combined with $u_j[t - 6]$ for $j = \{0, 1, \dots, 7\}$ but are omitted for simplicity.

However, we will see in simulations that the packet loss rate for PRC-MDS codes, as well as other codes that use block-MDS codes as constituent codes, is noticeably higher. Thus the smaller field size ultimately comes with a performance penalty.

B. Example - PRC-MDS

We discuss an example of a PRC-MDS code construction designed for $\mathcal{C}(K, N, B, W)$ with parameters, $K = 1$, $N = 1$, $B = 3$, $W = 17$ and a delay constraint of $T = 7$ symbols. We note that these parameters are the same as that in Section III-C. However, the rate achieved using a PRC-MDS code is $R = \frac{12}{22}$ which is slightly smaller.

In our construction, we again select $\Delta = 6$. Also note that the number of source symbols is $(T - \Delta + 1)\Delta = 12$. Thus

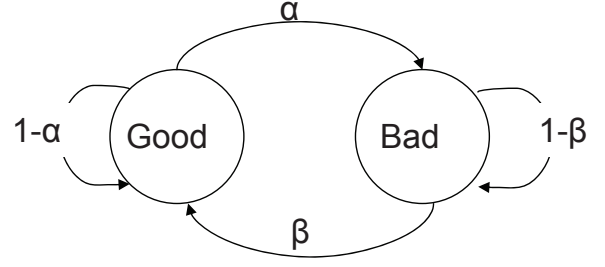


Fig. 4: The Gilbert-Elliott Channel Model

we split $\mathbf{s}[i] \in \mathbb{F}_q^{12}$ into two groups:

$$\mathbf{s}[i] = \underbrace{(u_0[i], \dots, u_7[i])}_{=\mathbf{u}[i]}, \underbrace{(v_0[i], \dots, v_3[i])}_{=\mathbf{v}[i]}. \quad (30)$$

We apply systematic $(\Delta, \Delta - B - 1) = (6, 2)$ MDS block codes with an interleaving factor of $T - \Delta + 1 = 2$ to the $\mathbf{v}[\cdot]$ sub-packets to generate the parity-check packets $\mathbf{p}^1[\cdot] = (p_0^1[\cdot], \dots, p_7^1[\cdot])$. The codewords of the two MDS block codes are:

$$c_0^1[i] = (v_0[i], v_2[i + 1], p_0^1[i + 2], p_2^1[i + 3], p_4^1[i + 4], p_6^1[i + 5])$$

$$c_1^1[i] = (v_1[i], v_3[i + 1], p_1^1[i + 2], p_3^1[i + 3], p_5^1[i + 4], p_7^1[i + 5])$$

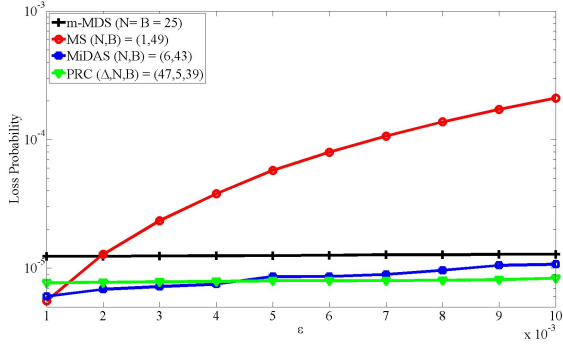
We then combine a shifted version of the $\mathbf{u}[\cdot]$ sub-packets with the $\mathbf{p}^1[\cdot]$ parity-checks, i.e., $\mathbf{q}[t] = \mathbf{p}^1[t] + \mathbf{u}[t - \Delta]$. We additionally apply $(T - \Delta + 2, T - \Delta + 1) = (3, 2)$ MDS block codes with an interleaving factor of $\Delta - B - 1 = 2$ to the $\mathbf{v}[\cdot]$ sub-packets to generate another set of parity-check packets, $\mathbf{p}^2[\cdot] = (p_0^2[\cdot], p_1^2[\cdot])$. The codewords of the two codes is given by,

$$c_0^2[i] = (v_0[i], v_2[i + 1], p_0^2[i + 2]) \quad (31)$$

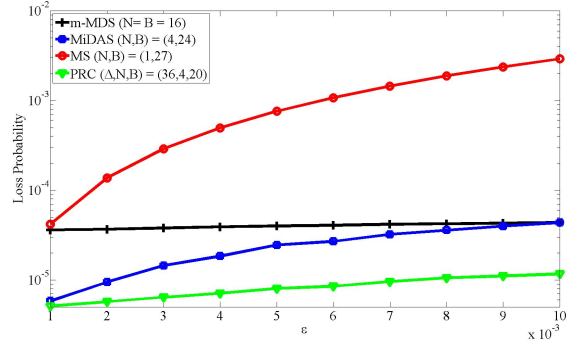
$$c_1^2[i] = (v_1[i], v_3[i + 1], p_1^2[i + 2]) \quad (32)$$

We finally concatenate the two parity-check packets with the source packet to generate the channel packet, i.e., $\mathbf{x}[t] = (\mathbf{s}[t], \mathbf{q}[t], \mathbf{p}^2[t])$. The construction is shown in Table III.

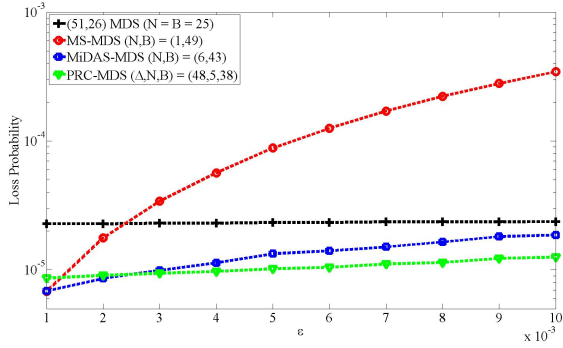
We discuss the decoding steps, for the case when a burst length $B = 3$ spans the interval $[i, i + 2]$ and an isolated erasure happens at time $t_i = i + \Delta = i + 6$. Without loss of generality we assume $i = 0$. The decoder first recovers the parity-check packets $\mathbf{p}^1[t]$ in the interval $t \in [3, 5]$ by subtracting the combined $\mathbf{u}[t - 6]$ since they are not erased. The diagonal codewords $c_0^1[t]$ and $c_1^1[t]$ that span the interval $t \in [0, 2]$ are sufficient to recover erased $\mathbf{v}[\cdot]$ sub-packets by time $t = 5$. It can be verified by inspection that all the erased $\mathbf{v}[t]$ for $t \in [0, 2]$ are recovered by $t = 5$. In a similar fashion, by using the codewords $c_0^2[t]$ and $c_1^2[t]$ for $t \in \{5, 6\}$ it is possible to recover the erased $\mathbf{v}[6]$ by time $t = 8$. By time $t = 8$ all the erased $\mathbf{v}[\cdot]$ sub-packets have been recovered. Hence, the decoder can go back and compute all the parity-check packets $\mathbf{p}^1[\cdot]$ and subtract them to recover the combined $\mathbf{u}[1]$, $\mathbf{u}[2]$ and $\mathbf{u}[6]$ sub-packets at time 8, 9 and 13 respectively. The analysis of other erasure patterns can be done in a similar fashion, but will be omitted due to space constraints.



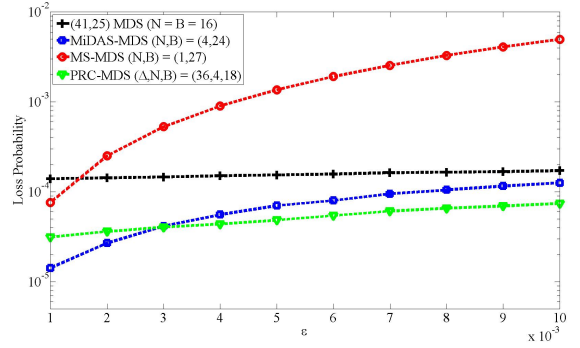
(a) All codes are evaluated using a decoding delay of $T = 50$ symbols and a rate of $R = 50/99 \approx 0.51$ and use m -MDS as the constituent codes.



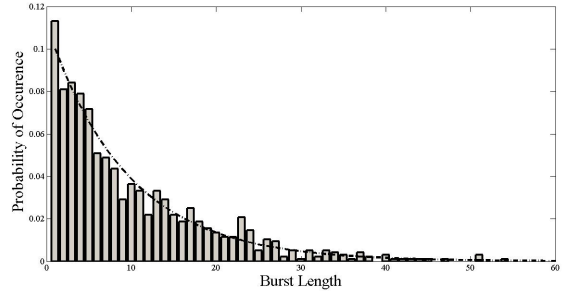
(a) All codes are evaluated using a decoding delay of $T = 40$ symbols and a rate of $R = 40/67 \approx 0.6$ and use m -MDS as the constituent codes.



(b) All codes are evaluated using a decoding delay of $T = 50$ symbols and a rate of $R = 50/99 \approx 0.51$ and use block MDS codes as the constituent codes

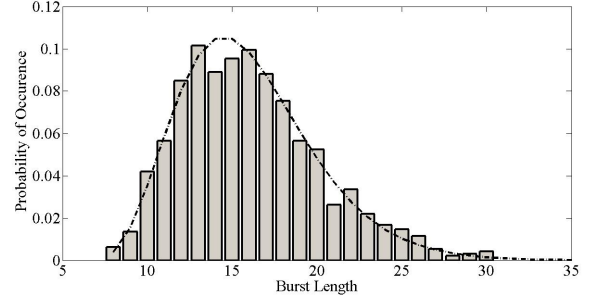


(b) All codes are evaluated using a decoding delay of $T = 40$ symbols and a rate of $R = 40/67 \approx 0.6$ and use block MDS codes as the constituent codes.



(c) Burst Histogram at $\beta = 0.1$ which approximates a geometric distribution (shown dotted) with the same success probability.

Fig. 5: Simulation Experiments for Gilbert-Elliott Channel Model with $(\alpha, \beta) = (10^{-5}, 0.1)$.



(c) Burst Histogram at $\beta = 0.5$ in a $\mathcal{N} + 1 = 9$ -States Fritchman Channel. The distribution follows a negative binomial distribution (shown dotted) of $\mathcal{N} = 8$ failures and a success probability of 0.5.

Fig. 6: Simulation Experiments for Fritchman Channel Model with $(\mathcal{N} + 1, \alpha, \beta) = (9, 2 \times 10^{-5}, 0.5)$.

V. SIMULATION RESULTS

We compare different codes over the Gilbert-Elliott (GE) and Fritchman channels models. As shown in Fig. 4, the GE channel model has two states, bad-state and good-state. The probability that the packet is erased in the bad-state is 1 while that in the good-state is ϵ . In other words, the channel introduces bursts in the bad-state and i.i.d. erasures in the good-state. The transition probability from the good-state to the bad-state is α whereas the transition probability from the bad state to the good state equals β . A Fritchman channel model [15] is similar to a GE channel, but it consists of $\mathcal{N} + 1$ states, one good-state and \mathcal{N} bad-states. The burst length follows a hyper-geometric distribution in the Fritchman

model, but follows a geometric distribution in a GE model. The inter-burst gap distribution in both models is a geometric distribution. Such channels are used to model packet losses in many practical systems (see e.g. [3] and references therein).

The channel and code parameters used in these experiments are given in Table. IV. We also indicate the achievable values of N and B for the different codes considered in the simulations. Note that the delays in the two simulations are taken to be 40 and 50 packets respectively. These choices are motivated by some realistic delay bounds in practical video streaming systems. For example, a delay of 40 packets corresponds to 200 ms in a 2 Mbps video consisting of 1250 Byte packets.

(α, β)	Fig. 5		Fig. 6	
	$(5 \times 10^{-5}, 0.1)$		$(5 \times 10^{-4}, 0.5)$	
Channel Length	10^8		10^8	
Rate R	$50/99 \approx 0.51$		$40/67 \approx 0.6$	
Delay T	50		40	
	N	B	N	B
m -MDS	25	25	16	16
MS	1	49	1	27
MiDAS	6	43	4	24
PRC	5	39	4	20

TABLE IV: Channel & Code Parameters used in Simulations.

In Fig. 5a, 5b, 6a and 6b, we plot the residual packet loss rate for different codes vs the erasure probability ε in the good state whereas Fig. 5c and 6c indicate the burst length histograms for the two channels. Fig. 5a and 6a use codes with m -MDS as constituent codes, whereas Fig. 5b and 6b use block-MDS codes, as discussed in Section IV. The range of ε considered is $[10^{-3}, 10^{-2}]$.

The plots indicated by the black line and marked with '+' correspond to the m -MDS codes in Section II-B. These codes are optimal for the $\mathcal{C}_I(N, W)$ channel and achieve the largest value of N among all codes in Table IV. We note that the performance of these codes is relatively constant as ε is increased. This is because the value of N is sufficiently large that these codes can correct all isolated erasures in the good state. The bottleneck for these codes are long erasure bursts. In particular, note that in Table IV these codes achieve a much smaller value of B and hence incur significant packet losses due to long bursts.

The plots marked with circles, and coloured red, correspond to maximally short (MS) codes proposed in [4], [5]. These codes are optimal for the $\mathcal{C}_{II}(B, W)$ channel and achieve the largest value of B among all codes in Table IV. However, they only achieve $N = 1$ and hence the performance is very sensitive to isolated losses in the good state. In particular, note that as we increase ε the performance in Fig. 5a, 5b, 6a and 6b quickly deteriorates.

The plots marked with squares, and coloured dark blue, correspond to the MiDAS codes in Section II-C. These codes achieve near optimal performance over the $\mathcal{C}_{III}(N, B, W)$ channel and are able to correct both isolated erasures in the good state and longer burst losses in the bad state. However, we notice that in Fig. 5a, 5b, 6a and 6b the performance deteriorates noticeably as ε is increased. This is due to the fact that these codes are sensitive to error events consisting of a burst erasure followed (or preceded) by an isolated erasure as discussed in Section II-D. When the channel transitions from a bad state to a good state, and incurs an isolated erasure in the good state, the entire erasure burst can be lost. For the delays of interest, we see through simulations that such events are significant and must be accounted for.

The plots marked with triangles and coloured green are the PRC codes. While they achieve a slightly smaller value of N and/or B than the MiDAS codes in Table IV, they can correct erasure patterns involving a burst erasure followed (or

preceded) by an isolated erasure. We see that the performance of these codes is relatively constant as ε is increased and is significantly better than MiDAS codes, particularly over the Fritchman channel in Fig. 6a and 6b. As a final point we note that while the codes based on block-MDS codes in Fig. 5b and 6b show similar trend as the codes based on m -MDS in Fig. 5a and 6a, there is a noticeable increase in packet loss rate in the former. Thus the reduction in the field size comes at a penalty in performance loss.

VI. CONCLUSION

In this paper, we study streaming codes that can recover from patterns involving both burst and isolated erasures within the same decoding window. Such patterns arise during state transitions in models such as Gilbert-Elliott or Fritchman channels. Instead of directly analyzing such statistical models, we propose a sliding window erasure channel model to capture such error patterns. We observe that previously proposed streaming codes can lead to significant losses over such channels. We propose a class of Partial Recovery Codes (PRC) which can recover all but one source packet from any pattern of burst plus isolated erasure. Our construction is based on a layering approach where a code for the burst-erasure channel is extended by adding additional layer of parity-checks to correct from isolated erasures. The construction and analysis are non-trivial due to the complexity of the burst plus isolated erasure pattern. We also propose an alternative construction which uses block MDS codes as constituent codes and requires a much reduced field-size. Simulation results over the Gilbert-Elliott and Fritchman channels indicate that the ability to correct burst and isolated losses within the same window leads to significant performance improvements over previously proposed code constructions. We also observe that constructions based on block-MDS codes incur a higher loss rate than the codes based on m -MDS and hence the reduction in the field-size in the former comes with a performance penalty. As part of our future work we are investigating the optimality of these codes by establishing new upper bounds on the capacity of the $\mathcal{C}(K, N, B, W)$ channel.

APPENDIX A PROOF OF LEMMA 1

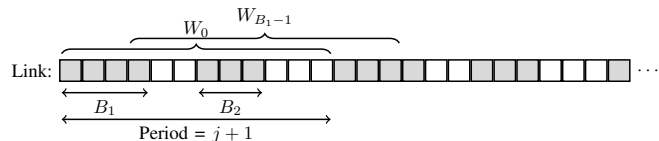


Fig. 7: The periodic erasure channel used in proving P3 in Lemma 1, and indicating the first and last windows of interest, W_0 and W_{B_1-1} , respectively. Grey and white squares denote erased and unerased symbols respectively.

As mentioned before the proof of Property P1 and P2 is already provided in [6]. We present the proof of property P3. Consider a periodic erasure channel of period length $j+1$ as shown in Fig. 7. In each period, the channel introduces an erasure burst of length B_1 packets followed by another burst

of length B_2 packets starting r packets from the start of the first burst and the rest of the period is not erased. In the first period starting at time $t = 0$, the two bursts of length B_1 and B_2 span the intervals $[0, B_1 - 1]$ and $[r, r + B_2 - 1]$ respectively, where $r < \frac{B_1}{1-R}$.

We consider the intervals $\mathcal{W}_i = [i, i + j - 1]$ for $i = \{0, 1, \dots, B_1 - 1\}$. In each such interval, the total number of erasures is no more than $B_1 + B_2 \leq \lfloor (1-R)(j+1) \rfloor$. Thus using property P1 in Lemma 1, we can recover $\mathbf{s}[0], \mathbf{s}[1], \dots, \mathbf{s}[B_1 - 1]$ by time j , since $\mathbf{x}[j+1], \mathbf{x}[j+2], \dots, \mathbf{x}[j+B_1 - 1]$ are erased.

For the second burst, we consider the interval $[r, j]$ of length $j - r + 1$. At this point all the source packets in the first burst have been recovered and only a total of B_2 erasures remain. The property P2 in Lemma 1 can be used for recovering $\mathbf{s}[r], \dots, \mathbf{s}[r + B_2 - 1]$ by time j since,

$$\begin{aligned} B_2 &\leq (1-R)(j+1) - B_1 \\ &\leq (1-R)(j+1) - (1-R)r = (1-R)(j-r+1). \end{aligned} \quad (33)$$

where the second step uses (4). At this point all the erased packets in the first period have been recovered by time j and the claim follows.

APPENDIX B

DECODING ANALYSIS OF PARTIAL RECOVERY CODES

To establish Theorem 2 we divide the erasure patterns into two main categories. In the first case, the erasure burst is followed by an isolated erasure whereas in the second case an isolated erasure precedes the erasure burst.

A. Erasure Burst followed by an Isolated Erasure

Without loss of generality assume that the channel introduces an erasure burst in the interval $[0, B - 1]$ and that the isolated erasure occurs at time $t \geq B$. Since the associated isolated erasure follows the erasure burst from Def. 2 it must occur in the interval $[B, T + B - 1]$. This implies that the interval $[-T, -1]$ is free of any erasure so that there is only one burst and isolated erasure in the interval $[-T, T + B - 1]$, which is of length $2T + B$. Since the memory of the code equals T , any erased packets before $t < -T$ will not affect the decoder. Thus we assume that there are no erasures before $t = 0$.

We further consider two cases as stated below.

1) Burst and Isolated Erasures Recovered Simultaneously:

In this case, the burst and the isolated erasures are close enough such that all the $\mathbf{v}[\cdot]$ sub-packets are recovered simultaneously. This case is illustrated in Table Va. The isolated erasure happens at time t where $B \leq t < \frac{B}{B+1}\Delta$. The recovery of the erased packets proceeds as follows:

- 1) Recover $\{\mathbf{v}[0], \dots, \mathbf{v}[B - 1], \mathbf{v}[t]\}$ at time $\tau = \Delta - 1$ using the $(k^v + k^u + k^r, k^v)$ m -MDS code \mathcal{C}_{12} in the interval $[0, \Delta - 1]$.
- 2) Recover $\{\mathbf{u}[0], \dots, \mathbf{u}[B - 1], \mathbf{u}[t]\}$ at time $\tau = \Delta, \dots, \Delta + B - 1$ and $\tau = t + \Delta$ respectively from the associated parity-checks $\mathbf{q}[\cdot]$.

To justify the recovery of $\mathbf{v}[0], \dots, \mathbf{v}[B - 1], \mathbf{v}[t]$ in the first step we consider the available parity-checks of \mathcal{C}_{12} in the interval $[0, \Delta - 1]$. We first note that the interfering sub-packets $\mathbf{u}[\cdot]$ in this interval are not erased and can be cancelled out from $\mathbf{q}[\cdot]$ to recover the parity-checks $\mathbf{p}^1[\cdot]$. We apply P3 in Lemma 1 with $B_1 = B$, $B_2 = 1$, $r = t$ and $R = R_{12}$ and $j = \Delta - 1$. Note that $t < \frac{B}{B+1}\Delta$ also satisfies $t < \frac{B_1}{1-R_{12}}$ since $R_{12} = \frac{\Delta - B - 1}{\Delta}$ from (12). Thus the first condition in (4) in P3 in Lemma 1 is satisfied. Furthermore note that $(1 - R_{12})(j + 1) = B + 1$ and thus the second condition in (4) in P3 in Lemma 1 is also satisfied. Thus P3 in Lemma 1 applies and the recovery of $\mathbf{v}[0], \dots, \mathbf{v}[B - 1], \mathbf{v}[t]$ at time $\Delta - 1$ follows.

To justify the recovery of $\{\mathbf{u}[0], \dots, \mathbf{u}[B - 1], \mathbf{u}[t]\}$, recall that $\mathbf{q}[i] = \mathbf{u}[i - \Delta] + \mathbf{p}^1[i]$. Since all the $\mathbf{v}[\cdot]$ sub-packets have been recovered in step (1), the associated parity-checks $\mathbf{p}^1[i]$ can be computed and cancelled by the decoder to recover the $\mathbf{u}[\cdot]$ sub-packets as claimed.

As a final remark we note that all the erased packets are recovered for the above erasure pattern.

2) *Burst and Isolated Erasures Recovered Separately:* In this case, there is a sufficiently large gap between the burst and the isolated erasures so that the $\mathbf{v}[\cdot]$ sub-packets of the erasure burst are recovered before the isolated erasure takes place. This case is illustrated in Table Vb. The isolated erasure happens at time $t \geq B \frac{\Delta}{B+1}$. The recovery of the erased packets proceeds as follows:

- 1) Recover $\{\mathbf{v}[0], \dots, \mathbf{v}[B - 1]\}$ by time $\tau = \left\lceil B \frac{\Delta}{B+1} \right\rceil - 1$ using the $(k^v + k^u + k^r, k^v)$ m -MDS code \mathcal{C}_{12} in the interval $[0, \tau]$.
- 2) Recover $\{\mathbf{u}[0], \dots, \mathbf{u}[t - \Delta - 1]\}$ from $\mathbf{q}[\Delta], \dots, \mathbf{q}[t - 1]$ respectively by cancelling the interfering $\mathbf{p}^1[\cdot]$ packets.
- 3) Recover $\mathbf{v}[t]$ by time $\tau = t + T - \Delta + 1$ using the $(k^v + k^r, k^v)$ m -MDS code \mathcal{C}_2 in the interval $[t, t + T - \Delta + 1]$.
- 4) Recover $\{\mathbf{u}[t - \Delta + 1], \dots, \mathbf{u}[B - 1], \mathbf{u}[t]\}$ from $\mathbf{q}[t + 1], \dots, \mathbf{q}[B + \Delta - 1], \mathbf{q}[t + \Delta]$ by cancelling the interfering $\mathbf{p}^1[\cdot]$ packets.

To justify the recovery of $\{\mathbf{v}[0], \dots, \mathbf{v}[B - 1]\}$ in the first step above, we consider the available parity-checks of \mathcal{C}_{12} in the interval $[0, \tau]$. Note that the interfering $\mathbf{u}[\cdot]$ sub-packets in this interval are not erased and can be cancelled out from $\mathbf{q}[\cdot]$ to recover the underlying parity-checks $\mathbf{p}^1[\cdot]$. Furthermore,

$$(1 - R_{12})(\tau + 1) \geq (1 - R_{12}) \left(\frac{B}{B+1} \Delta \right) = B \quad (34)$$

where we substituted (12) for R_{12} . Thus using property P2 in Lemma 1 we recover the $\mathbf{v}[0], \dots, \mathbf{v}[B - 1]$ by time τ as stated. To justify step (2) above note that $\mathbf{q}[i] = \mathbf{u}[i - \Delta] + \mathbf{p}^1[i]$ and the interfering $\mathbf{p}^1[i]$ are only functions of $\mathbf{v}[\cdot]$ sub-packets that have either been recovered in step (1) or are not erased. To justify (3) we consider the interval $[t, t + T - \Delta + 1]$ and consider the parity-checks of \mathcal{C}_2 in this interval. Note that using (13) we have:

$$(1 - R_2)(T - \Delta + 2) \geq 1 \quad (35)$$

holds and hence using property P2 in Lemma 1 we recover $\mathbf{v}[t]$ by time $t + T - \Delta + 1$. To justify step (4), note that once $\mathbf{v}[t]$ is recovered in step (3), the parity-checks $\mathbf{p}^1[t + 1], \dots, \mathbf{p}^1[B + \Delta], \mathbf{p}^1[t + \Delta]$ can be computed and cancelled

	[0]	...	[B-1]	[B]	...	[t]	[t+1]	...	[Δ-1]	[Δ]	...	[Δ+B-1]
u	u[0]	...	u[B-1]	u[B]	...	u[t]	u[t+1]	...	u[Δ-1]	u[Δ]	...	u[Δ+B-1]
v	v[0]	...	v[B-1]	v[B]	...	v[t]	v[t+1]	...	v[Δ-1]	v[Δ]	...	v[Δ+B-1]
u	p¹[0]+	...	p¹[B-1]+	p¹[B]+	...	p¹[t]+	p¹[t+1]+	...	p¹[Δ-1]	p¹[Δ]	...	p¹[Δ+B-1]
s	u[-Δ]	...	u[B-Δ-1]	u[B-Δ]	...	u[t-Δ]	u[t-Δ+1]	...	u[-1]	u[0]	...	u[B-1]
s	p²[0]	...	p²[B-1]	p²[B]	...	p²[t]	p²[t+1]	...	p²[Δ-1]	p²[Δ]	...	p²[Δ+B-1]
	Burst Erasure			Recover $v[0], \dots, v[B-1], v[t]$						$u[0]$...	$u[B-1]$

(a) Burst followed by Isolated Erasure: Burst and Isolated Erasures Recovered Simultaneously.

	[0]	...	[B-1]	...	[τ]	...	[Δ]	...	[t]	...	[t+T-Δ-1]	...	[Δ+B-1]	
u	u[0]	...	u[B-1]	...	u[τ]	...	u[Δ]	...	u[t]	...	u[t+T-Δ-1]	...	u[Δ+B-1]	
v	v[0]	...	v[B-1]	...	v[τ]	...	v[Δ]	...	v[t]	...	v[t+T-Δ-1]	...	v[Δ+B-1]	
u	p¹[0]+	...	p¹[B-1]+	...	p¹[τ]+	...	p¹[Δ]	...	p¹[t]+	...	p¹[t+T-Δ-1]	...	p¹[Δ+B-1]	
s	u[-Δ]	...	u[B-Δ-1]	...	u[τ-Δ]	...	u[0]	...	u[t-Δ]	...	u[t+T-2Δ-1]	...	u[B-1]	
s	p²[0]	...	p²[B-1]	...	p²[τ]	...	p²[Δ]	...	p²[t]	...	p²[t+T-Δ-1]	...	p²[Δ+B-1]	
	Burst Erasure			$v[0], \dots, v[B-1]$		$u[0]$...	Recover $v[t]$				$u[t+T-2Δ+1]$...	$u[B-1]$

(b) Burst followed by Isolated Erasure: Burst and Isolated Erasures Recovered Separately. $\tau = \left\lceil B \frac{\Delta}{B+1} \right\rceil - 1$.

	[0]	...	[t]	...	[t+B-1]	[t+B]	...	[Δ-1]	[Δ]	...	[t+Δ]	...	[t+Δ+B-1]
u	u[0]	...	u[t]	...	u[t+B-1]	u[t+B]	...	u[Δ-1]	u[Δ]	...	u[t+Δ]	...	u[t+Δ+B-1]
v	v[0]	...	v[t]	...	v[t+B-1]	v[t+B]	...	v[Δ-1]	v[Δ]	...	v[t+Δ]	...	v[t+Δ+B-1]
u	p¹[0]+	...	p¹[t]+	...	p¹[t+B-1]+	p¹[t+B]+	...	p¹[Δ-1]	p¹[Δ]	...	p¹[t+Δ]	...	p¹[t+Δ+B-1]
s	u[-Δ]	...	u[t-Δ]	...	u[t-Δ+B-1]	u[t-Δ+B]	...	u[-1]	u[0]	...	u[t]	...	u[t+B-1]
s	p²[0]	...	p²[t]	...	p²[t+B-1]	p²[t+B]	...	p²[Δ-1]	p²[Δ]	...	p²[t+Δ]	...	p²[t+Δ+B-1]
	Recover $v[0], v[t], \dots, v[t+B-1]$						$u[0]$...	$u[t]$...	$u[t+B-1]$		

(c) Isolated Erasure followed by Burst: Isolated and Burst Erasures Recovered Simultaneously.

	[0]	...	[τ]	...	[t]	...	[t+B-1]	...	[Δ]	...	[t+Δ]	...	[t+Δ+B-1]
u	u[0]	...	u[τ]	...	u[t]	...	u[t+B-1]	...	u[Δ]	...	u[t+Δ]	...	u[t+Δ+B-1]
v	v[0]	...	v[τ]	...	v[t]	...	v[t+B-1]	...	v[Δ]	...	v[t+Δ]	...	v[t+Δ+B-1]
u	p¹[0]+	...	p¹[τ]+	...	p¹[t]+	...	p¹[t+B-1]+	...	p¹[Δ]	...	p¹[t+Δ]	...	p¹[t+Δ+B-1]
s	u[-Δ]	...	u[τ-Δ]	...	u[t-Δ]	...	u[t-Δ+B-1]	...	u[0]	...	u[t]	...	u[t+B-1]
s	p²[0]	...	p²[τ]	...	p²[t]	...	p²[t+B-1]	...	p²[Δ]	...	p²[t+Δ]	...	p²[t+Δ+B-1]
	Recover $v[0]$			Burst Erasure				Recover $v[t], \dots, v[t+B-1]$			$u[t]$...	$u[t+B-1]$

(d) Isolated Erasure followed by Burst: Isolated and Burst Erasures Recovered Separately. $\tau = \left\lceil \frac{\Delta}{B+1} \right\rceil - 1$.

TABLE V: The decoding analysis of PRC codes for various erasure patterns in Channel $\mathcal{C}(K=1, N=1, B, W \geq 2T+B)$. The erasures are shaded grey boxes whereas the parity check symbols used to recover the $v[\cdot]$ symbols are marked using bold borders. The $u[\cdot]$ symbols are recovered by subtracting the combined $p^1[\cdot]$ parity checks. For simplicity we let $u = k^u$, $v = k^v$ and $s = k^r$.

from the associated $\mathbf{q}[\cdot]$ packets, and the claim follows.

As a final remark, we note that when $t \in [\Delta, \Delta+B-1]$, the sub-packet $\mathbf{u}[t-\Delta]$ which is erased in the first burst, cannot be recovered as its repeated copy at time t is also erased. This is the only packet that cannot be recovered.

B. Isolated Erasure followed by an Erasure Burst

We assume without loss of generality that the isolated erasure happens at time zero and that the burst erasure happens at time $t > 0$. Since the isolated erasure precedes the erasure burst, it follows that the erasure burst must begin in the interval $t \in [1, T]$ from Def. 2. This implies that there cannot be any erasure in the interval $[-T, -1]$ since the interval $[-T, T+B-1]$ must have only one isolated erasure and one erasure burst. Since the memory of the code equals T , any erased packet before time $t = -T$ will not affect the decoder. Thus in what follows we assume that there are no erasures before time 0.

This class of patterns is sub-divided into two cases discussed below.

1) *Isolated and Burst Erasures Recovered Simultaneously:* In this case the burst erasure and the isolated erasure are close enough so that all the $v[\cdot]$ sub-packets are simultaneously recovered. This case is illustrated in Table Vc. The burst erasure begins at time $t < \frac{\Delta}{B+1}$. The recovery of the erased packets proceeds as follows:

- 1) Recover $\{v[0], v[t], \dots, v[t+B-1]\}$ using the $(k^v + k^u + k^r, k^v)$ m -MDS code \mathcal{C}_{12} in the interval $[0, \Delta-1]$.
- 2) Recover $\{u[0], u[t], \dots, u[t+B-1]\}$ at time $\tau = \Delta, t+\Delta, \dots, t+B+\Delta-1$ respectively from the associated parity-checks $\mathbf{q}[\cdot]$.

To justify step (1) we note that the interfering $u[\cdot]$ sub-packets in $\mathbf{q}[\cdot]$ in the interval $[0, \Delta-1]$ are not erased and can be cancelled to recover $p^1[\cdot]$. We apply P3 in Lemma 1 to code \mathcal{C}_{12} in the interval $[0, \Delta-1]$ using $B_1 = 1$, $B_2 = B$ and $r = t$. Note that by assumption on t and from (12) we

have that

$$r < \frac{\Delta}{B+1} = \frac{1}{1-R_{12}} \quad (36)$$

and thus the first condition in (4) holds. Furthermore from (12) we also have that $(1-R_{12})\Delta = B+1$ and thus the second condition in (4) also holds. Thus P3 in Lemma 1 guarantees the recovery of $\{\mathbf{v}[0], \mathbf{v}[t], \dots, \mathbf{v}[t+B-1]\}$ by time $\tau = \Delta - 1$.

To justify step (2), note that there are no further erasures in the interval $[\Delta, \Delta + t + B - 1]$. Since all the erased $\mathbf{v}[\cdot]$ sub-packets are recovered in step (1), the decoder can compute $\mathbf{p}^1[\Delta], \mathbf{p}^1[i+\Delta], \dots, \mathbf{p}^1[i+B+\Delta-1]$ and subtract them from the corresponding $\mathbf{q}[\cdot]$ packets to recover $\mathbf{u}[0], \mathbf{u}[i], \dots, \mathbf{u}[i+B-1]$, respectively with a delay of $\Delta \leq T$.

As a final remark we note that all the erased packets are fully recovered in this erasure pattern.

2) *Isolated and Burst Erasures Recovered Separately:* In this case, the gap between the isolated erasure and the burst erasure is sufficiently large so that $\mathbf{v}[0]$ is recovered before the burst erasure begins. This case is illustrated in Table Vd. In this case, we have that $t \geq \frac{\Delta}{B+1}$. The recovery of the erased packets proceeds as follows:

- 1) Recover the sub-packet $\mathbf{v}[0]$ by time $\tau = \left\lceil \frac{\Delta}{B+1} \right\rceil - 1$ using the $(k^v + k^u + k^r, k^v)$ m -MDS code \mathcal{C}_{12} in the interval $[0, \tau]$.
- 2) Recover the sub-packets $\mathbf{v}[t], \dots, \mathbf{v}[t+B-1]$ by time $t + \Delta - 1$ using the $(k^v + k^u + k^r, k^v)$ m -MDS code \mathcal{C}_{12} in the interval $[t, t + \Delta - 1]$.
- 3) Recover $\mathbf{u}[t], \dots, \mathbf{u}[t+B-1]$ from $\mathbf{q}[t+\Delta], \dots, \mathbf{q}[t+B+\Delta-1]$ respectively by cancelling the associated $\mathbf{p}^1[\cdot]$ packets.

To justify the above steps note the interfering $\mathbf{u}[\cdot]$ sub-packets in $\mathbf{q}[\cdot]$ for $t \in [0, \Delta-1]$ are not erased and can be cancelled out to recover $\mathbf{p}^1[\cdot]$. In step (1), it suffices to use P1 in Lemma 1 and show that $\mathbf{v}[0]$ is recovered by time $\tau = \left\lceil \frac{\Delta}{B+1} \right\rceil - 1$. Note that

$$(1-R_{12})(\tau+1) \geq (1-R_{12})\frac{\Delta}{B+1} = 1 \quad (37)$$

where we substitute (12) for R_{12} above. Since by assumption on t , $\mathbf{v}[0]$ is the only sub-packet erased in the interval $[0, \tau]$ it follows that $\mathbf{v}[0]$ is recovered by this time.

To justify step (2), consider the interval $[t, t + \Delta - 1]$ and recall that the erasure burst spans $[t, t + B - 1]$. Furthermore even though $\mathbf{v}[0]$ has been recovered in step (1) and its effect can be cancelled out, the sub-packet $\mathbf{u}[0]$ appears in $\mathbf{q}[\Delta]$ and may contribute to one additional erasure when $t \leq \Delta$. In this case, we assume that a total of $B+1$ erasures occur in the above stated interval. We use P3 in Lemma 1 applied to the code \mathcal{C}_{12} with $B_1 = B$ and $B_2 = 1$, in order to show the recovery of $\mathbf{v}[t], \dots, \mathbf{v}[t+B-1]$. Note that the first condition in (4) is satisfied since

$$\Delta - t \leq \Delta \frac{B}{B+1} = \frac{B}{1-R_{12}} \quad (38)$$

is satisfied and the second condition is satisfied as well since $(1-R_{12})\Delta = B+1$. By time $t + \Delta - 1$, the decoder has recovered all the erased $\mathbf{v}[\cdot]$ sub-packets. If instead we had $t > \Delta$ then $\mathbf{u}[0]$ can be recovered at time $t = \Delta$ and there remain only B erasures in the interval $[t, t + \Delta - 1]$, so the recovery of $\mathbf{v}[t], \dots, \mathbf{v}[t+B-1]$ again follows.

Finally to recover the $\mathbf{u}[\cdot]$ sub-packets in the interval $[t, t + \Delta - 1]$, we compute the parity-check packets $\mathbf{p}^1[\cdot]$ in the interval $[t + \Delta, t + B + \Delta - 1]$, subtract them from the corresponding $\mathbf{q}[\cdot]$ packets, and recover $\mathbf{u}[t], \dots, \mathbf{u}[t+B-1]$ respectively as stated in step (4).

As a final remark we note that the sub-packet $\mathbf{u}[0]$ may not be recovered if its repeated copy at time Δ is erased as part of the erasure burst. Thus we may have one unrecovered packet for the above erasure pattern.

This completes the proof of the decoder in Theorem 2.

REFERENCES

- [1] X. Zhang, Y. Xu, H. Hu, Y. Liu, Z. Guo, and Y. Wang, "Modeling and analysis of Skype video calls: Rate control and video quality," *IEEE Transactions on Multimedia*, vol. 15, no. 6, pp. 1446–1457, Oct 2013.
- [2] —, "Profiling Skype video calls: Rate control and video quality," in *Proc. International Conference on Computer Communications (INFOCOM)*, March 2012, pp. 621–629.
- [3] A. Konrad, B. Y. Zhao, A. D. Joseph, and R. Ludwig, "A markov-based channel model algorithm for wireless networks," *Wireless Networks*, vol. 9, no. 3, pp. 189–199, 2003.
- [4] E. Martinian and C. W. Sundberg, "Burst erasure correction codes with low decoding delay," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2494–2502, 2004.
- [5] E. Martinian and M. Trott, "Delay-optimal burst erasure code construction," in *Proc. International Symposium on Information Theory (ISIT)*, Nice, France, July 2007.
- [6] A. Badr, P. Patil, A. Khisti, W. Tan, and J. Apostolopoulos, "Layered constructions for low-delay streaming codes," *CoRR*, vol. abs/1308.3827, 2013.
- [7] O. Tekin, T. Ho, H. Yao, and S. Jaggi, "On erasure correction coding for streaming," in *Information Theory and Applications Workshop (ITA)*, 2012, pp. 221–226.
- [8] D. Leong and T. Ho, "Erasure coding for real-time streaming," in *Proc. International Symposium on Information Theory (ISIT)*, 2012.
- [9] G. Joshi, Y. Kochman, and G. Wornell, "On playback delay in streaming communications," in *Proc. International Symposium on Information Theory (ISIT)*, 2012.
- [10] D. Vasudevan, V. G. Subramanian, and D. J. Leith, "On arq for packet erasure channels with bernoulli arrivals," *International Symposium on Information Theory*, pp. 1793–1797, 2010.
- [11] S. Mehrotra, J. Li, and Y. Huang, "Optimizing FEC transmission strategy for minimizing delay in lossless sequential streaming," *IEEE Transactions on Multimedia*, vol. 13, no. 5, pp. 1066–1076, 2011.
- [12] R. Sukhvasi and B. Hassibi, "Linear error correcting codes with anytime reliability," in *Proc. International Symposium on Information Theory (ISIT)*, 2011, pp. 1748–1752.
- [13] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Systems Technical Journal*, vol. 39, pp. 1253–1265, September 1960.
- [14] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Systems Technical Journal*, vol. 42, pp. 1977–1997, September 1963.
- [15] B. D. Fritchman, "A binary channel characterization using partitioned markov chains," *IEEE Transactions on Information Theory*, vol. 13, pp. 221–227, 1967.
- [16] J. Justesen and L. Hughes, "On maximum-distance-separable convolutional codes," *IEEE Transactions on Information Theory*, vol. 20, no. 2, p. 288, 1974.
- [17] E. M. Gabidulin, "Convolutional codes over large alphabets," in *Proc. International Workshop on Algebraic Combinatorial and Coding Theory*, Varna, Bulgaria, 1988, pp. 80–84.
- [18] H. Gluesing-Luerssen, J. Rosenthal, and R. Smarandache, "Strongly-MDS convolutional codes," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 584–598, 2006.
- [19] A. Badr, A. Khisti, W. Tan, and J. Apostolopoulos, "Robust streaming erasure codes based on deterministic channel approximations," in *Proc. International Symposium on Information Theory (ISIT)*, Istanbul, Turkey, 2013.
- [20] A. Badr, "Error-correcting codes for low-delay streaming communications," Ph.D. dissertation, University of Toronto, 2014.
- [21] R. Blahut, *Algebraic Codes for Data Transmission*. Cambridge, UK: Cambridge University Press, 2003.