An Explicit Rate-Optimal Streaming Code for Channels with Burst and Arbitrary Erasures

Elad Domanovitz, Silas L. Fong and Ashish Khisti

***Abstract*—This paper considers the transmission of an infinite sequence of messages (a streaming source) over a packet erasure channel, where every source message must be recovered perfectly at the destination subject to a fixed decoding delay. While the capacity of a channel that introduces only bursts of erasures has been known for more than fifteen years, only recently, the capacity of a channel with either one burst of erasures or multiple arbitrary erasures in any fixed-sized sliding window has been established. However, explicit codes shown to achieve this capacity for any admissible set of parameters require a large field size that scales exponentially with the delay. Recently, it has been shown that there exist codes that achieve the capacity with field size which scales quadratically with the delay. However, explicit constructions were shown only to specific combinations of parameters. This work describes an explicit rate-optimal construction for all admissible channel and delay parameters over a field size that scales quadratically with the delay.**

***Index Terms*—Streaming codes, Delays, Forward error correction**

I. Introduction

Real-time interactive video streaming became part of the day-to-day life of many people in the world. Unlike traditional traffic, which is not extremely sensitive to latency, real-time interactive video streaming is very sensitive to latency. According to [1], while all IP video traffic will account for 82 percent of the overall traffic by 2022, the portion of real-time interactive video streaming is expected to grow dramatically in the upcoming years.

One of the fundamental requirements of a communication system is to handle interruptions that occur during the transmission of information. Such interruptions occur either due to the physical nature of the channel (for example, fading) or due to packet drops in an interim point of the link (for example, due to congestion or overload). In general, there are two main error control schemes in use: Automatic repeat request (ARQ) and forward error correction (FEC).

When ARQ is used, the receiver acknowledges every received packet (or group of packets). If the transmitter does not receive an acknowledgment (after some predefined timeout), it sends this packet again. An inherent benefit of ARQ (compared to using FEC) is that the packets can be composed only of payload (no additional parity symbols are added). However, when considering low latency applications, it may not be an acceptable solution as its latency is at least three times the one-way delay of the link.

An alternative method for handling errors in the transmission is FEC. By "sacrificing" some throughput in advance, FEC has the potential to lower the recovery latency as it does not require signaling back to the transmitter. Traditionally, FEC is designed to maximize the error-correcting capabilities while ignoring the impact it may have on the latency. Two commonly used codes are Low-density parity-check (LDPC) [2], [3] and digital fountain codes [4], [5]. The typical block length of these codes is very long (usually a few hundreds of symbols), resulting in high latency, which means they are not suitable for real-time interactive video streaming.

Martinian and Sunderberg first presented the challenge of finding optimal low-latency codes in [6]. In their work, a new class of encoders was shown to have the shortest possible decoding delay required to correct all bursts of a given size with fixed redundancy. In [7], a rate-optimal streaming code with a field size that is linear in the delay parameter was introduced for the burst-only case. The construction suggested used diagonally interleaving of judicially chosen block-codes. Leong et al. later showed in [8] that diagonally interleaved codes derived from specific systematic block codes are asymptotically optimal. Adler et al. derived bounds on the average delay of reconstruction in [9], and (different) codes which achieve it were suggested..

While assuming errors will happen in bursts is a realistic assumption for several real-life sources of errors (for example, an overflow of a buffer in any interim point of the transmission), it turns out that the error-correction capability of the burst-optimal codes deteriorates significantly even when a smaller amount of arbitrary erasures is introduced. Therefore, it is desirable that the code could handle both bursts and arbitrary erasures efficiently.

Motivated by this goal, Badr et al., [10] studied codes for a class of channels that introduce both burst and isolated erasures. They further presented the sliding-window burst erasure model and developed an upper bound for a case in which there are either a burst of B erasures or a maximum of N arbitrary erasures in any window of length W. However, an achievable scheme that achieves the upper bound was presented only for $R = 1/2$. Later, [11] extended the analysis over the sliding-window model to variable-size arrivals (where both the message-size and the packet-size can vary).

Recently, two achievable schemes which achieve the upper bound defined in [10] (and thus show it is the capacity of this channel) were presented independently by Fong et al. in [12] and Krishnan et al. in [13] for any rate. The proof

E. Domanovitz and A. Khisti are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (E-mails: elad.domanovitz@utoronto.ca,akhisti@ece.utoronto.ca).

S. L. Fong is with Qualcomm Flarion Technologies, NJ 08807, USA (E-mail: silas.fong@ieee.org).

The material in this paper was presented in part at the 2019 IEEE Information Theory Workshop, Visby, Gotland, Sweden.

in [12] is existential (proving the existence of an appropriate generator matrix), while the field size requirements are large ($O\binom{T}{N}$). Recently, Dudzicz et al. in [14] showed an explicit (and systematic) construction for all rates greater than or equal $1/2$, albeit the required field size is larger than the non-explicit constructions of [12].

In [13], an explicit construction based on linearized polynomials is presented. However, except for a small range of parameters, the field-size requirements are still large (scale exponentially with T). The field size was further addressed in [15], in which a new rate-optimal code construction which covers all channel and delay parameters. The field size of this construction was showed to grow quadratically with the maximal delay constraint ($O(T^2)$). However, explicit constructions were presented only for specific cases. [16] provided explicit constructions with field size which scales linearly with the maximal delay constraint only when $T = N+1 \pmod B$. In this paper, we present an explicit construction that requires a field size which scales quadratically with the maximal delay constraint for any delay T, burst size B, or N arbitrary erasures such the $T \geq B \geq N$.

The rest of this paper is organized as follows. Section I-B provides basic definitions of streaming codes. Section I-C outlines the channel model used in this paper. Section I-D provides basic definitions and properties of block codes that will be used to show that the suggested construction achieves capacity. Section II describes the process of designing the generator matrix of a block code that we later show is used to achieve the streaming capacity, with a field size that scales quadratically with the latency constraint. This construction is based on MDS codes; therefore, some of their properties are recalled in Section II-A. Section II-C provides an example for capacity-achieving streaming code for a channel with $B = 4,\ N = 3$ and a maximal decoding delay of $T = 6$. Section III provides proof that the block code built using the procedure described in Section II can recover from any burst of B erasures or N arbitrary erasures with maximal decoding delay of T. Finally, we note that [15] designed capacity-achieving streaming codes by specifying the parity-check matrix of these codes rather than describing its generator matrix. In section IV, we show that the parity-check matrix of the systematic version of the construction described in Section II meets the set of requirements defined in [15] for a parity-check matrix of a capacity-achieving streaming code.

A. Notation

The set of non-negative integers is denoted by $\mathbb{Z}_+ \cup \{0\}$. The set of k-dimensional row vectors over $\mathbb{F}$ is denoted by $\mathbb{F}^k$, and the set of $k \times n$ matrices over $\mathbb{F}$ is denoted by $\mathbb{F}^{k\times N}$. For any matrix $\mathbf{G}$, we let $\mathbf{G}^t$ and $\mathrm{rank}(\mathbf{G})$ denote respectively the transpose and the rank of G.

We denote by $\mathbf{G}_{\mathcal{A},\mathcal{B}}$ the sub-matrix generated from $\mathbf{G}$ by taking rows with indices in $\mathcal{A}$ and columns with indices in $\mathcal{B}$. We denote by $\mathbf{G}_{r_1:r_2,c_1:c_2}$ the $(r_2 - r_1 + 1 \times c_2 - c_1 + 1)$ sub-matrix generated from taking rows r_1 up to r_2 and columns c_1 up to c_2 from $\mathbf{G}$ where $r_1 \geq 1$ and $c_1 \geq 1$, i.e., we denote the index of the first row and column as "1". We denote by $\mathbf{G}_{:,c_1:c_2}$ the sub-matrix generated from taking columns c_1 up to c_2 from $\mathbf{G}$, and with $\mathbf{G}_{r_1:r_2,:}$ the sub-matrix formed from taking rows r_1 up to r_2 from $\mathbf{G}$. Further, denoting $\mathbf{G}^{-1}_{r_1:r_2,c_1:c_2}$ means that we take the sub-matrix defined above from $\mathbf{G}^{-1}$.

B. Streaming Codes

A source node wants to send a sequence of messages $\{\mathbf{s}_i\}_{i=0}^{\infty}$ to a destination node through an erasure channel. Each $\mathbf{s}_i$ is an element in $\mathbb{F}^k$ where $\mathbb{F}$ is some finite field. Let $\mathbf{y}_i$ be the packet received by the destination at time i for each $i \in \{0, 1, \ldots, i+T\}$ where $\mathbf{y}_i$ equals either $\mathbf{x}_i$ or the erasure symbol "$*$". Every source message has to be recovered perfectly at the destination within a delay constraint of T time slots.

Definition 1 ([17], Sec. II-B)**.** *An $(n, k, T)_{\mathbb{F}}$-streaming code consists of the following:*

1) *A sequence of source messages $\{\mathbf{s}_i\}_{i=0}^{\infty}$ where $\mathbf{s}_i \in \mathbb{F}^k$.*
2) *An encoding function $f_i : \underbrace{\mathbb{F}^k \times \ldots \times \mathbb{F}^k}_{i+1 \text{ times}} \to \mathbb{F}^n$ for each $i \in \mathbb{Z}_+ \cup \{0\}$, where f_i is used by the source at time i to encode $\mathbf{s}_i$ according to*
$$\mathbf{x}_i = f_i\left(\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_i\right).$$
3) *A decoding function $\phi_{i+T} : \underbrace{\mathbb{F}^n \cup \{*\} \times \ldots \times \mathbb{F}^n \cup \{*\}}_{i+T+1 \text{ times}} \to \mathbb{F}^n$ for each $i \in \mathbb{Z}_+ \cup \{0\}$ is used by the destination at time $i+T$ to estimate $\mathbf{s_i}$ according to*
$$\hat{\mathbf{s}}_i = \phi_{i+T}\left(\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_{i+T}\right). \tag{1}$$

Definition 2. *An $(n, k, m, T)_{\mathbb{F}}$-convolutional code is an $(n, k, T)_{\mathbb{F}}$-streaming code constructed as follows:*

1) *Let $\mathbf{G}_0^{\mathrm{conv}}, \mathbf{G}_1^{\mathrm{conv}}, \ldots, \mathbf{G}_m^{\mathrm{conv}}$ be $m+1$ generator matrices in $\mathbb{F}^{k\times n}$.*
2) *Then, for each $i \in \mathbb{Z}_+ \cup \{0\}$*
$$\mathbf{x}_i = \sum_{l=0}^{m} \mathbf{s}_{i-l}\mathbf{G}_l^{\mathrm{conv}} \tag{2}$$
where $\mathbf{s}_{-1} = \mathbf{s}_{-2} = \ldots = \mathbf{s}_{-m} = \mathbf{0}^{1\times k}$ by convention.

C. Channel Model

The channel model considered in this work is the sliding-window burst erasure channel that was introduced by Badr et al. in [10]. This model introduces up to B consecutive erasures, or N arbitrarily positioned isolated erasures in any window of size W among the sequence of transmitted packets $\mathbf{x}[t]$.

Definition 3. *An erasure sequence is a binary sequence denoted by $e^{\infty} \triangleq \{e_i\}_{i=0}^{\infty}$ where $e_i = \mathbf{1}\{\text{erasure occurs at time i}\}$.*

Definition 4. *The mapping $g_n : \mathbb{F}^n \times \{0, 1\} \to \mathbb{F}^n \cup \{*\}$ of an erasure channel is defined as*
$$g_n(\mathbf{x}, e) = \begin{cases} \mathbf{x} & \text{if } e = 0, \\ * & \text{if } e = 1. \end{cases} \tag{3}$$

Definition 5. *A (W,B,N)-erasure sequence is an erasure sequence e^∞ that satisfies the following: For each $i \in \mathbb{Z}_+ \cup \{0\}$ and any window*

$$\mathcal{W}_i \triangleq \{i, i+1, \ldots, i+W-1\}, \tag{4}$$

either $N < \sum_{l\in\mathcal{W}_i} e_l \leq B$ holds with all the 1's in $(e_i, e_{i+1}, \ldots, e_{i+W-1})$ are located at consecutive positions or $\sum_{l\in\mathcal{W}_i} e_l \leq N$ hold with no restrictions on the locations of 1's.

In other words, a (W,B,N)-erasure sequence introduces either one burst erasure with length no longer than B or multiple arbitrary erasures with a total count no larger than N in any window $\mathcal{W}_i$, $\forall i \in \mathbb{Z}_+ \cup \{0\}$. The set of (W,B,N)-erasure sequences is denoted by $\Omega^\infty(W,B,N)$.

We further assume that $W \geq T+1$.[1] Thus, we can assume without loss of generality that

$$W > T \geq B \geq N \geq 1. \tag{5}$$

For further details, refer to Section I-B of [12].

The next definition defines a streaming code which is (W,B,N)-achievable.

Definition 6. *An $(n,k,T)F$-streaming code is said to be (W,B,N)-achievable if the following holds for any (W,B,N)-erasure sequence $e^\infty \in \Omega^\infty(W,B,N)$: For all $i \in \mathbb{Z}_+ \cup \{0\}$ and all $\mathbf{s}_i \in \mathbb{F}^k$, we have*

$$\hat{\mathbf{s}}_i = \mathbf{s}_i \tag{6}$$

where

$$\begin{aligned}\hat{\mathbf{s}}_i &= \phi_{i+T}\left(\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_{i+T}\right) \\ &= \phi_{i+T}\left(g_n(\mathbf{x}_0, e_0), \ldots, g_n(\mathbf{x}_{i+T}, e_{i+T})\right)\end{aligned} \tag{7}$$

Definition 7. *Fix any (W,T,B,N) that satisfies* (5)*. The (W,T,B,N)-capacity, denoted by $C_{W,T,B,N}$, is the supremum of the rates attained by $(n,k,T)_\mathbb{F}$-streaming codes that are (W,B,N)-achievable, i.e.,*

$$C_{W,T,B,N} \triangleq \sup\Big\{\frac{k}{n} \;\Big|\; \text{There exists an } (W,B,N)- \tag{8}$$

$$\text{achievable } (n,k,T)_\mathbb{F}-\text{streaming} \tag{9}$$

$$\text{code for some } \mathbb{F}\Big\} \tag{10}$$

Recently, independent works in [12], [13] established that the capacity of $\mathcal{C}(W,B,N)$ is given by:

$$C_{W,T,B,N} = \frac{T-N+1}{T-N+B+1}. \tag{11}$$

D. Block Codes

As we show next, linear block codes serve as a building block for capacity-achieving streaming codes. Therefore, we recall several definitions and properties of these codes that will be of use in the following Sections.

[1]In case where $B < W \leq T+1$ we can achieve the capacity by reducing the effective delay to $T_{\text{eff}} = W-1$ as discussed in [17]. Furthermore, the capacity is trivially zero if $W \leq B$ as an erasure sequence that erases all the channel packets becomes admissible.

Let $y[i]$ be the symbol received by the destination at time i for each $i \in \{0,1,\ldots,n-1\}$.

Definition 8. *A point-to-point $(n,k)_\mathbb{F}$-block code consists of the following*

1) *A sequence of k symbols $\{u[l]\}_{l=0}^{k-1}$ where $u[l] \in \mathbb{F}$.*
2) *A generator matrix $\mathbf{G} \in \mathbb{F}^{k\times n}$. The source codeword is generated according to*

$$[x[0]\ x[1]\ \ldots\ x[n-1]] = [u[0]\ u[1]\ \ldots\ u[k-1]]\,\mathbf{G} \tag{12}$$

3) *A decoding function $\varphi_{l+n} : \mathbb{F}\cup\{*\} \times \ldots \mathbb{F}\cup\{*\} \to \mathbb{F}$ for each $l \in \{0,1,\ldots,k-1\}$, where φ_{l+n} is used by the destination at time to estimate $u[l]$ according to*

$$\hat{u}[l] = \varphi_{l+n}(y[0], y[1], \ldots, y[n-1]) \tag{13}$$

When the code is systematic, the generator matrix associated with it can be expressed in the form of

$$\mathbf{G} = \left[\mathbf{I}_{k\times k} \;\middle|\; \mathbf{P}_{k\times(n-k)}\right] \tag{14}$$

Definition 9. Punctured Code

Let $\mathcal{C}$ be an $(n,k)_\mathbb{F}$ linear code. Given a subset $\mathcal{P}$ of $[0 : n-1]$, the code $\mathcal{C}$ punctured on the coordinates in $\mathcal{P}$, is the linear code of length $(n-|\mathcal{P}|)$ obtained from $\mathcal{C}$ by deleting all the coordinates in $\mathcal{P}$. We denote the punctured code as $\mathcal{C}\,\big|_{\mathcal{P}}$.

Definition 10. Shortened Code

Let $\mathcal{C}$ be an $(n,k)_\mathbb{F}$ linear code. Given a subset $\mathcal{P}$ of $[0 : n-1]$, Consider the set $\mathcal{C}(\mathcal{P})$ of codewords which are $\mathbf{0}$ on $\mathcal{P}$; this set is a subcode of $\mathcal{C}$. Puncturing $\mathcal{C}(\mathcal{P})$ on $\mathcal{P}$ gives a code over $\mathbb{F}$ of length $n-|\mathcal{P}|$ called the code shortened on $\mathcal{P}$ and denoted as $\mathcal{C}^\mathcal{P}$.

Corollary 1. *For a systematic code with generator matrix $\mathbf{G}$, shortening on any coordinate $i \in \{0,\ldots,k-1\}$ results in a systematic code with generator matrix which results from erasing the ith row and the ith column in $\mathbf{G}$.*

Definition 11. Parity-check matrix *A parity-check matrix, $\mathbf{H}$ of a $(n,k)_\mathbb{F}$ linear code $\mathcal{C}$, is a generator matrix of the dual code, $\mathcal{C}^\perp$. This means that a codeword $\mathbf{c}$ is in $\mathcal{C}$ if and only if the matrix-vector product $\mathbf{H}\mathbf{c}^T = 0$. When $\mathbf{G}$ is systematic the parity-check matrix can be expressed as*

$$\mathbf{H} = \left[-\mathbf{P}^T_{k\times(n-k)} \;\middle|\; \mathbf{I}_{(n-k)\times(n-k)}\right]. \tag{15}$$

We recall the following Theorem and Lemma with respect to the parity-check matrix.

Theorem 1 ([18], Theorem 1.5.7)**.** *Let $\mathcal{P} \in [0 : n-1]$. Then*

$$(\mathcal{C}\,\big|_{\mathcal{P}})^\perp = (\mathcal{C}^\perp)^\mathcal{P} \tag{16}$$

and

$$(\mathcal{C}^\mathcal{P})^\perp = (\mathcal{C}^\perp)\,\big|_{\mathcal{P}}. \tag{17}$$

Alternatively, this Theorem states that puncturing the generator matrix of $\mathcal{C}$ is equivalent to shortening its parity-check matrix and that shortening the generator matrix of $\mathcal{C}$ is equivalent to puncturing its parity-check matrix.

In the sequel, we use the following Lemma to derive a set of conditions on the parity-check matrix.

Lemma 1 ([15], Lemma II.6)**.** *Consider an $(n,k)_{\mathbb{F}}$ block code $\mathcal{C}$ having a parity-check matrix $\mathbf{H}$. Let the subset $\mathcal{P}$ of $[0:n-1]$ be erased from $\mathcal{C}$ and let $i \in \mathcal{P}$. Then the i-th code symbol in a codeword can be recovered from the code symbols of the same codeword corresponding to coordinates in $\mathcal{P}^c$ if:*

$$\mathbf{h}_i \notin \operatorname{span}\left\langle \{\mathbf{h}_j\}_{j\in\mathcal{P}\setminus\{i\}} \right\rangle$$

where $\mathbf{h}_j$ denoted the j-th column of $\mathbf{H}$.

In order to generate capacity-achieving streaming codes, a subclass of linear block codes is required, which are block codes that conform to a (stricter) decoding constraint.

Definition 12. *A point-to-point $(n,k,T)_{\mathbb{F}}$-block code consists of the following*

1) *A sequence of k symbols $\{u[l]\}_{l=0}^{k-1}$ where $u[l] \in \mathbb{F}$.*
2) *A generator matrix $\mathbf{G} \in \mathbb{F}^{k\times n}$. The source codeword is generated according to*

$$[x[0]\ x[1]\ \dots\ x[n-1]] = [u[0]\ u[1]\ \dots\ u[k-1]]\,\mathbf{G} \tag{18}$$

3) *A decoding function $\varphi_{l+T} : \mathbb{F}\cup\{*\}\times\dots\mathbb{F}\cup\{*\}\to\mathbb{F}$ for each $l\in\{0,1,\dots,k-1\}$, where φ_{l+T} is used by the destination at time $\min\{l+T,n-1\}$ to estimate $u[l]$ according to*

$$\hat{u}[l] = \begin{cases} \varphi_{l+T}(y[0],y[1],\dots,y[l+T]) & \text{if } l+T\le n-1 \\ \varphi_{l+T}(y[0],y[1],\dots,y[n-1]) & \text{if } l+T > n-1 \end{cases} \tag{19}$$

Definition 13. *An $(n,k,T)_{\mathbb{F}}$-block code is said to be (W,B,N)-achievable if the following holds for any (W,B,N)-erasure sequence $e^{\infty}\in\Omega^{\infty}(W,B,N)$: Let*

$$y[i] \triangleq g_1(x[i],e_i) \tag{20}$$

be the symbol received by the destination at time i for each $i\in\{0,1,\dots,n-1\}$ where g_1 is as defined in (3)*. For the $(n,k,T)_{\mathbb{F}}$-block code, we have*

$$\hat{u}[i] = u[i] \tag{21}$$

for all $i\in\{0,1,\dots,k-1\}$ and all $u[i]\in\mathbb{F}$ where $\hat{u}$ is constructed according to (19) *and* (20)*.*

E. Generating streaming $(n,k,T)_{\mathbb{F}}$-streaming code which is (W,B,N)-achievable from $(n,k,T)_{\mathbb{F}}$-block code which is (W,B,N)-achievable

The following Lemma shows that given an $(n,k,T)_{\mathbb{F}}$-block code which is (W,B,N)-achievable, a corresponding $(n,k,n-1,T)_{\mathbb{F}}$-convolutional code (which is an $(n,k,T)_{\mathbb{F}}$-streaming code) which is (W,B,N)-achievable can be constructed.

Lemma 2 (Lemma 1 in [12])**.** *Given an $(n,k,T)_{\mathbb{F}}$-block code which is (W,B,N)-achievable, we can construct an $(n,k,n-1,T)_{\mathbb{F}}$ convolutional code which is (W,B,N)-achievable. More specifically, given that $\mathbf{G}$ is the generator matrix of the $(n,k,T)_{\mathbb{F}}$-block code where $g_{i,j}$ is the entry situated in row i and column j of $\mathbf{G}$, we can construct the $n-1$ generator matrices of the $(n,k,n-1,T)_{\mathbb{F}}$-convolutional code as follows. For each $l\in\{0,1,\dots,n-1\}$, construct*

- *If $0\le l\le n-k$*

$$\mathbf{G}_l^{\text{conv}} \triangleq \left[\ \mathbf{0}^{k\times l}\ \ \operatorname{diag}\left(g_{0,l}, g_{1,l+1},\dots,g_{k-1,l+k-1}\right)\ \ \mathbf{0}^{k\times(n-k-l)}\ \right]$$

- *If $n-k\le l\le n-1$*

$$\mathbf{G}_l^{\text{conv}} \triangleq \begin{bmatrix} \mathbf{0}^{k\times l} & \operatorname{diag}\left(g_{0,l}, g_{1,l+1},\dots,g_{n-1,l+n-1}\right) \\ & \mathbf{0}^{k-n+l\times(n-l)} \end{bmatrix}$$

We note that $\mathbf{G} = \sum_{l=1}^{n-1}\mathbf{G}_l^{\text{conv}}$. In particular, if we let $\mathbf{s}_i \triangleq [s_i[0]\ s_i[1]\ \cdots\ s_i[k-1]]$ and let

$$\begin{aligned}&[x_i[0]\ x_{i+1}[1]\ \cdots\ x_{i+n-1}[n-1]] \triangleq \\ &[s_i[0]\ s_{i+1}[1]\ \cdots\ s_{i+k-1}[k-1]]\,\mathbf{G},\end{aligned} \tag{22}$$

i.e., we apply diagonal interleaving for all $i\in\mathbb{Z}_+\cup\{0\}$, then the symbols generated at time i by the $(n,k,n-1,T)_{\mathbb{F}}$ convolutional code (which is an $(n,k,T)_{\mathbb{F}}$-streaming code) are

$$\mathbf{x}_i \triangleq [x_i[0]\ x_i[1]\ \cdots\ x_i[n-1]]\,. \tag{23}$$

F. Main result

The following Theorem is proved in Section III.

Theorem 2. *For any (W,T,B,N) that meets* (5)*, there exists an explicit $(n,k,T)_{\mathbb{F}}$-block code with*

$$\begin{aligned} R &= \frac{k}{n} \\ &= \frac{T-N+1}{T+B-N+1} \\ &= C_{W,T,B,N} \end{aligned} \tag{24}$$

which is (W,B,N)-achievable where the field size scales quadratically with the delay constraint ($O(T^2)$).

Recalling Lemma 2, it follows that using an $(n,k,T)_{\mathbb{F}}$-block code which is (W,B,N)-achievable with field size that scales quadratically with the delay constraint ($O(T^2)$), an $(n,k,T)_{\mathbb{F}}$-streaming code which is (W,B,N)-achievable can be generated with a field size that scales quadratically with the delay constraint ($O(T^2)$).

Thus, proving Theorem 2 means that for any (W,T,B,N) that meets (5), there exists an $(n,k,T)_{\mathbb{F}}$-streaming code which is (W,B,N)-achievable can be generated where the field size scales quadratically with the delay constraint ($O(T^2)$). Thus, for any T,B,N, which meet (5), the capacity $C_{W,T,B,N}$ can be achieved with field size that scales quadratically with the delay.

II. Explicit Construction of capacity-achieving streaming code

We start by recalling some standard definitions of MDS codes. We then present an explicit method of generating a generator matrix of a (W,B,N)-achievable block code with field size which scales quadratically with the delay constraint

which is based on MDS codes and concludes that it can be used to generate a (W, B, N)-achievable streaming code with the same field size.

A. MDS codes

We show next that a building block in generating capacity-achieving streaming code is a maximum distance separable (MDS) block code. Linear block codes that achieve equality in the Singleton bound are called MDS codes. A $k \times n$ matrix $\mathbf{G}$ over a finite field $\mathbb{F}_q$, with $k \leq n$, will be referred to as a generator matrix of an MDS matrix if any k distinct columns of $\mathbf{G}$ form a linearly independent set.

We recall the following properties with respect to the generator matrix of MDS code.

Property 1 ([19])**.** *$\mathbf{G}$ generates a (systemtic) $(n, k)_{\mathbb{F}}$ MDS code if and only if every square submatrix of $\mathbf{P}_{k\times(n-k)}$ (defined in* (14)*) is non-singular, i.e. each square submatrix is invertible and has full rank.*

Corollary 2. *We note that this Property means that any rectangular submatrix of $\mathbf{P}_{k\times(n-k)}$ also has a full rank which equals the minimal dimension.*

Another Corollary we use is the following.

Corollary 3. *Define $\mathcal{A} \subset \{1, \ldots, k\}$ and $\mathcal{B} \subset \{1, \ldots, B\}$. The sub matrix $\mathbf{P}_{\mathcal{A},\mathcal{B}}$ which is generated from taking rows with indices which match the values in $\mathcal{A}$ and columns with indices which match the values in $\mathcal{B}$ is non-singular. Thus,*

$$\operatorname{rank}(\mathbf{P}_{\mathcal{A},\mathcal{B}}) = \min\left(|\mathcal{A}|, |\mathcal{B}|\right). \tag{25}$$

This corollary holds since performing column swap or row swap on the parity matrix of an MDS code results in an MDS code. We further note that

Corollary 4. *From Lemma 1 and Property 1 it follows that any set of $n-k$ column vectors from the parity matrix $\mathbf{H}$ of an (n, k) MDS code are linearly independent.*

Next, we state the following corollaries, which outline the outcome of puncturing and shortening an MDS code.

Corollary 5. *When $\mathcal{C}$ is an $(n, k)_{\mathbb{F}}$ MDS code, puncturing it on $\mathcal{P}$ results in an $(n - |\mathcal{P}|, k)_{\mathbb{F}}$ MDS code.*

Corollary 6. *When $\mathcal{C}$ is an $(n, k)_{\mathbb{F}}$ MDS code, shortening it on $\mathcal{P}$ results in an $(n - |\mathcal{P}|, k - |\mathcal{P}|)_{\mathbb{F}}$ MDS code.*

We conclude with the following Proposition on the minimal size required to generate an MDS code.

Proposition 1 (Main conjecture on MDS codes([20]))**.** *The maximal length n of a MDS code of dimension k over $\mathbb{F}_q$ is*

$$n \leq \begin{cases} q+2 & q \text{ even and } k = 3 \text{ or } k = q-1 \\ q+1 & \text{otherwise} \end{cases}$$

This conjecture means that for any k and n ($n \geq k$) there exists an MDS codes with field size which scales as $O(n)$.

B. Explicit construction of (W, B, N)-achievable $(n, k, T)_{\mathbb{F}}$-block code

Next we describe how to generate (an explicit) $\mathbf{G}$ which is the generator matrix of $(n, k, T)_{\mathbb{F}}$-block code with a field size which scales quadratically with the delay T. In Section III we show that this code is (W, B, N)-achievable for any (W, T, B, N) which meet (5).

We define

$$\begin{aligned} k &= T - N + 1 \\ n &= k + B \end{aligned} \tag{26}$$

in the same manner as was defined in [12]. The construction of the generator matrix is done in the following steps:

1) We start with an $(n, k)_{\mathbb{F}_q}$ MDS code $\mathcal{C}''$ which has the following generator matrix[2]

$$\mathbf{G}'' = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & \multicolumn{6}{c}{\multirow{6}{*}{$\mathbf{P}''_{k\times B}$}} \\ 0 & 1 & 0 & 0 & \cdots & 0 & & & & & & \\ 0 & 0 & \ddots & 0 & \cdots & 0 & & & & & & \\ \vdots & \vdots & & 1 & & \vdots & & & & & & \\ \vdots & \vdots & & & \ddots & 0 & & & & & & \\ 0 & 0 & \cdots & \cdots & 0 & 1 & & & & & & \end{bmatrix}. \tag{27}$$

   First we note that following Property 1, the minimal field size for this code scales linearly with the delay T, i.e. $q = O(T)$. We further note that following Corollary 2, each submatrix of matrix $\mathbf{P}''$ has a full rank.
2) We generate code $\mathcal{C}'$ with the generator matrix

$$\mathbf{G}' = \mathbf{M}\mathbf{G}'' = \begin{bmatrix} 1 & X & \cdots & X & 0 & 0 & 0 & \cdots & 0 & X & \cdots & X \\ 0 & 1 & X & \cdots & X & 0 & 0 & \cdots & 0 & X & \cdots & \vdots \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & 0 & \cdots & 0 & X & \cdots & X \\ \vdots & \vdots & & 1 & \ddots & \ddots & \ddots & & \vdots & X & \cdots & X \\ \vdots & \vdots & & & \ddots & X & \cdots & X & 0 & \vdots & & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & 1 & X & \cdots & X & X & \cdots & X \\ \underbrace{}_{k} & & & & & & \underbrace{}_{N-1} & & & \underbrace{}_{B-N+1} & & \end{bmatrix}, \tag{28}$$

[2]We note that Y is not a constant element, but rather represent (potentially different) element taken from $\mathbb{F}_q$. See Section II-C for a specific example.

where matrix $\mathbf{M}$ is an upper triangular matrix which is denoted as

$$\mathbf{M} = \begin{bmatrix} 1 & \overbrace{Y' & \cdots & Y'}^{N-1} & 0 & \cdots & \cdots & 0 \\ 0 & 1 & Y' & \cdots & Y' & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & Y' & \cdots & Y' & 0 \\ 0 & \cdots & 0 & 0 & 1 & Y' & \cdots & Y' \\ 0 & \cdots & 0 & 0 & 0 & \ddots & \ddots & Y' \\ 0 & \cdots & 0 & 0 & 0 & 0 & 1 & Y' \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (29)$$

where $Y' = a \cdot Y$, $a \in \mathbb{F}_q$ denotes a function of one of the (placeholder) Y symbols used in (27).[3]

The goal is to "spread" $N-1$ parity symbols diagonally with the information symbols. As it is easy to see that code $\mathcal{C}'$ (and also $\mathcal{C}''$) can recover from a burst of size B starting at time 0 only at time $T' = k+B-1$. It follows that if $B > N$ we have $T' > T$. Therefore, this code does not meet the required maximal delay constraint. Nevertheless, as we show next, it is an important interim step.

Matrix $\mathbf{M}$ is an upper-triangular matrix. Thus, it is a full-rank matrix, and hence it is invertible. Since an erasure of any l columns in $\mathbf{G}'$ can be translated to erasure of l columns in $\mathbf{G}''$ (by multiplying with the inverse of $\mathbf{M}$), the following property holds.

Property 2. *Block code $\mathcal{C}'$ with generator matrix $\mathbf{G}'$ is an $(n,k)_{\mathbb{F}_q}$ MDS code.*

3) Finally, denoting with $\mathbb{F}_{q^2}$ the extension field of $\mathbb{F}_q$, we replace the $(B-N+1)\times(B-N+1)$ upper right matrix with $\alpha \cdot \mathbf{I}_{B-N+1}$ where $\alpha \in \mathbb{F}_{q^2} \setminus \mathbb{F}_q$ which generates the code $\mathcal{C}$ with the generator matrix

$$\mathbf{G} = \begin{bmatrix} 1 & X & \cdots & X & X & 0 & \cdots & 0 & \cdots & 0 & \alpha & \cdots & 0 \\ 0 & 1 & X & \cdots & X & X & \cdots & 0 & \cdots & 0 & 0 & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & \cdots & 0 & 0 & \cdots & \alpha \\ \vdots & \vdots & & 1 & \ddots & \ddots & \ddots & \ddots & & \vdots & X & \cdots & X \\ \vdots & \vdots & & & \ddots & X & \cdots & X & X & 0 & \vdots & & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & 1 & X & \cdots & X & X & X & \cdots & X \end{bmatrix}. \quad (30)$$

$$\underbrace{\qquad\qquad}_{k} \quad \underbrace{\qquad\qquad\qquad}_{N-1} \quad \underbrace{\qquad\qquad}_{B-N+1}$$

We note that this operation increased the field size of the code, thus the minimal required field size now scales as $O(T^2)$.

The generator matrix $\mathbf{G}$ is composed of the following three blocks

- $\mathbf{G}_1$ - The upper left $k \times (k+N-1)$ sub-matrix of $\mathbf{G}$.
- $\mathbf{G}_2$ - The lower right $(k-(B-N+1)) \times (n-(B-N+1))$ sub-matrix of $\mathbf{G}$.
- $\mathbf{G}_3$ - The upper right $(B-N+1)\times(B-N+1)$ sub-matrix of $\mathbf{G}$.

These blocks are depicted in (31) below.

$$\mathbf{G} = \begin{bmatrix} 1 & X & \cdots & X & X & 0 & \cdots & 0 & \cdots & 0 & \alpha & \cdots & 0 \\ 0 & 1 & X & \cdots & X & X & \cdots & 0 & \cdots & 0 & 0 & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & \cdots & 0 & 0 & \cdots & \alpha \\ \vdots & \vdots & & 1 & \ddots & \ddots & \ddots & \ddots & & \vdots & X & \cdots & X \\ \vdots & \vdots & & & \ddots & X & \cdots & X & X & 0 & \vdots & & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & 1 & X & \cdots & X & X & X & \cdots & X \end{bmatrix}. \quad (31)$$

$\mathbf{G}_1$ $\mathbf{G}_3$ $\mathbf{G}_2$

Before showing that generator matrix $\mathbf{G}$ generates an $(n,k,T)_{\mathbb{F}_{q^2}}$-block code (where $q = O(T)$) which is (W,B,N)-achievable, we denote the following properties of $\mathbf{G}_1$ and $\mathbf{G}_2$.

Property 3. *Block $\mathbf{G}_1$ is a generator matrix of a $(k+N-1,k)_{\mathbb{F}_q}$ MDS code.*

This can be viewed from Corollary 5 and noting that $\mathbf{G}_1$ is generated by of puncturing $(B-N+1)$ columns from G' which is a generator matrix of (n,k) MDS code.

Property 4. *Block $\mathbf{G}_2$ is a generator matrix of an $(n-(B-N+1),k-(B-N+1))_{\mathbb{F}_q}$ MDS code.*

This property holds since it can be assumed that $\mathbf{G}_2$ is generated from $\mathbf{G}'$ by assuming that the first $(B-N+1)$ symbols encodes using $\mathbf{G}'$ ($\{x[0],\ldots,x[B-N]\}$) were received without errors (equivalently assume that $\{u[0],\ldots,u[B-N]\} = \{0,\ldots,0\}$). Even though $\mathbf{G}'$ is not systematic, due to its structure receiving $\{x[0],\ldots,x[B-N]\}$ without errors means that $\{u[0],\ldots,u[B-N]\}$ can be decoded correctly and hence can be cancelled from the other received symbols. Therefore, the remaining code is an $(n-(B-N+1),k-(B-N+1))$ MDS code over $\mathbb{F}_q$.

Following Properties 3 and 4 we denote the codes induced by $\mathbf{G}_1$ and $\mathbf{G}_2$ as MDS_1 and MDS_2.

Remark 1. *Using (26) it follows that MDS_2 is a $(T,T-B)_{\mathbb{F}_q}$ MDS code.*

Remark 2. *In case $T > B$, symbols $\{x[B],\ldots,x[T-1]\}$ are composed only from $\{u[B-N+1],\ldots,u[k-1]\}$, i.e., these are the only symbols of MDS_2 without contribution from $\{u[0],\ldots,u[B-N]\}$ (which are not information symbols of MDS_2).*

This can be viewed from noting that the last $B-N+1$ columns of $\mathbf{G}$ carry information on $\{u[0],\ldots,u[B-N]\}$ (due to $\mathbf{G}_3$). Further, Since $\mathbf{G}_1$ is upper triangular (with N diagonals), the first B columns also carry information on $\{u[0],\ldots,u[B-N]\}$.

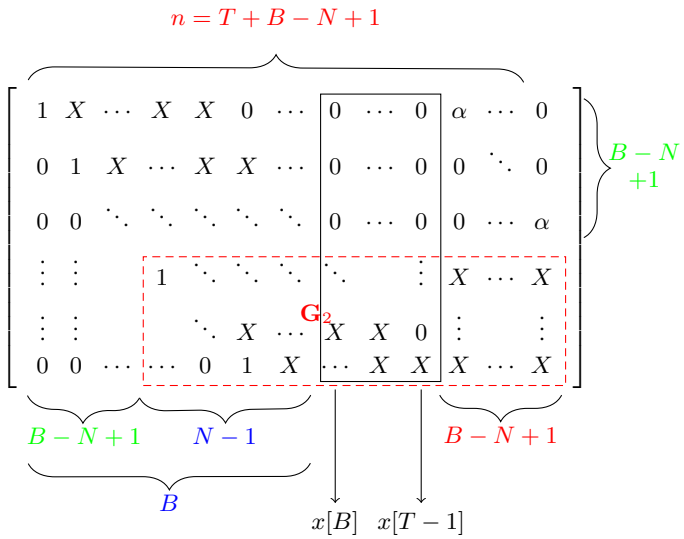
Alternatively, we note that symbols $\{x[B],\ldots,x[T-1]\}$ are the outcome of multiplying $u[0],\ldots,u[k-1]$ with $G_{:,B+1:T}$

[3] Similar to Y, Y' is not a constant element but rather represent (potentially different) element taken from $\mathbb{F}_q$.

(which is marked with a solid frame in (32) *below). From the construction of* $\mathbf{G}$*, sub-matrix* $\mathbf{G}_{1:B-N+1,B+1:T}$ *is a matrix of zeros hence these symbols does not contain any information about* $\{u[0],\ldots,u[B-N]\}$.

Recalling that $k-1=T-N$ *we have that the* $T-B$ *symbols* $\{x[B],\ldots,x[T-1]\}$ *are composed only of* $T-B$ *information symbols* $\{u[B-N+1],\ldots,u[T-N]\}$.

$\mathbf{G}=$



(32)

Remark 3. *Although* $\mathbf{G}$ *is not systematic, since* $\mathbf{G}_{:,1:k}$ *is upper triangular, shortening over the set of coordinates* $\{1,\ldots,i\}$ *where* $i\leq k$ *results in an MDS code with a generator matrix which can be generated from* $\mathbf{G}$ *by deleting the first* i *rows and first* i *columns. The above holds since assuming that given* $\mathbf{G}$*, assuming coordinates* $\{1,\ldots,i\}$ *are zero is equivalent to assuming that* $u[0]=\cdots u[i]=0$ *which is equivalent to delete the first* i *rows and columns in* $\mathbf{G}$.

Remark 4. *Although* $\mathbf{G}$ *is not a systematic generator matrix, we note that* $\tilde{\mathbf{G}}=\mathbf{M}^{-1}\mathbf{G}$ *is a systematic generator matrix. Since* $\mathbf{M}$ *is an invertible (upper triangular) matrix, we prove in Section IV that the parity-check matrix of* $\tilde{\mathbf{G}}$ *meets a set of conditions defined in [15] on the parity-check matrix of a capacity-achieving code.*

C. Example

As an example, we look at the case where $B=4$, $N=3$ and $T=6$. The first phases of the generator matrix are generated over $GF(11)$. In the final stage we take an element from $GF(121)\setminus GF(11)$. The resulting generator matrix of the $(8,4,6)_{GF(121)}$ code is:

$$\left(\begin{array}{c|cccccccc} T= & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline & 1 & 10 & 9 & 0 & 0 & 0 & \alpha & 0 \\ & 0 & 1 & 9 & 1 & 0 & 0 & 0 & \alpha \\ & 0 & 0 & 1 & 6 & 9 & 0 & 4 & 8 \\ & 0 & 0 & 0 & 1 & 4 & 1 & 9 & 8 \end{array}\right), \tag{33}$$

where

- MDS_1 (marked in blue and solid frame in (33)) is a $(6,4)_{\mathrm{GF}(11)}$ MDS code.
- MDS_2 (marked in red and dashed frame in (33)) is a $(6,2)_{\mathrm{GF}(11)}$ MDS code.

We set (for example) $\alpha=0\cdot 11+1\cdot 11$ as the element from $GF(121)\setminus GF(11)$.[4] Next, we demonstrate the decoding process for several cases of erasures. We focus on decoding symbol $u[0]$ and show for that for any burst of $B=4$ symbols or $N=2$ random erasures, $u[0]$ can be recovered with maximal delay of $T=6$.

As decoding symbol $u[0]$ when symbol $x[0]$ is not erased is trivial, we focus only on cases where $x[0]$ is erased:

- A burst of size $B=4$ starting at time 0

$$\left(\begin{array}{c|cccccccc} T= & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline & 1 & 10 & 9 & 0 & 0 & 0 & \alpha & 0 \\ & 0 & 1 & 9 & 1 & 0 & 0 & 0 & \alpha \\ & 0 & 0 & 1 & 6 & 9 & 0 & 4 & 8 \\ & 0 & 0 & 0 & 1 & 4 & 1 & 9 & 8 \end{array}\right),$$

  Using MDS_2, $u[2]$ and $u[3]$ can be decoded at time 5 since we have two linear independent equations from MDS_2 which is $(6,2)_{\mathrm{GF}(11)}$ MDS code. Thus, $u[2]$ and $u[3]$ can be recovered at time 5. Cancelling $u[2]$ and $u[3]$ from $x[6]$ results in decoding $u[0]$ at time 6 as required.
- $N=3$ arbitrary erasures where $x[6]$ is erased:

$$\left(\begin{array}{c|cccccccc} T= & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline & 1 & 10 & 9 & 0 & 0 & 0 & \alpha & 0 \\ & 0 & 1 & 9 & 1 & 0 & 0 & 0 & \alpha \\ & 0 & 0 & 1 & 6 & 9 & 0 & 4 & 8 \\ & 0 & 0 & 0 & 1 & 4 & 1 & 9 & 8 \end{array}\right).$$

  Using MDS_1, which is $(6,4)_{\mathrm{GF}(11)}$ MDS code, all information symbols can be decoded at time 4 thus meeting the delay constraint T.
- $N=3$ arbitrary erasures where $x[6]$ is not erased:

$$\left(\begin{array}{c|cccccccc} T= & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline & 1 & 10 & 9 & 0 & 0 & 0 & \alpha & 0 \\ & 0 & 1 & 9 & 1 & 0 & 0 & 0 & \alpha \\ & 0 & 0 & 1 & 6 & 9 & 0 & 4 & 8 \\ & 0 & 0 & 0 & 1 & 4 & 1 & 9 & 8 \end{array}\right). \tag{34}$$

  We note that the (3×5) lower right matrix of $\mathbf{G}_1$ (marked in dashed blue in (34)) is a generator matrix of $(5,3)_{\mathrm{GF}(11)}$ MDS code which is the outcome of shortening $\mathbf{G}_1$ over $\mathcal{P}=\{0\}$ (denoted as MDS_1^1). Receiving symbols $\{x[1],x[2],x[3]\}$ with no erasures means that bringing MDS_1^1 to an equivalent row echelon form results in three linear combinations of $u[j]\in\{u[1],u[2],u[3]\}$ and $u[0]$ (with coefficients taken from $\mathrm{GF}(11)$) where each combination consists only of $u[j]$ and $u[0]$. Denoting

[4] As one of the most common examples for field extension is generating the complex field from the real field, we borrow the complex field notation to indicate that α is from the extension field with no intersection with the base field (i.e., in the example of complex fields, it is equivalent to taking a pure imaginary number).

these combination as $\{\tilde{u}[1] = 10u[0] + u[1], \tilde{u}[2] = 7u[0]+u[2], \tilde{u}[3] = 3u[0]+u[3]\}$, using $\tilde{u}[2]$ and $\tilde{u}[3]$, $u[2]$ and $u[3]$ can be cancelled from $x[6]$. Since we assume that the element $\alpha = 0 \cdot 11 + 1 \cdot 11 \in \mathbb{F}_{\mathrm{GF}(121)\setminus\mathrm{GF}(11)}$, it is guaranteed that it is not nulled out (since the column operations are performed over $\mathrm{GF}(11)$). Hence, $u[0]$ can be decoded at time 6 as required.
Equivalently, the dashed part of $\mathbf{g}_6$ (the part of $x[6]$ which is a function of $\{u[1], u[2], u[3]\}$) is in the span of MDS_1^1. Since we have three symbols from MDS_1^1, the dashed part of $\mathbf{g}_6$ can be cancelled. Since we assume that the element $\alpha = 0{\cdot}11{+}1{\cdot}11 \in \mathbb{F}_{\mathrm{GF}(121)\setminus\mathrm{GF}(11)}$, it is guaranteed that it is not nulled out (since column operations are performed over $\mathrm{GF}(11)$). Hence, $u[0]$ can be recovered.

The decoding of $u[i] \in \{u[1], u[2], u[3]\}$ is done in a similar manner where we assume by induction that $\{u[0], \ldots, u[i{-}1]\}$ have already been recovered by time $T + i$.

We thus showed that the suggested $(8, 4, 6)_{GF(121)}$ block code is $(7, 4, 3)$-achievable. Therefore, using Lemma 2, a corresponding $(8, 4, 6)_{GF(121)}$-streaming code which is $(7, 4, 3)$-achievable is generated by apply diagonal interleaving. An example of a single diagonal is given in Table I below.

III. Proof Of Theorem 2

We prove next that the $(n, k, T)_{\mathbb{F}_{q^2}}$ block code described above is (W, B, N) achievable. The field size of this construction scales as $O(T^2)$. We therefore describe the decoding function φ_{l+T} defined in Definition 12 and prove it can recover any $u[l]$ (for any $l \in \{0, \ldots, k-1\}$) with maximal delay of T.

Decoding $\{u[0], \ldots, u[B-N]\}$:
We analyze the two different types of erasures:

- A Burst of length B starting at time i
  Decoding $u[0]$:
  Following Property 4 and Remark 1 we recall that MDS_2 is a $(T, T-B)_{\mathbb{F}_q}$ MDS code. A burst of B symbols starting at time 0 means that symbols $\{x[B], \ldots, x[T-1]\}$ are not erased. Recalling Remark 2, these symbols don't have an interference from information symbols $\{u[0], \ldots, u[B-N]\}$. Therefore, information symbols $\{u[B-N+1], \ldots, u[k-1]\}$ can be recovered using MDS_2.
  Noting that $x[T]$ is composed of $u[0]$ and $\{u[B-N+1], \ldots, u[k-1]\}$, after symbols $\{u[B-N+1], \ldots, u[k-1]\}$ are decoded, they can cancelled from $x[T]$ thus $u[0]$ can be decoded.[5]
  Decoding $u[i] \in \{u[1], \ldots, u[B-N]\}$:
  Noting that $x[i{+}T]$ is composed of $u[i]$ and $\{u[B{-}N{+}1], \ldots, u[k-1]\}$, we again perform the decoding in two steps:
  - Decoding of $\{u[B-N+1], \ldots, u[k-1]\}$ using MDS_2.
  - Decoding of $u[i]$ from $x[i{+}T]$ by canceling $\{u[B{-}N{+}1], \ldots, u[k-1]\}$ from it.

[5]In case $T = B$ it can be shown that $\mathbf{G}_3 = \alpha \cdot \mathbf{I}_{B-N+1}$ hence $s[0]$ can be recovered directly from $x[T]$.

We first assume by induction that symbols $\{u[0], \ldots, u[i-1]\}$ have already been recovered by time $T + i$, thus we cancel information symbols $\{u[0], \ldots, u[i-1]\}$ from $\{x[i+B], \ldots, x[i+T-1]\}$. Denoting $\{\tilde{x}[i{+}B], \ldots, \tilde{x}[i{+}T{-}1]\}$ as the symbols after the cancellation, we note that $\{\tilde{x}[i{+}B], \ldots, \tilde{x}[i{+}T{-}1]\}$ belong to MDS_2.
Therefore, using MDS_2 (which we recall again that following Remark 1 is $(T, T{-}B)_{\mathbb{F}_q}$ MDS code), information symbols $\{u[B-N+1], \ldots, u[k]\}$ can be recovered and cancelled from $x[i+T]$ and thus $u[i]$ can be recovered.

- N arbitrary erasures
  Decoding $u[0]$:
  First, we note that we assume that symbol $x[0]$ is one of the erased symbols otherwise decoding is trivial.[6] We further differentiate between the following two cases
  - Symbol $x[T]$ is erased
    We note that in this case, in MDS_1 we have a maximum of $N-1$ erasures. Following Property 3, MDS_1 can correct any $N-1$ erasures with maximal delay of T and hence information symbol $u[0]$ can be recovered with a maximal delay of T.
  - Symbol $x[T]$ is not erased

    Shortening MDS_1 on coordinate $\{0\}$ results in an $(k-1+N-1, k-1)_{\mathbb{F}_q}$ MDS code with a generator matrix that is generated from $\mathbf{G}_1$ by deleting the first row and column. Denoting this code as MDS_1^1, its generator matrix is depicted in (35) as the dashed $(k-1) \times (k-1+N-1)$ lower right submatrix of $\mathbf{G}_1$.

$$\mathbf{G}_1 = \left[\begin{array}{c|cccccccc|c} 1 & X & \cdots & X & 0 & 0 & 0 & \cdots & 0 & \alpha \\ 0 & 1 & X & \cdots & X & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & 1 & \ddots & \ddots & \ddots & & \vdots & X \\ \vdots & \vdots & & & \ddots & X & \cdots & X & 0 & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & 1 & X & \cdots & X & X \end{array}\right] \tag{35}$$

    Symbols $\{x[1], \ldots, x[T-1]\}$ have up to $N-1$ erasures (since we assumed $x[0]$ is erased). Therefore, using MDS_1^1, we can generate $k-1$ linear combinations of the form

$$\tilde{u}[j] = a_j \cdot u[j] + b_j \cdot u[0]$$

    where $j \in \{1, \ldots, k{-}1\}$ and $a_j, b_j \in \mathbb{F}_q$ and $a_j \neq 0$. Since $\alpha \in \mathbb{F}_{q^2} \setminus \mathbb{F}_q$, $\{u[1], \ldots, u[B-N]\}$ can be cancelled from symbol $x[T]$ using $\{\tilde{u}[1], \ldots, \tilde{u}[B-$

[6]If $N = 1$ it means that the only erasure is that of symbol $x[0]$ and hence decoding is done as described next for the case when symbol $x[T]$ is not erased.

$\mathbf{x}_i$	$\mathbf{x}_{i+1}$	$\mathbf{x}_{i+2}$	$\mathbf{x}_{i+3}$	$\mathbf{x}_{i+4}$	$\mathbf{x}_{i+5}$	$\mathbf{x}_{i+6}$	$\mathbf{x}_{i+7}$	$\mathbf{x}_{i+8}$
$s_i[0]$	$s_{i+1}[0]$							
	$10s_i[0]$ $+s_{i+1}[1]$	$10s_{i+1}[0]$ $+s_{i+2}[1]$						
		$9s_i[0]+9s_{i+1}[1]$ $+s_{i+2}[2]$	$9s_{i+1}[0]+9s_{i+2}[1]$ $+s_{i+3}[2]$					
			$s_{i+1}[1]+6s_{i+2}[2]$ $+s_{i+3}[3]$	$s_{i+2}[1]+6s_{i+3}[2]$ $+s_{i+4}[3]$				
				$9s_{i+2}[2]$ $+4s_{i+3}[3]$	$9s_{i+3}[2]$ $+4s_{i+4}[3]$			
					$s_{i+3}[3]$	$s_{i+4}[3]$		
						$11s_i[0]+4s_{i+2}[2]$ $+9s_{i+3}[3]$	$11s_{i+1}[0]+4s_{i+3}[2]$ $+9s_{i+4}[3]$	
							$11s_{i+1}[1]+8s_{i+2}[2]$ $+8s_{i+3}[3]$	$11s_{i+2}[1]+8s_{i+3}[2]$ $+8s_{i+4}[3]$

TABLE I
AN EXAMPLE OF APPLYING DIAGONAL INTERLEAVING TO GENERATE $(8,4,6)_{GF(121)}$-STREAMING CODE FOR A $(8,4,6)_{GF(121)}$ BLOCK CODE.

$N]\}$ while it is guaranteed that α is not nulled out[7], symbol $u[0]$ can be recovered with a delay of T.

Decoding $u[i] \in \{u[1],\ldots,u[B-N]\}$:
We first assume by induction that $\{u[0],\ldots,u[i-1]\}$ have already been recovered by time $T+i$. We further assume that their contribution on symbols $\{x[i],\ldots,x[i+T]\}$ is cancelled. After cancelling symbols $\{u[0],\ldots,u[i-1]\}$ from MDS_1 we are left with a $(k-i+N-1,k-i)_{\mathbb{F}_q}$ MDS code (which can recover any $N-1$ erasures). Equivalently, this can be viewed as shortening MDS_1 on coordinates $\{0,\ldots,i-1\}$. We denote it as MDS_1^i.
Since $x[i]$ is composed of (a sub group) of $\{u[0],\ldots,u[i]\}$ and we assumed that $\{u[0],\ldots,u[i-1]\}$ have been decoded correctly, if $x[i]$ is not erased, $u[i]$ can be trivially decoded.
Therefore, assuming $x[i]$ is erased, the decoder does the following:

- In case symbol $x[i+T]$ is erased
  Assuming $x[i+T]$ is erased means that there are at most $N-1$ erasures in MDS_1^i. Recalling that MDS_1^i can recover any $N-1$ erasures, all information symbols can be decoded up to time $T+i$.
- Symbol $x[i+T]$ not erased
  We note that the lower right $k-(i+1)\times k-(i+1)+N-1$ sub-matrix of $\mathbf{G}_1$ is a generator matrix of a $(k-(i+1)+N-1,k-(i+1))_{\mathbb{F}_q}$ MDS code (can be also viewed as the outcome of shortening MDS_1^i on coordinate $\{i\}$). We denote it as MDS_1^{i+1}. We may assume that symbols $\{x[i+1],\ldots,x[i+T-1]\}$ have up to $N-1$ erasures. Hence, using MDS_1^{i+1}, we can generate $k-i-1$ linear combinations of the form

$$\tilde{u}[j]=a_j\cdot u[j]+b_j\cdot u[0]$$

  where $j\in\{i+1,\ldots,k-1\}$ and $a_j,b_j\in\mathbb{F}_q$ and $a_j\neq 0$. Next, $\{\tilde{u}[i+1],\ldots,\tilde{u}[k-1]\}$ are cancelled from symbol $x[i+T]$ using operations over $\mathbb{F}_q$ and, again, it is guaranteed that $\alpha\in\mathbb{F}_{q^2}\setminus\mathbb{F}_q$ is not cancelled, thus $u[i]$ can be recovered.

[7]Since all elements in MDS_2 belong to $\mathbb{F}_q$, the cancellation is done by multiplying $\{\tilde{u}[1],\ldots,\tilde{u}[B-N]\}$ with coefficients $\mathbb{F}_q$.

Decoding $\{u[B-N+1],\ldots,u[k-1]\}$:
We first assume by induction that $\{u[0],\ldots,u[B-N]\}$ have already been recovered by time $T+B-N$ and cancelled from the received symbols. This means that we are left with MDS_2. Recalling Property 4 MDS_2, is a $(k-(B-N+1)\times(n-(B-N+1))_{\mathbb{F}_q}$ MDS code which means it can correct any B erasures within the delay constraint ($\forall i\in\{B-N+1,\ldots,k-1\}$ the information symbols transmitted at time i is recovered at time $T+B-N+1\leq i+T$) . Recalling that $B\geq N$ it means that either a burst of B erasures or arbitrary N erasures can be correctly decoded within the delay constraint.

IV. ALTERNATIVE PROOF OF THEOREM 2 - ANALYSIS OF THE PARITY-CHECK MATRIX

In [15], a set of requirements on the parity-check matrix of $(n,k,T)_{\mathbb{F}}$-block code which is (W,N,B)-achievable were defined. In this section we first recall this set of requirements and than we show that the parity-check matrix of the code suggested in Section II meet these requirements.

A. Requirements on the parity-check matrix (Section II.F in [15])

Let i denote an erased coordinate. Due to the delay constraint, all symbols in time $[i+T+1:n-1]$ are unavailable to the decoder. We assume though that during the decoding of symbol $x[i]$ all symbols $x[0],\ldots,x[i-1]$ are available. Therefore we have:

$$\underbrace{x[0],\ldots,x[i-1]}_{\text{Known}},\quad\underbrace{x[i]}_{\text{Symbol to be recovered}},\qquad(36)$$

$$\underbrace{x[i+1],\ldots,x[i+1],\ldots,x[i+T]}_{\text{All the non-erased symbols are accessible}},\underbrace{x[i+T+1],\ldots,x[n-1]}_{\text{Inaccessible symbols}}\qquad(37)$$

Let $\mathcal{K}$ denote the set of coordinates of $\{0:i-1\}$ and $\mathcal{U}$ denote the set of coordinates of $\{i+T+1:n-1\}$. Thus when decoding $x[i]$, we assume that any $x[j]$, $j\in\mathcal{K}$ are known, any $x[j]$, $j\in\mathcal{U}$ are inaccessible and that non-erased code symbol $x[j]$, $j\in\{i+1:t+T\}$ are also accessible.

The effective code used to decode $x[i]$ is obtained from $\mathcal{C}$ by shortening it on the coordinates in $\mathcal{K}$ and puncturing it on coordinates in $\mathcal{U}$. Recalling Theorem 2, puncturing of the

generator matrix is equivalent to the shortening of the parity-check matrix, and the shortening of the generator matrix is equivalent to the puncturing of the parity-check matrix.

Let $\mathbf{H}$ be the parity-check matrix of the code $\mathcal{C}$. For $0 \leq l \leq B-N$, set

$$\mathbf{H}^{(l)} = \left[\mathbf{h}_0^{(l)}, \ldots, \mathbf{h}_{l+T}^{(l)}\right], \tag{38}$$

be the parity-check of the code $\mathcal{C}$ restricted to the coordinates $[0 : l+T]$ (alternatively, we can say that it is the parity-check of a code which is the outcome of puncturing coordinates $[l+T+1 : n]$ from the original code). This means that $\mathbf{H}^{(l)}$ is the derived from $\mathbf{H}$ by shortening coordinates $[l+T+1 : n]$. Therefore, $\mathbf{H}^{(l)}$ is a $N + l \times T + l + 1$ matrix.

Recalling Lemma 1, this means that in order to guarantee recovery of all information symbols with delay constraint T from any admissible erasure event the following conditions on $\mathbf{H}$ and $\left\{\mathbf{H}^{(l)},\ 0 \leq l \leq B-N\right\}$ are sufficient:

1) **Condition B1** For $0 \leq l \leq B-N$, the l-th column, $\mathbf{h}_l^{(l)}$ of $\mathbf{H}^{(l)}$ should be linearly independent of the set of $(B-1)$ columns
$$\left\{\mathbf{h}_j^{(l)},\ l+1 \leq j \leq l+B-1\right\}$$
2) **Condition R1** For $0 \leq l \leq B-N$, the l-th column, $\mathbf{h}_l^{(l)}$ of $\mathbf{H}^{(l)}$ should be linearly independent of the any set of $(N-1)$ columns taken from the set
$$\left\{\mathbf{h}_j^{(l)},\ l+1 \leq j \leq l+T\right\}$$
3) **Condition B2** For $B-N+1 \leq l \leq T-N+1$, the set
$$\{\mathbf{h}_j,\ l \leq j \leq l+B\}$$
of columns of $\mathbf{H}$ should be linearly independent.
4) **Condition R2** Any set of N columns from the set
$$\{\mathbf{h}_j,\ B-N+1 \leq j \leq T+B-N+1\}$$
of columns of $\mathbf{H}$ should be linearly independent.

We show next that all these conditions hold for the parity-check matrix of the suggested code, which is at a rate that equals the capacity. Therefore, it proves that the suggested code achieves that capacity.

B. Rank of matrix multiplication

In analyzing the parity matrix of the suggested code, we use the following properties of the rank of matrix multiplication.

Lemma 3. *Let* $\mathbf{A}$ *be* $K \times L$ *matrix and* $\mathbf{B}$ *an* $L \times M$ *matrix. Then*

$$\operatorname{rank}(\mathbf{AB}) \leq \min\left(\operatorname{rank}(\mathbf{A}), \operatorname{rank}(\mathbf{B})\right) \tag{39}$$

The proof of Lemma 3 is given in Appendix B.

Lemma 4 (Sylvester's rank inequality, [21])**.** *Let* $\mathbf{A}$ *be* $K \times L$ *matrix and* $\mathbf{B}$ *an* $L \times M$ *matrix. Then*

$$\operatorname{rank}(\mathbf{AB}) \geq \operatorname{rank}(\mathbf{A}) + \operatorname{rank}(\mathbf{B}) - L \tag{40}$$

Corollary 7. *If* $\mathbf{A}$ *is a square* $(L \times L)$ *full rank matrix, Sylvester's rank inequality means that*

$$\operatorname{rank}(\mathbf{AB}) = \operatorname{rank}(\mathbf{B}). \tag{41}$$

If $\mathbf{B}$ *is a* $(L \times L)$ *full rank matrix, from Sylvester's rank inequality we have*

$$\operatorname{rank}(\mathbf{AB}) = \operatorname{rank}(\mathbf{A}). \tag{42}$$

C. Analysis of the parity-check matrix of the suggested block code

As mentioned in Remark 4, the suggested code can be transformed to a systematic block code. The generator matrix of the systematic code can be denoted as

$$\tilde{\mathbf{G}} = \mathbf{M}^{-1}\mathbf{G} \tag{43}$$

where $\mathbf{M}$ and $\mathbf{G}$ are defined in (29) and (31).

Since $\mathbf{G}$ is generated from $\mathbf{G}'$ (defined in (28)) by replacing the $(B-N+1) \times (B-N+1)$ upper right matrix with $\alpha \cdot \mathbf{I}_{B-N+1}$ where $\alpha \in \mathbb{F}_{q^2} \setminus \mathbb{F}_q$ and recalling that $\mathbf{G}' = \mathbf{M}\mathbf{G}''$ (where $\mathbf{G}''$ is a systematic matrix defined in (27)) we have that

$\tilde{\mathbf{G}} =$

$$\left[\begin{array}{cccccccccccccccc}
1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & Y & \cdots & Y & f_{1,1}(\alpha) & \cdots & \cdots & f_{1,B-N+1}(\alpha) \\
0 & 1 & 0 & \cdots & \cdots & \cdots & 0 & Y & \cdots & Y & Y_\alpha & f_{2,2}(\alpha) & \cdots & f_{2,B-N+1}(\alpha) \\
0 & 0 & \ddots & \cdots & \vdots & \vdots & 0 & Y & \cdots & Y & \vdots & \ddots & \ddots & \vdots \\
\vdots & \vdots & & \ddots & \vdots & \vdots & 0 & Y & \cdots & Y & Y_\alpha & \cdots & Y_\alpha & f_{B-N+1,B-N+1}(\alpha) \\
\vdots & \vdots & & & \ddots & \vdots & 0 & Y & \cdots & Y & Y & \cdots & Y & Y \\
0 & 0 & \cdots & \cdots & 0 & 1 & 0 & Y & \cdots & Y & Y & \cdots & Y & Y \\
0 & 0 & \cdots & \cdots & \cdots & 0 & 1 & Y & \cdots & Y & Y & \cdots & Y & Y
\end{array}\right].$$

$$\underbrace{\qquad\qquad\qquad\qquad}_{k \qquad\qquad N-1} \qquad \underbrace{\qquad\qquad\qquad\qquad}_{B+N-1}$$

(44)

where $f_{i,j}(\alpha) = a + b \cdot \alpha$ where $a, b \in \mathbb{F}_q$ is a function of α (where in general, $f(\alpha) = 0$ is no precluded, besides for specific functions which are defined in Property 5 below), and Y_α stands for elements from the base field ($\mathbb{F}_q$) which are now different from the original elements in $\mathbf{P}''$. More details on the generation of $\tilde{\mathbf{G}}$ and each of its elements are given in Appendix A.

We note the following property

Property 5. *For any* $i \in \{0, \ldots, B-N+1\}$, $f_{i,i}(\alpha) \neq 0$.

This property holds since $f_{i,i}(\alpha)$ is the outcome of multiplying the i'th row in $\mathbf{M}^{-1}$ with the $k+i$ column of $\mathbf{G}$. Since $\mathbf{M}_{i,i}^{-1} = 1$ and this is the only element which multiplies α it follows that $f_{i,i}(\alpha) \neq 0$ (summing α with elements from $\mathbb{F}_q$ cannot null it).

Therefore, the parity-check matrix can be written as We note that several examples of $\mathbf{H}^{(l)}$ are marked (each in different color and different type of frame) in (45).

1) **Condition B1**

$$\mathbf{H} = \left[-\tilde{\mathbf{P}}^T \,\middle|\, \mathbf{I}_{B\times B}\right] = \begin{array}{c} N-1 \\ \\ B-N+1 \end{array} \left[\begin{array}{ccccccccccccc} -Y & \cdots & \cdots & \cdots & -Y & \cdots & -Y & 1 & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & 1 & 0 & \cdots & \cdots & 0 \\ -Y & \cdots & \cdots & \cdots & -Y & \cdots & -Y & 0 & 0 & \ddots & \cdots & \vdots & 0 \\ f_{1,1}(\alpha) & -Y_\alpha & \cdots & -Y_\alpha & -Y & \cdots & -Y & \vdots & \vdots & & 1 & \ddots & 0 \\ f_{1,2}(\alpha) & f_{2,2}(\alpha) & \ddots & \vdots & \vdots & \cdots & -Y & \vdots & \vdots & & 0 & \ddots & 0 \\ \vdots & \vdots & \ddots & -Y_\alpha & \vdots & \vdots & -Y & \vdots & \vdots & & & 0 & \ddots & 0 \\ f_{1,B-N+1}(\alpha) & f_{2,B-N+1}(\alpha) & \cdots & f_{B-N+1,B-N+1}(\alpha) & -Y & \cdots & -Y & 0 & 0 & \cdots & \cdots & 0 & 0 & 1 \end{array}\right] \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \end{array} \quad (45)$$

$$\underbrace{\qquad\qquad}_{B-N+1} \quad \underbrace{\qquad\qquad}_{T-B} \quad \underbrace{\qquad}_{B}$$

Before showing condition B1 holds, we show the following properties which will be used in the proof. Denote the lower right $(T-B \times B)$ matrix of $\mathbf{G}$

$$\mathbf{G}_{k-(T-B)+1:k,n-B+1:n} \triangleq \mathbf{G}_{T-B,B}$$

and the upper right $(B \times T-B)$ matrix of $\mathbf{M}^{-1}$

$$\mathbf{M}^{-1}_{1:B,n-(T-B)+1:n} \triangleq \mathbf{M}^{-1}_{B,T-B}.$$

Property 6. *$\mathbf{G}_{T-B,B}$ has rank of $\min(T-B,B)$. Further, any sub-matrix formed by taking w columns of $\mathbf{G}_{T-B,B}$ has a rank of $\min(T-B,w)$.*

Proof. See Appendix C □

Property 7. *The first $B-N+1$ rows of $\mathbf{M}^{-1}_{B,T-B}$ are a linear function of the last $N-1$ rows.*

Proof. See Appendix D. □

Condition B1 requires that $\mathbf{h}_l^{(l)}$ is linearly independent of the set of $(B-1)$ columns which follows it. Therefore, we first limit $\mathbf{H}^{(l)}$ to contain a maximum of B columns starting for its l'th column. Further, we note the following

- If $B+l<k$. We denote the the group of B columns starting from the l'th column of $\mathbf{H}^{(l)}$ as $\hat{\mathbf{H}}^{(l)} = \mathbf{H}^{(l)}_{:,l:l+B}$. It follows the dimensions of $\hat{\mathbf{H}}^{(l)}$ are $N+l \times B$.
- If $B+l \geq k$. We note than when $B+l \geq k$, the group of B columns starting from the l'th column of $\mathbf{H}^{(l)}$ contains $B+l-k$ vectors from $\mathbf{I}_{B\times B}$ (the unit matrix). These $B+l-k$ are part of the basis that spans this group of B vectors. Hence, its suffices to analyze this group without the top $B+l-k$ rows and last $B+l-k$ columns. Therefore, we denote by $\hat{\mathbf{H}}_{(l)}$ the group of B vectors starting from l'th column without the top $B+l-k$ rows and last $B+l-k$. Hence $\hat{\mathbf{H}}_{(l)} = \mathbf{H}^{(l)}_{B+l-k+1:N+l+1,l+1:k}$. It follows that the dimensions of $\hat{\mathbf{H}}_{(l)}$ are $N+(k-B)\times k-l$.

Some examples of $\hat{\mathbf{H}}^{(l)}$ (where $l=0$ and $l=k-B$) and for $\hat{\mathbf{H}}_{(l)}$ (where $l=B-N$) are given in (46) where each $\hat{\mathbf{H}}^{(l)}$ is denoted with different color and different type of frame.

- $B+l<k$: analyzing $\hat{\mathbf{H}}^{(l)}$.

Proving condition B1 is equivalent to prove that column $\hat{\mathbf{h}}_0^{(l)}$ (the first column in $\hat{\mathbf{H}}^{(l)}$) is linearly independent of the rest of the $B-1$ columns of $\hat{\mathbf{H}}^{(l)}$ which we denote as $\hat{\mathbf{H}}^{(l)}_{2:B}$. This in turn is equivalent to show that

$$\text{rank}(\hat{\mathbf{H}}^{(l)}_{2:B}) \leq \text{rank}(\hat{\mathbf{H}}^{(l)}) - 1.$$

Hence, we carefully analyze next the rank of these matrices.

Recalling that $\hat{\mathbf{H}}^{(l)}$ has $N+l$ rows and B columns, an example of $\hat{\mathbf{H}}^{(l)}$ (marked in solid) is

$$\hat{\mathbf{H}}^{(l)} = \mathbf{H}^{(l)}_{1:N+l,l+1:l+B} = \begin{array}{c} N+l \\ \\ \\ \\ \\ \\ \\ \end{array} \left[\begin{array}{ccccccc} -Y & \cdots & \cdots & \cdots & \cdots & \cdots & -Y \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -Y & \cdots & \cdots & \cdots & \cdots & \cdots & -Y \\ -Y_\alpha & \cdots & \cdots & -Y_\alpha & -Y & \cdots & -Y \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ f_{l+1,l+1}(\alpha) & -Y_\alpha & \cdots & -Y_\alpha & -Y & \cdots & -Y \end{array}\right] \quad (47)$$

$$\underbrace{\qquad\qquad}_{B-N+1-l} \quad \underbrace{\qquad\qquad}_{N+l-1}$$

$$\underbrace{\qquad\qquad\qquad\qquad}_{B}$$

We start with analyzing the rank of $\hat{\mathbf{H}}^{(l)}$. First, we note that since we want to show condition B1 for $0 \leq l \leq B-N$ it follows that $N+l \leq B$ thus the maximal rank of $\hat{\mathbf{H}}^{(l)}$ is $N+l$.

We note that when $B+l \leq k$, each $\hat{\mathbf{H}}^{(l)}$ contains a $(N+l-1 \times N+l-1)$ sub-matrix of $\mathbf{P}''$ marked in dashed in (47). Following Corollary 2, each sub-matrix of $\mathbf{P}''$ has a full rank thus the rank of this sub-matrix is $N+l-1$. This means we can transform $\hat{\mathbf{H}}^{(l)}$ to an equivalent row echelon form of

$$\left[\begin{array}{cccccc} \mathbf{0}_{N+l-1\times B-(N+l-1)} & & & & \mathbf{I}_{N+l-1} & \\ \tilde{f}_{l+1,l+1}(\alpha) & -\tilde{Y}_\alpha & \cdots & -\tilde{Y}_\alpha & -\tilde{Y} & \cdots & -\tilde{Y} \end{array}\right], \quad (48)$$

where $\tilde{f}_{l+1,l+1}(\alpha)$,each $-\tilde{Y}_\alpha$ and each $-\tilde{Y}$ are the outcome of the column operations performed on $\hat{\mathbf{H}}^{(l)}$ (over $\mathbb{F}_q$). Following Property 5 we have that $f_{l+1,l+1}(\alpha) \neq 0$ and since all

$$\mathbf{H} = \left[-\tilde{\mathbf{P}}^T \,\middle|\, \mathbf{I}_{B\times B}\right] = \begin{matrix} N-1 \\ \\ B-N+1 \end{matrix} \left[\begin{array}{cccccccccccccc} -Y & \cdots & \cdots & \cdots & -Y & \cdots & -Y & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & 1 & 0 & \cdots & \cdots & \cdots & 0 \\ -Y & \cdots & \cdots & \cdots & -Y & \cdots & -Y & 0 & 0 & \ddots & \cdots & \vdots & \vdots & 0 \\ f_{1,1}(\alpha) & -Y_\alpha & \cdots & -Y_\alpha & -Y & \cdots & -Y & \vdots & \vdots & & 1 & \ddots & \vdots & 0 \\ f_{1,2}(\alpha) & f_{2,2}(\alpha) & \ddots & \vdots & \vdots & \cdots & -Y & \vdots & \vdots & & 0 & \ddots & \vdots & 0 \\ \vdots & \vdots & \ddots & -Y_\alpha & \vdots & \vdots & -Y & \vdots & \vdots & & & 0 & \ddots & 0 \\ f_{1,B-N+1}(\alpha) & f_{2,B-N+1}(\alpha) & \cdots & f_{B-N+1,B-N+1}(\alpha) & -Y & \cdots & -Y & 0 & 0 & \cdots & \cdots & 0 & 0 & 1 \end{array}\right] \tag{46}$$

$\hat{\mathbf{H}}^{(0)}$, $\hat{\mathbf{H}}^{(k-B)}$, $\hat{\mathbf{H}}_{(B-N)}$; $B-N+1$, $T-B$

column operations were performed over $\mathbb{F}_q$ it follows that $\tilde{f}_{l+1,l+1}(\alpha) \neq 0$. Thus, we have that

$$\text{rank}(\hat{\mathbf{H}}^{(l)}) = N + l. \tag{49}$$

Next, we analyze the rank of $\hat{\mathbf{H}}^{(l)}_{2:B}$. We note that $\hat{\mathbf{H}}^{(l)}$ is generated by taking the transpose of multiplying $\mathbf{M}^{-1}_{(l+1:B+l,l+1:k)}$ with $\mathbf{G}_{(l+1:k,k+1:k+N+l)}$ which we denote as $\mathbf{G}^{(l)}$ Hence,

$$\hat{\mathbf{H}}^{(l)} = \left(\mathbf{M}^{-1}_{(l+1:B+l,l+1:k)} \times \mathbf{G}^{(l)}\right)^T. \tag{50}$$

In (51) below we show examples of $\mathbf{G}^{(l)}$ when we assume $B + l \leq k$ for all $0 \leq l \leq B - N$.

$\mathbf{G} =$

$$\left[\begin{array}{cccccc|cccc} 1 & X & \cdots & X & 0 & 0 & 0 & \cdots & 0 & \alpha \\ 0 & 1 & X & \cdots & X & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & 1 & \ddots & \ddots & \ddots & & \vdots & X \\ \vdots & \vdots & & & \ddots & X & \cdots & X & 0 & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & 1 & X & \cdots & X & X \end{array}\right] \begin{matrix} \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \mathbf{G}^{(0)} \\ \mathbf{G}^{(1)} \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ B-N \\ +1 \\ \\ T-B \end{matrix}$$

k; $N-1$, $B-N+1$ (51)

We note that since in generating $\mathbf{G}$ we replaced the upper right $(B-N+1)\times(B-N+1)$ matrix of $\mathbf{G}'$ with $\alpha\mathbf{I}_{B-N+1}$ it means that there are only $T-B+1$ non-zero rows in $\mathbf{G}^{(l)}_{(l+2:k,:)}$ (the first row and the last $T-B$ rows).

Therefore, denoting with $\mathbf{G}^{(l)}_{T-B,:}$ the last $T-B$ rows of $\mathbf{G}^{(l)}$, we have

$$\hat{\mathbf{H}}^{(l)}_{2:B} = \left(\mathbf{M}^{-1}_{(l+2:B+l,2+B-N:k)} \times \mathbf{G}^{(l)}_{T-B,:}\right)^T, \tag{52}$$

where $\mathbf{M}^{-1}_{(l+2:B+l,2+B-N:k)}$ is acquired by taking the $B-1\times T-B$ lower right matrix of $\mathbf{M}^{-1}_{(l+2:B+l,1:k)}$. Several examples of $\mathbf{M}^{-1}_{(l+2:B+l,2+B-N:k)}$ and $\mathbf{M}^{-1}_{(l+2:B+l,1:k)}$ for different values of l (where the full matrix is marked in solid and the partial matrix is marked in dashed) are shown in (53).

$\mathbf{M}^{-1} =$

$$\left[\begin{array}{cccccccc} 1 & Y'' & \cdots & Y'' & Y'' & \cdots & \cdots & Y'' \\ 0 & 1 & Y'' & \cdots & Y'' & Y'' & \cdots & Y'' \\ \vdots & \cdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & Y'' & \cdots & Y'' & Y'' \\ 0 & \cdots & 0 & 0 & 1 & Y'' & \cdots & Y'' \\ \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \ddots & Y'' \\ 0 & \cdots & 0 & 0 & 0 & 0 & 1 & Y'' \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ N-1, \\ \\ \\ \end{matrix}$$

B; $T-B$

$\mathbf{M}^{-1} =$

$$\left[\begin{array}{cccccccc} 1 & Y'' & \cdots & Y'' & Y'' & \cdots & \cdots & Y'' \\ 0 & 1 & Y'' & \cdots & Y'' & Y'' & \cdots & Y'' \\ \vdots & \cdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & Y'' & \cdots & Y'' & Y'' \\ 0 & \cdots & 0 & 0 & 1 & Y'' & \cdots & Y'' \\ \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \ddots & Y'' \\ 0 & \cdots & 0 & 0 & 0 & 0 & 1 & Y'' \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ N-1, \\ \\ \\ \end{matrix}$$

B; $T-B$ (53)

Noting that $\mathbf{M}^{-1}_{(l+2:B+l,2+B-N:k)}$ contains $B-N-l$ rows from the first $B-N+1$ rows of $\mathbf{M}^{-1}_{B,T-B}$ (depicted in (94)), and recalling Property 7, it follows that these rows are a linear function of the last $N-1$ rows of $\mathbf{M}^{-1}_{B,T-B}$ (which are also a part of $\mathbf{M}^{-1}_{(l+2:B+l,2+B-N:k)}$). Further, the lower $N-1+l$ rows of $\mathbf{M}^{-1}_{(l+1:B+l,2+B-N:k)}$ are an upper triangular matrix

therefore,

$$\text{rank}(\mathbf{M}^{-1}_{(l+1:B+l,2+B-N:k)}) = N-1+l. \tag{54}$$

Since taking the last $T-B$ rows of $\mathbf{G}^{(l)}$ results in a matrix which is a sub-matrix of $\mathbf{G}_{T-B,B}$ (by taking only $N+l$ columns out of the B columns of $\mathbf{G}_{T-B,B}$), recalling Property 6, it follows that

$$\text{rank}(\mathbf{G}^{(l)}_{T-B,:}) = \min(T-B, N+l). \tag{55}$$

Therefore, following Lemma 3, (54) and (55) we have

$$\begin{aligned}\text{rank}(\hat{\mathbf{H}}^{(l)}_{2:B}) \leq &\min\left(\text{rank}(\mathbf{M}^{-1}_{(l+2:B+l,2+B-N:k)}), \text{rank}(\mathbf{G}^{(l)}_{T-B,:})\right)\\ &\min(T-B, N-1+l)\end{aligned} \tag{56}$$

and since we assume

$$\begin{aligned}B+l &\leq k\\ &= T-N+1,\end{aligned} \tag{57}$$

we have $N-1+l \leq T-B$. Thus,

$$\text{rank}(\hat{\mathbf{H}}^{(l)}_{2:B}) \leq N-1+l \tag{58}$$

Therefore, since $\text{rank}(\hat{\mathbf{H}}^{(l)}) = N+l$ and $\text{rank}(\hat{\mathbf{H}}^{(l)}_{2:B}) \leq N-1+l$, it follows that $\hat{\mathbf{h}}^{(l)}_0$ is linearly independent with the following $B-1$ columns.

- $B+l \geq k$: analyzing $\hat{\mathbf{H}}_{(l)}$.

Recalling that $\hat{\mathbf{H}}_{(l)}$ has $N+k-B$ rows and $k-l$ columns, proving condition B1 is equivalent to prove that column $\hat{\mathbf{h}}^{(l)}_0$ (the first column in $\hat{\mathbf{H}}^{(l)}$) is linearly independent of the rest of the $k-l-1$ columns of $\hat{\mathbf{H}}^{(l)}$ which we denote as $\hat{\mathbf{H}}^{(l)}_{2:k-l}$. This in turn is equivalent to show that

$$\text{rank}(\hat{\mathbf{H}}^{(l)}_{2:k-l}) \leq \text{rank}(\hat{\mathbf{H}}^{(l)}) - 1.$$

Hence, we carefully analyze next the rank of these matrices.

An example of $\hat{\mathbf{H}}_{(l)}$ is

$$\hat{\mathbf{H}}_{(l)} = \mathbf{H}^{(l)}_{B+l+k+1:N+l,l+1:k}$$

$$= \begin{matrix} \\ N+l \end{matrix} \left\{ \begin{matrix} B+l \\ -k \\ \\ N+k \\ -B \end{matrix} \right. \left[\begin{matrix} -Y & \cdots & \cdots & \cdots & \cdots & \cdots & -Y \\ \vdots & & & & & & \vdots \\ -Y & \cdots & \cdots & \cdots & \cdots & \cdots & -Y \\ -Y_\alpha & \cdots & \cdots & -Y_\alpha & -Y & \cdots & -Y \\ \vdots & & & & & & \vdots \\ f_{l+1,l+1}(\alpha) & -Y_\alpha & \cdots & -Y_\alpha & -Y & \cdots & -Y \end{matrix} \right] \begin{matrix} \\ \\ \\ \\ \\ \\ \end{matrix}$$

$$\begin{matrix} & & & & B-N+1-l & & T-B \\ & & & & \underbrace{\qquad\qquad\qquad\qquad}_{k-l} & & \end{matrix} \tag{59}$$

We first note that each $\hat{\mathbf{H}}_{(l)}$ contains a $(N+k-B-1 \times T-B)$ sub-matrix of $\mathbf{P}''$ (marked in dashed in (59)). We note that

$$\begin{aligned}N+k-B-1 &= N+(T-N+1)-B-1\\ &= T-B\end{aligned} \tag{60}$$

Hence, the rank of this matrix is $T-B$.

This means we can transform $\hat{\mathbf{H}}_{(l)}$ to an equivalent row echelon form of

$$\hat{\mathbf{H}}_{(l)} = \left[\begin{matrix} \mathbf{0}_{T-B\times B-N+1-l} & & & & \mathbf{I}_{T-B} & & \\ \tilde{f}_{l+1,l+1}(\alpha) & -\tilde{Y}_\alpha & \cdots & -\tilde{Y}_\alpha & -\tilde{Y} & \cdots & -\tilde{Y} \end{matrix} \right] \tag{61}$$

where $\tilde{f}_{l+1,l+1}(\alpha)$,each $-\tilde{Y}_\alpha$ and each $-\tilde{Y}$ are the outcome of the column operations performed on $\hat{\mathbf{H}}_{(l)}$. Following Property 5 we have that $f_{l+1,l+1}(\alpha) \neq 0$ and since all column operations were performed over $\mathbb{F}_q$ it follows that $\tilde{f}_{l+1,l+1}(\alpha) \neq 0$. Thus, we have that

$$\text{rank}(\hat{\mathbf{H}}_{(l)}) = T-B+1. \tag{62}$$

Proving condition B1 is equivalent to prove that column $\hat{\mathbf{h}}_{(l),0}$ (the first column in $\hat{\mathbf{H}}_{(l)}$) is linearly independent of the rest of the $k-1-l$ columns of $\hat{\mathbf{H}}_{(l)}$ which we denote as $\hat{\mathbf{H}}_{(l),2:k-l}$. We note that $\hat{\mathbf{H}}_{(l)}$ is generated by taking the transpose of multiplying $\mathbf{M}^{-1}_{l+1:k,l+1:k}$ with $\mathbf{G}_{l+1+l:k,k+1:k+N+l}$ which we denoted as $\mathbf{G}_{(l)}$.

Hence,

$$\hat{\mathbf{H}}_{(l)} = \left(\mathbf{M}^{-1}_{(l+1:k,l+1:k)} \times \mathbf{G}_{(l)}\right)^T. \tag{63}$$

In (64) below we show examples of $\mathbf{G}_{(l)}$ when we assume $B+l > k$ for all $0 \leq l \leq B-N$.

$$\mathbf{G} = \begin{matrix} \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \overset{\mathbf{G}_{(0)}}{\overbrace{}} \overset{\mathbf{G}_{(1)}}{\overbrace{}} \\ \\ \\ \\ \\ \\ \end{matrix}$$

$$\left[\begin{matrix} 1 & X & \cdots & X & 0 & 0 & 0 & \cdots & 0 & \alpha & \cdots & 0 \\ 0 & 1 & X & \cdots & X & 0 & 0 & \cdots & 0 & 0 & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & 0 & \cdots & 0 & 0 & \cdots & \alpha \\ \vdots & \vdots & & 1 & \ddots & \ddots & \ddots & & \vdots & X & \cdots & X \\ \vdots & \vdots & & & \ddots & X & \cdots & X & 0 & \vdots & & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & 1 & X & \cdots & X & X & \cdots & X \end{matrix} \right] \begin{matrix} \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} B-N+1 \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} T-B \end{matrix}$$

$$\begin{matrix} \underbrace{\qquad\qquad}_{k} & \underbrace{\qquad\qquad\qquad}_{N-1} & \underbrace{\qquad\qquad}_{B-N+1} \end{matrix} \tag{64}$$

We note that

$$\hat{\mathbf{H}}_{(l),2:k-l} = \left(\mathbf{M}^{-1}_{T-B,T-B}\mathbf{G}_{(l),T-B,:}\right)^T \tag{65}$$

where $\mathbf{M}^{-1}_{T-B,T-B}$ is acquired by taking the lower right $(T-B \times T-B)$ matrix of $\mathbf{M}^{-1}$, and $\mathbf{G}_{(l),T-B,:}$ is acquired by taking the last $T-B$ rows of $\mathbf{G}_{(l)}$. Since $\mathbf{M}^{-1}$ is an upper triangular matrix we have

$$\text{rank}(\mathbf{M}^{-1}_{T-B,T-B}) = T-B. \tag{66}$$

Further, since taking the last $T-B$ rows of $\mathbf{G}_{(l)}$ results in a matrix which a sub-matrix of $\mathbf{G}_{T-B,B}$ (by taking only $T-B$ columns out of the B columns of $\mathbf{G}_{T-B,B}$), recalling Property 6, it follows that

$$\text{rank}(\mathbf{G}_{(l),T-B,:}) = T-B. \tag{67}$$

Therefore, following Lemma 3, (66) and (67), we have that

$$\text{rank}(\hat{\mathbf{H}}_{(l),2:k-l}) \leq \min\left(\text{rank}(\mathbf{M}^{-1}_{T-B,T-B}), \text{rank}(\mathbf{G}_{(l),T-B,:})\right)$$

$$= T - B. \tag{68}$$

Therefore, since $\text{rank}(\hat{\mathbf{H}}_{(l)}) = T - B + 1$ and $\text{rank}(\hat{\mathbf{H}}_{(l),2:k-l}) \leq T - B$, it follows that $\hat{\mathbf{h}}_{(l),0}$ is linearly independent with the following $k - l + 1$ columns.

2) **Condition R1**

This condition requires that for for $0 \leq l \leq B - N$, the l-th column, $\mathbf{h}_l^{(l)}$ of $\mathbf{H}^{(l)}$ should be linearly independent of the any set of $(N-1)$ columns taken from the set

$$\left\{\mathbf{h}_j^{(l)},\ l+1 \leq j \leq l+T\right\}.$$

To simplify notations, we note that for $0 \leq l \leq B - N$, we denote by $\bar{\mathbf{H}}^{(\mathbf{l})}$ as taking the last $T - l$ columns of $\mathbf{H}^{(l)}$. Hence, showing that the l-th column, $\mathbf{h}_l^{(l)}$ of $\mathbf{H}^{(l)}$ is linearly independent of the any set of $(N-1)$ columns taken from the set $\left\{\mathbf{h}_j^{(l)},\ l+1 \leq j \leq l+T\right\}$ is equivalent to show that $\bar{\mathbf{h}}_0$ is linearly independent of any set of $(N-1)$ columns taken from the other $T-l-1$ columns of $\bar{\mathbf{H}}^{(\mathbf{l})}$. $\bar{\mathbf{H}}^{(\mathbf{l})}$ can be denoted as

Denote $\mathcal{A} \subset \{2, \ldots, T\}$ the set of $N-1$ coordinates taken from $\bar{\mathbf{H}}^{(l)}_{:,2:T}$. Hence, we denote as $\bar{\mathbf{H}}^{(l)}_{:,\mathcal{A}}$ as the sub matrix formed from taking the columns matching these coordinates (i.e. the matrix of $N-1$ columns we need to show that $\bar{\mathbf{h}}_0$ is linearly independent of).

Proving condition R1 is thus equivalent to prove that column $\bar{\mathbf{H}}^{(l)}_{:,1}$ (the first columns in $\bar{\mathbf{H}}^{(l)}_{:,\{1,\mathcal{A}\}}$) is linearly independent of the other $N-1$ columns. This in turn is equivalent to show that

$$\text{rank}(\bar{\mathbf{H}}^{(l)}_{:,\mathcal{A}}) \leq \text{rank}(\bar{\mathbf{H}}^{(l)}_{:,\{1,\mathcal{A}\}}) - 1.$$

Hence, we carefully analyze next the rank of these matrices.

We note that $\bar{\mathbf{H}}^{(l)}$ is a $N+l \times T+1$ matrix therefore $\bar{\mathbf{H}}^{(l)}_{:,\mathcal{A}}$ is a $N+l \times N-1$ matrix. Hence

$$\text{rank}(\bar{\mathbf{H}}^{(l)}_{:,\mathcal{A}}) \leq N - 1. \tag{70}$$

We note that the $N-1 \times T-l-1$ upper right sub matrix of $\bar{\mathbf{H}}^{(l)}$, can be denoted as

$$\bar{\mathbf{H}}^{(l)}_{1:N-1,2:,T-l} = \begin{bmatrix}\bar{\mathbf{P}} & I_{N-1}\end{bmatrix}, \tag{71}$$

where $\bar{\mathbf{P}}$ is a $N-1 \times k-1-l$ sub matrix of $\mathbf{P}''$. This is the dash-dot matrix in (69).

Denote the set of w_1 vectors taken from indices $\{2, \ldots, k-l\}$ as $\mathcal{W}_1 \subset \{2, \ldots, k-l\}$, the set of w_2 vectors taken from indices $\{k-l+1, \ldots, T-l\}$ as $\mathcal{W}_2 \subset \{k-l+1, \ldots, T-l\}$. We note that $\mathcal{W} = \mathcal{W}_1 \cup \mathcal{W}_2$ is the set of $w = w_1 + w_2$ vectors taken from indices $\{2, \ldots, T-l\}$. We further denote the set $\tilde{\mathcal{W}}_2 \subset \{2, \ldots, N-1\}$ where this set indicates the location of the elements in $\mathcal{W}_2$ with respect to column $K - l + 1$.

Taking columns in $\mathcal{W}$ from $\bar{\mathbf{H}}^{(l)}_{1:N-1,2:,T-l-1}$ can be denoted as

$$\bar{\mathbf{H}}^{(l)}_{1:N-1,\mathcal{W}} = \begin{bmatrix}\bar{\mathbf{P}}_{\mathcal{W}_1} & I_{N-1,\mathcal{W}_2}\end{bmatrix}. \tag{72}$$

We show next that

$$\text{rank}(\bar{\mathbf{H}}^{(l)}_{1:N-1,\mathcal{W}}) = w. \tag{73}$$

Applying column operations, we can zero the w_2 rows corresponds to indices $\tilde{\mathcal{W}}_2$ in $\bar{\mathbf{P}}_{\mathcal{W}_1}$. Thus, we are left with a matrix composed taking columns in indices $\mathcal{W}_1$ and rows in $\{1 : N-1\} \cap \tilde{\mathcal{W}}_2$. Following Corollary 3, this matrix has rank which equals $\min(N - 1 - w_2, w_1)$.

An example where $w_1 = 3$ and $w_2 = 2$ is given in (74) (after applying some row swaps). Therefore, applying column operations (and some row swaps), we get that

$$\bar{\mathbf{H}}^{(l)}_{:,\mathcal{W}} = \begin{bmatrix} I_w \\ \tilde{\mathbf{H}}^{(l)}_{w+1:N+l,\mathcal{W}} \end{bmatrix}, \tag{75}$$

where $\tilde{\mathbf{H}}^{(l)}_{w+1:N-1,\mathcal{W}}$ are the $N + l - w$ rows resulting from the column operations (over $\mathbb{F}_q$).

We note that if $w = N - 1$ (i.e., none of the erasures occurred in indices $\{T - l + 1, \ldots, T + 1\}$) we have $\mathcal{A} = \mathcal{W}$, thus we have

$$\bar{\mathbf{H}}^{(l)}_{:,\{1,\mathcal{A}\}} = \begin{bmatrix} \mathbf{0}_{N-1\times 1} & I_{N-1} \\ \tilde{\mathbf{h}}_{0,l+1\times 1} & \tilde{\mathbf{H}}^{(l)}_{w+1:N+l,\mathcal{W}} \end{bmatrix} \tag{76}$$

where $\tilde{\mathbf{h}}_{0,l+1\times 1}$ is the $l+1 \times 1$ lower part of $\tilde{\mathbf{h}}_\mathbf{0}$ which is the outcome of the column operations that zeroed the upper $N-1$ part of this vector. Following Property 5 we have that $f_{l+1,l+1}(\alpha) \neq 0$ and since all column operations were performed over $\mathbb{F}_q$ it follows that $\tilde{f}_{l+1,l+1}(\alpha) \neq 0$. Thus, we have that

$$\text{rank}(\bar{\mathbf{H}}^{(l)}_{:,\{1,\mathcal{A}\}}) = N. \tag{77}$$

We note that if $w < N-1$ (i.e. some of the erasure occurred at the last $l+1$ indices), when applying column operations, not all upper $N-1$ elements of $\tilde{\mathbf{h}}_{0,l+1\times 1}$ are nulled. Since these elements are taken from $\mathbf{P}''$, it follows each one of them does not equal 0. Thus, again we get that

$$\text{rank}(\bar{\mathbf{H}}^{(l)}_{:,\{1,\mathcal{A}\}}) = N. \tag{78}$$

Therefore, since $\text{rank}(\bar{\mathbf{H}}^{(l)}_{:,\{1,\mathcal{A}\}}) = N$ and $\text{rank}(\bar{\mathbf{H}}^{(l)}_{:,\mathcal{A}}) \leq N-1$, it follows that $(\bar{\mathbf{H}}^{(l)}_{:,1})$ is linearly independent with any other $N-1$ columns taken from $\bar{\mathbf{H}}$.

3) **Condition B2**

We note that for $B - N + 1 \leq l \leq T - N + 1$, the set

$$\{\mathbf{h}_j,\ l \leq j \leq l + B - 1\}$$

can be viewed as set of B columns taken from the parity-check matrix associated with the (n, k) MDS code $\mathcal{C}''$ with generator matrix $\mathbf{G}''$. Examples of groups of B columns to be analyzed are depicted in (79) below for $l = B-N+1$ and $l = B-N+2$. Following Property 1, we note that $\mathbf{H}_{:,\{B-N+2,\ldots,n\}}$ (marked in solid in (79)), can be viewed as the parity matrix of $(B+N-1, N-1)$ MDS code. Following Corollary 4, any B columns from the parity of this code matrix are independent, it follows that for any $B - N + 1 \leq l \leq T - N + 1$, the set $\{\mathbf{h}_j,\ l \leq j \leq l + B - 1\}$ is independent.

4) **Condition R2**

Following the explanation for condition B2, and since $B \geq N$, it follows that condition R2 also holds.

$$\tilde{\mathbf{H}}^{(1)} = \mathbf{H}^{(l)}_{1:N+l,l:l+T}$$

$$= \begin{matrix} N+l \\ \left\{ \right. \end{matrix} \begin{matrix} N-1 \\ \left\{ \right. \end{matrix} \left[\begin{array}{ccccccccccc} -Y & \cdots & \cdots & \cdots & \cdots & \cdots & -Y & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & \cdots & 0 \\ -Y & \cdots & \cdots & \cdots & \cdots & \cdots & -Y & 0 & \cdots & 1 & 0 \\ -Y_\alpha & \cdots & \cdots & -Y_\alpha & -Y & \cdots & -Y & \vdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & 0 \\ f_{l+1,l+1}(\alpha) & -Y_\alpha & \cdots & -Y_\alpha & -Y & \cdots & -Y & \cdots & \cdots & \cdots & 1 \end{array} \right] .$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad}_{B-N+1-l} \quad \underbrace{\qquad\qquad}_{T-B} \quad \underbrace{\qquad}_{N-1} \quad \underbrace{\qquad}_{l+1}$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{T-l}$$

(69)

$$\tilde{\mathbf{H}}^{(l)}_{1:N-1,\mathcal{W}} = \begin{bmatrix} -Y & -Y & -Y & 1 & 0 \\ \vdots & \vdots & \vdots & 0 & 1 \\ \vdots & \vdots & \vdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -Y & \cdots & -Y & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ -Y & -Y & -Y & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -Y & \cdots & -Y & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \vdots & \vdots \\ 0 & 0 & 1 & \vdots & \vdots \\ -\tilde{Y} & \cdots & -\tilde{Y} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\tilde{Y} & \cdots & -\tilde{Y} & 0 & 0 \end{bmatrix} .$$

(74)

$$\mathbf{H} = \left[-\tilde{\mathbf{P}}^T \,\middle|\, \mathbf{I}_{B\times B} \right]$$

$$= \left[\begin{array}{cccc|ccccccccc} -Y & \cdots & \cdots & \cdots & \cdots & \cdots & -Y & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & 1 & 0 & \cdots & \cdots & \cdots & 0 \\ -Y & \cdots & \cdots & \cdots & \cdots & \cdots & -Y & 0 & 0 & \ddots & \cdots & \vdots & \vdots & 0 \\ f_{1,1}(\alpha) & -Y_\alpha & \cdots & -Y_\alpha & \cdots & \cdots & -Y & \vdots & \vdots & & 1 & \ddots & \vdots & 0 \\ f_{1,2}(\alpha) & f_{2,2}(\alpha) & -Y_\alpha & \vdots & \cdots & \cdots & -Y & \vdots & \vdots & & 0 & \ddots & \vdots & 0 \\ \vdots & \vdots & \ddots & -Y_\alpha & \vdots & \vdots & -Y & \vdots & \vdots & & & 0 & \ddots & 0 \\ f_{1,B-N+1}(\alpha) & f_{2,B-N+1}(\alpha) & \cdots & f_{B-N+1,B-N+1}(\alpha) & -Y & \cdots & -Y & 0 & 0 & \cdots & \cdots & 0 & 0 & 1 \end{array} \right]$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{N-1}$$

(79)

V. Concluding remarks

In this paper, we study streaming codes over an erasure channel whose erasure pattern in every sliding window of size W is either a burst erasure of maximum length B or multiple arbitrary erasures of maximum total count N where these parameters meet (5). While the capacity of this channel was derived (separately) by Fong et al. in [12] and Krishnan et al. [13], no explicit code construction which achieves this capacity with low field size (for any coding parameters) was suggested.

In this paper, we presented an explicit code construction that achieves the capacity of the channel mentioned above with a field size that scales quadratically in the delay constraint. The construction relies on properties of MDS codes from a base field (whose size scales linearly with the delay constraint) while replacing some of the elements in the generator matrix with elements for an extension field, thus resulting in a final field size that scales quadratically with the delay constraint.

An open question is whether there exists an explicit construction for capacity-achieving streaming codes (for all admissible parameters) with smaller field size (i.e., a field size

the scales linearly with the delay constraint). While this seems possible to some specific values of (W, B, N), finding a coding scheme for any (admissible) values is an interesting avenue.

Appendix A
Explicit description of $\tilde{\mathbf{G}}$

Writing (43) explicitly is described in (80). Recalling that $\mathbf{M}$ is an upper triangular matrix it follows that $\mathbf{M}^{-1}$ is also upper triangular. Since $\mathbf{G}_{:,1:T} = \mathbf{G}'_{:,1:T}$ it follows that the first T columns of $\tilde{\mathbf{G}}$ equal the first T columns of $\mathbf{G}'$.

We denote with $f_{i,j} = \mathbf{M}^{-1}_{i,:}\mathbf{G}_{:,j+T}$ where $1 \leq i \leq B+N-1$ and $i \leq j \leq B-N+1$. We note that only these elements in $\mathbf{G}$ include $\alpha \in \mathbb{F}_{q^2} \setminus \mathbb{F}_q$. We further note that since $\mathbf{M}^{-1}_{i,i} = 1$ (for any $i \in \{1, \ldots, k\}$, this element is not nulled in this multiplication.

We denote with $Y_\alpha = \mathbf{M}^{-1}_{i,:}\mathbf{G}_{:,j+T}$ where $2 \leq i \leq B+N-1$ and $1 \leq j \leq i-1$, i.e., these elements does not equal to the elements Y at this location in $\mathbf{G}'$ (due to replacing the upper right $B-N+1 \times B-N+1$ matrix in $\mathbf{G}'$ with $\mathbf{I}\cdot\alpha$).[8] We note though, that since $\mathbf{M}^{-1}$ is upper triangular, these elements do not contain α.

Appendix B
Proof of Lemma 3

The space spanned by the columns of $\mathbf{AB}$ is the space $\mathcal{S}$ of all vectors s that can be written as linear combinations of the columns of $\mathbf{AB}$:

$$s = (\mathbf{AB})\, v \tag{81}$$

where v is the $M \times 1$ vector of coefficients of the linear combination. We can also write

$$s = \mathbf{A}\,(\mathbf{B}v) \tag{82}$$

where $\mathbf{B}v$ is an $L \times 1$ vector (being a product of an $L \times M$ matrix and an $M \times 1$ vector). Thus, any vector $s \in \mathcal{S}$ can be written as a linear combination of the columns of $\mathbf{A}$, with coefficients taken from the vector $\mathbf{B}v$. As a consequence, the space $\mathcal{S}$ is no larger than the span of the columns of $\mathbf{A}$, whose dimension is $\mathrm{rank}(\mathbf{A})$. This implies that the dimension of $\mathcal{S}$ is less than or equal to $\mathrm{rank}(\mathbf{A})$. Since the dimension of $\mathcal{S}$ is the rank of $\mathbf{AB}$, we have

$$\mathrm{rank}(\mathbf{AB}) \leq \mathrm{rank}(\mathbf{A}) \tag{83}$$

Now, the space spanned by the rows of AB is the space $\mathcal{T}$ of all vectors t that can be written as linear combinations of the rows of $\mathbf{AB}$:

$$t = u\,(\mathbf{AB}) \tag{84}$$

where u is the $1 \times K$ vector of coefficients of the linear combination. We can also write

$$t = (u\mathbf{A})\,\mathbf{B} \tag{85}$$

[8]Again, we note that Y and Y_α are not a constant element but rather represent (potentially different) element taken from $\mathbb{F}_q$.

where $u\mathbf{A}$ is a $1 \times L$ vector (being a product of a $1 \times K$ vector and a $K \times L$ matrix). Thus, any vector $t \in \mathcal{T}$ can be written as a linear combination of the rows of $\mathbf{B}$, with coefficients taken from the vector $u\mathbf{A}$. As a consequence, the space $\mathcal{T}$ is no larger than the span of the rows of $\mathbf{B}$, whose dimension is $\mathrm{rank}(\mathbf{B})$. This implies that the dimension of $\mathcal{T}$ is less than or equal to $\mathrm{rank}(\mathbf{B})$. Since the dimension of T is the rank of $\mathbf{AB}$, we have

$$\mathrm{rank}(\mathbf{AB}) \leq \mathrm{rank}(\mathbf{B}) \tag{86}$$

(83) and (86) imply that

$$\mathrm{rank}(\mathbf{AB}) \leq \min\left(\mathrm{rank}(\mathbf{A}), \mathrm{rank}(\mathbf{B})\right) \tag{87}$$

Appendix C
Proof of Proposition 6

Proof. Recall that $\mathbf{G}_{T-B,B}$ the lower right $(T-B \times B)$ matrix of $\mathbf{G}$ marked in solid in (88).

$$\mathbf{G} = \left[\begin{array}{cccccccccccc}
1 & X & \cdots & X & 0 & 0 & 0 & \cdots & 0 & \alpha & \cdots & 0 \\
0 & 1 & X & \cdots & X & 0 & 0 & \cdots & 0 & 0 & \ddots & 0 \\
0 & 0 & \ddots & \ddots & \ddots & \ddots & 0 & \cdots & 0 & 0 & \cdots & \alpha \\
\vdots & \vdots & & 1 & \ddots & \ddots & \ddots & & \vdots & X & \cdots & X \\
\vdots & \vdots & & & \ddots & X & \cdots & X & 0 & \vdots & & \vdots \\
0 & 0 & \cdots & \cdots & 0 & 1 & X & \cdots & X & X & \cdots & X
\end{array}\right] \Bigg\} T-B$$
$$\mathbf{G}_{T-B,B} \tag{88}$$

First step in analyzing the rank of $\mathbf{G}_{T-B,B}$ is to note that $\mathbf{G}_{T-B,B} = \mathbf{G}'_{T-B,B}$. Now, from the definition of $\mathbf{G}'$ we have that $\mathbf{G}' = \mathbf{M}\mathbf{G}''$. Hence, $\mathbf{G}_{T-B,B}$ can be viewed as multiplying the $(T-B \times T-B)$ lower right part of $\mathbf{M}$ (which we denote as $\mathbf{M}_{T-B,T-B}$) with the $(T-B \times B)$ lower right part of $\mathbf{G}''$ (which we denote as $\mathbf{G}''_{T-B,B}$).

$$\mathbf{G}_{T-B,B} = \mathbf{M}_{T-B,T-B} \times \mathbf{G}''_{T-B,B}$$
$$= \underbrace{\left[\begin{array}{cccccccc}
1 & Y' & \cdots & Y' & 0 & \cdots & \cdots & 0 \\
0 & 1 & Y' & \cdots & Y' & 0 & \cdots & 0 \\
\vdots & \cdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & 1 & Y' & \cdots & Y' & 0 \\
0 & \cdots & 0 & 0 & 1 & Y' & \cdots & Y' \\
0 & \cdots & 0 & 0 & 0 & \ddots & \ddots & Y' \\
0 & \cdots & 0 & 0 & 0 & 0 & 1 & Y' \\
0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right]}_{\mathbf{M}} \times$$

$$
\begin{aligned}
\tilde{\mathbf{G}} &= \mathbf{M}^{-1}\mathbf{G} \\
&= \begin{bmatrix} \mathbf{M}_{1,:}^{-1} \\ \mathbf{M}_{2,:}^{-1} \\ \vdots \\ \mathbf{M}_{k,:}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{G}_{:,1} \cdots \mathbf{G}_{:,n} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{M}_{1,:}^{-1} \\ \mathbf{M}_{2,:}^{-1} \\ \vdots \\ \mathbf{M}_{k,:}^{-1} \end{bmatrix} \begin{bmatrix} 1 & X & \cdots & X & 0 & 0 & 0 & \cdots & 0 & \overbrace{\alpha & \cdots & 0}^{B-N+1} \\ 0 & 1 & X & \cdots & X & 0 & 0 & \cdots & 0 & 0 & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & 0 & \cdots & 0 & 0 & \cdots & \alpha \\ \vdots & \vdots & & 1 & \ddots & \ddots & \ddots & & \vdots & X & \cdots & X \\ \vdots & \vdots & & & \ddots & X & \cdots & X & 0 & \vdots & & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & 1 & X & \cdots & X & X & \cdots & X \end{bmatrix} \Bigg\} B-N+1 \\
&= \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & Y & \cdots & Y & f_{1,1}(\alpha) & \cdots & \cdots & f_{1,B-N+1}(\alpha) \\ 0 & 1 & 0 & \cdots & \cdots & \cdots & 0 & Y & \cdots & Y & Y_\alpha & f_{2,2}(\alpha) & \cdots & f_{2,B-N+1}(\alpha) \\ 0 & 0 & \ddots & \cdots & \vdots & \vdots & 0 & Y & \cdots & Y & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \vdots & \vdots & 0 & Y & \cdots & Y & Y_\alpha & \cdots & Y_\alpha & f_{B-N+1,B-N+1}(\alpha) \\ \vdots & \vdots & & & \ddots & \vdots & 0 & Y & \cdots & Y & Y & \cdots & Y & Y \\ 0 & 0 & \cdots & \cdots & 0 & 1 & 0 & Y & \cdots & Y & Y & \cdots & Y & Y \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 1 & Y & \cdots & Y & Y & \cdots & Y & Y \end{bmatrix}. \\
&\qquad\qquad \underbrace{\qquad\qquad}_{k} \quad \underbrace{\qquad\qquad}_{N-1} \qquad\qquad \underbrace{\qquad\qquad\qquad}_{B+N-1}
\end{aligned} \tag{80}
$$

$$
\underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 & Y & \cdots & Y \\ 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & Y & \cdots & Y \\ \vdots & \cdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & Y & \cdots & Y \\ 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & Y & \cdots & Y \\ 0 & \cdots & 0 & 0 & 0 & \ddots & \ddots & 0 & Y & \cdots & Y \\ 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 0 & Y & \cdots & Y \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 & Y & \cdots & Y \end{bmatrix}}_{\mathbf{G}''} \quad \overbrace{\qquad\qquad}^{\mathbf{P}''_{k\times B}} \tag{89}
$$

Since $\mathbf{G}''_{T-B,B}$ is a sub-matrix of $\mathbf{P}''$, following Corollary 2, it has a full rank. Thus,

$$\operatorname{rank}(\mathbf{G}''_{T-B,B}) = \min(T-B, B). \tag{90}$$

Matrix $\mathbf{M_{T-B,T-B}}$ is a square upper triangular matrix and thus we have

$$\operatorname{rank}(\mathbf{M}_{T-B,T-B}) = T-B. \tag{91}$$

Following Lemma 3 we have that

$$\begin{aligned}\operatorname{rank}(\mathbf{G}_{T-B,B}) \leq &\min\left(\operatorname{rank}(\mathbf{M}_{T-B,T-B}), \operatorname{rank}(\mathbf{G}''_{T-B,B})\right) \\ &\min(T-B, B).\end{aligned} \tag{92}$$

Further, following Corollary 7 (and recalling that $\mathbf{M}_{T-B,T-B}$ is a square, full rank matrix) we have

$$\begin{aligned}\operatorname{rank}(\mathbf{G}_{T-B,B}) &\geq \operatorname{rank}(\mathbf{G}''_{T-B,B}) \\ &= \min(T-B, B).\end{aligned} \tag{93}$$

Combining (92) and (93), we conclude that $\operatorname{rank}(\mathbf{G}_{T-B,B}) = \min(T-B, B)$.

Further, we note that the sub-matrix of $\mathbf{G}_{T-B,B}$ which is formed by taking any w columns of $\mathbf{G}_{T-B,B}$ has a rank of $\min(T-B, w)$ since taking w columns is equivalent to multiply $\mathbf{M}_{T-B,T-B}$ (which is a square full matrix) with w columns from $\mathbf{G}''_{T-B,B}$.

□

Appendix D
Proof of Proposition 7

Proof. Recall that $\mathbf{M}^{-1}_{B,T-B}$ is the upper right $(B \times T-B)$ matrix of $\mathbf{M}^{-1}$ of marked in solid in (94).

$$\mathbf{M}^{-1} = \begin{bmatrix} 1 & Y'' & \cdots & \overbrace{Y''}^{\mathbf{M}_{B,T-B}^{-1}} & Y'' & \cdots & \cdots & Y'' \\ 0 & 1 & Y'' & \cdots & Y'' & Y'' & \cdots & Y'' \\ \vdots & \cdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & Y'' & \cdots & Y'' & Y'' \\ 0 & \cdots & 0 & 0 & 1 & Y'' & \cdots & Y'' \\ \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \ddots & Y'' \\ 0 & \cdots & 0 & 0 & 0 & 0 & 1 & Y'' \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ \} B-N+1 \\ \\ \} N-1 \\ \\ \\ \end{matrix} \tag{94}$$

Since $\mathbf{M}\mathbf{M}^{-1} = \mathbf{I}$ we have

$$\underbrace{\begin{bmatrix} \overbrace{1}^{\mathbf{M}_{1:B-N+1,1:B}} & Y' & \cdots & Y' & 0 & \cdots & \cdots & 0 \\ 0 & 1 & Y' & \cdots & Y' & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & Y' & \cdots & Y' & 0 \\ 0 & \cdots & 0 & 0 & 1 & Y' & \cdots & Y' \\ 0 & \cdots & 0 & 0 & 0 & \ddots & \ddots & Y' \\ 0 & \cdots & 0 & 0 & 0 & 0 & 1 & Y' \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{M}} \times \underbrace{\begin{bmatrix} 1 & Y'' & \cdots & \overbrace{Y''}^{\mathbf{M}_{B,T-B}^{-1}} & Y'' & \cdots & \cdots & Y'' \\ 0 & 1 & Y'' & \cdots & Y'' & Y'' & \cdots & Y'' \\ \vdots & \cdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & Y'' & \cdots & Y'' & Y'' \\ 0 & \cdots & 0 & 0 & 1 & Y'' & \cdots & Y'' \\ 0 & \cdots & 0 & 0 & 0 & \ddots & \ddots & Y'' \\ 0 & \cdots & 0 & 0 & 0 & 0 & 1 & Y'' \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{M}^{-1}}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & \overbrace{0}^{\mathbf{M}_{B,T-B}^{-1}} & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{95}$$

which means

$$\mathbf{M}_{1:B-N+1,1:B} \times \mathbf{M}_{B,T-B}^{-1} = \mathbf{0}_{B-N+1 \times B} \tag{96}$$

which means that the first $B-N+1$ rows of $\mathbf{M}_{B,T-B}^{-1}$ can be expressed as linear combination of the last $N-1$ rows. □

REFERENCES

[1] V. Cisco, "Cisco visual networking index: Forecast and trends, 2017–2022," *White Paper*, vol. 1, 2018.

[2] R. Gallager, "Low-density parity-check codes," *IRE Transactions on information theory*, vol. 8, no. 1, pp. 21–28, 1962.

[3] D. J. MacKay and R. M. Neal, "Near shannon limit performance of low density parity check codes," *Electronics letters*, vol. 32, no. 18, pp. 1645–1646, 1996.

[4] M. Luby, "LT codes," in *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.* IEEE, 2002, pp. 271–280.

[5] A. Shokrollahi, "Raptor codes," *IEEE transactions on information theory*, vol. 52, no. 6, pp. 2551–2567, 2006.

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

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

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

[9] N. Adler and Y. Cassuto, "Burst-erasure correcting codes with optimal average delay," *IEEE Transactions on Information Theory*, vol. 63, no. 5, pp. 2848–2865, 2017.

[10] A. Badr, A. Khisti, W.-T. Tan, and J. Apostolopoulos, "Streaming codes for channels with burst and isolated erasures," in *2013 Proceedings IEEE INFOCOM*. IEEE, 2013, pp. 2850–2858.

[11] M. Rudow and K. Rashmi, "Streaming codes for variable-size arrivals," in *2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2018, pp. 733–740.

[12] S. L. Fong, A. Khisti, B. Li, W.-T. Tan, X. Zhu, and J. Apostolopoulos, "Optimal streaming codes for channels with burst and arbitrary erasures," *IEEE Transactions on Information Theory*, 2019.

[13] M. N. Krishnan and P. V. Kumar, "Rate-optimal streaming codes for channels with burst and isolated erasures," in *2018 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2018, pp. 1809–1813.

[14] D. Dudzicz, S. L. Fong, and A. Khisti, "An explicit construction of optimal streaming codes for channels with burst and arbitrary erasures," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 12–25, 2020.

[15] M. N. Krishnan, D. Shukla, and P. V. Kumar, "Rate-optimal streaming codes for channels with burst and random erasures," *IEEE Transactions on Information Theory*, vol. 66, no. 8, pp. 4869–4891, 2020.

[16] M. Nikhil Krishnan, V. Ramkumar, M. Vajha, and P. Vijay Kumar, "Simple streaming codes for reliable, low-latency communication," *IEEE Communications Letters*, vol. 24, no. 2, pp. 249–253, 2020.

[17] A. Badr, P. Patil, A. Khisti, W.-T. Tan, and J. Apostolopoulos, "Layered constructions for low-delay streaming codes," *IEEE Transactions on Information Theory*, vol. 63, no. 1, pp. 111–141, 2017.

[18] W. C. Huffman and V. Pless, *Fundamentals of Error-Correcting Codes*. Cambridge University Press, 2003.

[19] R. M. Roth and A. Lempel, "On mds codes via cauchy matrices," *IEEE transactions on information theory*, vol. 35, no. 6, pp. 1314–1319, 1989.

[20] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes*. Elsevier, 1977, vol. 16.

[21] F. E. Hohn, *Elementary matrix algebra*, 2013.

Elad Domanovitz (IEEE member) received the B.Sc. degree (cum laude), M.Sc. and Ph.D. degrees in 2005, 2011, and 2020 respectively, in Electrical Engineering from Tel Aviv University, Israel. He is currently a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada. His research interests include information theory, communication theory and statistical signal processing.

Slias L. Fong (IEEE member) received the B.Eng., M.Phil., and Ph.D. degrees in Information Engineering from The Chinese University of Hong Kong (CUHK), Hong Kong, in 2005, 2007, and 2011, respectively. From 2011 to 2013, he was a Post-Doctoral Fellow with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong. From 2013 to 2014, he was a Post-Doctoral Associate with the Department of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA. From 2014 to 2017, he was a Research Fellow with the Department of Electrical and Computer Engineering, National University of Singapore (NUS), Singapore. From 2017 to 2019, he was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada. He is currently a Staff Engineer with Qualcomm Technologies, Inc., NJ, USA. His research interests include information theory and its applications to communication systems such as relay networks, wireless networks, and energy-harvesting channels.

Ashish Khisti (IEEE member) received the B.A.Sc. degree from the Engineering Science Program, University of Toronto, in 2002, and the master's and Ph.D. degrees from the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2004 and 2008, respectively. Since 2009, he has been on the Faculty in the Electrical and Computer Engineering (ECE) Department, University of Toronto, where he was an Assistant Professor from 2009 to 2015, an Associate Professor from 2015 to 2019, and is currently a Full Professor. He also holds the Canada Research Chair of information theory with the ECE Department. His current research interests include theory and applications of machine learning and communication networks. He is also interested in interdisciplinary research involving engineering and healthcare.