New Bounds and Constructions for Variable Packet-Error Coding

Xiangliang Kong, Xin Wang, Ron M. Roth, and Itzhak Tamo

***Abstract*—In this paper, we consider the problem of variable packet-error coding, which emerges in network communication scenarios where a source transmits information to a destination through multiple disjoint paths. The objective is to design codes with dynamic error-correcting capabilities that adapt to varying numbers of errors. Specifically, we first provide several bounds on the rate–distortion trade-off for general variable packet-error coding schemes. Then, we present two explicit constructions of variable packet-error coding schemes. The first construction uses higher-order MDS codes and provides a coding scheme that achieves a better rate–distortion trade-off compared to known results for general parameter regimes. The second construction is based on a variant of the repetition code and yields a coding scheme with an optimal rate–distortion trade-off, with respect to our bound, for certain parameter regimes.**

I. Introduction

Since the publication of [1], network coding theory has emerged as one of the most influential and dynamic branches of coding theory and information theory. Over the years, researchers from different fields have shown significant interest in network coding theory due to its broad range of applications in practical scenarios. Within this field, network error-correcting (NEC) codes were first introduced in [2] to provide resistance against adversarial errors in network communications. Since their introduction, NEC codes have been extensively studied; see, for example, [3]–[5] and the references in the survey [6].

In this paper, we consider a new type of NEC coding problem introduced and studied in [7] and [8], called *variable packet-error coding* (VPEC). This problem arises in a communication scenario where a source transmits information to a destination over multiple non-intersecting paths in a network (which can be viewed as multiple transmission channels). During transmission, messages sent through these paths may encounter interference, resulting in adversarial errors, which are referred to as *packet errors*. Classical network error-correcting codes typically focus on ensuring robustness against the maximum possible number of packet errors. However, in many real-world systems, such as those where errors arise from adversarial jammers, attacks may occur only sporadically, and the network often operates without errors. This makes the worst-case approach overly pessimistic. Thus, the goal of variable packet-error coding is to design a coding system that meets performance objectives when the maximum number of errors are present while ensuring improved performance when there are fewer or no errors.

In VPEC, one needs to construct codes whose performance dynamically improves as the number of errors decreases. This is not achievable with the conventional approach, where maximum distance separable (MDS) codes with certain minimum distances are typically employed. Since the code parameters are designed to recover the source when the maximum number of errors occur, MDS codes offer no performance improvement when fewer errors are present. Similar problems have also been studied recently for storage codes; see, for example, [9]–[12], where the code rate is required to vary over time in response to the variations in failure rates of storage devices. The VPEC problem is also closely related to the multiple-description (MD) problem [13] in network information theory (see also [14], [15]), and coding approaches in cryptography, such as message authentication codes (MACs) [16, Sec. 6.1] (see also [7]).

Based on prior works [17] and [7], which showed that source–channel separation is not optimal for the VPEC problem, Fan, Kosut, and Wagner appropriately addressed the VPEC problem using rate–distortion theory in [8]. Specifically, they consider the scenario where a source sequence is encoded into N packets (or messages) at a given rate R, with up to T packets potentially

This work was supported in part by the European Research Council (ERC) under Grant 852953 and in part by the Israel Science Foundation under Grant 2369/24.

X. Kong (rongxlkong@gmail.com) is with the State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. He was with the School of Electrical & Computer Engineering, Tel Aviv University, Tel Aviv 6139001, Israel, when this manuscript was submitted. X. Wang (xinw@suda.edu.cn) is with the Department of Mathematics, Soochow University, Suzhou 215005, Jiangsu, China. Ron M. Roth (ronny@cs.technion.ac.il) is with the Computer Science Department, Technion, Haifa 3200003, Israel. I. Tamo (tamo@tauex.tau.ac.il) is with the School of Electrical & Computer Engineering, Tel Aviv University, Tel Aviv-Yafo 6997801, Israel.

Part of the work in this paper has been accepted for presentation at the 2025 IEEE International Symposium on Information Theory (ISIT).

being adversarially altered by the network. The decoder receives N packets without any knowledge of how many or which packets were altered, only knowing that the total number of altered packets is at most T. The decoder then reconstructs the source sequence, and a distortion measure is applied to quantify the discrepancy between the source and its reconstruction. Hence, from this perspective, the goal in VPEC is to design codes that guarantee a specified distortion level when T errors occur, with progressively lower distortion levels as the number of errors decreases.

As in [8], we study the VPEC problem under the erasure distortion measure. In this setting, the per-symbol distortion is zero if the reconstruction symbol matches the corresponding source symbol, one if the reconstruction symbol is an "erasure" symbol, and infinity otherwise. Hence, as long as the decoder does not incorrectly guess any source symbol, the overall distortion of a sequence is defined as the fraction of erasures in the reconstruction.

In the same spirit as [7] and [8], this paper focuses on the trade-off between the rate R of the messages encoded in each packet and the (erasure) distortion D that quantifies the reconstruction accuracy in the VPEC problem. We not only study the general achievability of rate–distortion pairs, but also investigate the performance of rate–distortion pairs for certain classes of codes. Specifically, our contributions are summarized next:

- **Bounds on the rate–distortion trade-off:** We establish several lower bounds on the trade-off between the rate R of each packet and the distortion D, in terms of the number of packets N and the number of packet errors T. These bounds include the first part of Theorem 1 in [8] as a special case, and one of them is shown to be optimal for certain parameter regimes based on our construction.
- **Constructions:** We present two explicit constructions of VPEC codes. The first construction utilizes higher-order MDS codes introduced in [18] (see also [19]) and yields VPEC codes that achieve improved rate–distortion trade-offs compared to the polytope code in [8] for general parameter regimes. The second construction is based on a variant of repetition codes and leads to VPEC codes with $N = 2T + 1$ packets, capable of correcting T packet errors for any positive integer T. These codes achieve the bound on the rate–distortion trade-off for $0 \leqslant D \leqslant T/N$. Moreover, this construction includes a decoding algorithm with a time complexity of $O(N^3)$.

The rest of the paper is structured as follows. In Section II, we provide a formal description of the VPEC problem along with our main results, and include a comparison with previous works. Section III is devoted to proving our bounds on the rate–distortion trade-off for the VPEC problem. In Section IV, we present our constructions of coding schemes. Finally, Section V concludes the paper by highlighting several open problems for future research.

II. Problem formulation and main results

A. Notation

Let $\mathbb{Z}$ denote the set of all integers. We use $\mathbb{Z}^+$ to denote the set of all positive integers, and for $L \in \mathbb{Z}^+$, we use $\mathbb{Z}/L$ to denote the set of rationals whose denominator divides L. For integers $1 \leqslant m \leqslant n$, we denote $[m:n] \triangleq \{m, m+1, \ldots, n\}$ and $[n] \triangleq [1:n]$, and the notation $\mathbb{Z}_m$ will stand for the ring of integers modulo m. For a subset A of $[n]$ and a positive integer t, we use $\binom{A}{\leqslant t}$ to denote the family of all subsets of A with size at most t. Given an alphabet Σ of size at least 2, for words $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ in Σ^n, we use the notation

$$d_H(\mathbf{x}, \mathbf{y}) \triangleq |\{i \in [n] : x_i \neq y_i\}|$$

for the Hamming distance between $\mathbf{x}$ and $\mathbf{y}$, and we denote by $B(\mathbf{x}, r)$ the Hamming ball in Σ^n centered at $\mathbf{x}$ with radius r. Given a code $\mathcal{C} \subseteq \Sigma^n$, we denote $d_H(\mathcal{C})$ as the minimum distance of the code $\mathcal{C}$. For a codeword $\mathbf{c} \in \mathcal{C}$ and a subset $S \subseteq [n]$, let $\mathbf{c}|_S$ denote the word obtained by projecting the coordinates of $\mathbf{c}$ onto S, and define $\mathcal{C}|_S \triangleq \{\mathbf{c}|_S : \mathbf{c} \in \mathcal{C}\}$. Given $d \in [n]$, we call $\mathcal{C} \subseteq \Sigma^n$ an (n, d)-anticode if $d_H(\mathbf{c}, \mathbf{c}') \leqslant d$ for any $\mathbf{c}, \mathbf{c}' \in \mathcal{C}$.

Let q be a prime power. We use $\mathbb{F}_q$ to denote the finite field of size q and $\mathbb{F}_q^n$ to denote the n-dimensional linear space over $\mathbb{F}_q$. We call a code $\mathcal{C} \subseteq \mathbb{F}_q^n$ a linear $[n, k]$ code over $\mathbb{F}_q$ if $\mathcal{C}$ is a k-dimensional linear subspace of $\mathbb{F}_q^n$.

Throughout the paper, we use standard asymptotic notation, as follows. Let $f(\cdot)$ and $g(\cdot)$ be two non-negative functions defined on the positive integers. We say that $f = O(g)$ (or $g = \Omega(f)$) if there is some constant $c \geqslant 0$ such that $f(n) \leqslant cg(n)$ for sufficiently large n.

B. Problem formulation and previous results

Let Σ denote the alphabet of the source and let N denote the number of channels, i.e., the number of packets that are sent in each transmission. Given a source sequence $\mathbf{x} \in \Sigma^k$ of length k, the encoder creates N packets via the following encoding functions

$$f_\ell : \Sigma^k \longrightarrow \Sigma^{kR_\ell}, \ \ell \in [N],$$

for some $R_\ell > 0$. We refer to R_ℓ as the *rate of the ℓ-th packet (ℓ-th channel)*. Then, the encoder sends the following list of N packets

$$\mathbf{f}_N(\mathbf{x}) \triangleq (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_N(\mathbf{x})), \tag{1}$$

and we sometimes use the notation C_j for $f_j(\mathbf{x})$. After transmission through the channels, for any $j \in [N]$, the content C_j of the j-th packet may be altered to some different value $\tilde{C}_j \in \Sigma^{kR_\ell}$. Let $\tilde{\Sigma}$ denote the reconstruction alphabet, which hereafter will be $\Sigma \cup \{e\}$, where e denotes the erasure symbol. Then, the decoder employs a function

$$g: \prod_{\ell=1}^{N} \Sigma^{kR_\ell} \longrightarrow \tilde{\Sigma}^k \tag{2}$$

to reproduce the source given the received packets. The fidelity of the reproduction is measured using distortion measure

$$\Delta: \Sigma^k \times \tilde{\Sigma}^k \to [0, \infty],$$
$$\Delta(\mathbf{x}, \tilde{\mathbf{x}}) = \frac{1}{k} \sum_{j=1}^{k} \Delta(x_j, \tilde{x}_j),$$

where $\Delta(x_j, \tilde{x}_j)$ is the single-letter distortion between x_j and $\tilde{x}_j$. As in [8], we limit our focus to the *erasure distortion measure* [20, Page 338]: for $x \in \Sigma$ and $\tilde{x} \in \tilde{\Sigma}$,

$$\Delta(x, \tilde{x}) \triangleq \begin{cases} 0, & \text{if } x = \tilde{x}; \\ 1, & \text{if } \tilde{x} = e; \\ \infty, & \text{otherwise.} \end{cases} \tag{3}$$

The distortion of the coding scheme defined by the pair $(\mathbf{f}_N, g)$ under at most T packet errors is denoted by

$$D_T(\mathbf{f}_N, g) \triangleq \max_{\mathbf{x} \in \Sigma^k} \max_{A \in \binom{[N]}{\leqslant T}} \max_{\tilde{C}_A} \Delta\big(\mathbf{x}, g(C_{[N]\setminus A}, \tilde{C}_A)\big).$$

Here, $g(C_{[N]\setminus A}, \tilde{C}_A)$ denotes the decoder's output when it receives $C_\alpha = f_\alpha(\mathbf{x})$ for all $\alpha \in [N] \setminus A$ and $\tilde{C}_\alpha \in \Sigma^{kR_\alpha}$ for all $\alpha \in A$. We are interested in the rate–distortion trade-off when T can take two different values, T_0 and T_1.

Definition II.1 ([8, Definition 1])**.** The rate–distortion vector $(R_1, \ldots, R_N, D_0, D_1)$ is *achievable* with respect to (T_0, T_1), if for all $\varepsilon > 0$, there exists a coding scheme $(\mathbf{f}_N, g)$ for a source of blocklength k, for some k, satisfying

$$\begin{aligned} &\frac{1}{k} \log_{|\Sigma|} |f_\ell| \leqslant R_\ell + \varepsilon, \ \ell \in [N]; \\ &D_{T_0}(\mathbf{f}_N, g) \leqslant D_0 + \varepsilon; \\ &D_{T_1}(\mathbf{f}_N, g) \leqslant D_1 + \varepsilon. \end{aligned} \tag{4}$$

In [8], Fan, Kosut and Wagner focused on the case where $T_0 = 0$, $T_1 = T$, and all R_ℓ are equal. Specifically, in their setting, the coding scheme has rate R for each packet, allows lossless reconstruction when there are $T_0 = 0$ packet errors, and ensures an erasure distortion of at most D when there are at most $T_1 = T$ packet errors. We refer to this requirement of the coding scheme as the T-VPEC. We say the rate–distortion pair (R, D) is *achievable* for T-VPEC if the rate–distortion vector $(R, \ldots, R, 0, D)$ is achievable according to Definition II.1 with respect to $(T_0, T_1) = (0, T)$.

Given N, T, a source of blocklength k, and a coding scheme defined by the encoder-decoder pair $(\mathbf{f}_N, g)$ of the form (1) and (2), suppose it satisfies (4) with parameters $R_\ell = R$, $T_0 = D_0 = \varepsilon = 0$, $D_1 = D$, and $T_1 = T$. Let $\mathcal{C}_{(\mathbf{f}_N, g)}$ denote the image of $\mathbf{f}_N$. We refer to such a set $\mathcal{C}_{(\mathbf{f}_N, g)} \subseteq (\Sigma^{kR})^N$ as a *T-VPEC code* over Σ with parameters (N, k, R, D). When $\Sigma = \mathbb{F}_q$ and each f_ℓ is a linear map from $\mathbb{F}_q^k$ to $\mathbb{F}_q^{kR}$ for every $\ell \in [N]$, we say that $\mathcal{C}_{(\mathbf{f}_N, g)}$ is a linear T-VPEC code over $\mathbb{F}_q$. For simplicity, we usually omit the subscript when the coding scheme $(\mathbf{f}_N, g)$ is clear from the context. Moreover, in the following sections, we use the notation $\mathcal{C}(\mathbf{x})$ to denote the codeword produced by encoding the message word $\mathbf{x} \in \Sigma^k$ using the encoding map $\mathbf{f}_N$ that defines $\mathcal{C}$.

The following theorem states the main result from [8].

Theorem II.1 ([8, Theorem 1])**.** *Let N and T be positive integers with $N > T$.*

1) *If $0 \leqslant R < \frac{1}{N-T}$, then there is no finite D for which (R, D) is achievable for T-VPEC.*
2) *Let $F(T)$ denote $T + \left\lfloor \frac{T^2}{4} \right\rfloor + 1$ and suppose that $N \geqslant F(T) + 1$. Then for any $\frac{1}{N-T} \leqslant R \leqslant \frac{1}{N-2T}$, the rate–distortion pair*

$$\left(R, \frac{(N-2T)^{-1} - R}{(N-2T)^{-1} - (N-T)^{-1}} \frac{F(T)}{N}\right)$$

is achievable.

In this paper, following the approach of [8], we focus on the achievability of rate–distortion pairs (R, D), given Σ, N, and T, and assuming a blocklength k.

C. Higher-order MDS codes

To present our construction of VPEC codes, we need the notion of higher-order MDS codes defined in [18], which builds upon the theory of list decoding.

Let $\mathcal{C}$ be a code in Σ^n, and let $L, \tau \in \mathbb{Z}^+$ be given. We say that $\mathcal{C}$ is *(τ, L)-list decodable* if for every $\mathbf{y} \in \Sigma^n$, there are no $L+1$ distinct codewords of $\mathcal{C}$ within Hamming distance at most τ from $\mathbf{y}$. Given a positive $\tau \in \mathbb{Z}/(L+1)$, we say that $\mathcal{C}$ is *strongly-(τ, L)-list decodable* (also called *average-radius (τ, L)-list decodable*) if for every $\mathbf{y} \in \Sigma^n$, the sum of distances from $\mathbf{y}$ of any $L+1$ distinct codewords of $\mathcal{C}$ exceeds $(L+1)\tau$. Note that strong (τ, L)-list decodability implies ordinary $(\lfloor \tau \rfloor, L)$-list decodability.

As a generalization of unique decoding, list decoding was introduced independently by Elias [21] and Wozencraft [22] in the 1950s. In contrast to unique decoding, which outputs at most one codeword, list decoding

allows multiple candidate codewords. This flexibility enables it to correct more adversarial errors. Over the years, list decoding has found numerous applications in information theory (see [23]–[26]) and theoretical computer science (see [27]–[31]).

Let $\mathcal{C}$ be a linear $[n,k]$ code over $\mathbb{F}_q$, let $L \leqslant \min\{q-1, \binom{n-1}{k-1}\}$, and let $\tau \in \mathbb{Z}/(L+1)$. It was shown in [18] that $\mathcal{C}$ is strongly-(τ, L)-list decodable only if

$$\tau \leqslant \frac{L(n-k)}{L+1}. \tag{5}$$

A linear code is said to be *L-MDS* if it attains the bound in (5), and *$[L]$-MDS* if it is L'-MDS for every $L' \in [L]$. As shown in [18], every L-MDS code is also an MDS code.

D. Main results

Our first result is the following theorem, which provides several bounds on the trade-off between the rate R of each packet and the distortion D for the T-VPEC problem.

Theorem II.2. *Let N and T be positive integers with $N > T$. Then, the following holds.*

1. *If the rate–distortion pair (R, D) is achievable for T-VPEC, then we have*
$$R \geqslant \begin{cases} \frac{1}{N-T}, & 0 \leqslant D \leqslant 1; \\ \frac{1}{N-2T}, & D = 0. \end{cases} \tag{6}$$
Moreover, $D = 0$ holds only if $N \geqslant 2T+1$.
2. *Suppose that $N \geqslant 2T+1$ and let $Ant_q(k,d)$ denote the maximum size of a (k,d)-anticode over Σ with $|\Sigma| = q$. Given $0 \leqslant D < 1$, if there is a T-VPEC code over Σ with parameters (N,k,R,D), then we have*
$$R \geqslant \left(1 - \frac{\log_q(Ant_q(k, kD))}{k}\right) \frac{1}{N-2T}. \tag{7}$$
3. *Suppose that $N \geqslant 2T+1$ and that there exists a linear T-VPEC code $\mathcal{C}$ over $\mathbb{F}_q$ with parameters (N,k,R,D), where $D \in \mathbb{Z}/k$. Then, we have*
$$R \geqslant \frac{1-D}{N-2T}.$$

Note that the bound on R in (6) of Theorem II.2 contains the result from the first part of Theorem II.1. Moreover, the value of $Ant_q(k,d)$ is characterized by the well-known Diametric Theorem of Ahlswede and Khachatrian [32]. Therefore, when $q \geqslant 2(k-kD)/3+2$, a closed-form expression for the bound in the second part of Theorem II.2 can be derived. This is stated in the following corollary; see Appendix A for a detailed description of the Diametric Theorem and the proof of the corollary.

Corollary II.1. *Given $0 \leqslant D \leqslant 1$, let k, N, T, and q be positive integers such that $N \geqslant 2T+1$ and $q \geqslant 2k(1-D)/3+2$. If there is a T-VPEC code over $[q]$ with parameters (N,k,R,D), then*

$$R \geqslant \max\left\{\frac{1-D}{N-2T}, \frac{1}{N-T}\right\}. \tag{8}$$

Our second result includes two explicit constructions of VPEC codes. The first construction employs L-MDS codes and yields T-VPEC codes that achieve improved rate–distortion trade-offs compared to existing results for general parameters $N \geqslant 2T+1$. The construction is detailed in Construction 1 in Section IV-A, and the rate–distortion performance of the resulting T-VPEC code is summarized in the following theorem.

Theorem II.3. *Let N, T and L be positive integers such that $N \geqslant 2T+1$, $2 \leqslant L \leqslant N/T$ and $L \mid T$. If there exists an $[N, \rho N]$ L-MDS code with*

$$\rho = 1 - \left(1 + \frac{1}{L}\right)\frac{T}{N}, \tag{9}$$

then there is an explicit construction of T-VPEC code with parameters

$$\left(N, \rho N^2, \frac{1}{\rho N}, \frac{LT}{N}\right).$$

By dividing the message to be transmitted into M sub-messages of equal length, one can, for any $0 < c < 1$, use an $[N, \rho N]$ L-MDS code to encode the first cM sub-messages, and the polytope code in [8] with rate–distortion pair $\left((N-T)^{-1}, F(T)/N\right)$ to encode the remaining $(1-c)M$ sub-messages. This results in a coding scheme that achieves the rate–distortion pair

$$\left(\frac{c}{\rho N} + \frac{1-c}{N-T}, \frac{cLT}{N} + \frac{(1-c)F(T)}{N}\right).$$

Such an operation is referred to as *time-sharing* in [8]. Similarly, by using an $[N, N-2T]$ MDS code and an $[N, \rho N]$ L-MDS code, one can show that the rate–distortion pair

$$\left(\frac{c}{N-2T} + \frac{1-c}{\rho N}, \frac{(1-c)LT}{N}\right)$$

is achievable via an similar time-sharing argument. Consequently, Theorem II.3 yields the following achievable rate–distortion pairs.

Corollary II.2. *Let N, T and L be positive integers such that $N \geqslant F(T)+1$, $2 \leqslant L \leqslant N/T$ and $L \mid T$, where $F(T) = T + \left\lfloor \frac{T^2}{4} \right\rfloor + 1$. If there exists an $[N, \rho N]$ L-MDS code with ρ defined as in (9), then for any $\frac{1}{N-T} \leqslant R \leqslant \frac{1}{N-2T}$, the rate–distortion pair (R, D) is achievable for any $D \leqslant 1$ satisfying*

$$D \geqslant \frac{1 - T/N - \rho(N-T)R}{N - T - \rho N} \cdot F(T) + \frac{\rho(N-T)R - \rho}{N - T - \rho N} \cdot LT,$$

when $R \leqslant (\rho N)^{-1}$, *and*

$$D \geqslant \frac{\rho - \rho(N-2T)R}{\rho N - (N-2T)} \cdot LT,$$

when $R > (\rho N)^{-1}$.

Remark II.1. 1) *When* $R = (N-(1+1/L)T)^{-1}$, *the distortion achieved by Theorem II.1 is given by*

$$\begin{aligned}&\frac{(N-2T)^{-1}-R}{(N-2T)^{-1}-(N-T)^{-1}} \cdot \frac{F(T)}{N}\\ =&\frac{(L-1)(N-T)}{LN-(L+1)T} \cdot \frac{F(T)}{N}\\ =&\frac{(L-1)(N-T)}{L(N-T)-T} \cdot \frac{F(T)}{LT} \cdot \frac{LT}{N}.\end{aligned}$$

This is strictly larger than LT/N *when* $T \geqslant 4(L+1)$, *as*

$$\begin{aligned}\frac{(L-1)(N-T)}{L(N-T)-T} \cdot \frac{F(T)}{LT} &> \frac{L-1}{L} \cdot \frac{F(T)}{LT}\\ &\geqslant \frac{L-1}{L} \cdot \frac{1+\frac{T}{4}}{L}\\ &\geqslant \frac{(L-1)(L+2)}{L^2} > 1,\end{aligned}$$

where the second inequality follows by $F(T) \geqslant T + \frac{T^2}{4}$, *and the third inequality follows by* $T \geqslant 4(L+1)$. *Thus, Theorem II.3 provides a better achievable rate–distortion pair than that in Theorem II.1 for* $T \geqslant 4(L+1)$.

2) *In [18], explicit constructions of* $[N, N-3T/2]$ 2-*MDS codes over fields of size* $q = O(N^{(3T/2)^{3T}})$ *were provided, which is polynomial in* N *for a fixed* T. *Whereas the polytope code construction in [8] relies on infinite fields, specifically the reals. Thus, Theorem II.3 imposes a more relaxed requirement on the field size compared to that in [8].*

In Figure 1, we compare our bounds on the rate–distortion trade-off from Theorem II.2 and the result of Theorem II.3 with known results in [8] for the case $N = 128,\ T = 18$, and it has the following features:

- The blue line is obtained by connecting the points $\big((N-2T)^{-1}, 0\big)$ and $\big((N-T)^{-1}, 1\big)$, which correspond to the rate–distortion pairs achieved by $[N, N-2T]$ and $[N, N-T]$ MDS codes, respectively. As shown in [8], the rate–distortion pairs corresponding to other points on this line is achieved via time-sharing.
- The yellow line corresponds to the second part of Theorem II.1, and is obtained by connecting the points $\big((N-T)^{-1}, F(T)/N\big)$ and $\big((N-2T)^{-1}, 0\big)$. In fact, Fan, Kosut and Wagner [8] showed that the rate–distortion pair $\big((N-T)^{-1}, F(T)/N\big)$ is achieved by polytope codes for $N \geqslant F(T)+1$. The rate–distortion pairs corresponding to other points are also achieved via time-sharing.
- The dark red line corresponds to our bound from Theorem II.2 for linear T-VPEC codes. The dark green line corresponds to the result of Theorem II.3, obtained by connecting the following points:

$$\left(\frac{1}{N-T}, \frac{F(T)}{N}\right), \left(\frac{1}{N-\frac{4}{3}T}, \frac{3T}{N}\right),$$

and

$$\left(\frac{1}{N-\frac{3}{2}T}, \frac{2T}{N}\right), \left(\frac{1}{N-2T}, 0\right).$$

Among these, the first and last points correspond to the rate–distortion pairs achieved by the polytope codes in [8] and the $[N, N-2T]$ MDS codes, respectively. The parameters $T = 18$ and $N = 128$ satisfy

$$N \geqslant F(T)+1 = 101, \quad N/T > 3, \quad \text{and} \quad 2, 3 \mid T,$$

which meet the conditions of Theorem II.3 for both $L = 2$ and $L = 3$. Therefore, the second and third points represent the rate–distortion pairs achieved by the T-VPEC codes in Theorem II.3 (see Construction 1) obtained using 2-MDS and 3-MDS codes, respectively. Clearly, the dark green line achieves a better rate–distortion trade-off than that achieved by the polytope codes in [8].

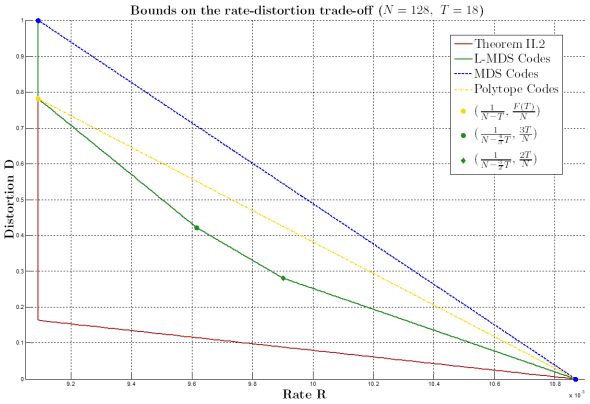


Figure 1. Comparison of the results on the rate–distortion trade-off from [8], Theorem II.2, and Theorem II.3, for $N = 128$ and $T = 18$.

Remark II.2. *For a coding scheme defined by an encoder-decoder pair* $(\mathbf{f}_N, g)$ *of the form* (1) *and* (2), *we define its* overall rate *as the sum of the per-packet rates:*

$$R^O \triangleq \sum_{i=1}^{N} R_i.$$

In the case of T-VPEC, where each packet is required to have the same rate, we have $R^O = NR$. We say that a pair (R^O, D) is achievable *for T-VPEC if the rate-distortion pair (R, D) is achievable, and the converse holds naturally.*

As in classical coding problems, we can study the asymptotic behavior of achievable pairs (R^O, D) in the regime where the error fraction $T/N = \theta$ is fixed for some constant $0 < \theta < 1$, and $N \to \infty$. In this setting, the result for polytope codes from Theorem II.1 becomes trivial, whereas L-MDS codes continue to outperform MDS codes. For detailed comparisons, we refer interested readers to Appendix B.

Our second construction employs a variant of the repetition code and yields a T-VPEC code with an optimal rate–distortion trade-off in the case where $N = 2T + 1$ and $(2T + 1) \mid k$. The code construction is presented in Construction 2 in Section IV-B, and the rate–distortion performance of the resulting T-VPEC code is summarized in the following theorem.

Theorem II.4. *Let k, s and T be positive integers satisfying $(2T + 1) \mid k$ and $s \leqslant T$. Then, for any given alphabet Σ, there is a T-VPEC code over Σ with parameters*

$$\left(2T+1, k, 1 - \frac{s}{2T+1}, \frac{s}{2T+1}\right)$$

and a decoding algorithm with a time complexity of $O(T^3)$.

As an immediate corollary, Theorem II.4 provides the following achievable rate–distortion pair.

Corollary II.3. *Given positive integers T, for any $(T + 1)/(2T + 1) \leqslant R \leqslant 1$, the rate–distortion pair $(R, 1 - R)$ is achievable for the T-VPEC problem with $2T + 1$ packets.*

Remark II.3. *Note that for any $0 \leqslant D \leqslant s/(2T + 1)$, it always holds that*

$$\begin{aligned} 1 - D &\geqslant 1 - \frac{s}{2T+1} \\ &> \frac{1}{2} \geqslant \frac{1}{T+1}. \end{aligned}$$

For $0 \leqslant D \leqslant T/(2T + 1)$ and $k = N = 2T + 1$, the bound in Corollary II.1 implies that

$$R \geqslant 1 - D,$$

whenever $|\Sigma| \geqslant 2k/3 + 2 = 4(T + 2)/3$. This confirms that the T-VPEC code in Theorem II.4 achieves the optimal rate–distortion trade-off when $|\Sigma| \geqslant 4(T+2)/3$.

In Figure 2, we compare our bound in Corollary II.1 and the result of Theorem II.4 with known results in [8] for the cases $N = 3$, $T = 1$ and $N = 5$, $T = 2$. The sub-figures in Figure 2 have the following common features:

- As in Figure 1, the dark red line corresponds to the rate–distortion bound from Corollary II.1, the blue and yellow lines correspond to the rate–distortion pairs achieved by MDS codes and the polytope codes in [8], respectively.
- The dark green line corresponds to the result of Theorem II.4. In the regime where $(N - T)/N \leqslant R \leqslant 1$, it coincides with the dark red line. Since Theorem II.4 guarantees achievability only for distortions up to $D \leqslant T/(2T + 1)$, the segment of the dark green line corresponding to larger distortions (i.e., $D > T/(2T + 1)$, or equivalently $R < (T + 1)/(2T + 1)$) is obtained by connecting the points $\big((T + 1)^{-1}, F(T)/(2T + 1)\big)$ and $\big((T + 1)^{-1}, T/(2T + 1)\big)$. These two points represent the rate–distortion pairs achieved by the polytope codes in [8] and the T-VPEC code in Theorem II.4 (see Construction 2), respectively.

III. Bounds on the Rate–Distortion Trade-off

In this section, we present the proof of Theorem II.2. Let q, k, T, and $n > T$ be positive integers, and let $|\Sigma| = q$. For $0 \leqslant R \leqslant 1$, denote $Q \triangleq \Sigma^{kR}$. We start with the following lemma which characterizes the necessary and sufficient conditions for $\mathcal{C}$ being a T-VPEC code.

Lemma III.1. *Let $\mathcal{C}$ be the image of an encoder $\mathbf{f}_N : \Sigma^k \longrightarrow Q^N$. Then, $\mathcal{C}$ is a T-VPEC code with parameters (N, k, R, D) for $0 \leqslant D \leqslant 1$ if and only if the following conditions hold:*

1. *The minimum Hamming distance of $\mathcal{C}$ satisfies $d_H(\mathcal{C}) \geqslant T + 1$.*
2. *For any $\mathbf{y} \in Q^N$, there exists a subset $S \subseteq [k]$ of at least $(1-D)k$ coordinates with the property that for any two codewords $\mathcal{C}(\mathbf{x})$ and $\mathcal{C}(\tilde{\mathbf{x}})$ in $B(\mathbf{y}, T)$, corresponding to message words $\mathbf{x}, \tilde{\mathbf{x}} \in \Sigma^k$, it holds that $\mathbf{x}|_S = \tilde{\mathbf{x}}|_S$.*

Proof: We begin with the proof of the "if" part. Given a code $\mathcal{C} \subseteq Q^N$ satisfying conditions 1 and 2, we define the decoder $g : Q^N \to \tilde{\Sigma}^k$ as follows: for any received word $\mathbf{y} \in Q^N$, suppose that $\mathcal{C} \cap B(\mathbf{y}, T) = \{\mathcal{C}(\mathbf{x}_1), \mathcal{C}(\mathbf{x}_2), \ldots, \mathcal{C}(\mathbf{x}_m)\}$ for some $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in \Sigma^k$, then

$$g(\mathbf{y})(j) = \begin{cases} \mathbf{x}_1(j), & \text{if } j \in S; \\ e, & \text{otherwise,} \end{cases}$$

where $S \triangleq \{j \in [k] : \mathbf{x}_1(j) = \mathbf{x}_2(j) = \cdots = \mathbf{x}_m(j)\}$ is the set of coordinates that all $\mathbf{x}_i$s agree.

If $\mathbf{y} \in \mathcal{C}$, i.e., $\mathbf{y} = \mathcal{C}(\mathbf{x}_0)$ for some $\mathbf{x}_0 \in \Sigma^k$. Since $d_H(\mathcal{C}) \geqslant T + 1$, it follows that $\mathcal{C} \cap B(\mathbf{y}, T) = \{\mathcal{C}(\mathbf{x}_0)\}$.

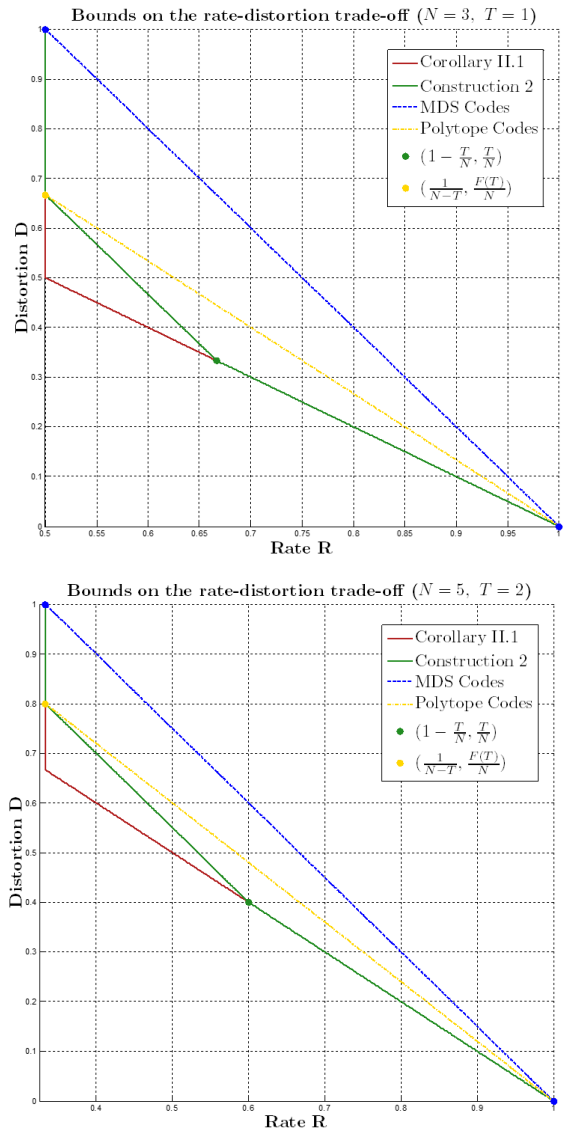


Figure 2. Comparison of the results on the rate–distortion trade-off from [8], Corollary II.1, and Theorem II.4, for $N = 3$, $T = 1$ and $N = 5$, $T = 2$.

Thus, $g(\mathbf{y}) = \mathbf{x}_0$, and we have $\Delta(\mathbf{x}_0, g(\mathbf{y})) = 0$ in this case.

If $\mathbf{y} \notin \mathcal{C}$, then $\mathbf{y}$ can only be obtained from $\mathcal{C}(\mathbf{x}_1), \ldots, \mathcal{C}(\mathbf{x}_m)$ by altering at most T packets. Then, by condition 2, $|S| \geqslant (1-D)k$. Thus, for each $i \in [m]$, we have $\Delta(\mathbf{x}_i, g(\mathbf{y})) \leqslant D$. Hence, the encoder-decoder pair $(\mathbf{f}_N, g)$ satisfies (4) with parameters $R_\ell = R$, $T_0 = D_0 = \varepsilon = 0$, $D_1 = D$, and $T_1 = T$. By definition, $\mathcal{C}$ is a T-VPEC code with parameters (N, k, R, D). This confirms the "if" part.

Next, suppose that $\mathcal{C}$ is a T-VPEC code with parameters (N, k, R, D) defined by an encoder-decoder pair $(\mathbf{f}_N, g)$, and we proceed with the proof for the "only if" part.

Assume, towards a contradiction, that $d_H(\mathcal{C}) \leqslant T$. Then, there exist two distinct messages words $\mathbf{x}, \tilde{\mathbf{x}}$ such that the Hamming distance between their encodings satisfies $d_H(\mathcal{C}(\mathbf{x}), \mathcal{C}(\tilde{\mathbf{x}})) \leqslant T$. Suppose that the codeword $\mathcal{C}(\mathbf{x})$ was transmitted, but the codeword $\mathcal{C}(\tilde{\mathbf{x}})$ was received. Since $\mathcal{C}$ allows lossless reconstruction when no packet errors occur, it follows that $g(\mathcal{C}(\tilde{\mathbf{x}})) = \tilde{\mathbf{x}}$. However, by (3), this implies

$$\Delta(\mathbf{x}, g(\mathcal{C}(\mathbf{x}))) = \Delta(\mathbf{x}, \tilde{\mathbf{x}}) = \infty,$$

which contradicts the assumption that $D \leqslant 1$. This shows that $\mathcal{C}$ satisfies condition 1.

Next, fix $\mathbf{y} \in Q^N$. By definition, $g(\mathbf{y}) \in \tilde{\Sigma}^k$. For any $\mathcal{C}(\mathbf{x}) \in B(\mathbf{y}, T)$, since $\mathcal{C}$ is a T-VPEC code, it holds that $\Delta(\mathbf{x}, g(\mathbf{y})) \leqslant D$. In particular, $\mathbf{x}$ agrees with $g(\mathbf{y})$ on all those coordinates where $g(\mathbf{y})$ is not equal to the erasure symbol e, and the number of such coordinates is at least $(1-D)k$. This shows that $\mathcal{C}$ satisfies condition 2 and the "only if" part follows. ∎

Corollary III.1. *Suppose that $\mathcal{C}$ is a T-VPEC code with parameters (N, k, R, D). Then*

$$d_H(\mathcal{C}) \geqslant \begin{cases} T+1, & 0 < D \leqslant 1; \\ 2T+1, & D = 0. \end{cases}$$

Proof: It remains to prove the case when $D = 0$.

Suppose $D = 0$. By the condition 2 in Lemma III.1, the ball $B(\mathbf{y}, T)$, for any $\mathbf{y} \in Q^N$, contains at most one codeword. This immediately implies that the balls of radius T centered at the codewords are disjoint. Consequently, the minimum distance of the code must be at least $2T + 1$. ∎

Proof of Theorem II.2: Let $\mathcal{C}$ be a T-VPEC code in Q^N with parameters (N, k, R, D) defined by an encoder-decoder pair $(\mathbf{f}_N, g)$.

By the Singleton bound, we have

$$q^k = |\mathcal{C}| \leqslant |Q|^{N - d_H(\mathcal{C}) + 1} = q^{kR(N - d_H(\mathcal{C}) + 1)}.$$

This implies that $R \geqslant (N - d_H(\mathcal{C}) + 1)^{-1}$. Then, part 1 of the theorem follows directly by Corollary III.1.

Next, we proceed to the proof of the second result.

Since $N \geqslant 2T + 1$, then, by the pigeonhole principle, $\mathcal{C}$ has a subset of codewords $\mathcal{C}'$ of size at least

$$\frac{|\mathcal{C}|}{|Q|^{N-2T}}$$

such that they all agree on their first $N - 2T$ coordinates. Let $\mathcal{C}(\mathbf{x}), \mathcal{C}(\tilde{\mathbf{x}}) \in \mathcal{C}'$ be two such codewords. Then there exists a word $\mathbf{y} \in Q^N$ such that $d_H(\mathbf{y}, \mathcal{C}(\mathbf{x})), d_H(\mathbf{y}, \mathcal{C}(\tilde{\mathbf{x}})) \leqslant T$. By Lemma III.1, $\mathbf{x}$ and $\tilde{\mathbf{x}}$ agree on at least $(1-D)k$ coordinates. Thus, $d_H(\mathbf{x}, \tilde{\mathbf{x}}) \leqslant kD$. This implies that the set of message words

$$\{\mathbf{x} \in \Sigma^k : \mathcal{C}(\mathbf{x}) \in \mathcal{C}'\}$$

is a (k, kD)-anticode. Thus, we have $|\mathcal{C}'| \leqslant Ant_q(k, kD)$. Then, part 2 follows directly by $|\mathcal{C}'| \geqslant \frac{|\mathcal{C}|}{|Q|^{N-2T}}$ and $|\mathcal{C}| = q^k$.

Finally, we proceed to the proof of part 3.

Recall that the subcode $\mathcal{C}'$ consists of all the codewords in $\mathcal{C}$ that agree on the first $N-2T$ coordinates. Let $\mathbf{c}_0 \in \mathcal{C}'$. Then, denote

$$\mathcal{C}'' \triangleq \{\mathbf{c} - \mathbf{c}_0 : \ \mathbf{c} \in \mathcal{C}'\}.$$

Since $\mathcal{C}$ is linear, $\mathcal{C}''$ is the subcode of $\mathcal{C}$ consisting of all codewords in $\mathcal{C}$ that are zero on the first $N-2T$ coordinates. Clearly, $\mathcal{C}''$ is a linear subcode of $\mathcal{C}$ in $(\mathbb{F}_q^{kR})^N$. This implies that the set $\{\mathbf{x} \in \mathbb{F}_q^k : \mathcal{C}(\mathbf{x}) \in \mathcal{C}''\}$ is a linear subspace of $\mathbb{F}_q^k$.

On the other hand, according to the proof of part 2, the set $\{\mathbf{x} \in \mathbb{F}_q^k : \mathcal{C}(\mathbf{x}) \in \mathcal{C}''\}$ is also a (k, kD)-anticode. Hence, its dimension is at most kD. Noting that $|\mathcal{C}''| = |\mathcal{C}'|$, we have

$$\begin{aligned} q^{kD} &\geqslant \left|\{\mathbf{x} \in \mathbb{F}_q^k : \mathcal{C}(\mathbf{x}) \in \mathcal{C}'\}\right| \\ &= |\mathcal{C}'| \geqslant q^{k-(N-2T)kR}, \end{aligned}$$

which leads to the inequality $R \geqslant (1-D)/(N-2T)$.

This completes the proof. ∎

IV. Constructions of VPEC coding schemes

In this section, we present two explicit constructions of T-VPEC codes. The first construction employs L-MDS codes together with an interleaving technique, resulting in a T-VPEC code that achieves a better rate–distortion pair than the polytope code in [8]. The second construction yields T-VPEC codes of length $N = 2T+1$ that achieve the optimal rate–distortion trade-off, along with a decoding algorithm of time complexity $O(N^3)$.

A. VPEC codes from Higher-Order MDS codes

In this section, we prove Theorem II.3 by constructing T-VPEC codes using L-MDS codes. To start, we introduce the notion of *interleaved codes*.

Given $\ell, n \in \mathbb{Z}^+$, for $i \in [\ell]$ let $\mathcal{C}_i$ be a code in Σ_i^n. The *interleaved code* $\otimes_{i\in[\ell]}\mathcal{C}_i$ over the alphabet $\otimes_{i\in[\ell]}\Sigma_i = \Sigma_1 \times \Sigma_2 \times \cdots \times \Sigma_\ell$ is defined as the set

$$\left\{\Gamma = (c_{i,j})_{i\in[\ell],j\in[n]} \ : \ (c_{i,j})_{j\in[n]} \in \mathcal{C}_i, \ \text{for } i \in [\ell]\right\},$$

namely, it consists of all $\ell \times n$ arrays Γ in which each row i is a codeword in $\mathcal{C}_i$ and each column is viewed as a symbol in $\otimes_{i\in[\ell]}\Sigma_i$. When all the constituent codes $\mathcal{C}_i$ are identical to $\mathcal{C}$, we use the notation $\mathcal{C}^{\otimes\ell}$ and refer to it as the *ℓ-level interleaving* of $\mathcal{C}$ (see, e.g., [33, Exercise 320] and [34, Problem 2.21]). If $\mathcal{C}$ is a linear $[n,k]$ code over $\mathbb{F}_q$ with a generator matrix $\mathbf{G}$ and parity-check matrix $\mathbf{H}$, then we can view $\mathcal{C}^{\otimes\ell}$ as a linear $[n,k]$ code over the extension field $\Phi \triangleq \mathbb{F}_{q^\ell}$: We do this by interpreting each column in an $\ell \times n$ code array Γ as the representation of an element of Φ, expressed as a column vector in $\mathbb{F}_q^\ell$ with respect to some fixed basis of Φ over $\mathbb{F}_q$. The matrices $\mathbf{G}$ and $\mathbf{H}$ then function as the generator matrix and parity-check matrix, respectively, of the code $\mathcal{C}^{\otimes\ell}$.

Proposition IV.1. *Let $\mathcal{C}$ be a linear $[n,k]$ code over $\mathbb{F}_q$ and suppose that $\mathcal{C}$ is (τ, L)-list decodable, for some $L, \tau \in \mathbb{Z}^+$. If $q > \binom{L+1}{2}$ then $\mathcal{C}^{\otimes\ell}$ is (τ, L)-list decodable, for every $\ell \in \mathbb{Z}^+$.*

Proof: Suppose, to the contrary, that $\mathcal{C}^{\otimes\ell}$ is not (τ, L)-list decodable. That is, there exists an array $\mathbf{Y} \in \mathbb{F}_q^{\ell\times n}$ and $L+1$ distinct codewords $\Gamma_1, \Gamma_2, \ldots, \Gamma_{L+1} \in \mathcal{C}^{\otimes\ell}$ such that their Hamming distances satisfy

$$d_H(\mathbf{Y}, \Gamma_r) \leqslant \tau, \quad \text{for all } r \in [L+1],$$

when viewed as vectors of length n over $\mathbb{F}_{q^\ell}$. Then for every row vector $\mathbf{u} \in \mathbb{F}_q^\ell$,

$$d_H(\mathbf{u}\cdot\mathbf{Y}, \mathbf{u}\cdot\Gamma_r) \leqslant \tau, \quad \text{for } r \in [L+1].$$

Since $\mathbf{u}\cdot\mathbf{Y} \in \mathbb{F}_q^n$ and $\mathbf{u}\cdot\Gamma_r \in \mathcal{C}$ for all $r \in [L+1]$, we will reach a contradiction once we find some $\mathbf{u} \in \mathbb{F}_q^\ell$ for which the codewords $\mathbf{u}\cdot\Gamma_r$ are distinct for all $r \in [L+1]$. Indeed, for any two distinct $r, s \in [L+1]$ we have $\mathbf{u}\cdot\Gamma_r = \mathbf{u}\cdot\Gamma_s$, if and only if $\mathbf{u}$ belongs to the left kernel of (the nonzero matrix) $\Gamma_r - \Gamma_s$; the size of this kernel, in turn, is at most $q^{\ell-1}$. Ranging over all unordered pairs $\{r, s\} \subseteq [L+1]$, by a simple union bound argument, there exist at least

$$q^\ell - \binom{L+1}{2} q^{\ell-1} > 0$$

vectors $\mathbf{u} \in \mathbb{F}_{q^\ell}$ that belong to none of these kernels. Therefore, the result follows since $q > \binom{L+1}{2}$. ∎

A (τ, L)-list decoder for the code $\mathcal{C}^{\otimes\ell}$ in Proposition IV.1 (which decodes any pattern of up to τ column errors by returning a list of codewords of size $\leqslant L$) can be obtained by iteratively applying a (τ, L)-list decoder for $\mathcal{C}$ on the rows of a received array $\mathbf{Y} \in \mathbb{F}_q^{\ell\times n}$, provided that $\mathbf{Y}$ contains at most τ column errors. Indeed, assume by induction on ℓ that we are able to decode the $(\ell-1)\times n$ array $\mathbf{Y}'$ formed by the first $\ell-1$ rows of $\mathbf{Y}$, ending up with at most L codewords $\Gamma'_1, \Gamma'_2, \ldots \in \mathcal{C}^{\otimes(\ell-1)}$. We then decode the ℓth row of $\mathbf{Y}$ using a (τ, L)-list decoder for $\mathcal{C}$, ending up with at most L codewords $\mathbf{c}_1, \mathbf{c}_2, \ldots \in \mathcal{C}$. Among all the (at most) L^2 codewords of $\mathcal{C}^{\otimes\ell}$ of the form

$$\Gamma_{r,s} = \begin{pmatrix} \Gamma'_r \\ \mathbf{c}_s \end{pmatrix} \in \mathcal{C}^{\otimes\ell}, \quad r, s \in [L],$$

we are guaranteed that there are at most L pairs (r, s) for which $d_H(\mathbf{Y}, \Gamma_{r,s}) \leqslant \tau$.

Remark IV.1. *Regarding the list decoding of generic interleaved codes, the only work we are aware of is [35]. In [35], it was shown that the list decoding radius of $\mathcal{C}^{\otimes \ell}$ remains unchanged compared to that of $\mathcal{C}$; see Theorem 2.5 in [35] for a detailed description. This result does not imply Proposition IV.1. In fact, the proof of Lemma 3.7 in [35] shows that Proposition IV.1 is generally false if we do not put any restrictions on the field size q (in terms of L). For example, if we take $\mathcal{C}$ to be the $[n, 1]$ repetition code over $\mathbb{F}_q$, then $\mathcal{C}$ is clearly (n, L)-list decodable for $L = q$, but its ℓ-level interleaving is not (n, L')-list decodable for any $L' < q^\ell$.*

Remark IV.2. *Following a similar proof, one can show that the result of Proposition IV.1 also holds when $\mathcal{C}$ is an $\mathbb{F}_q$-linear code over a finite vector space Σ defined over $\mathbb{F}_q$, provided that $q > \binom{L+1}{2}$. Moreover, we show that strong $(\tau, L = 2)$-list decodability is preserved under interleaving, where τ is of the form $a/(L+1)$ for some $a \in \mathbb{Z}^+$. Interested readers are referred to Appendix C for the detailed proof.*

It is known that for sufficiently large field size q (specifically $q = n^{\Theta(\min(k, n-k)L)}$), there always exists a linear $[n, k]$ code $\mathcal{C}$ over $\mathbb{F}_q$ that is $[L]$-MDS (and hence, L-MDS) [36], [37]. This follows from the fact that the $[L]$-MDS property translates into algebraic conditions on the $(n-k) \times n$ parity-check matrix $\mathbf{H}$ of $\mathcal{C}$: these conditions require a collection of square matrices to be nonsingular, and they hold for most matrices $\mathbf{H}$ if the underlying field is sufficiently large (in fact, they hold for most GRS codes). By Proposition IV.1, it follows that the interleaved code $\mathcal{C}^{\otimes \ell}$ is $[L]$-MDS as well. This conclusion can be reached also from recalling that $\mathbf{H}$ is also a parity-check matrix of the code $\mathcal{C}^{\otimes \ell}$ when seen over $\Phi = \mathbb{F}_{q^\ell}$. Since the aforementioned algebraic conditions hold also over Φ, it follows that $\mathcal{C}^{\otimes \ell}$ is $[L]$-MDS.

Next, we present our code construction.

Construction 1. *Let N, T and L be positive integers such that $N \geqslant 2T+1$, $2 \leqslant L \leqslant N/T$ and $L \mid T$. Let $\mathcal{C}$ be an L-MDS $[N, \rho N]$ code over a (sufficiently large) field $\mathbb{F}_q$, where*

$$\rho = 1 - \left(1 + \frac{1}{L}\right)\frac{T}{N}. \tag{10}$$

Then the ℓ-level interleaving $\mathcal{C}^{\otimes \ell}$ is an L-MDS code and is (τ, L)-list decodable, where

$$\tau = \frac{L(N - \rho N)}{L+1} \overset{(10)}{=} T. \tag{11}$$

Let $\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_N$ be generator matrices of $\mathcal{C}$, where each $\mathbf{G}_i$ contains the identity matrix in columns i through $i + \rho N - 1$ (with indices taken modulo N if they exceed N). Since $\mathcal{C}$ is MDS by Theorem 3 in [18], such generator matrices indeed exist. Letting $\Sigma = \mathbb{F}_q$ and $k = \rho N^2$, we write each information vector $\mathbf{x} \in \Sigma^{\rho N^2}$ as a concatenation of N row vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N \in \Sigma^{\rho N}$ and define the encoder $\mathbf{f}_N : \Sigma^k \to \Sigma^{N \times N}$ by

$$\mathbf{f}_N(\mathbf{x}) = \begin{pmatrix} \mathbf{x}_1 \cdot \mathbf{G}_1 \\ \mathbf{x}_2 \cdot \mathbf{G}_2 \\ \vdots \\ \mathbf{x}_N \cdot \mathbf{G}_N \end{pmatrix}. \tag{12}$$

Clearly, $\mathbf{f}_N(\mathbf{x})$ is a codeword in $\mathcal{C}^{\otimes N}$.

Next, we show that the code $\mathcal{C}^{\otimes N}$ with the encoder $\mathbf{f}_N$ defined above is a T-VPEC code over $\mathbb{F}_q$ with parameters

$$\left(N, \rho N^2, \frac{1}{\rho N}, \frac{LT}{N}\right).$$

Recall that the rate–distortion pairs $(1/(N-T), F(T)/N)$ and $(1/(N-2T), 0)$ are achievable by Theorem II.1 and by $[N, N-2T]$ MDS codes, respectively. Thus, through a similar time-sharing argument as in [8, Section II.B], the rate–distortion pair (R, D) that lies on the broken line connecting

$$\left(\frac{1}{N-T}, \frac{F(T)}{N}\right), \left(\frac{1}{N-(1+L^{-1})T}, \frac{LT}{N}\right),$$

and

$$\left(\frac{1}{N-2T}, 0\right)$$

can also be achieved. This confirms Corollary II.2.

Proof of Theorem II.3: For every information vector $\mathbf{x} \in \Sigma^k$ and every $j \in [N]$, take the j-th column of $\mathbf{f}_N(\mathbf{x})$ as the j-th packet. That is,

$$C_j = \left(\mathbf{x}_1 \cdot \mathbf{G}_1(j),\, \mathbf{x}_2 \cdot \mathbf{G}_2(j),\, \ldots,\, \mathbf{x}_N \cdot \mathbf{G}_N(j)\right)^\top,$$

where $\mathbf{G}_i(j)$ denotes the j-th column of $\mathbf{G}_i$. Recall that the code $\mathcal{C}^{\otimes N}$ is an $N \times N$ array code of dimension $k = \rho N^2$; thus, the rate of each packet is equal to

$$\frac{N}{\rho N^2} \overset{(10)}{=} \frac{1}{N-(1+L^{-1})T}. \tag{13}$$

According to the definition of $\mathbf{f}_N$ in (12), when there are no packet errors, one can recover $\mathbf{x}_i$ by reading the i-th to $(i + \rho N - 1)$-st entries of the i-th row of the received word, and thus recover the original information vector $\mathbf{x}$. Thus, it remains to show that the distortion is at most D in the case where there are at most T packet errors.

Given $\mathbf{Y} \in \Sigma^{N \times N}$ such that

$$d_H(\mathbf{Y}, \mathbf{f}_N(\mathbf{x})) \leqslant T,$$

a list decoder for $\mathcal{C}^{\otimes N}$ will generate at most L codewords $\Gamma \in \mathcal{C}^{\otimes N}$ such that $d_H(\mathbf{Y}, \Gamma) \leqslant T$. Then, by erasing from $\mathbf{Y}$ all the columns in which there is disagreement

among those codewords, at most LT columns will be erased. The remaining columns are those where all the outputted codewords agree with $\mathbf{Y}$, and they are the same as the correct codeword. Due to the specific structure of $\mathbf{f}_N(\mathbf{x})$ in (12), each column of $\mathbf{f}_N(\mathbf{x})$ contains exactly ρN information symbols. Therefore, the total number of information symbols in the erased columns is at most

$$\rho N \cdot LT = k \cdot \frac{LT}{N},$$

which is at most k since $L \leqslant N/T$. Hence, the distortion of the code in the case where there are at most T packet errors is at most $\frac{LT}{N}$. This concludes the proof. ∎

Remark IV.3.

1. *We note that the requirement that the generator matrices $\mathbf{G}_1, \ldots, \mathbf{G}_N$ be in systematic form is essential in Construction 1. Otherwise, given $\mathbf{Y} \in \Sigma^{N \times N}$ as the received word, the proof of Theorem II.3 still implies that at least $N - LT$ columns of $\mathbf{Y}$ coincide with those of the correct codeword. However, if $\mathbf{G}_1, \ldots, \mathbf{G}_N$ are not all systematic, we can no longer guarantee that each column of the correct codeword contains exactly ρN information symbols. Consequently, the desired upper bound on the distortion may no longer hold.*
2. *If we relax (10) by requiring only that*

$$\rho \geqslant 1 - \left(1 + \frac{1}{L}\right)\theta - \varepsilon,$$

for some (arbitrarily small yet) fixed small $\varepsilon > 0$, this would translate into subtracting ε from the denominator in the right-hand side of (13). On the other hand, this allows us to select the alphabet Σ to be a much smaller vector space over some $\mathbb{F}_q$ and take $\mathcal{C}$ to be an $\mathbb{F}_q$-linear folded Reed–Solomon code over Σ.
3. *In the final part of the proof of Theorem 5, we present a decoding algorithm for the code obtained in Construction 1. The decoding complexity of this algorithm is essentially the same as that of the list decoder for the underlying L-MDS code. However, to the best of our knowledge, the design of an efficient explicit list decoding algorithm for general L-MDS codes remains an open problem. Nevertheless, efficient randomized list decoding algorithms are known for folded Reed–Solomon codes and multiplicity codes; see [38], [39].*

B. VPEC codes with optimal rate–distortion trade-off

Our second coding scheme can be regarded as a variant of the repetition code. Before delving into the formal description, we first introduce some additional notation and a toy example, which serves as a special case of the general coding scheme.

We use $\{\{\cdot\}\}$ to denote multisets. Let B be a multiset of symbols in Σ of size b, denoted by

$$B = \{\{x_1, x_2, \ldots, x_b\}\}.$$

For each $y \in B$, we denote $F_B(y) \triangleq |\{i \in [b] : x_i = y\}|$ as the frequency of y in B. Then, we use MF_B to denote the most frequent elements in B.

Example 1. *Let Σ be an arbitrary alphabet of size at least 2 and let T be a positive integer. For any message word $\mathbf{x} = (x_1, x_2, \ldots, x_{2T+1}) \in \Sigma^{2T+1}$ from the source, we encode it into the following $2T+1$ packets. For each $j \in [2T+1]$, the j-th packet is encoded as:*

$$C_j = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_{2T+1})^\top,$$

which is a word in Σ^{2T}.

For decoding, we use the majority decoder to recover the value of x_j for each $j \in [2T+1]$ and if it fails for some j, we then output C_j. Specifically, for each $j \in [2T+1]$, let

$$\tilde{C}_j = (\tilde{x}_{1,j}, \ldots, \tilde{x}_{j-1,j}, \tilde{x}_{j+1,j}, \ldots, \tilde{x}_{2T+1,j})^\top$$

denote the received j-th packet, which may or may not be equal to C_j, where $\tilde{x}_{i,j} \in \Sigma$ for each $i \in [2T+1] \setminus \{j\}$. We view $\tilde{x}_{i,j}$ as the candidate of x_i from the received j-th packet $\tilde{C}_j$. Note that for each $i \in [2T+1]$, there is only one packet C_i which does not contain x_i. We denote $A_i = \{\{\tilde{x}_{i,j} : j \in [2T+1] \setminus \{i\}\}\}$ as the multiset of all the candidates of x_i. Then, the decoding algorithm is detailed in Algorithm 1.

Algorithm 1 Decoding Algorithm for codes in Example 1

Input: $(\tilde{C}_1, \ldots, \tilde{C}_{2T+1})$, parameter T
Output: $\mathbf{y} = (y_1, y_2, \ldots, y_{2T+1})$
Initialize $I = \{1, 2, \ldots, 2T+1\}$
for each $i \in I$ **do**
  Let $v \leftarrow MF_{A_i}$ (Most frequent element in A_i)
  if $F_{A_i}(v) \geqslant T+1$ **then**
    $y_i \leftarrow v$
    $I \leftarrow I \setminus \{i\}$
  end if
end for
if $|I| \leqslant 0$ **then**
  return $\mathbf{y} = (y_1, \ldots, y_{2T+1})$
else
  Pick any $i \in I$
  for each $j \in [2T+1] \setminus \{i\}$ **do**
    $y_j \leftarrow \tilde{x}_{j,i}$
  end for
  $y_i \leftarrow e$ (Assign erasure symbol)
end if

For each message word $\mathbf{x} \in \Sigma^{2T+1}$, the encoder in Example 1 produces $2T+1$ packets $C_j \in \Sigma^{2T}$, which can be viewed as the columns in an $2T \times (2T+1)$ array over Σ. We claim that it achieves the rate–distortion pair $(2T/(2T+1), 1/(2T+1))$, which is optimal with respect to the bound by Corollary II.1.

Clearly, the rate of each packet is $2T/(2T+1)$. Since there are at most T altered packets, for each $i \in [2T+1]$, at least T packets contain the correct x_i. Thus, by the decoding algorithm, if the most frequent element MF_{A_i} in A_i has frequency at least $T+1$, it must be correct. Otherwise, MF_{A_i} shall have frequency T. Then, this implies that there are T altered packets containing x_i. Thus, the received packet $\tilde{C}_i$, which does not contain x_i, must be correct. According to the decoding algorithm, we either recover all the correct x_i or obtain $2T$ correct message symbols x_i along with one erasure symbol, resulting in a distortion of at most $\frac{1}{2T+1}$.

According to Algorithm 1, determining the most frequent element is required for at most $N = 2T+1$ sets A_i, each of size $|A_i| = N-1 = 2T$. Using the Boyer-Moore majority vote algorithm [40], which runs in $O(N)$ time per set, the overall time complexity of the decoding algorithm in Example 1 is $O(N^2)$.

Next, we present the formal description of our general construction.

Construction 2. *Let Σ be an arbitrary alphabet of size at least 2. Let T and s be positive integers such that $s \leqslant T$. For each $j \in [2T+1]$, we denote $S_j \triangleq \{j-s+1, j-s+2, ..., j\}$, where we subtract $2T+1$ from an index if it exceeds $2T+1$, and $\overline{S}_j = [2T+1] \setminus S_j$.*

For any message word $\mathbf{x} = (x_1, x_2, \ldots, x_{2T+1}) \in \Sigma^{2T+1}$ from the source, we encode it into the following $2T+1$ packets. For each $j \in [2T+1]$, the j-th packet is encoded as the following word in Σ^{2T+1-s}:

$$C_j = (x_1, \ldots, x_{j-1}, x_{j+s}, \ldots, x_{2T+1})^\top = \left(\mathbf{x}|_{\overline{S}_j}\right)^\top. \tag{14}$$

For each $j \in [2T+1]$, we denote

$$\tilde{C}_j = (\tilde{x}_{1,j}, \ldots, \tilde{x}_{j-1,j}, \tilde{x}_{j+s,j}, \ldots, \tilde{x}_{2T+1,j})^\top \tag{15}$$

as the received j-th packet, where $\tilde{x}_{i,j} \in \Sigma$ for each $i \in \overline{S}_j$. Note that for each $i \in [2T+1]$, there are exactly s packets—namely, $C_i, C_{i-1}, \ldots, C_{i-s+1}$—that do not contain x_i. Therefore, symbols in $A_i = \{\{\tilde{x}_{i,j} : j \in \overline{S}_i\}\}$ are viewed as a candidate for x_i from the received packet $\tilde{C}_j$. Similar to Example 1, we select an appropriate symbol as y_i from these candidates using a majority decoding approach. The full decoding procedure is detailed in Algorithm 2.

Remark IV.4. *Construction 2 can be naturally extended to the case where $k = m(2T+1)$ for some $m \in \mathbb{Z}^+$. Specifically, we view $\mathbf{x} \in \Sigma^{m(2T+1)}$ as a concatenation of $2T+1$ words $\mathbf{x}_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,m}) \in \Sigma^m$ for $i \in [2T+1]$, and perform m independent transmissions. In the ℓ-th transmission, send a batch of $2T+1$ packets encoding the message tuple $(x_{1,\ell}, x_{2,\ell}, \ldots, x_{2T+1,\ell})$. The decoder then runs the decoding algorithm separately on each batch of received words. This extension preserves both the per-packet rate and the distortion between the original message word $\mathbf{x} \in \Sigma^{m(2T+1)}$ and the final output.*

Algorithm 2 Decoding Algorithm for codes by Construction 2

Input: $(\tilde{C}_1, \ldots, \tilde{C}_{2T+1})$, parameters T, s
Output: $\mathbf{y} = (y_1, y_2, \ldots, y_{2T+1})$
Initialize $I = \{1, 2, \ldots, 2T+1\}$
for each $i \in I$ **do**
  Let $v \leftarrow MF_{A_i}$ (Most frequent element in A_i)
  if $F_{A_i}(v) \geqslant T+1$ **then**
    $y_i \leftarrow v$
    $I \leftarrow I \setminus \{i\}$
  end if
end for
while $|I| > s$ **do**
  Find $i_1, i_2 \in I$ such that:
  $i_2 - i_1 \in \{s, s+1, \ldots, 2T+1-s\} \mod 2T+1$
  Set $\tilde{A}_1 = \{\{\tilde{x}_{i_1,j} : j \in \overline{S}_{i_1-s+1}\}\}$, $\tilde{A}_2 = \{\{\tilde{x}_{i_2,j} : j \in \overline{S}_{i_2-s+1}\}\}$, $\xi_1 = MF_{\tilde{A}_1}$, $\xi_2 = MF_{\tilde{A}_2}$, $B_1 = \{\{\tilde{x}_{i_1,j} : j \in S_{i_2-s+1}\}\}$, and $B_2 = \{\{\tilde{x}_{i_2,j} : j \in S_{i_1-s+1}\}\}$
  for $\ell \in [2]$ **do**
    Define Boolean expressions:
    $$U_\ell \triangleq \left[F_{B_\ell}(\xi_\ell) \leqslant T - F_{\tilde{A}_{3-\ell}}(\xi_{3-\ell})\right]$$
    $$Z_\ell \triangleq \left[F_{B_\ell}(\xi_\ell) \geqslant s + F_{\tilde{A}_{3-\ell}}(\xi_{3-\ell}) - T\right]$$
    if U_ℓ is TRUE **then**
      $y_{i_{3-\ell}} \leftarrow MF_{B_{3-\ell}}$
      $I \leftarrow I \setminus \{i_{3-\ell}\}$
    else if Z_ℓ is TRUE **or** $Z_{3-\ell}$ is FALSE **then**
      $y_{i_\ell} \leftarrow \xi_\ell$
      $I \leftarrow I \setminus \{i_\ell\}$
    end if
  end for
end while
for $i \in I$ **do**
  $y_i \leftarrow e$ (Assign erasure symbol)
end for

Before giving the detailed proof that Construction 2 yields T-VPEC codes with the parameters stated in Theorem II.4, we first examine the special case $T = 3$, $s = 2$, and $N = 7$. This example is intended to

provide intuition by illustrating the structures of A_i and B_i in Algorithm 2, as well as the associated decoding mechanism.

Example 2. *Let Σ be an arbitrary alphabet of size at least 2. Let $T = 3$, $s = 2$, and $N = 7$. Then, for each $j \in [7]$, we have $S_j \triangleq \{j, j+1\}$ and $\overline{S}_j = [7] \setminus S_j$. Moreover, for any message word $\mathbf{x} = (x_1, x_2, \ldots, x_7) \in \Sigma^7$ and $j \in [7]$, the j-th packet is encoded as: $C_j = (x_1, \ldots, x_{j-1}, x_{j+2}, \ldots, x_7)^\top$.*

Assume that $(\tilde{C}_1, \ldots, \tilde{C}_7)$ are the received packets, which admit the following form

$$\begin{array}{ccccccc}
\tilde{C}_1 & \tilde{C}_2 & \tilde{C}_3 & \tilde{C}_4 & \tilde{C}_5 & \tilde{C}_6 & \tilde{C}_7 \\
* & \tilde{x}_{1,2} & \tilde{x}_{1,3} & \tilde{x}_{1,4} & \tilde{x}_{1,5} & \tilde{x}_{1,6} & * \\
* & * & \tilde{x}_{2,3} & \tilde{x}_{2,4} & \tilde{x}_{2,5} & \tilde{x}_{2,6} & \tilde{x}_{2,7} \\
\tilde{x}_{3,1} & * & * & \tilde{x}_{3,4} & \tilde{x}_{3,5} & \tilde{x}_{3,6} & \tilde{x}_{3,7} \\
\tilde{x}_{4,1} & \tilde{x}_{4,2} & * & * & \tilde{x}_{4,5} & \tilde{x}_{4,6} & \tilde{x}_{4,7} \\
\tilde{x}_{5,1} & \tilde{x}_{5,2} & \tilde{x}_{5,3} & * & * & \tilde{x}_{5,6} & \tilde{x}_{5,7} \\
\tilde{x}_{6,1} & \tilde{x}_{6,2} & \tilde{x}_{6,3} & \tilde{x}_{6,4} & * & * & \tilde{x}_{6,7} \\
\tilde{x}_{7,1} & \tilde{x}_{7,2} & \tilde{x}_{7,3} & \tilde{x}_{7,4} & \tilde{x}_{7,5} & * & *
\end{array}$$

We next apply Algorithm 2 to perform the decoding. Recall that $A_i = \{\{\tilde{x}_{i,j} : j \in [7] \setminus \{i, i+1\}\}\}$. Since $T = 3$, if the most frequent element MF_{A_i} appears with frequency at least 4, then it must be the correct symbol. The first "for" loop in Algorithm 2 outputs all correct symbols satisfying this condition.

After the first "for" loop, the set I consists precisely of those indices satisfying $F_{A_i}(M_{F_{A_i}}) \leqslant 3$. If $|I| \leqslant 2$, then setting $y_i = e$ for all $i \in I$ yields an output word $\mathbf{y}$ with erasure distortion at most $2/7$, as claimed in Theorem II.4. Otherwise, by Lemma IV.2, there exist $i_1, i_2 \in I$ such that $i_2 - i_1 \in \{2, 3, 4, 5\}$. Without loss of generality, assume $i_1 = 1$ and $i_2 = 3$. Then $\tilde{A}_1 = \{\{\tilde{x}_{1,j} : j \in \{2, 3, 4, 5, 6\}\}\}$, $\tilde{A}_2 = \{\{\tilde{x}_{3,j} : j \in \{1, 4, 5, 6, 7\}\}\}$, and $B_1 = \{\{\tilde{x}_{1,2}, \tilde{x}_{1,3}\}\}$, $B_2 = \{\{\tilde{x}_{3,1}, \tilde{x}_{3,7}\}\}$.

Next, we show that the "for" loop within the "while" loop further reduces the size of I. Recall that the correct symbol x_3 appears in $\tilde{A}_2$ at least twice, since there are at most $T = 3$ packet errors. Hence, $F_{\tilde{A}_2}(\xi_2) \geqslant 2$. Moreover, $F_{\tilde{A}_2}(\xi_2) = 2$ only if exactly three of $\{\tilde{C}_j : j \in \{1, 4, 5, 6, 7\}\}$ are altered. On the other hand, $F_{B_1}(\xi_1) \geqslant 1$, and $F_{B_1}(\xi_1) = 1$ only if exactly one of $\{\tilde{C}_j : j \in \{2, 3\}\}$ is altered. Thus, $F_{\tilde{A}_2}(\xi_2) = 2$ and $F_{B_1}(\xi_1) = 1$ cannot hold simultaneously. Hence, U_1 (and similarly U_2) is always FALSE *in this example.*

Note that $F_{B_1}(\xi_1) = 2$ whenever $F_{\tilde{A}_2}(\xi_2) = 2$, and in this case $\xi_1 = x_1$. For the case $F_{\tilde{A}_2}(\xi_2) \geqslant 3$, Z_1 is TRUE *only if $F_{B_1}(\xi_1) = 2$ and $F_{\tilde{A}_2}(\xi_2) = 3$. Suppose ξ_1 is incorrect. Then both $\tilde{C}_1$ and $\tilde{C}_2$ have to be altered, which implies that the correct symbol x_3 appears at least four times in $\tilde{A}_2$. This contradicts the assumption that $F_{\tilde{A}_2}(\xi_2) = 3$. Thus, we can conclude that $\xi_1 = x_1$ whenever Z_1 is* TRUE.

Since Z_ℓ, $\ell \in [2]$, is either TRUE *or* FALSE*, the size of I strictly decreases after each round of the "for" loop within the "while" loop. Consequently, the algorithm always terminates with an output $\mathbf{y}$ whose erasure distortion is at most $2/7$ relative to the correct message $\mathbf{x}$.*

For the proof of Theorem II.4, we need the following lemmas.

Lemma IV.2. *Let J be a subset of $\mathbb{Z}_{2T+1}$. If $|J| \geqslant s+1$ for some positive integer $s \leqslant T$, then there are $a, b \in J$ such that*

$$a - b \in \{s, s+1, \ldots, 2T+1-s\} \subseteq \mathbb{Z}_{2T+1}.$$

Proof: We define a graph G over $\mathbb{Z}_{2T+1}$ with edge set

$$E(G) \triangleq \left\{(a, b) \in (\mathbb{Z}_{2T+1})^2 : a - b \in \{s, s+1, \ldots, 2T+1-s\}\right\}.$$

We claim that G has independence number at most s. Then, for every vertex set J of size at least $s+1$, there must be an edge among vertices in J.

Clearly, each vertex $a \in \mathbb{Z}_{2T+1}$ has $2T+2-2s$ different neighbors $a+s, a+s+1, \ldots, a+2T+1-s$. Assume that I is an independent set containing vertex a. Then, we know that

$$\begin{aligned} I \setminus \{a\} \subseteq & \{a+1, \ldots, a+s-1\} \cup \\ & \{a+2T+2-s, \ldots, a+2T\}. \end{aligned}$$

Now, notice that for $i \in [s-1]$ we have $(a+2T+1-i) - (a+s-i) = 2T+1-s$; that is, $(a+s-i, a+2T+1-i)$ is an edge in G. Thus, for each $1 \leqslant i \leqslant s-1$, at most one of $\{a+s-i, a+2T+1-i\}$ is contained in I. Thus, we have $|I \setminus \{a\}| \leqslant s-1$ and this concludes the proof. ∎

Lemma IV.3. *Let $i_1, i_2 \in [2T+1]$ such that $i_2 - i_1 \in \{s, s+1, \ldots, 2T+1-s\} \bmod 2T+1$. Denote $A_1 = \{\{\tilde{x}_{i_1,j} : j \in \overline{S}_{i_1-s+1}\}\}$, $A_2 = \{\{\tilde{x}_{i_2,j} : j \in \overline{S}_{i_2-s+1}\}\}$, $\xi_1 = MF_{A_1}$, $\xi_2 = MF_{A_2}$, $B_1 = \{\{\tilde{x}_{i_1,j} : j \in S_{i_2-s+1}\}\}$, and $B_2 = \{\{\tilde{x}_{i_2,j} : j \in S_{i_1-s+1}\}\}$. For $\ell \in [2]$, define the Boolean expressions:*

$$\begin{aligned} U_\ell &\triangleq \left[F_{B_\ell}(\xi_\ell) \leqslant T - F_{A_{3-\ell}}(\xi_{3-\ell})\right], \\ Z_\ell &\triangleq \left[F_{B_\ell}(\xi_\ell) \geqslant s + F_{A_{3-\ell}}(\xi_{3-\ell}) - T\right]. \end{aligned}$$

Then, for each $\ell \in [2]$, the correctness of ξ_ℓ can be characterized by Table I. Moreover, if both ξ_1 and ξ_2 are incorrect, then it must hold that $U_1 = U_2 =$ TRUE.

Proof: First, we need the following observations. Note that for any $i_1, i_2 \in [2T+1]$ such that $i_2 - i_1 \in \{s, s+1, \ldots, 2T+1-s\} \mod 2T+1$, we also have

Table I
TRUTH TABLE FOR ξ_ℓ

U_ℓ	Z_ℓ	$Z_{3-\ell}$	Status of ξ_ℓ
TRUE	–	–	Incorrect
FALSE	TRUE	TRUE	Correct
FALSE	TRUE	FALSE	Correct
FALSE	FALSE	FALSE	Correct

$i_1 - i_2 \in \{s, s+1, \ldots, 2T+1-s\} \mod 2T+1$. Thus, it holds that $S_{i_1-s+1} \cap S_{i_2-s+1} = \emptyset$ and

$$S_{i_2-s+1} \subseteq \overline{S}_{i_1-s+1},$$
$$S_{i_1-s+1} \subseteq \overline{S}_{i_2-s+1}.$$

Recall that for each $i \in [2T+1]$, there are exactly s packets $C_i, C_{i-1}, \ldots, C_{i-s+1}$ that do not contain $\mathbf{x}_i$. Therefore, the received packets $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_{2T+1}$ have the following form:

$$\begin{array}{rccccc}
\tilde{C}_{i_2-s+1} = & (\cdots & \tilde{x}_{i_1,i_2-s+1} & \cdots & * & \cdots)^\top, \\
\vdots & & \vdots & & \vdots & \\
\tilde{C}_{i_2} = & (\cdots & \tilde{x}_{i_1,i_2} & \cdots & * & \cdots)^\top, \\
\tilde{C}_{i_1-s+1} = & (\cdots & * & \cdots & \tilde{x}_{i_2,i_1-s+1} & \cdots)^\top, \\
\vdots & & \vdots & & \vdots & \\
\tilde{C}_{i_1} = & (\cdots & * & \cdots & \tilde{x}_{i_2,i_1} & \cdots)^\top,
\end{array} \tag{16}$$

where $*$ denotes the null symbol, and

$$\tilde{C}_j = (\cdots \quad \tilde{x}_{i_1,j} \quad \cdots \quad \tilde{x}_{i_2,j} \quad \cdots)^\top \tag{17}$$

for the rest $j \in [2T+1] \setminus (S_{i_1-s+1} \cup S_{i_2-s+1})$.

Now, consider statement 1). Suppose that $U_1 =$ TRUE. If ξ_1 were correct, that is, $\xi_1 = x_{i_1}$. Then, by the definition of A_ℓ and B_ℓ, we have

$$\begin{aligned}
T+1 &\leqslant |\{j \in [2T+1] : \tilde{C}_j = C_j\}| \\
&\leqslant F_{B_1}(x_{i_1}) + F_{A_2}(x_{i_2}) \\
&= F_{B_1}(\xi_1) + F_{A_2}(x_{i_2}) \\
&\leqslant F_{B_1}(\xi_1) + F_{A_2}(\xi_2) \leqslant T,
\end{aligned}$$

which leads to a contradiction. Similarly, we have that ξ_2 is incorrect if $U_2 =$ TRUE.

For statement 2), suppose both ξ_1 and ξ_2 are incorrect. Then, we have

$$F_{B_1}(\xi_1) + F_{A_2}(\xi_2) \leqslant |\{j \in [2T+1] : \tilde{C}_j \neq C_j\}| \leqslant T,$$
$$F_{B_2}(\xi_2) + F_{A_1}(\xi_1) \leqslant |\{j \in [2T+1] : \tilde{C}_j \neq C_j\}| \leqslant T,$$

which directly implies $U_1 = U_2 =$ TRUE.

For statement 3), suppose that $U_1 =$ FALSE and that $Z_1 =$ TRUE. If ξ_1 were incorrect then ξ_2 had to be correct by statement 2). Among the packets that are sampled in B_1, there would be at most $s - F_{B_1}(\xi_1)$ correct packets. Hence,

$$\begin{aligned}
T+1 &\leqslant \left|\left\{j \in [2T+1] : \tilde{C}_j = C_j\right\}\right| \\
&\leqslant s - F_{B_1}(\xi_1) + F_{A_2}(\xi_2) \\
&\leqslant T,
\end{aligned}$$

which leads to a contradiction, where the last inequality follows from $Z_1 =$ TRUE. Similarly, we have that ξ_2 is correct if $U_2 =$ FALSE and $Z_2 =$ TRUE.

For statement 4), suppose that $U_1 = Z_2 =$ FALSE. If ξ_1 were incorrect then ξ_2 had to be correct by statement 2). Among the packets that are sampled in B_2, there would be at least $s - F_{B_2}(\xi_2)$ incorrect packets. Thus,

$$\begin{aligned}
T &\geqslant |\{j \in [2T+1] : \tilde{C}_j \neq C_j\}| \\
&\geqslant s - F_{B_2}(\xi_2) + F_{A_1}(\xi_1) > T,
\end{aligned}$$

which leads to a contradiction, where the last inequality follows from $Z_2 =$ FALSE. Similarly, we have that ξ_2 is correct if $U_2 = Z_1 =$ FALSE.

This completes the proof. ∎

Now, we proceed with the proof of Theorem II.4

Proof of Theorem II.4: By Remark IV.4, it suffices to show that Construction 2 yields T-VPEC codes that achieve the rate–distortion pair $(1 - s/(2T+1), s/(2T+1))$.

For each message word $\mathbf{x} \in \Sigma^{2T+1}$, the encoder in Construction 2 produces $2T+1$ packets $C_j \in \Sigma^{2T+1-s}$. Clearly, the rate of each packet C_j is $1 - s/(2T+1)$. Next, we show that the output $\mathbf{y} \in \tilde{\Sigma}^{2T+1}$ of the Algorithm 2 is the correct message word $\mathbf{x}$ when there are no errors, and the erasure distortion between $\mathbf{y}$ and $\mathbf{x}$ satisfies $\Delta(\mathbf{x}, \mathbf{y}) \leqslant s/(2T+1)$ when there are at most T erroneous packets.

First, in each iteration of the "while" loop of Algorithm 2, we note the following two facts:

- By Lemma IV.2, one can always find $i_1, i_2 \in I$ such that $i_2 - i_1 \in \{s, s+1, \ldots, 2T+1-s\} \mod 2T+1$.
- In the inner "for" loop, the "if–else if" condition will be satisfied for at least one $\ell \in \{1, 2\}$.

Thus, the "while" loop in Algorithm 2 can always proceed as long as $|I| > s$, and the value of $|I|$ decreases by at least one in each iteration. This ensures that Algorithm 2 always terminates.

Now, consider the case where there are no errors; that is, $\tilde{C}_j = C_j$ for every $j \in [2T+1]$. Then, it holds that $x_i = MF_{A_i}$ and $F_{A_i}(x_i) = |A_i| = 2T+1-s$, for each $i \in [2T+1]$. Since $s \leqslant T$, we have $F_{A_i}(x_i) \geqslant T+1$ for every $i \in [2T+1]$. Therefore, we would have $|I| = 0$ after completing the first "for" loop. The algorithm then terminates and outputs $\mathbf{y} = (x_1, x_2, \ldots, x_{2T+1})$, which is the correct message word.

Next, we consider the case where there are at most T erroneous packets.

Fix an index $i_1 \in [2T+1]$ such that $y_{i_1} \neq e$. The proof consists of the following two parts:

- Suppose y_{i_1} is output in the first "for" loop of Algorithm 2. Then we have $F_{A_{i_1}}(y_{i_1}) \geqslant T+1$. Since there are at most T erroneous packets, this implies $y_{i_1} = x_{i_1}$.
- Suppose y_{i_1} is output in the inner "for" loop of the "while" loop. Then there exists some $i_2 \in [2T+1]$ such that $i_2 - i_1 \in \{s, s+1, \ldots, 2T+1-s\} \bmod 2T+1$. Furthermore, one of the following three conditions must hold: 1) $U_2 =$ TRUE, 2) $U_1 =$ FALSE and $Z_1 =$ TRUE, 3) $U_1 = Z_2 =$ FALSE. In cases 2) and 3), the correctness of $y_{i_1} = \xi_1$ follows directly from the third and fourth statements of Lemma IV.3, respectively. For case 1), by the first statement of Lemma IV.3, ξ_2 is incorrect. This implies that at least $(2T+1-s)/2$ members of $\tilde{A}_2$ are incorrect. Consequently, at most $(s-1)/2$ members of B_1 can be incorrect, which in turn implies that MF_{B_1} is correct.

Therefore, each $y_i \neq e$ in the output is correct. Since there are at least $2T+1-s$ non-erasure symbols in the output, the erasure distortion between the output $\mathbf{y}$ and $\mathbf{x}$ satisfies $\Delta(\mathbf{x}, \mathbf{y}) \leqslant s/(2T+1)$.

Finally, we conclude the proof by analyzing the time complexity of Algorithm 2.

According to its description, in the first "for" loop, one needs to determine the most frequent element of A_i for all $i \in [2T+1]$. In the inner "for" loop of the "while" loop, one needs to determine the most frequent element for at most $|I|(2T+2-2s)$ different set pairs (B_1, B_2). Moreover, finding $i_1, i_2 \in I$ such that $i_2 - i_1 \in \{s, s+1, \ldots, 2T+1-s\} \bmod 2T+1$ requires $O(|I|^2)$ time, and computing the values of U_1, U_2, Z_1, Z_2 for each such pair (i_1, i_2) incurs an additional constant computational cost.

Note that each set A_i has size $2T+1-s$, and for each pair (i_1, i_2) satisfying the condition, both B_1 and B_2 have size s. Hence, by the well-known Boyer-Moore majority vote algorithm [40], which operates with a time complexity linear in the size of the set, the overall time complexity of the decoding algorithm is at most

$$(2T+1) \cdot O(2T+1-s) + |I|(2T+2-2s) \cdot O(s) + O(T^2) = O(T^3).$$

This completes the proof. ∎

V. Conclusion and Further Research

The study presented in this paper is inspired by [7] and [8], which explore achievable rate–distortion pairs and the construction of corresponding VPEC coding schemes. Hence, our primary focus is on investigating the rate–distortion trade-off and constructing VPEC codes with good rate–distortion performance. In the following, we highlight several directions for future research.

1. The primary open question remains the improvement of bounds on the rate–distortion trade-off for the T-VPEC problem. Although we have shown that the bound in Theorem II.2 is tight for the case $N = 2T+1$ via Construction 2, there remains a gap between this bound and the rate–distortion pairs achieved by Construction 1 for general parameter regimes (see Figure 1 for an illustration).
2. In Construction 2 (or Remark IV.4), a message word is equally partitioned into $2T+1$ fragments, every such fragment is contained in exactly $2T+1-s$ packets, and every packet consists of exactly $2T+1-s$ message fragments. Such kinds of combinatorial structures are commonly seen in combinatorial design theory. Thus, a possible direction is to employ similar combinatorial structures from this field to design VPEC codes with better rate–distortion performance.
3. In the original setting of the variable packet-error code problem in [7], the rates of different packets can be different. This fits the communication scenario where different paths in the network have different channel environments. Thus, it's also interesting to investigate the achievable rate–distortion vectors and constructions of corresponding coding schemes in this general setting.

Acknowledgment

We are grateful to Prof. Aaron B. Wagner for introducing us to the question of variable packet-error coding. We also thank the anonymous reviewers for their comments and suggestions, which have improved the quality and presentation of this manuscript.

Appendix A
The Diametric Theorem and the Proof of Corollary II.1

Theorem A.1 ([32, Diametric Theorem])**.** *For positive integers n, d, and $q \geqslant 2$ satisfying $n - d \geqslant 1$, let $r \geqslant 0$ be the largest integer such that*

$$2r \leqslant \min\left\{d, \frac{2(n-d)-q}{q-2}\right\}, \tag{18}$$

where we set $\frac{1}{q-2} = \infty$ for $q = 2$. Denote $\mathscr{K}_r$ as the following set of words in $[q]^n$:

$$\mathscr{K}_r = \{\mathbf{x} \in [q]^n : |\{i \in [n-d+2r] : x_i = 1\}| \geqslant n-d+r\}. \tag{19}$$

Then, $Ant_q(n,d) = |\mathscr{K}_r|$.

Proof of Corollary II.1: Note that in Theorem A.1, when $q \geqslant \frac{2}{3}(n-d)+2$, we have

$$\begin{aligned}\frac{2(n-d)-q}{q-2} &= \frac{2(n-d)-2}{q-2} - 1 \\ &\leqslant \frac{2(n-d)-2}{\frac{2}{3}(n-d)} - 1 \\ &< 2.\end{aligned}$$

Thus, in this case, the largest integer $r \geqslant 0$ satisfying (18) is $r = 0$. This implies that $Ant_q(n,d) = |\mathscr{K}_0|$ when $q \geqslant \frac{2}{3}(n-d)+2$. Moreover, from (19), we know that $\mathscr{K}_0 = \{\mathbf{x} \in [q]^n : x_i = 1 \text{ for all } i \in [n-d]\}$. Consequently, the result follows directly from $|\mathscr{K}_0| = q^d$ and the first two parts of Theorem II.2. ∎

Appendix B
Asymptotic behaviors of (R^O, D)

By Theorem II.1, when $T/N = \theta$, the overall rate–distortion pair (R^O, D) achieved by polytope codes becomes

$$\left(R^O, \frac{(1-2\theta)^{-1} - R^O}{(1-2\theta)^{-1} - (1-\theta)^{-1}} \cdot \frac{F(T)}{N}\right), \tag{20}$$

for $\frac{1}{1-\theta} \leqslant R^O \leqslant \frac{1}{1-2\theta}$. Since $F(T) = T + \left\lfloor \frac{T^2}{4} \right\rfloor + 1$, we have $F(T)/N \to \infty$ as $N \to \infty$, causing the distortion in (20) to also diverge. Meanwhile, the overall rate–distortion pair achieved by MDS codes is

$$\left(R^O, \frac{1-\theta}{\theta}\left(1-(1-2\theta)R^O\right)\right), \tag{21}$$

which is clearly better than (20).

Moreover, by Theorem II.3, the following overall rate–distortion pair is achievable by L-MDS codes:

$$\left(\frac{L}{L-(L+1)\theta}, L\theta\right).$$

When $R^O = \frac{L}{L-(L+1)\theta}$, the distortion in (21) becomes

$$\begin{aligned}&\frac{1-\theta}{\theta}\left(1 - \frac{1-2\theta}{1-(1+\frac{1}{L})\theta}\right) \\ =&\frac{1-\theta}{\theta} \cdot \frac{(1-\frac{1}{L})\theta}{1-(1+\frac{1}{L})\theta} \\ =&\left(1-\frac{1}{L}\right) \cdot \frac{1-\theta}{1-(1+\frac{1}{L})\theta} \\ =&\left(1-\frac{1}{L}\right) \cdot \left(\frac{L}{L+1} + \frac{1/(L+1)}{1-(1+\frac{1}{L})\theta}\right) \\ =&\frac{L-1}{L+1} + \frac{1}{L+1} \cdot \frac{L-1}{L-(L+1)\theta}.\end{aligned}$$

Consider the function

$$f(x) = Lx - \frac{L-1}{L+1} - \frac{1}{L+1} \cdot \frac{L-1}{L-(L+1)x},$$

we have

$$f'(x) = L - \frac{(L-1)(L+1)^2}{(L-(L+1)x)^2},$$

which is at least $L-1$ for $L \geqslant 1$ and $x \leqslant 1/(L+1)$. That is, for $L \geqslant 2$, the function $f(x)$ is increasing in x over the interval $x \leqslant 1/(L+1)$. Therefore, we obtain

$$f(\theta) \leqslant f\left(\frac{1}{L+1}\right) = 0,$$

whenever $\theta \leqslant 1/(L+1)$. This implies that L-MDS codes achieve a better overall rate–distortion trade-off than MDS codes when $\theta \leqslant 1/(L+1)$. For illustrative comparisons, see Figure 3.

Appendix C
Preserving strong list decodability under interleaving

Let ℓ, n, d be given positive integers such that $d \leqslant n$ and let $\tau \in \mathbb{Z}/3$ be positive such that

$$\tau \leqslant \frac{2d-1}{3}. \tag{22}$$

For $i \in [\ell]$, let Σ_i be a finite alphabet (which, without loss of generality, is assumed to be an Abelian group) and let $\mathcal{C}_i$ be a code in Σ_i^n of minimum distance $\geqslant d$ that is strongly-$(\tau, 2)$-list decodable.

Proposition C.1. *The interleaved code $\otimes_{i\in[\ell]}\mathcal{C}_i$ is strongly-$(\tau,2)$-list decodable over $\otimes_{i\in[\ell]}\Sigma_i$.*

Proof: It suffices to prove the claim for $\ell = 2$, as the general case follows by induction.

Given a $2 \times n$ array

$$\mathbf{Y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix},$$

with rows $\mathbf{y}_1 \in \Sigma_1^n$ and $\mathbf{y}_2 \in \Sigma_2^n$, we show that there can be at most two code arrays $\Gamma \in \mathcal{C}_1 \otimes \mathcal{C}_2$ such that $d_H(\mathbf{Y}, \Gamma) \leqslant \tau$.

Suppose to the contrary that there exist three distinct code arrays $\Gamma_1, \Gamma_2, \Gamma_3 \in \mathcal{C}_1 \otimes \mathcal{C}_2$ such that

$$\sum_{r\in[3]} d_H(\mathbf{Y}, \Gamma_r) \leqslant 3\tau. \tag{23}$$

For $r \in [3]$, write

$$\Gamma_r = \begin{pmatrix} \mathbf{c}_1^{(r)} \\ \mathbf{c}_2^{(r)} \end{pmatrix},$$

where $\mathbf{c}_i^{(r)} \in \mathcal{C}_i$, for $i \in [2]$; also, let

$$\mathcal{A}_i^{(r)} = \mathsf{Supp}\big(\mathbf{y}_i - \mathbf{c}_i^{(r)}\big).$$

Then (23) is equivalent to

$$\sum_{r\in[3]} \left|\mathcal{A}_1^{(r)} \cup \mathcal{A}_2^{(r)}\right| \leqslant 3\tau. \tag{24}$$

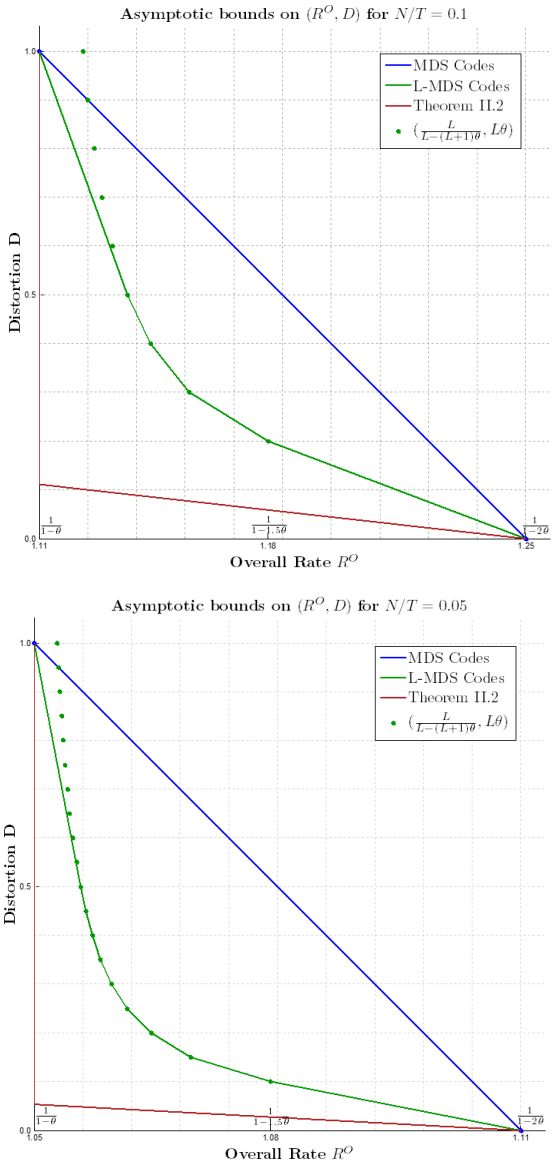


Figure 3. Asymptotic bounds on the overall rate–distortion trade-off from Theorem II.2, MDS codes and L-MDS codes (Corollary II.2), for $T/N = 0.1$, and $T/N = 0.05$.

Yet, since $\mathcal{C}_1$ is strongly-$(\tau, 2)$-list decodable, at least two among the codewords $\mathbf{c}_1^{(1)}, \mathbf{c}_1^{(2)}, \mathbf{c}_1^{(3)}$ must be the same. Without loss of generality assume that $\mathbf{c}_1^{(2)} = \mathbf{c}_1^{(3)}$, which means that

$$\mathcal{A}_1^{(2)} = \mathcal{A}_1^{(3)} \tag{25}$$

and that $\mathbf{c}_2^{(2)} \neq \mathbf{c}_2^{(3)}$. Similarly, at least two among the codewords $\mathbf{c}_2^{(1)}, \mathbf{c}_2^{(2)}, \mathbf{c}_2^{(3)}$ must be the same, say $\mathbf{c}_2^{(1)} = \mathbf{c}_2^{(2)}$, which means that

$$\mathcal{A}_2^{(1)} = \mathcal{A}_2^{(2)} \tag{26}$$

and that $\mathbf{c}_1^{(1)} \neq \mathbf{c}_1^{(2)}$.

Write

$$\begin{aligned} \mathcal{V} &= \mathcal{A}_1^{(1)}, \\ \mathcal{X} &= \mathcal{A}_1^{(2)} \setminus \mathcal{V} \overset{(25)}{=} \mathcal{A}_1^{(3)} \setminus \mathcal{V}, \\ \mathcal{Y} &= \mathcal{A}_2^{(3)}, \\ \mathcal{Z} &= \mathcal{A}_2^{(1)} \setminus \mathcal{Y} \overset{(26)}{=} \mathcal{A}_2^{(2)} \setminus \mathcal{Y}. \end{aligned}$$

From (24) we get:

$$\big|\mathcal{V} \cup \mathcal{Z}\big| + \big|\mathcal{X} \cup \mathcal{Z}\big| + \big|\mathcal{Y}\big| \leqslant 3\tau,$$

i.e.,

$$\big|\mathcal{Z} \setminus \mathcal{V}\big| + \big|\mathcal{V}\big| + \big|\mathcal{Z} \setminus \mathcal{X}\big| + \big|\mathcal{X}\big| + \big|\mathcal{Y}\big| \leqslant 3\tau.$$

Since $\mathcal{X} \cap \mathcal{V} = \emptyset$, we have

$$\big|\mathcal{Z}\big| \leqslant \big|\mathcal{Z} \setminus \mathcal{X}\big| + \big|\mathcal{Z} \setminus \mathcal{V}\big|$$

and, so, the last two inequalities result in

$$\big|\mathcal{V}\big| + \big|\mathcal{Z}\big| + \big|\mathcal{X}\big| + \big|\mathcal{Y}\big| \leqslant 3\tau.$$

On the other hand, since $\mathcal{C}_1$ and $\mathcal{C}_2$ have minimum distance $\geqslant d$, we also have

$$\big|\mathcal{V}\big| + \big|\mathcal{X}\big| = \big|\mathcal{A}_1^{(1)} \cup \mathcal{A}_1^{(2)}\big| \geqslant \big|\mathsf{Supp}\big(\mathbf{c}_1^{(1)} - \mathbf{c}_1^{(2)}\big)\big| \geqslant d$$

and

$$\big|\mathcal{Y}\big| + \big|\mathcal{Z}\big| = \big|\mathcal{A}_2^{(3)} \cup \mathcal{A}_2^{(2)}\big| \geqslant \big|\mathsf{Supp}\big(\mathbf{c}_2^{(3)} - \mathbf{c}_2^{(2)}\big)\big| \geqslant d.$$

Combining the last three inequalities leads to

$$2d \leqslant \big|\mathcal{V}\big| + \big|\mathcal{X}\big| + \big|\mathcal{Y}\big| + \big|\mathcal{Z}\big| \leqslant 3\tau \overset{(22)}{\leqslant} 2d - 1,$$

which is a contradiction. ■

References

[1] R. Ahlswede, N. Cai, S.-Y. Li, and R. W. Yeung, "Network Information Flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[2] N. Cai and R. W. Yeung, "Network Coding and Error Correction," in *Proceedings of the IEEE Information Theory Workshop*. IEEE, 2002, pp. 119–122.

[3] ——, "Network Error Correction, I: Basic Concepts and Upper Bounds," *Communications in Information and Systems*, vol. 6, no. 1, pp. 19–35, 2006.

[4] ——, "Network Error Correction, II: Lower Bounds," *Communications in Information and Systems*, vol. 6, no. 1, pp. 37 – 54, 2006.

[5] H. Yao, D. Silva, S. Jaggi, and M. Langberg, "Network Codes Resilient to Jamming and Eavesdropping," *IEEE/ACM Transactions on Networking*, vol. 22, no. 6, pp. 1978–1987, 2014.

[6] R. Bassoli, H. Marques, J. Rodriguez, K. W. Shum, and R. Tafazolli, "Network Coding Theory: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 1950–1978, 2013.

[7] E. Ahmed and A. B. Wagner, "Coding for the Large-Alphabet Adversarial Channel," *IEEE Transactions on Information Theory*, vol. 63, no. 10, pp. 6347–6363, 2017.

[8] X. Fan, O. Kosut, and A. B. Wagner, "Variable Packet-Error Coding," *IEEE Transactions on Information Theory*, vol. 64, no. 3, pp. 1530–1547, 2018.

[9] S. Kadekodi, K. V. Rashmi, and G. R. Ganger, "Cluster Storage Systems Gotta Have HeART: Improving Storage Efficiency by Exploiting Disk-Reliability Heterogeneity," in *Proceedings of the 17th USENIX Conference on File and Storage Technologies*, ser. FAST'19. USA: USENIX Association, 2019, pp. 345–358.

[10] S. Kadekodi, F. Maturana, S. J. Subramanya, J. Yang, K. Rashmi, and G. R. Ganger, "PACEMAKER: Avoiding HeART Attacks in Storage Clusters with Disk-adaptive Redundancy," in *14th USENIX Symposium on Operating Systems Design and Implementation (OSDI 20)*, 2020, pp. 369–385.

[11] F. Maturana and K. V. Rashmi, "Convertible Codes: Enabling Efficient Conversion of Coded Data in Distributed Storage," *IEEE Transactions on Information Theory*, vol. 68, no. 7, pp. 4392–4407, 2022.

[12] F. Maturana and K. Rashmi, "Code Conversions in Storage Systems," *IEEE BITS the Information Theory Magazine*, 2024.

[13] V. K. Goyal, "Multiple Description Coding: Compression Meets the Network," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 74–93, 2001.

[14] A. Gamal and T. Cover, "Achievable Rates for Multiple Descriptions," *IEEE Transactions on Information Theory*, vol. 28, no. 6, pp. 851–857, 1982.

[15] E. Ahmed and A. B. Wagner, "Erasure Multiple Descriptions," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1328–1344, 2012.

[16] O. Goldreich, *Foundations of Cryptography: Volume 2, Basic Applications*. USA: Cambridge University Press, 2009.

[17] E. Ahmed and A. B. Wagner, "Lossy Source Coding with Byzantine Adversaries," in *2011 IEEE Information Theory Workshop*. IEEE, 2011, pp. 462–466.

[18] R. M. Roth, "Higher-Order MDS Codes," *IEEE Transactions on Information Theory*, vol. 68, no. 12, pp. 7798–7816, 2022.

[19] J. Brakensiek, S. Gopi, and V. Makam, "Lower Bounds for Maximally Recoverable Tensor Codes and Higher Order MDS Codes," *IEEE Transactions on Information Theory*, vol. 68, no. 11, pp. 7125–7140, 2022.

[20] T. M. Cover and J. A. Thomas, *Rate Distortion Theory*. John Wiley & Sons, Ltd, 2005, ch. 10, pp. 301–346.

[21] P. Elias, *List Decoding for Noisy Channels*. Research Laboratory of Electronics, Massachusetts Institute of Technology, 1957.

[22] J. M. Wozencraft, "List Decoding," *Quarterly Progress Report*, vol. 48, pp. 90–95, 1958.

[23] R. Ahlswede, "Channel Capacities for List Xodes," *Journal of Applied Probability*, vol. 10, no. 4, pp. 824–836, 1973.

[24] V. M. Blinovskii, "Bounds for Codes in Decoding by a List of Finite Length," *Problemy Peredachi Informatsii*, vol. 22, no. 1, pp. 11–25, 1986.

[25] V. Blinovsky, *Asymptotic Combinatorial Coding Theory*. Springer Science & Business Media, 1997, vol. 415.

[26] P. Elias, "Error-Correcting Codes for List Decoding," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 5–12, 1991.

[27] J.-Y. Cai, A. Pavan, and D. Sivakumar, "On the Hardness of Permanent," in *Annual Symposium on Theoretical Aspects of Computer Science*. Springer, 1999, pp. 90–99.

[28] V. Guruswami, C. Umans, and S. Vadhan, "Unbalanced Expanders and Randomness Extractors from Parvaresh–Vardy Codes," *Journal of the ACM (JACM)*, vol. 56, no. 4, pp. 1–34, 2009.

[29] D. Sivakumar, "On Membership Comparable Sets," *Journal of Computer and System Sciences*, vol. 59, no. 2, pp. 270–280, 1999.

[30] M. Sudan, L. Trevisan, and S. Vadhan, "Pseudorandom Generators without the XOR Lemma," in *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, 1999, pp. 537–546.

[31] V. Guruswami, A. Rudra, and M. Sudan, "Essential coding theory," *Draft available at http://cse. buffalo. edu/faculty/atri/courses/coding-theory/book*, 2019.

[32] R. Ahlswede and L. H. Khachatrian, "The Diametric Theorem in Hamming Spaces-Optimal Anticodes," *Advances in Applied mathematics*, vol. 20, no. 4, pp. 429–449, 1998.

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

[34] R. Roth, *Introduction to Coding Theory*. Cambridge University Press, 2006.

[35] P. Gopalan, V. Guruswami, and P. Raghavendra, "List decoding tensor products and interleaved codes," *SIAM Journal on Computing*, vol. 40, no. 5, pp. 1432–1462, 2011. [Online]. Available: https://doi.org/10.1137/090778274

[36] J. Brakensiek, S. Gopi, and V. Makam, "Generic Reed-Solomon Codes Achieve List-decoding Capacity," in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, 2023, pp. 1488–1501.

[37] J. Brakensiek, M. Dhar, and S. Gopi, "Improved Field Size Bounds for Higher Order MDS Codes," *IEEE Transactions on Information Theory*, 2024.

[38] I. Tamo, "Tighter List-Size Bounds for List-Decoding and Recovery of Folded Reed-Solomon and Multiplicity Codes," *IEEE Transactions on Information Theory*, 2024.

[39] Y. Chen and Z. Zhang, "Explicit Folded Reed-Solomon and Multiplicity Codes Achieve Relaxed Generalized Singleton Bounds," *arXiv preprint arXiv:2408.15925*, 2024.

[40] R. S. Boyer and J. S. Moore, "MJRTY-A Fast Majority Vote Algorithm," in *Automated reasoning: essays in honor of Woody Bledsoe*. Springer, 1991, pp. 105–117.

Xiangliang Kong (Member, IEEE) received the B.Sc. degree in mathematics from Xi'an Jiaotong University, China, in 2014, and the Ph.D. degree in mathematics from Capital Normal University, China, in 2021. From 2021 to 2025, he was a postdoctoral researcher with the School of Electrical & Computer Engineering, Tel Aviv University, Israel. From 2025 to 2026, he was a postdoctoral researcher with the School of Electrical and Electronics Engineering, Bilkent University, Turkey. He is currently with the State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China. His research interests include coding theory, information theory, combinatorics, and their interactions.

Xin Wang

Ron M. Roth (Life Fellow, IEEE) received the B.Sc. degree in computer engineering, the M.Sc. in electrical engineering, and the D.Sc. in computer science from Technion—Israel Institute of Technology, Haifa, Israel, in 1980, 1984, and 1988, respectively. Since 1988 he has been with the Computer Science Department at Technion, where he now holds the General Yaakov Dori Chair in Engineering. During the academic years 1989–91 he was a Visiting Scientist at IBM Research Division, Almaden Research Center, San Jose, California. During 1996–97, 2004–05, and 2011–2012 he was on sabbatical leave at Hewlett–Packard Laboratories, Palo Alto, California, and in 2022 at Hewlett Packard Enterprise, Milpitas, California. He is the author of the book *Introduction to Coding Theory*, published by Cambridge University Press in 2006. Dr. Roth was an associate editor for coding theory in IEEE TRANSACTIONS ON INFORMATION THEORY from 1998 till 2001, and he is now serving as an associate editor in *SIAM Journal on Discrete Mathematics*. His research interests include coding theory, information theory, and their application to storage, computation, and the theory of complexity.

Itzhak Tamo (Senior Member, IEEE) received the B.A. degree in mathematics and the B.Sc. and Ph.D. degrees in electrical engineering from Ben-Gurion University of the Negev, Israel, in 2008 and 2012, respectively. From 2012 to 2014, he was a Post-Doctoral Researcher with the Institute for Systems Research, University of Maryland, College Park. Since 2015, he has been with the Electrical Engineering Department, Tel Aviv University, Israel, where he was promoted to an Associate Professor in 2020. His research interests include storage systems and devices, coding, information theory, and combinatorics. He was a recipient of the IEEE Communication Society Data Storage Technical Committee 2013 Best Paper Award. He also received the 2015 IEEE Information Theory Society Paper Award and the Krill Prize in 2018.