Bounds and Constructions for Granular Media Coding

Artyom Sharov and Ron M. Roth, *Fellow, IEEE*

***Abstract*—Bounds on the rates of grain-correcting codes are presented. The lower bounds are Gilbert–Varshamov-like ones, whereas the upper bounds improve on the previously known result by Mazumdar *et al.* Constructions of t-grain-correcting codes of length n for certain values of n and t are discussed. Finally, an infinite family of codes of rate approaching 1 that can detect an arbitrary number of grain errors is shown to exist.**

***Index Terms*—convex optimization, Gilbert–Varshamov bound, grain-correcting codes, granular media, lower bounds, magnetic recording, Markov chain, upper bounds.**

I. Introduction

The essential building blocks of conventional magnetic recording media are so-called *grains* which are arbitrarily-shaped two-dimensional magnetizable units assuming one of two possible types of polarity. In modern technologies, the writing medium is partitioned into cells, typically larger in size than the grains, thereby determining how the process of setting a value to a cell is carried out, namely, the process boils down to magnetizing all the grains within the boundaries of this cell. Recently, a novel mechanism was proposed by Wood *et al.* [23] that enables to magnetize areas that are proportionate in their size to the size of grains effectively creating a different type of medium where the grain polarity is determined by the last bit written into the grain. However, recording with areal densities this high introduced new errors to the grains located in the immediate vicinity of the grain being written, which made the case for a technique called *bit-patterned media recording*. This technique makes use of regularly-shaped magnetic units insulated from one another by a nonmagnetic substance to circumvent the aforementioned problem brought on by the high-density writing. However, this technique is not without its own flaws, as it necessitates a faultless synchronization of the write head position over the magnetic units. This, coupled with the specific geometry of the write head whose magnetic field extends across several units, may cause overlapping patterns of errors in shingled writing on the bit-patterned media [5]. Iyengar *et al.* [7] modeled the one-dimensional versions of both media as write channels and studied their information-theoretic properties.

Mazumdar *et al.* [17] considered a combinatorial error model describing the one-dimensional granular medium. In what follows, we will define a somewhat generalized version of that model by augmenting it with the overlapping error patterns occurring during shingled writing on bit-patterned media. Let $\langle s\rangle$ denote the set $\{0,1,\ldots,s-1\}$ for any positive integer s. Let $\Sigma=\langle q\rangle$ be an alphabet for an integer $q\geq 2$.[1] A *grain (of length* 2*)* ending at location $e\in\langle n\rangle\setminus\{0\}$ in a word $\boldsymbol{x}=(x_i)_{i\in\langle n\rangle}$ of length n over Σ causes the value of x_e to equal that of x_{e-1}. Given n consecutive positions on the medium (where words of length n over Σ are to be written), define a *grain pattern* as a set $\mathcal{S}\subseteq\langle n\rangle\setminus\{0\}$ containing all the locations in these n positions where grains end. We will commonly refer to the elements of $\mathcal{S}$ (which indicate grain locations) simply as grains. Thus, a grain pattern $\mathcal{S}$ inflicts errors to a word $\boldsymbol{x}=(x_i)_{i\in\langle n\rangle}$ over Σ by means of the smearing operator $\sigma_{\mathcal{S}}$ that yields an output word $\boldsymbol{y}=(y_i)_{i\in\langle n\rangle}=\sigma_{\mathcal{S}}(\boldsymbol{x})$ over Σ in the following way: for any index $e\in\langle n\rangle\setminus\{0\}$,

$$y_e=\begin{cases}x_{e-1} & \text{if } e\in\mathcal{S}\\ x_e & \text{otherwise}\end{cases}.$$

We will say that a grain pattern has *overlaps* if there exist two grains $e,e'\in\mathcal{S}$ such that $e'=e+1$; otherwise the grain pattern will be called *nonoverlapping*. It should be pointed out that we use the term "overlaps" with respect to grains in a borrowed sense to uniformize two similar error models; the physical notion of grains does not really apply to the application of [5] and [7].

Example 1.1: Let $\Sigma=\langle 3\rangle$ ($q=3$), $n=6$, $\boldsymbol{x}=102022$, $\mathcal{S}=\{1,3,5\}$ and $\mathcal{S}'=\{1,2\}$. Then $\sigma_{\mathcal{S}}(\boldsymbol{x})=112222$ and $\sigma_{\mathcal{S}'}(\boldsymbol{x})=110022$. The grain pattern $\mathcal{S}$ is nonoverlapping whereas the grain pattern $\mathcal{S}'$ has overlaps. □

For a positive integer t and $\boldsymbol{x},\boldsymbol{y}\in\Sigma^n$, we say that $\boldsymbol{x}$ and $\boldsymbol{y}$ are *t-confusable* if there exist grain patterns $\mathcal{S},\mathcal{S}'$ of size at most t for which $\sigma_{\mathcal{S}}(\boldsymbol{x})=\sigma_{\mathcal{S}'}(\boldsymbol{y})$. Words $\boldsymbol{x}$ and $\boldsymbol{y}$ are *confusable* if they are t-confusable for some finite t; otherwise, we say that they are *non-confusable*. A code $\mathcal{C}$ of length n over Σ (namely, a nonempty subset of Σ^n) is called *t-grain-correcting* if no two distinct codewords in $\mathcal{C}$ are t-confusable. A code $\mathcal{C}$ will be called *∞-grain-correcting* if any pair of distinct codewords in $\mathcal{C}$ are non-confusable. Let $M_q^{(\mathrm{N})}(n,t)$ and $M_q^{(\mathrm{O})}(n,t)$ denote the largest size of any t-grain-correcting

Artyom Sharov and Ron M. Roth are with the Computer Science Department, Technion, Haifa 32000, Israel.

Emails: {sharov, ronny}@cs.technion.ac.il.

This work was supported in part by Grants 1280/08 and 1092/12 from the Israel Science Foundation. Part of this work was presented at the IEEE International Symposium on Information Theory, St. Petersburg, Russia, July–August 2011.

[1]The application of [5] and [7] supports our model for $q=2$; the motivation for studying the nonbinary case is mainly theoretical at this stage.

code of length n over Σ when overlaps are disallowed and allowed, respectively (readily, $M_q^{(\mathrm{N})}(n,t) \geq M_q^{(\mathrm{O})}(n,t)$ for any q, n, and t). For $\tau \in [0,1]$ and $j \in \{\mathrm{N},\mathrm{O}\}$, define the (asymptotic) *rate* of $\lceil \tau n \rceil$-grain-correcting codes over Σ as

$$R_q^{(j)}(\tau) = \limsup_{n\to\infty} \frac{1}{n} \log_q M_q^{(j)}(n, \lceil \tau n \rceil).$$

The main objective of this paper is obtaining lower and upper bounds on $M_q^{(j)}(n,t)$ and on $R_q^{(j)}(\tau)$ for $j \in \{\mathrm{N},\mathrm{O}\}$. In Section II, we compute asymptotic Gilbert–Varshamov-like lower bounds on $R_q^{(j)}(\tau)$ for different values of q, using several results from [11] and [15]. The main effort in this method will be to estimate the size of a set of ordered pairs of t-confusable words of length n taken from some subset $\mathcal{X}$ of Σ^n. To this end, we will make a reduction from ordered pairs of t-confusable words of length n to cycles of length n in a specifically designed directed graph, and the growth rate of the number of these cycles will then be assessed using the tools of the Markov-chain machinery.

In Section III, we find an upper bound on $M_2^{(j)}(n,t)$ using a general technique of Abdel-Ghaffar and Weber [1] and improve the best known upper bound on $R_2^{(\mathrm{N})}(\tau)$. This technique, like many of its counterparts, is essentially based on sphere-packing argument. However, unlike the traditional Hamming-metric setting, the sizes of spheres of some fixed radius around binary words with respect to the granular error model, as defined above, are not all equal, thereby invalidating the simple upper bound, obtained in the Hamming-metric setting, as the size of the space divided by the size of a sphere around, say, the all-zero word 0^n. Bounding the sizes of all the spheres from below does not lend itself to a good upper bound either, as there is a large discrepancy between the sizes of spheres and some of them (e.g., those around the words 0^n and 1^n) might be very small. The technique of Abdel-Ghaffar and Weber suggests to overcome these problems by sorting the spheres around all the binary words in an increasing order of their sizes and by packing the space $\langle 2 \rangle^n$ with spheres in that order as long as their collective size is less than 2^n.

In Section IV, we present constructions of binary t-grain-correcting codes of length n for some values of n and t and show the optimality and the uniqueness of some of these codes. Most of the constructions are based on a simple construction from [17, Sec. 2] for ∞-grain-correcting codes. Finally, in Section V, we demonstrate how to obtain codes of rate approaching 1 (as $n \to \infty$) that can detect an arbitrary number of grain errors for any alphabet size q.

We mention that the grain error model somewhat resembles *bitshift errors*, typical in magnetic recording systems with peak detection [12], [13], [16, Sec. 7.1], [20]; such errors occur due to the misdetection of recorded pairs of 01's as 10's and vice versa, thereby shifting the read 1 in those pairs from its designated position one position to the left or to the right. Another similar error model appears to be that of *overreach errors*, which may be found in phase-change memories [8, Sec. 3]; these errors are characterized by smearing the recorded 1's to the adjacent (from the left and/or from the right) 0's causing misdetection of those 0's as 1's.

II. Gilbert–Varshamov-like bounds

We start by stating the main result of this section, the proof of which appears in Section II-B.

Let

$$H_q(p) = -p\log_q p - (1-p)\log_q(1-p) + p\log_q(q-1) \quad (1)$$

be the q-ary entropy function and let $\lambda(\cdot)$ denote the spectral radius of a square real matrix.

Theorem 2.1: Let $q \geq 2$ be an integer, and let

$$\mathcal{A}^{(\mathrm{N})} = \begin{array}{c|ccc} & 0 & 1 & 2 \\ \hline 0 & 1+(q-1)h^2 & 2(q-1)hz & (q-1)(q-2)h^2z^2 \\ 1 & 2h+(q-2)h^2 & (q-1)h^2z & 0 \\ 2 & 2h+(q-2)h^2 & 0 & 0 \end{array} \quad (2)$$

and

$$\mathcal{A}^{(\mathrm{O})} = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1+(q-1)h^2 & 2(q-1)hz+(q-1)(q-2)h^2z^2 \\ 1 & 2h+(q-2)h^2 & (2q-3)h^2z \end{array} \quad (3)$$

be parametric real matrices. Then for $j \in \{\mathrm{N},\mathrm{O}\}$,

$$R_q^{(j)}(\tau) \geq \sup_{\substack{p\in[0,1],\, z\in(0,1] \\ h\in(0,\infty)}} \Big\{ 2H_q(p) - \log_q \lambda(\mathcal{A}^{(j)}(z,h)) + 2\tau\log_q z + 2p\log_q h \Big\} . \quad (4)$$

This section is organized as follows. In Section II-A, we list the definitions (following the notation in [15, Sec. 3]) and the known results which will be of use later in Section II. In Section II-B, we prove Theorem 2.1, using the tools mentioned in Section II-A. In a nutshell, the proof relies on counting cycles of a certain type in a specifically designed finite directed graph. The cycles will be shown to correspond to pairs of words $(\boldsymbol{x}, \boldsymbol{y})$ along with a (minimal) grain pattern $\mathcal{S}$ that makes them confusable (i.e., $\mathcal{S}$ is a minimal grain pattern for which $\sigma_{\mathcal{S}}(\boldsymbol{x}) = \sigma_{\mathcal{S}}(\boldsymbol{y})$). Section II-C compares the lower bounds obtained from Theorem 2.1 with known results. Finally, we conclude by considering in Section II-D the more general error model where grains can be of length larger than 2 and describe how the technique of Section II-B can be adapted to yield lower bounds on the rate of binary codes that correct such error patterns.

A. Definitions and useful tools

1) Graphs and Markov chains: Let $G = (V_G, E_G)$ be a (finite) directed graph with state set V_G and edge set $E_G \subseteq V_G \times V_G$, without parallel edges. The graph G will be called *primitive* if there exists a positive integer s such that for any ordered pair of states (v, v') of V_G there is a path of length s from v to v' in G.

Let $P : E_G \to [0,1]$ be a probability distribution on E_G, namely, $\sum_{e\in E_G} P(e) = 1$. A *stationary Markov chain* is a probability distribution P on E_G such that

$$\sum_{v' : e=(v',v)\in E_G} P(e) = \sum_{v' : e=(v,v')\in E_G} P(e)$$

for any $v \in V_G$. The (stationary) probability $\pi(v)$ to be in a state $v \in V_G$ in a random walk on G according to a stationary Markov chain P is

$$\pi(v) = \sum_{v' : e=(v,v') \in E_G} P(e).$$

For a path $\gamma = (v_i)_{i \in \langle n+1 \rangle}$ of length[2] n in G, let $\mathsf{P}_\gamma : E_G \to [0,1]$ be the *empirical probability distribution* of γ, namely, for $e \in E_G$,

$$\mathsf{P}_\gamma(e) = \frac{1}{n} \Big| \{ i \in \langle n \rangle : (v_i, v_{i+1}) = e \} \Big| ;$$

it is readily verified that when γ is a cycle (namely, when $v_0 = v_n$), P_γ is a stationary Markov chain on E_G.

For a probability distribution $P : E_G \to [0,1]$, a positive integer k and a vector function $f : E_G \to \mathbb{R}^k$, denote by $\mathsf{E}_P \{f\}$ the expected value of f with respect to P, that is,

$$\mathsf{E}_P \{f\} = \sum_{e \in E_G} P(e) f(e). \tag{5}$$

Finally, for an integer $q \geq 2$, the *entropy rate* of a stationary Markov chain P is defined as

$$\mathsf{H}_q(P) = - \sum_{\substack{v \in V_G: \\ \pi(v) > 0}} \; \sum_{\substack{v' : e = (v,v') \in E_G \\ \text{s.t. } P(e) > 0}} P(e) \log_q \frac{P(e)}{\pi(v)} \, .$$

2) Optimizing concave functions: We proceed by citing special cases of [15, Lemma 2] and [15, Lemma 5] which are consequences of well-known results on optimizing convex (concave) functions subject to linear equality and linear inequality constraints (also see [3, Lemma 2], [14, pp. 312–316], [18, Ch. 2, Th. 25], and [19, Sec. 28]) and which are to be employed in the following subsection. In both lemmas, $\mathcal{M}(f; U)$ denotes the set of all stationary Markov chains P on a graph G such that $\mathsf{E}_P \{f\} \in U \subseteq \mathbb{R}^k$, for a given function $f : E_G \to \mathbb{R}^k$.

Lemma 2.2: Let $G = (V_G, E_G)$ be a primitive directed graph and $f : E_G \to \mathbb{R}^k$ be a function. Let U be an open rectangular parallelepiped $\prod_{i \in \langle k \rangle} (\tilde{s}_i, s_i)$ and let Γ_n denote the set of all cycles of length n in G. Then

$$\lim_{n \to \infty} \frac{1}{n} \log_q \left| \{ \gamma \in \Gamma_n : \mathsf{E}_{\mathsf{P}_\gamma} \{f\} \in U \} \right| = \sup_{P \in \mathcal{M}(f;U)} \mathsf{H}_q(P).$$

Let k be a positive integer. For a graph $G = (V_G, E_G)$, vectors of positive real indeterminates $\boldsymbol{z} = (z_i)_{i \in \langle k \rangle}$ and $\boldsymbol{h} = (h_i)_{i \in \langle k' \rangle}$, and functions $f = (f_i)_{i \in \langle k \rangle} : E_G \to \mathbb{R}^k$ and $f' = (f'_i)_{i \in \langle k' \rangle} : E_G \to \mathbb{R}^{k'}$, define the parametric real matrix $\mathsf{A}_G(\boldsymbol{z}, \boldsymbol{h})$ (whose rows and columns are indexed by V_G) as

$$[\mathsf{A}_G(\boldsymbol{z}, \boldsymbol{h})]_{v,v' \in V_G} = \begin{cases} \boldsymbol{z}^{f(e)} \boldsymbol{h}^{f'(e)} & \text{if } e = (v,v') \in E_G \\ 0 & \text{otherwise} \end{cases} , \tag{6}$$

where $\boldsymbol{z}^{f(e)} \boldsymbol{h}^{f'(e)} = \prod_{i \in \langle k \rangle} z_i^{f_i(e)} \cdot \prod_{i \in \langle k' \rangle} h_i^{f'_i(e)}$.

Lemma 2.3: Let $G = (V_G, E_G)$ be a directed graph. Let $\boldsymbol{p} = (p_i)_{i \in \langle k' \rangle} \in [0,1]^{k'}$ be a vector and let $f : E_G \to \mathbb{R}^k$, $f' : E_G \to \mathbb{R}^{k'}$ be functions. Let U be a closed rectangular parallelepiped $\prod_{i \in \langle k \rangle} [0, s_i]$. Then

$$\sup_{\substack{P \in \mathcal{M}(f;U): \\ \mathsf{E}_P\{f'\} = \boldsymbol{p}}} \mathsf{H}_q(P) = \inf_{\boldsymbol{z}, \boldsymbol{h}} \Big\{ \log_q \lambda(\mathsf{A}_G(\boldsymbol{z}, \boldsymbol{h})) - \sum_{i \in \langle k \rangle} s_i \log_q z_i - \sum_{i \in \langle k' \rangle} p_i \log_q h_i \Big\},$$

where $\lambda(\cdot)$ denotes the spectral radius of a square real matrix, $\boldsymbol{z} = (z_i)_{i \in \langle k \rangle}$ ranges over $(0,1]^k$ and $\boldsymbol{h} = (h_i)_{i \in \langle k' \rangle}$ ranges over $(0, \infty)^{k'}$.

3) Basic bound: For an integer $q \geq 2$, an alphabet $\Sigma = \langle q \rangle$, positive integers t and n, a subset $\mathcal{X} \subseteq \Sigma^n$, and a word $\boldsymbol{x} \in \mathcal{X}$, let $\mathcal{R}_t^{(\mathrm{N})}(\boldsymbol{x}; \mathcal{X})$ and $\mathcal{R}_t^{(\mathrm{O})}(\boldsymbol{x}; \mathcal{X})$ be the sets of all the words in $\mathcal{X}$ that are t-confusable with $\boldsymbol{x}$ when overlaps are disallowed and allowed, respectively.

Example 2.4: Let $\Sigma = \langle 2 \rangle$. It can be readily verified that

$$\mathcal{R}_2^{(\mathrm{N})}(0000; \Sigma^4) = \mathcal{R}_2^{(\mathrm{O})}(0000; \Sigma^4) = \{0000, 0100, 0010, 0001, 0101\}$$

and

$$\mathcal{R}_2^{(\mathrm{N})}(0101; \Sigma^4) = \mathcal{R}_2^{(\mathrm{O})}(0101; \Sigma^4) = 0\Sigma^3 = \{ 0\boldsymbol{y} : \boldsymbol{y} \in \Sigma^3 \} \, .$$

□

Using the standard Gilbert–Varshamov-type argument, one can obtain the following lower bound on $M_q^{(\mathrm{N})}(n,t)$ and $M_q^{(\mathrm{O})}(n,t)$:

$$M_q^{(j)}(n,t) \geq \frac{|\mathcal{X}|}{\max\{ |\mathcal{R}_t^{(j)}(\boldsymbol{x}; \mathcal{X})| : \boldsymbol{x} \in \mathcal{X} \}} \, , \tag{7}$$

where $j \in \{\mathrm{N}, \mathrm{O}\}$ and $\mathcal{X}$ is any subset of Σ^n. Namely, we pick up a word $\boldsymbol{x}$ of $\mathcal{X}$ and remove from $\mathcal{X}$ the words of $\mathcal{R}_t^{(j)}(\boldsymbol{x}; \mathcal{X})$; we repeat the process until $\mathcal{X}$ is empty. Note, however, that the sets $\mathcal{R}_t^{(j)}(\boldsymbol{x}; \mathcal{X})$ for various words $\boldsymbol{x} \in \mathcal{X}$ may have different sizes, as demonstrated in Example 2.4; therefore, replacing the size of a largest such set in the denominator of (7) with the size of the average set (when $\boldsymbol{x}$ ranges over $\mathcal{X}$) is clearly advantageous and, as we will see later on, produces better lower bounds on the rates of grain-correcting codes.

For $j \in \{\mathrm{N}, \mathrm{O}\}$, let

$$\mathsf{W}_t^{(j)}(\mathcal{X}) = \sum_{\boldsymbol{x} \in \mathcal{X}} \Big| \mathcal{R}_t^{(j)}(\boldsymbol{x}; \mathcal{X}) \Big| . \tag{8}$$

Namely, $\mathsf{W}_t^{(\mathrm{N})}(\mathcal{X})$ and $\mathsf{W}_t^{(\mathrm{O})}(\mathcal{X})$ are the number of ordered pairs of t-confusable words in $\mathcal{X}$ when overlaps are disallowed and allowed, respectively. The following lemma, whose goal is to replace the denominator of the right-hand side of (7) with the average set $\frac{1}{|\mathcal{X}|} \mathsf{W}_t^{(j)}(\mathcal{X})$, is essentially a reformulation of [11, Lemma 1] for grain-correcting codes (also see [12, Sec. 3]). The lemma is a centerpiece of this section, with most of the rest of the section devoted to developing tight upper bounds on the asymptotic growth rate of $\mathsf{W}_t^{(j)}(\mathcal{X})$ for $j \in \{\mathrm{N}, \mathrm{O}\}$.

[2] A length of a path γ is the number of edges along γ. Since G has no parallel edges, we will specify a path γ through the sequence of states along γ.

Lemma 2.5: Let n, t be positive integers and let $\mathcal{X} \subseteq \Sigma^n$. Then, for $j \in \{\mathrm{N}, \mathrm{O}\}$,

$$M_q^{(j)}(n,t) \geq \frac{|\mathcal{X}|^2}{4\mathsf{W}_t^{(j)}(\mathcal{X})}. \tag{9}$$

Proof: For any positive integer t and $j \in \{\mathrm{N}, \mathrm{O}\}$, let

$$\overline{\mathcal{R}}_t^{(j)}(\mathcal{X}) = \frac{\mathsf{W}_t^{(j)}(\mathcal{X})}{|\mathcal{X}|} = \frac{1}{|\mathcal{X}|} \sum_{\boldsymbol{x} \in \mathcal{X}} \left|\mathcal{R}_t^{(j)}(\boldsymbol{x}; \mathcal{X})\right|.$$

At most half of the sets $\mathcal{R}_t^{(j)}(\boldsymbol{x}; \mathcal{X})$ for $\boldsymbol{x}$ ranging over $\mathcal{X}$ have size greater than $2\overline{\mathcal{R}}_t^{(j)}(\mathcal{X})$, or else the average size of $\mathcal{R}_t^{(j)}(\boldsymbol{x}; \mathcal{X})$ over all $\boldsymbol{x} \in \mathcal{X}$ would exceed $\overline{\mathcal{R}}_t^{(j)}(\mathcal{X})$. Therefore, there are at least $|\mathcal{X}|/2$ words $\boldsymbol{x} \in \mathcal{X}$ such that $|\mathcal{R}_t^{(j)}(\boldsymbol{x}; \mathcal{X})| \leq 2\overline{\mathcal{R}}_t^{(j)}(\mathcal{X})$. Denote this subset of $\mathcal{X}$ by $\mathcal{X}'$. Clearly, for any $\boldsymbol{x} \in \mathcal{X}'$,

$$\left|\mathcal{R}_t^{(j)}(\boldsymbol{x}; \mathcal{X}')\right| \leq \left|\mathcal{R}_t^{(j)}(\boldsymbol{x}; \mathcal{X})\right| \leq 2\overline{\mathcal{R}}_t^{(j)}(\mathcal{X}).$$

Therefore, by iteratively picking up a word $\boldsymbol{x}$ of $\mathcal{X}'$ and removing the set $\mathcal{R}_t^{(j)}(\boldsymbol{x}; \mathcal{X}')$ from $\mathcal{X}'$, we can construct a t-grain-correcting code of size at least $|\mathcal{X}'|/(2\overline{\mathcal{R}}_t^{(j)}(\mathcal{X})) \geq |\mathcal{X}|/(4\overline{\mathcal{R}}_t^{(j)}(\mathcal{X}))$. □

Remark 2.6: We can get rid of the factor 4 in the denominator of the right-hand side of (9) by following the proof in [21], but it will have no effect on the asymptotic analysis we are about to do. □

Remark 2.7: One readily observes that $\mathcal{R}_t^{(j)}(\boldsymbol{x}; \langle q \rangle^n)$, for $j \in \{\mathrm{N}, \mathrm{O}\}$ and a word $\boldsymbol{x} \in \langle q \rangle^n$, is contained in the Hamming sphere of radius $2t$ around $\boldsymbol{x}$. Therefore the traditional Gilbert–Varshamov bound $\varrho_q^{(\mathrm{GV})}(\tau) = 1 - H_q(2\tau)$ (where $H_q(p)$ is the q-ary entropy function defined in (1) and $\tau = t/n$) on the rate $R_q^{(j)}(\tau)$ for $j \in \{\mathrm{N}, \mathrm{O}\}$ can be obtained from Lemma 2.5 by taking $\mathcal{X} = \langle q \rangle^n$ and bounding $|\mathcal{R}_t^{(j)}(\boldsymbol{x}; \langle q \rangle^n)|$ from above by the size of the Hamming sphere of radius $2t$. This specific selection of $\mathcal{X}$, for the fixed values of q and n, clearly maximizes the numerator in the right-hand side of (9), yet does not necessarily minimize the entire right-hand side. Indeed, as we will see later on, it will be beneficial to select sets $\mathcal{X}$ with smaller asymptotic growth rate than that of $\langle q \rangle^n$ (namely, rate smaller than 1), such that the growth rate of $\mathsf{W}_t^{(j)}(\mathcal{X})$ is smaller than $H_q(2\tau)$ and the growth rate of the right-hand side of (9) is greater than $1 - H_q(2\tau)$. □

B. Proof of Theorem 2.1

This subsection is structured as follows. First, we define two finite directed graphs $\mathcal{G}^{(\mathrm{N})}$ and $\mathcal{G}^{(\mathrm{O})}$ in Section II-B1 and then, in Section II-B2, make a reduction from the ordered pairs of t-confusable words of length n to certain cycles of length n in these two graphs. Next, in Section II-B3, we find upper bounds on the growth rate of $\mathsf{W}_t^{(\mathrm{N})}(\mathcal{X})$ and $\mathsf{W}_t^{(\mathrm{O})}(\mathcal{X})$ for certain sets $\mathcal{X}$ of words with prescribed empirical distribution of runs and, subsequently, lower bounds on $R_q^{(\mathrm{N})}(\tau)$ and $R_q^{(\mathrm{O})}(\tau)$. Throughout this subsection, $\Sigma = \langle q \rangle$ for some integer $q \geq 2$.

1) Graph presentations: Define two finite directed graphs $\mathcal{G}^{(\mathrm{N})} = (V^{(\mathrm{N})}, E^{(\mathrm{N})})$, $\mathcal{G}^{(\mathrm{O})} = (V^{(\mathrm{O})}, E^{(\mathrm{O})})$ corresponding to the scenarios without and with grain overlaps, respectively. The paths in each graph will correspond to pairs of words combined with the minimal grain pattern that makes them confusable (in the nonoverlapping and overlapping settings). Let $\underline{\Sigma} = \{\underline{a} : a \in \Sigma\}$ be a set where every element $\underline{a}$ designates a symbol whose original value $a \in \Sigma$ was overrun by a grain error. The set of states $V^{(\mathrm{N})} \subseteq (\Sigma \cup \underline{\Sigma})^2$ is defined as $V^{(\mathrm{N})} = V_0 \cup V_1 \cup V_2$ where

$$V_0 = \{\ell r : \ell = r \in \Sigma\},$$
$$V_1 = \left\{\ell \underline{r} : \ell r \in \Sigma^2,\ \ell \neq r\right\} \cup \left\{\underline{\ell} r : \ell r \in \Sigma^2,\ \ell \neq r\right\}, \text{ and}$$
$$V_2 = \left\{\underline{\ell}\,\underline{r} : \ell r \in \Sigma^2,\ \ell \neq r\right\}.$$

The states of V_0 correspond to the case where the symbols in the same position in two observed words are identical; the states of V_1 correspond to the case where such symbols in the two words are different and one grain (applied either to the first or to the second word as designated by the bar), which does not overlap with other grains, is sufficient to make the two symbols equal; for $q \geq 3$, the states of V_2 correspond to the case where such symbols in the two words are different and a pair of grains that do not overlap is necessary to make these symbols equal (for $q = 2$, we will disregard V_2 altogether). The set of states $V^{(\mathrm{O})} \subseteq \Sigma^2$ is defined as $V^{(\mathrm{O})} = V_0 \cup V_3$ where

$$V_3 = \left\{\ell r : \ell r \in \Sigma^2,\ \ell \neq r\right\}.$$

The states of V_3 correspond to the case where the symbols in the same position in two words are different and one grain (possibly overlapping with others) is sufficient to make those symbols equal. Observe that, in contrast to $V^{(\mathrm{N})}$, the states of $V^{(\mathrm{O})}$ do not encapsulate any information on the grain patterns due to the fact that, by definition, in the overlapping scenario, a grain may start before another grain has ended, whereas in the nonoverlapping setting we are forced to convey the information on where a grain ended from one state to the next.

Specifically, for $q = 2$ (which will be our running example throughout most of this section),

$$V_0 = \{00, 11\}, \quad V_1 = \{\underline{0}1, 0\underline{1}, \underline{1}0, 1\underline{0}\},$$
$$V_2 = \{\underline{0}\,\underline{1}, \underline{1}\,\underline{0}\}, \quad \text{and} \quad V_3 = \{01,\ 10\}$$

(the states of the set V_2 will have no incoming edges for $q = 2$ in $\mathcal{G}^{(\mathrm{N})}$, that is why for $q = 2$, we will disregard V_2 completely).

Next, we define the edge sets $E^{(\mathrm{N})}$ and $E^{(\mathrm{O})}$. Define the function $\phi : \Sigma \cup \underline{\Sigma} \to \Sigma$ as $\phi(a) = \phi(\underline{a}) = a$ for every $a \in \Sigma$. There is an edge in $E^{(\mathrm{N})}$ from state $v = \ell r$ to state $v' = \ell' r'$ if

- [N1] $v' \in V_0$; or
- [N2] $v \in V_0$, $v' \in V_1$, and either $\ell = \ell' \in \Sigma$ or $r = r' \in \Sigma$; or
- [N3] $v, v' \in V_1$, and either $\ell = r' \in \Sigma$ or $\ell' = r \in \Sigma$; or
- [N4] $v \in V_0$, $v' \in V_2$, $\ell \neq \phi(\ell')$, and $r \neq \phi(r')$.

Edges satisfying Condition [N1] correspond to two pairs of symbols (represented by $v = \ell r$ and $v' = \ell' r'$) seen at the same consecutive positions in two observed words,

$$\ldots \ell\ell' \ldots \quad \text{and} \quad \ldots rr' \ldots,$$

where no grain needs to be applied to equalize ℓ' to r'; edges satisfying Condition [N2] correspond to consecutive pairs v and v' with no grain ending at v and with one grain (which does not overlap with other grains) applied to equalize ℓ' to r'; edges satisfying Condition [N3] correspond to consecutive pairs v and v' with a grain ending at the first pair in one word forcing another grain, which does not overlap with others, to be applied to the other word to make the values of v' equal; and edges satisfying Condition [N4] correspond to locations where a pair of grains that do not overlap needs to be applied to both words to equalize ℓ' to r' (which can only occur when $q \geq 3$).

There is an edge in $E^{(\mathrm{O})}$ from $v = \ell r$ to $v' = \ell' r'$ if

[O1] $v' \in V_0$; or
[O2] $v \in V_0$ and $v' \in V_3$; or
[O3] $v, v' \in V_3$, $\ell r \neq r'\ell'$, and either $\ell = r'$ or $r = \ell'$; or
[O4] $v, v' \in V_3$ and $\ell r = r'\ell'$.

Edges satisfying Conditions [O1]–[O3] are similar in their description to their counterparts [N1]–[N3], respectively. Unlike Condition [N4], however, Condition [O4] corresponds to consecutive pairs where a grain overlapping with another grain might be applied. Notice that to equalize ℓ' to r', a grain ending at v' can be applied to either of the two words. Like Condition [N4], Condition [O3] can only occur when $q \geq 3$.

Example 2.8: For $\Sigma = \langle 3 \rangle$, consider the following path of length 5 in $\mathcal{G}^{(\mathrm{N})}$:

$$\gamma = (v_i)_{i \in \langle 6 \rangle} = 11 \quad 22 \quad 2\underline{0} \quad \underline{1}2 \quad 00 \quad \underline{1}\underline{2} \; .$$

The states v_0, v_1, and v_4 belong to the set V_0, the states v_2 and v_3 belong to the set V_1, whereas the state v_5 belongs to the set V_2. The edges (v_i, v_{i+1}) for $i = 0, 1, 2$ correspond to Conditions [N1]–[N3], respectively, and the edge (v_4, v_5) corresponds to Condition [N4]. Now, for the same alphabet Σ, consider the following path of length 7 in $\mathcal{G}^{(\mathrm{O})}$:

$$\gamma = (v_i)_{i \in \langle 8 \rangle} = 11 \quad 22 \quad 20 \quad 12 \quad 00 \quad 12 \quad 21 \quad 12 \; .$$

Here the states v_0, v_1, v_4 belong to the set V_0 whereas the states v_2, v_3, v_5, v_6, and v_7 belong to the set V_3. The edges (v_i, v_{i+1}) for $i = 0, 1, 2$ correspond to Conditions [O1]–[O3], respectively, and the edges (v_5, v_6) and (v_6, v_7) both correspond to Condition [O4]. □

Given a path $\gamma = (\ell_i r_i)_{i \in \langle n \rangle}$ of length $n-1$ in $\mathcal{G}^{(\mathrm{N})}$, define the sets

$$\mathsf{L}(\gamma) = \{ i : \ell_i \in \underline{\Sigma} \} \text{ and } \mathsf{R}(\gamma) = \{ i : r_i \in \underline{\Sigma} \} \, .$$

When the path γ is in $\mathcal{G}^{(\mathrm{O})}$, let

$$\mathsf{L}(\gamma) = \{ i : \ell_i \neq r_i, \; r_{i-1} \neq \ell_i \} \, ,$$
$$\mathsf{R}(\gamma) = \{ i : \ell_i \neq r_i, \; \ell_{i-1} \neq r_i \} \, .$$

In addition, for an edge $e \in (\ell r, \ell' r')$ in $\mathcal{G}^{(\mathrm{O})}$, define the function $\mu : E^{(\mathrm{O})} \to \langle 2 \rangle$ by

$$\mu(e) = \begin{cases} 1 & \text{if } e \text{ satisfies Condition [O4]} \\ 0 & \text{otherwise} \end{cases} ,$$

and extend this definition to any path $\gamma = (\ell_i r_i)_{i \in \langle n \rangle}$ in $\mathcal{G}^{(\mathrm{O})}$ by

$$\mu(\gamma) = \sum_{i \in \langle n-1 \rangle} \mu(\ell_i r_i, \ell_{i+1} r_{i+1}).$$

A path γ in $\mathcal{G}^{(\mathrm{N})}$ starting in V_0 represents a pair

$$(\boldsymbol{x} = (\phi(\ell_i))_{i \in \langle n \rangle}, \boldsymbol{y} = (\phi(r_i))_{i \in \langle n \rangle})$$

of confusable words in Σ^n, as well as grain patterns $\mathcal{S} = \mathsf{L}(\gamma), \mathcal{S}' = \mathsf{R}(\gamma)$ that cause the corresponding overrun words, $\sigma_{\mathcal{S}}(\boldsymbol{x})$ and $\sigma_{\mathcal{S}'}(\boldsymbol{y})$, to be equal. A path γ in $\mathcal{G}^{(\mathrm{O})}$ starting in V_0 represents a pair

$$(\boldsymbol{x} = (\ell_i)_{i \in \langle n \rangle}, \boldsymbol{y} = (r_i)_{i \in \langle n \rangle})$$

of confusable words in Σ^n and $2^{\mu(\gamma)}$ confusing grain patterns $\mathcal{S} = \mathsf{L}(\gamma) \cup \mathsf{M}(\gamma), \mathcal{S}' = \mathsf{R}(\gamma) \cup \mathsf{M}'(\gamma)$ where $(\mathsf{M}(\gamma), \mathsf{M}'(\gamma))$ is a partition of the set (of size $\mu(\gamma)$) of indices $i \in \langle n \rangle \setminus \{0\}$ of (the terminal states of) edges of γ that satisfy Condition [O4]. Indeed, $\mu(\gamma)$ is the number of positions along γ where overlapping grains, if switched from $\boldsymbol{x} = (\phi(\ell_i))_{i \in \langle n \rangle}$ to the corresponding position in $\boldsymbol{y} = (\phi(r_i))_{i \in \langle n \rangle}$ (or vice versa), will still make $\boldsymbol{x}$ and $\boldsymbol{y}$ confusable. For completeness, let $\mu(\gamma) = 0$ when γ is in $\mathcal{G}^{(\mathrm{N})}$.

Example 2.9: Continuing Example 2.8, consider again the path

$$\gamma = (v_i)_{i \in \langle 6 \rangle} = 11 \quad 22 \quad 2\underline{0} \quad \underline{1}2 \quad 00 \quad \underline{1}\underline{2}$$

in $\mathcal{G}^{(\mathrm{N})}$. This path corresponds to the pair of overrun words $122\underline{1}0\underline{1}$ and $12\underline{0}20\underline{2}$ (the grain-free words are $\boldsymbol{x} = 122101$ and $\boldsymbol{y} = 120202$), and the bars indicate the grain patterns $\mathcal{S} = \mathsf{L}(\gamma) = \{3, 5\}, \mathcal{S}' = \mathsf{R}(\gamma) = \{2, 5\}$ that make $\boldsymbol{x}$ and $\boldsymbol{y}$ confusable.

Next, consider the path

$$\gamma = (v_i)_{i \in \langle 8 \rangle} = 11 \quad 22 \quad 20 \quad 12 \quad 00 \quad 12 \quad 21 \quad 12$$

in $\mathcal{G}^{(\mathrm{O})}$. Here $\mathsf{L}(\gamma) = \{3, 5\}$, $\mathsf{R}(\gamma) = \{2, 5\}$, and $\mu(\gamma) = 2$, so that the pair of words 12210121 and 12020212 can be confused by any of the four pairs of grain patters $\mathcal{S} = \mathsf{L}(\gamma) \cup \mathsf{M}(\gamma)$, $\mathcal{S}' = \mathsf{R}(\gamma) \cup \mathsf{M}'(\gamma)$, where $(\mathsf{M}(\gamma), \mathsf{M}'(\gamma))$ is either $(\emptyset, \{6, 7\})$, $(\{6\}, \{7\})$, $(\{7\}, \{6\})$, or $(\{6, 7\}, \emptyset)$. □

The adjacency matrices $A_{\mathcal{G}}^{(\mathrm{N})}$ and $A_{\mathcal{G}}^{(\mathrm{O})}$ of the graphs $\mathcal{G}^{(\mathrm{N})}$ and $\mathcal{G}^{(\mathrm{O})}$ that are constructed as described above are shown in Table I for the case $q = 2$. The entries of these matrices are either 0 or 1 (since there are no parallel edges) and the subscript of each entry denotes the type of the corresponding edge. Notice that for $q = 2$, there are no edges satisfying Condition [N4] in $\mathcal{G}^{(\mathrm{N})}$, due to the aforementioned omission of V_2, and there are no edges satisfying Condition [O3] in $\mathcal{G}^{(\mathrm{O})}$. Also notice that for any q, the graphs $\mathcal{G}^{(\mathrm{N})}$ and $\mathcal{G}^{(\mathrm{O})}$ are primitive, because each entry in $(A_{\mathcal{G}}^{(\mathrm{N})})^2$ and in $(A_{\mathcal{G}}^{(\mathrm{O})})^2$ is strictly positive.

2) Reduction from pairs of t-confusable words to graph cycles: To make the presentation and the computation simpler, we will switch to a different criterion of confusability till the end of this subsection. Given a positive integer t, we will call two words $\boldsymbol{x}, \boldsymbol{y}$ *t-confusable in the wide sense* (or *t-cws*, in short) if there exist grain patterns $\mathcal{S}$ and $\mathcal{S}'$ such that

$$|\mathcal{S}| + |\mathcal{S}'| \leq 2t \quad \text{and} \quad \sigma_{\mathcal{S}}(\boldsymbol{x}) = \sigma_{\mathcal{S}'}(\boldsymbol{y}).$$

TABLE I
ADJACENCY MATRICES $A_{\mathcal{G}}^{(\mathrm{N})}$ AND $A_{\mathcal{G}}^{(\mathrm{O})}$, FOR $q=2$.

$A_{\mathcal{G}}^{(\mathrm{N})} =$

		V_0: 00	V_0: 11	V_1: $\underline{0}1$	V_1: $0\underline{1}$	V_1: $\underline{1}0$	V_1: $1\underline{0}$
V_0	00	1_{N1}	1_{N1}	0	1_{N2}	1_{N2}	0
V_0	11	1_{N1}	1_{N1}	1_{N2}	0	0	1_{N2}
V_1	$\underline{0}1$	1_{N1}	1_{N1}	0	0	0	1_{N3}
V_1	$0\underline{1}$	1_{N1}	1_{N1}	0	0	1_{N3}	0
V_1	$\underline{1}0$	1_{N1}	1_{N1}	0	1_{N3}	0	0
V_1	$1\underline{0}$	1_{N1}	1_{N1}	1_{N3}	0	0	0

$A_{\mathcal{G}}^{(\mathrm{O})} =$

		V_0: 00	V_0: 11	V_3: 01	V_3: 10
V_0	00	1_{O1}	1_{O1}	1_{O2}	1_{O2}
V_0	11	1_{O1}	1_{O1}	1_{O2}	1_{O2}
V_3	01	1_{O1}	1_{O1}	0	1_{O4}
V_3	10	1_{O1}	1_{O1}	1_{O4}	0

Since any t-confusable pair of words is also t-cws, it follows that any t-grain-correcting code in the wide sense is also t-grain-correcting in the ordinary sense. Our results will actually apply to the wide-sense notion of confusability.[3] Under the ordinary confusability criterion, a path γ in $\mathcal{G}^{(\mathrm{O})}$ starting in V_0 and representing a pair $(\boldsymbol{x},\boldsymbol{y})$ of t-confusable words in Σ^n has $2^{\mu(\gamma)}$ pairs of confusing grain patterns $\mathcal{S}$, $\mathcal{S}'$ such that $|\mathcal{S}|+|\mathcal{S}'|$ is minimal and at least one of these confusing pairs of patterns satisfies $|\mathcal{S}|, |\mathcal{S}'| \le t$; under the wide-sense confusability criterion, all of the $2^{\mu(\gamma)}$ confusing grain patterns $\mathcal{S}$, $\mathcal{S}'$ satisfy $|\mathcal{S}|+|\mathcal{S}'| \le 2t$. Due to this relaxed confusability notion, to determine that a path γ in $\mathcal{G}^{(j)}$, for $j \in \{\mathrm{N},\mathrm{O}\}$, represents a pair of t-cws words, it is sufficient to calculate the value of the expression $|\mathsf{L}(\gamma)|+|\mathsf{R}(\gamma)|+\mu(\gamma)$ (which equals $|\mathcal{S}|+|\mathcal{S}'|$ for any of the $2^{\mu(\gamma)}$ pairs of grain patterns $\mathcal{S}$, $\mathcal{S}'$, either nonoverlapping or overlapping, making $\boldsymbol{x}$ and $\boldsymbol{y}$ confusable).

Hereafter in this subsection, fix n to be a positive integer denoting the length of codewords. The following lemma (with proof given in Appendix A) establishes a correspondence between ordered pairs of t-cws words and paths in $\mathcal{G}^{(\mathrm{N})}$ or $\mathcal{G}^{(\mathrm{O})}$.

Lemma 2.10: Let t be a positive integer, $t \le n$. For $j \in \{\mathrm{N},\mathrm{O}\}$, let $\mathcal{W}_t^{(j)}$ denote the set of all t-cws (ordered) pairs $(\boldsymbol{x},\boldsymbol{y}) \in \Sigma^n \times \Sigma^n$ and let $\Pi_t^{(j)}$ be the following set of paths (of length $n{-}1$) in $\mathcal{G}^{(j)}$:

$$\Pi_t^{(j)} = \Big\{\gamma{=}(v_i)_{i\in\langle n\rangle} : v_0 \in V_0,\ |\mathsf{L}(\gamma)|{+}|\mathsf{R}(\gamma)|{+}\mu(\gamma) \le 2t\Big\}.$$

Then there exists a one-to-one[4] mapping from $\mathcal{W}_t^{(j)}$ to $\Pi_t^{(j)}$ that maps t-cws word pairs $((x_i)_{i\in\langle n\rangle},(y_i)_{i\in\langle n\rangle})$ to paths $(\ell_i r_i)_{i\in\langle n\rangle}$ such that $x_i = \phi(\ell_i)$ and $y_i = \phi(r_i)$, for all $i \in \langle n\rangle$.

For $j \in \{\mathrm{N},\mathrm{O}\}$, let $\Gamma^{(j)}$ denote the set of all the cycles in $\mathcal{G}^{(j)}$ of length n that start and terminate in the same state of V_0. Define the functions $f^{(\mathrm{N})} : E^{(\mathrm{N})} \to \langle 3\rangle^2$, $f^{(\mathrm{O})} : E^{(\mathrm{O})} \to \langle 3\rangle^2$ by

$$f^{(\mathrm{N})}(e) = (\nu(e)\ \ \chi(e)) \quad \text{and} \quad f^{(\mathrm{O})}(e) = (\omega(e)\ \ \chi(e))$$

for any edge e, where the functions $\nu : E^{(\mathrm{N})} \to \langle 3\rangle$, $\omega : E^{(\mathrm{O})} \to \langle 3\rangle$, $\chi : E^{(\mathrm{N})} \cup E^{(\mathrm{O})} \to \langle 3\rangle$ are defined next: for an edge $e = (\ell r, \ell' r')$,

$$\nu(e) = \begin{cases} 2 & \text{if } e \text{ satisfies Condition [N4]} \\ 1 & \text{if } e \text{ satisfies either Condition [N2], or [N3]} \\ 0 & \text{otherwise} \end{cases}, \tag{10}$$

$$\omega(e) = \begin{cases} 2 & \text{if } e \text{ satisfies Condition [O2] for } \ell{=}r{\notin}\{\ell',r'\} \\ 1 & \text{if } e \text{ satisfies Condition [O2] for } \ell{=}r{\in}\{\ell',r'\} \\ 1 & \text{if } e \text{ satisfies either Condition [O3], or [O4]} \\ 0 & \text{otherwise} \end{cases}, \tag{11}$$

$$\chi(e) = \begin{cases} 2 & \text{if } \phi(\ell) \ne \phi(\ell') \text{ and } \phi(r) \ne \phi(r') \\ 1 & \text{if } \phi(\ell) = \phi(\ell') \text{ and } \phi(r) \ne \phi(r'), \\ & \text{or } \phi(\ell) \ne \phi(\ell') \text{ and } \phi(r) = \phi(r') \\ 0 & \text{otherwise} \end{cases}. \tag{12}$$

The function $\nu(e)$ counts the smallest number of grains making $\ell\ell'$ and rr' confusable for $j=\mathrm{N}$; the function $\omega(e)$ counts the smallest number of overlapping grains making $\ell\ell'$ and rr' confusable for $j=\mathrm{O}$; and the function $\chi(e)$ counts the number of *transitions* (i.e., symbol changes) in $\ell\ell'$ and in rr', namely, we add 1 if $\phi(\ell) \ne \phi(\ell')$ and another 1 if $\phi(r) \ne \phi(r')$.

Now, set $\tau, p \in (0,1)$, let $\epsilon > 0$, and define

$$U_{\tau,p,\epsilon} = \{(u_1\ \ u_2) : -\epsilon < u_1 < 2\tau{+}\epsilon, |u_2 - 2p| < 2\epsilon\}\ .$$

Also, for $j \in \{\mathrm{N},\mathrm{O}\}$, let

$$\mathit{\Gamma}_{\tau,p,\epsilon}^{(j)} = \mathit{\Gamma}_{\tau,p,\epsilon}^{(j)}(n) = \{\gamma \in \Gamma^{(j)} : \mathsf{E}_{\mathsf{P}_\gamma}\{f^{(j)}\} \in U_{\tau,p,\epsilon}\},$$

where $\mathsf{E}_{\mathsf{P}_\gamma}\{f^{(j)}\}$ is the expected value of $f^{(j)}$ with respect to the empirical probability distribution P_γ, as defined in (5). The set $\mathit{\Gamma}_{\tau,p,\epsilon}^{(j)}$ for $j \in \{\mathrm{N},\mathrm{O}\}$ stands for all the cycles of length n in $\mathcal{G}^{(j)}$ representing pairs of words $(\boldsymbol{x},\boldsymbol{y})$ that can be confused by at most $(2\tau{+}\epsilon)n$ grains (either overlapping or not, depending on j) and whose total number of transitions is within $2(p{\pm}\epsilon)n$. Additionally, for $j \in \{\mathrm{N},\mathrm{O}\}$ and the same τ, p, ϵ, let

$$\mathit{\Pi}_{\tau,p,\epsilon}^{(j)} = \mathit{\Pi}_{\tau,p,\epsilon}^{(j)}(n) = \Big\{\gamma \in \Pi_{\lceil \tau(n-1)\rceil}^{(j)} : \big|\mathsf{E}_{\mathsf{P}_\gamma}\{\chi\}{-}2p\big| \le \epsilon\Big\}.$$

The set $\mathit{\Pi}_{\tau,p,\epsilon}^{(j)}$ includes all paths of length $n{-}1$ in $\mathcal{G}^{(j)}$ representing pairs of words $(\boldsymbol{x},\boldsymbol{y})$ that can be confused by at most $2\lceil \tau(n{-}1)\rceil$ grains (either overlapping or not, depending on the context) and whose total number of transitions is within $(2p{\pm}\epsilon)(n{-}1)$. The following lemma characterizes the relation between the sizes of $\mathit{\Pi}_{\tau,p,\epsilon}^{(j)}$ and $\mathit{\Gamma}_{\tau,p,\epsilon}^{(j)}$, for sufficiently large values of n.

Lemma 2.11: Let $\tau, p \in (0,1)$ and $\epsilon > 0$. Then, for $j \in \{\mathrm{N},\mathrm{O}\}$ and $n \ge 2/\epsilon$,

$$\Big|\mathit{\Pi}_{\tau,p,\epsilon}^{(j)}(n)\Big| \le \Big|\mathit{\Gamma}_{\tau,p,\epsilon}^{(j)}(n)\Big|\ .$$

[3]Though we do not currently have a general proof for this phenomenon, our numerical results show that the relaxation of the notion of confusability does not result in worse bounds while using our technique.

[4]In fact, one can prove that this mapping is also onto $\Pi_t^{(j)}$, but the one-to-one property will suffice for the forthcoming discussion.

Proof: We will prove the lemma for the case without overlaps; when overlaps are allowed, the proof is similar. For a path $\gamma \in \Pi^{(\mathrm{N})}_{\tau,p,\epsilon}$, one has

$$|\mathsf{L}(\gamma)| + |\mathsf{R}(\gamma)| \le 2\lceil \tau(n-1)\rceil \text{ and } \left|\mathsf{E}_{\mathsf{P}_\gamma}\{\chi\} - 2p\right| \le \epsilon.$$

We can draw an edge from the last state of γ to the first one (by the construction of $\mathcal{G}^{(\mathrm{N})}$, there is an edge to a state of V_0 from any state of $\mathcal{G}^{(\mathrm{N})}$) to create a cycle γ' of length n. Since

$$|\mathsf{L}(\gamma)| + |\mathsf{R}(\gamma)| = |\mathsf{L}(\gamma')| + |\mathsf{R}(\gamma')|,$$

we have $\mathsf{E}_{\mathsf{P}_{\gamma'}}\left\{f^{(\mathrm{N})}\right\} \in [0, 2\tau] \times [2p-2\epsilon, 2p+2\epsilon]$ for $n \ge 2/\epsilon$. Hence $|\Pi^{(\mathrm{N})}_{\tau,p,\epsilon}| \le |\Gamma^{(\mathrm{N})}_{\tau,p,\epsilon}|$. □

3) Lower bounds on the rates: In the subsequent lemma, we bound the growth rate of $|\Pi^{(j)}_{\tau,p,\epsilon}(n)|$ from above for $j \in \{\mathrm{N}, \mathrm{O}\}$.

Lemma 2.12: Let $\tau \in (0,1)$. Then for $j \in \{\mathrm{N}, \mathrm{O}\}$,

$$\limsup_{\epsilon\to 0}\limsup_{n\to\infty} \frac{1}{n}\log_q \left|\Pi^{(j)}_{\tau,p,\epsilon}(n)\right| \le K^{(j)}(\tau,p)\ , \tag{13}$$

where, for $j \in \{\mathrm{N}, \mathrm{O}\}$,

$$K^{(j)}(\tau,p) = \inf_{z\in(0,1],\ h\in(0,\infty)} \Big\{\log_q \lambda(\mathsf{A}^{(j)}_{\mathcal{G}}(z,h)) - 2\tau\log_q z - 2p\log_q h\Big\}. \tag{14}$$

Proof: For $z \in (0,1]$ and $h, m \in (0,\infty)$, let the matrices $\mathsf{A}^{(\mathrm{N})}_{\mathcal{G}}(z,h)$ and $\mathsf{A}^{(\mathrm{O})}_{\mathcal{G}}(z,h)$, with rows and columns indexed by the sets of states $V^{(\mathrm{N})}$ and $V^{(\mathrm{O})}$, respectively, be defined as a special case of (6):

$$\left[\mathsf{A}^{(\mathrm{N})}_{\mathcal{G}}(z,h)\right]_{v,v'\in V} = \begin{cases} z^{\nu(e)}h^{\chi(e)} & \text{if } e=(v,v')\in E^{(\mathrm{N})} \\ 0 & \text{otherwise} \end{cases}$$

and

$$\left[\mathsf{A}^{(\mathrm{O})}_{\mathcal{G}}(z,h)\right]_{v,v'\in V} = \begin{cases} z^{\omega(e)}h^{\chi(e)} & \text{if } e=(v,v')\in E^{(\mathrm{O})} \\ 0 & \text{otherwise} \end{cases},$$

where $\nu(e)$, $\omega(e)$, $\chi(e)$ are as defined in (10)–(12). Applying Lemma 2.2 with $G = \mathcal{G}^{(j)}$, $U = U_{\tau,p,\epsilon}$, and $f = f^{(j)}$, for $j \in \{\mathrm{N}, \mathrm{O}\}$, and combining the result with Lemma 2.11, we conclude that

$$\limsup_{n\to\infty}\frac{1}{n}\log_q|\Pi^{(j)}_{\tau,p,\epsilon}(n)| \le \sup_{P\in\mathcal{M}(f^{(j)};U_{\tau,p,\epsilon})} \mathsf{H}_q(P)\ .$$

By the continuity of the functions $P \mapsto \mathsf{E}_P(f^{(j)})$, for $j \in \{\mathrm{N}, \mathrm{O}\}$, and $P \mapsto \mathsf{H}_q(P)$,

$$\limsup_{\epsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log_q|\Pi^{(j)}_{\tau,p,\epsilon}(n)| \le \sup_{P\in\mathcal{M}(f^{(j)};U_{\tau,p})} \mathsf{H}_q(P)\ ,$$

where $U_{\tau,p} = \{(u\ \ 2p) : u \in [0,2\tau]\}$. Applying Lemma 2.3 first with $(G, f, f', U, \boldsymbol{p}) = (\mathcal{G}^{(\mathrm{N})}, \nu, \chi, [0,2\tau], 2p)$ and then with $(G, f, f', U, \boldsymbol{p}) = (\mathcal{G}^{(\mathrm{O})}, \omega, \chi, [0,2\tau], 2p)$ yields the upper bounds on the growth rate of $|\Pi^{(j)}_{\tau,p,\epsilon}(n)|$ for $j \in \{\mathrm{N}, \mathrm{O}\}$, as shown in (13) and (14). □

Now, we can find lower bounds on $R^{(j)}_q(\tau)$ for $j \in \{\mathrm{N}, \mathrm{O}\}$.

Proposition 2.13: Let $\tau \in (0,1)$. Then for $j \in \{\mathrm{N}, \mathrm{O}\}$,

$$R^{(j)}_q(\tau) \ge \varrho^{(j)}_q(\tau) \triangleq \sup_{p\in[0,1]} \Big\{2H_q(p) - K^{(j)}(\tau,p)\Big\},$$

where $H_q(p)$ is the q-ary entropy function defined in (1), and $K^{(j)}(\tau,p)$, for $j \in \{\mathrm{N}, \mathrm{O}\}$, are defined in (14).

Proof: For a word $\boldsymbol{x} = (x_i)_{i\in\langle n\rangle} \in \Sigma^n$ and symbols $a, a' \in \Sigma$, let

$$\kappa(\boldsymbol{x}; a, a') = \{i \in \langle n{-}1\rangle : (x_i, x_{i+1}) = (a, a')\}\,.$$

For $\epsilon > 0$, let $\mathcal{X}_{p,\epsilon}(n)$ be the set of all the words $\boldsymbol{x}$ in Σ^n such that for any $a, a' \in \Sigma$,

$$\left|\frac{\kappa(\boldsymbol{x};a,a')}{n-1} - \frac{p}{q(q-1)}\right| \le \frac{\epsilon}{2q(q-1)} \quad \text{for } a \ne a'$$

and

$$\left|\frac{\kappa(\boldsymbol{x};a,a')}{n-1} - \frac{1-p}{q}\right| \le \frac{\epsilon}{2q} \quad \text{for } a = a'.$$

It is well-known [2, Sec. 12.1] that

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\log_q|\mathcal{X}_{p,\epsilon}(n)| = H_q(p). \tag{15}$$

It follows from Lemma 2.10 that for $j \in \{\mathrm{N}, \mathrm{O}\}$,

$$\limsup_{\epsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log_q\left|\mathsf{W}^{(j)}_{\lceil\tau(n-1)\rceil}(\mathcal{X}_{p,\epsilon}(n))\right| \le \limsup_{\epsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log_q\left|\Pi^{(j)}_{\tau,p,\epsilon}(n)\right|, \tag{16}$$

where $\mathsf{W}^{(j)}_{\lceil\tau n\rceil}(\cdot)$ is defined in (8) (note that the one-to-one mapping $\mathcal{W}^{(j)}_{\lceil\tau n\rceil} \to \Pi^{(j)}_{\lceil\tau n\rceil}$ in Lemma 2.10 preserves the number of transitions for each preimage $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{W}^{(j)}_{\lceil\tau n\rceil}$). Therefore, for $j \in \{\mathrm{N}, \mathrm{O}\}$ and every $p \in (0,1)$,

$$\begin{aligned} R^{(j)}_q(\tau) &\overset{(9)}{\ge} \limsup_{\epsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\Big(2\log_q|\mathcal{X}_{p,\epsilon}(n)| - \log_q\left|\mathsf{W}^{(j)}_{\lceil\tau(n-1)\rceil}(\mathcal{X}_{p,\epsilon}(n))\right|\Big) \\ &\overset{(16)}{\ge} \lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{2}{n}\log_q|\mathcal{X}_{p,\epsilon}(n)| - \limsup_{\epsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log_q\left|\Pi^{(j)}_{\tau,p,\epsilon}(n)\right| \\ &\overset{(13),(15)}{\ge} 2H_q(p) - K^{(j)}(\tau,p). \end{aligned}$$

□

Now, we are in the position to prove the main theorem of this section.

Proof of Theorem 2.1: To simplify the computations of $\varrho^{(j)}_q(\tau)$ for $j \in \{\mathrm{N}, \mathrm{O}\}$, we merge states in $\mathcal{G}^{(j)}$ to reduce the order of the matrix $\mathsf{A}^{(j)}_{\mathcal{G}}$ while preserving its spectral radius for $j \in \{\mathrm{N}, \mathrm{O}\}$, as described in [16, Sec. 4.6]. This is similar to the standard procedure for reducing the number of states in a presentation of a constrained system using the Moore algorithm[5] [16, Sec. 2.6]. The states of V_0 can be merged into

[5]While the reduced graph may have parallel edges, we apply state merging just so that we can compute the spectral radius of a smaller matrix.

superstate 0, the states of V_1 in $\mathcal{G}^{(\mathrm{N})}$ and of V_3 in $\mathcal{G}^{(\mathrm{O})}$ — into superstate 1, whereas the states of V_2 — into superstate 2. The merging ends up with the reduced matrices $\mathcal{A}_{\mathcal{G}}^{(\mathrm{N})}$ and $\mathcal{A}_{\mathcal{G}}^{(\mathrm{O})}$, which are equal to the matrices $\mathcal{A}^{(\mathrm{N})}$ and $\mathcal{A}^{(\mathrm{O})}$ from (2) and (3), respectively, whose spectral radii equal those of $\mathsf{A}_{\mathcal{G}}^{(\mathrm{N})}$ and $\mathsf{A}_{\mathcal{G}}^{(\mathrm{O})}$, respectively.

The lower bound $\varrho_q^{(j)}(\tau)$ of Proposition 2.13 for $j \in \{\mathrm{N}, \mathrm{O}\}$ equals

$$\sup_{p\in[0,1]} \left\{2H_q(p) - K^{(j)}(\tau,p)\right\} = \sup_{\substack{p\in[0,1], z\in(0,1] \\ h\in(0,\infty)}} \Big\{2H_q(p) - \log_q \lambda(\mathcal{A}_{\mathcal{G}}^{(j)}(z,h)) + 2\tau\log_q z + 2p\log_q h\Big\},$$

which concludes the proof of the theorem. □

The supremum in the right-hand side of (4) is attained when

$$\begin{aligned} h &= \frac{p}{1-p} \\ \frac{\partial}{\partial z}(\lambda(\mathcal{A}_{\mathcal{G}}^{(j)}(z,h))) &= \frac{2\tau}{z}\lambda(\mathcal{A}_{\mathcal{G}}^{(j)}(z,h)) \\ \frac{\partial}{\partial h}(\lambda(\mathcal{A}_{\mathcal{G}}^{(j)}(z,h))) &= \frac{2p}{h}\lambda(\mathcal{A}_{\mathcal{G}}^{(j)}(z,h)) . \end{aligned}$$

Remark 2.14: For $q=2$, the matrices $\mathsf{A}_{\mathcal{G}}^{(\mathrm{N})}$ and $\mathsf{A}_{\mathcal{G}}^{(\mathrm{O})}$ can be further reduced to

$$\mathcal{A}_{\mathcal{G}}^{(\mathrm{N})} = \mathcal{A}_{\mathcal{G}}^{(\mathrm{O})} = \begin{array}{c|cc} & {\scriptstyle 0} & {\scriptstyle 1} \\ \hline {\scriptstyle 0} & 1+h^2 & 2hz \\ {\scriptstyle 1} & 2h & h^2 z \end{array} ,$$

whose spectral radius equals

$$\frac{1}{2}\left(1+h^2+h^2z+\alpha(z,h)\right),$$

where

$$\alpha(z,h) = \sqrt{(1+h^2+h^2z)^2 - 4h^2z(h^2-3)} ,$$

so that the lower bound $\varrho_2^{(\mathrm{N})}(\tau) = \varrho_2^{(\mathrm{O})}(\tau)$ of Theorem 2.1 is attained when

$$\begin{aligned} h &= \frac{p}{1-p} \\ \tau &= \frac{h^2 z}{2\alpha(z,h)} \, \frac{h^2z-h^2+7+\alpha(z,h)}{h^2z+h^2+1+\alpha(z,h)} \\ p &= \frac{h^2}{\alpha(z,h)}\left(1 + z\frac{h^2z-3h^2+7+\alpha(z,h)}{h^2z+h^2+1+\alpha(z,h)}\right) . \end{aligned}$$

It turns out that for $q=2$, $\varrho_2^{(\mathrm{N})}(\tau) = \varrho_2^{(\mathrm{O})}(\tau)$ for any $\tau \in [0,1]$. This phenomenon is due to the fact that when $q=2$, for any path $\gamma' \in \Pi_t^{(\mathrm{O})}$ that corresponds to a pair of t-cws words $(\boldsymbol{x},\boldsymbol{y})$, there exists a path $\gamma \in \Pi_t^{(\mathrm{N})}$ for the same pair of words: the path γ is obtained by moving overlapping grains from $\mathsf{M}(\gamma')$ to $\mathsf{R}(\gamma')$ and from $\mathsf{M}'(\gamma')$ to $\mathsf{L}(\gamma')$ until $\mathsf{M}(\gamma')$ and $\mathsf{M}'(\gamma')$ are empty. □

Remark 2.15: One can see that the set $\mathcal{X}$ from which the codewords are taken for the code attaining the bound of Theorem 2.1 is completely characterized by the prescribed frequencies with which consecutive pairs of symbols appear in the words of $\mathcal{X}$; these frequencies, in turn, are characterized only by the parameter p. We can further specify $\mathcal{X}$ by characterizing it by the prescribed frequencies with which consecutive k-tuples of symbols appear in the words of $\mathcal{X}$, for some positive integer $k \geq 3$. This may result in better lower bounds on the rates of grain-correcting codes at the expense of introducing more variables to the optimization process (due to the increase of the order of the adjacency matrices $\mathcal{A}_{\mathcal{G}}^{(\mathrm{N})}$ and $\mathcal{A}_{\mathcal{G}}^{(\mathrm{O})}$). However, at least for $q=2$ and $k=3$, while we obtain some gain in the lower bounds compared to the results of Theorem 2.1, such a gain is rather marginal in the range of values of τ where those lower bounds are above 0.5. □

C. Comparison with existing results

Theorem 2.1 strictly improves on the traditional (Hamming-distance) Gilbert–Varshamov bound,

$$\varrho_2^{(\mathrm{GV})}(\tau) = 1-H_2(2\tau),$$

on the entire interval $(0, 0.25]$; however, on the interval $[0.0566, 0.25]$ it falls short of the simple lower bound of 0.5 which is realized by an ∞-grain-correcting code (see Construction 4.2 in Section IV). The difference between $\varrho_2^{(\mathrm{N})}(\tau) = \varrho_2^{(\mathrm{O})}(\tau)$ and $\varrho_2^{(\mathrm{GV})}(\tau)$ on the interval $(0, 0.0566]$ does not exceed 0.012 (see Figure 2 in Section II-D and Figure 4 in Section III).

Figure 1 depicts the functions $\tau \mapsto \varrho_q^{(\mathrm{N})}(\tau)$ and $\tau \mapsto \varrho_q^{(\mathrm{O})}(\tau)$ for $q \in \{16, 1024\}$ along with the corresponding traditional Gilbert–Varshamov bounds $\varrho_q^{(\mathrm{GV})}(\tau) : \tau \mapsto 1 - H_q(2\tau)$. Both $\varrho_q^{(\mathrm{N})}(\tau)$ and $\varrho_q^{(\mathrm{O})}(\tau)$ strictly improve on $\varrho_q^{(\mathrm{GV})}(\tau)$ on the entire interval $(0, 0.5]$ (and $\varrho_q^{(\mathrm{N})}(\tau)$ is strictly above $\varrho_q^{(\mathrm{O})}(\tau)$). Moreover, both $\varrho_q^{(\mathrm{N})}(\tau)$ and $\varrho_q^{(\mathrm{O})}(\tau)$ converge to the line $\tau \mapsto 1-\tau$ on that interval when $q \to \infty$: this convergence readily follows from substituting $z = 1/\sqrt{q}$ and $h=1$ into (14) and noticing that $\lambda(\mathcal{A}_{\mathcal{G}}^{(\mathrm{N})}(z,h))$ and $\lambda(\mathcal{A}_{\mathcal{G}}^{(\mathrm{O})}(z,h))$ are bounded from below by the minimal row sum $(q+o(q))$ and from above by the maximal row sum $(2q+o(q))$ in the adjacency matrix [4, Ch. 13].

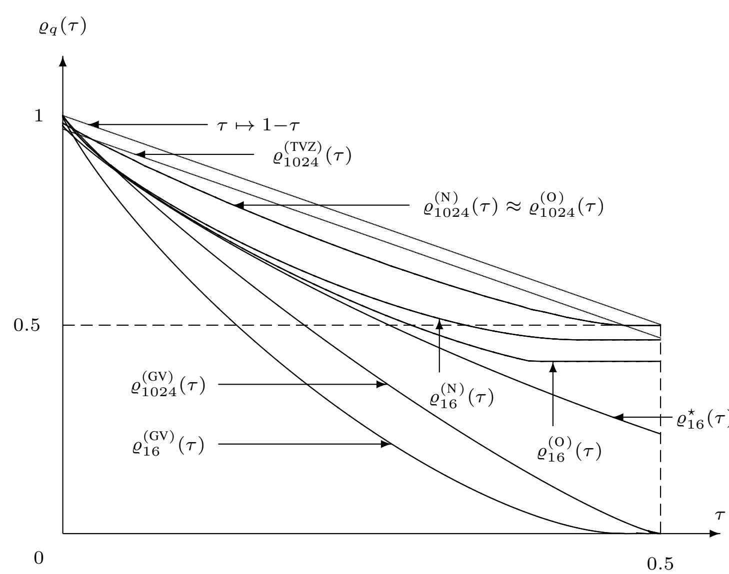

Fig. 1. Functions $\varrho_q^{(\mathrm{N})}(\tau)$, $\varrho_q^{(\mathrm{O})}(\tau)$, and $\varrho_q^{(\mathrm{GV})}(\tau)$ for $q \in \{16, 1024\}$.

For large values of q that are even powers of primes and when overlaps are disallowed, the lower bound $\varrho_q^{(\mathrm{N})}(\tau)$ is worse on nearly the entire interval $(0, 0.5)$ than the following construction based on the family of linear $[n, nR, \lceil \tau n \rceil + 1]$ algebraic geometry codes by Tsfasman *et al.* [22], with rate

$$R \geq 1 - \frac{1}{\sqrt{q}-1} - \tau - o(1) ,$$

where $o(1)$ goes to 0 for $n \to \infty$. By an averaging argument, there exists at least one coset of a code of this family whose intersection, $\mathcal{C}^{(\mathrm{TVZ})}$, with the set

$$\{\boldsymbol{c} = (c_i)_{i \in \langle n \rangle} \in \Sigma^n : c_i \neq c_{i+1} \text{ for any } i \in \langle n-1 \rangle\} \quad (17)$$

is of rate at least

$$\frac{\log_q \left|\mathcal{C}^{(\mathrm{TVZ})}\right|}{n} \geq R - 1 + \log_q (q-1).$$

Since adjacent symbols in each codeword in $\mathcal{C}^{(\mathrm{TVZ})}$ are different, grain errors become erasures, hence $\mathcal{C}^{(\mathrm{TVZ})}$ is a $\lceil \tau n \rceil$-grain-correcting code of rate at least

$$\varrho_q^{(\mathrm{TVZ})}(\tau) = \log_q (q-1) - \frac{1}{\sqrt{q}-1} - \tau - o(1) .$$

Remark 2.16: By the same token, when overlaps are allowed, one can construct a family of $\lceil \tau n \rceil$-grain-correcting codes of length n and rate at least

$$\frac{1}{2} \log_q \left\lceil \frac{q^2-1}{4} \right\rceil - \frac{1}{\sqrt{q}-1} - \tau - o(1) .$$

Specifically, instead of the set in (17) one can take the set $(\Sigma_1 \Sigma_2)^{n/2}$ where $\Sigma_1 = \langle \lfloor q/2 \rfloor \rangle$, $\Sigma_2 = \langle q \rangle \setminus \langle \lfloor q/2 \rfloor \rangle$, and n is even. □

A similar reasoning applied to the family of linear codes guaranteed by the Gilbert–Varshamov bound in the Hamming metric yields $\lceil \tau n \rceil$-grain-correcting codes (when overlaps are disallowed) of rate at least

$$\varrho_q^{\star}(\tau) = \log_q (q-1) - H_q(\tau) .$$

The functions $\tau \mapsto \varrho_{1024}^{(\mathrm{TVZ})}(\tau)$ and $\tau \mapsto \varrho_{16}^{\star}(\tau)$ are shown in Figure 1 alongside the other bounds (we have not drawn $\varrho_{16}^{(\mathrm{TVZ})}(\tau)$ as it is always worse than $\varrho_{16}^{\star}(\tau)$). It can be observed that $\varrho_{16}^{(\mathrm{N})}(\tau)$ and $\varrho_{16}^{(\mathrm{O})}(\tau)$ are strictly above $\varrho_{16}^{\star}(\tau)$, whereas $\varrho_{1024}^{(\mathrm{N})}(\tau)$ is above $\varrho_{1024}^{(\mathrm{TVZ})}(\tau)$ only in the interval $[0, 0.06] \cup [0.44, 0.5]$.

D. Generalization to arbitrary grain length when $q = 2$

In this subsection, we consider the generalization of the grain error model for $\Sigma = \langle 2 \rangle$, where we allow the grains to be of any length up to a prescribed integer g. In this case, since the grains are allowed to be of different lengths, the size of the confusing grain patterns alone does not make for a good definition of confusability. Instead, we suggest to call two words t-confusable if they can be confused by grain patterns such that the sum of the grain lengths in each grain pattern does not exceed $2t$ (the factor of 2 makes this definition coincide with our earlier definition of t-confusability for grains of length 2). Additionally, a generalization of the definition of a grain pattern is called for. As in the previous subsection, we will eventually switch to the wide-sense confusability, and due to an argument similar to that of Remark 2.14, the lower bounds of the nonoverlapping and the overlapping scenario will coincide in the current setting as well, hence throughout this subsection (unless explicitly stated otherwise) we will refer to the nonoverlapping scenario.

Refine the previous definition (see Section I) of a *grain pattern* as a set $\mathcal{S}(g) \subset \langle n-1 \rangle \times (\langle n \rangle \setminus \{0\})$ such that any pair $(b, e) \in \mathcal{S}(g)$, denoting the beginning and the end[6] of a grain, satisfies $b < e < b+g$. In a *nonoverlapping* grain pattern for any two pairs $(b_1, e_1), (b_2, e_2) \in \mathcal{S}(g)$ one has either $e_1 < b_2$ or $e_2 < b_1$. We prohibit grains from being nested, *viz.*, we disallow[7] the existence of grains (b_1, e_1), (b_2, e_2) such that $b_1 < b_2 < e_2 < e_1$. A grain pattern $\mathcal{S}(g)$ inflicts errors to a codeword $\boldsymbol{c} = (c_i)_{i \in \langle n \rangle}$ over an alphabet Σ of size q by means of the smearing operator $\sigma = \sigma_{\mathcal{S}(g)}$ that yields an output word $\boldsymbol{y} = (y_i)_{i \in \langle n \rangle} = \sigma(\boldsymbol{c})$ over Σ in the following way. For any index $i \in \langle n \rangle$,

$$y_i = \begin{cases} c_i & \text{if no pair } (b, e) \in \mathcal{S}(g) \text{ satisfies } b < i \leq e \\ c_b & \exists e : (b, e) \in \mathcal{S}(g) \text{ and } b < i \leq e \end{cases} .$$

Finally, for a positive integer t, two words $\boldsymbol{x}, \boldsymbol{y} \in \Sigma^n$ will be called *t-confusable* if there exist grain patterns $\mathcal{S} = \mathcal{S}(g)$ and $\mathcal{S}' = \mathcal{S}'(g)$ such that

$$\sum_{(b,e) \in \mathcal{S}} (e-b+1) \leq 2t, \qquad \sum_{(b,e) \in \mathcal{S}'} (e-b+1) \leq 2t,$$

and $\sigma_{\mathcal{S}}(\boldsymbol{x}) = \sigma_{\mathcal{S}'}(\boldsymbol{y})$; and, as before, a code of length n over Σ is called *t-grain-correcting* if no two distinct codewords in the code are t-confusable.

Given an alphabet A, let $\mathrm{v}_s(A) = \{a_{(s)} \mid a \in A\}$ for any $s \in \langle g \rangle$. The set of states $V^{(\mathrm{N})}$ of the graph $\mathcal{G}^{(\mathrm{N})}$ will now contain pairs from the alphabet $(\bigcup_{s \in \langle g \rangle} \mathrm{v}_s(\Sigma))^2$: the subscript (s) in an alphabet symbol $a_{(s)}$ will denote the distance from the beginning of the grain that overran symbol a at a given position. For brevity, we will write a instead of $a_{(0)}$ and $\underline{a}$ instead of $a_{(1)}$; thus, $\Sigma = \mathrm{v}_0(\Sigma)$ and $\underline{\Sigma} = \mathrm{v}_1(\Sigma)$ (compare with the counterparts in Section II-B). Specifically, we define $V^{(\mathrm{N})} = \bigcup_{s \in \langle g \rangle} V_s^{(\mathrm{N})}$ where $V_0^{(\mathrm{N})} = \{00, 11\}$ and for $s \in \langle g \rangle \setminus \{0\}$,

$$V_s^{(\mathrm{N})} = \{01_{(s)}\} \cup \{0_{(s)}1\} .$$

The new definition of the operator $\phi(\cdot)$ is $\phi(a_{(s)}) = a$ for every alphabet symbol a and every $s \in \langle g \rangle$; that is, the operator $\phi(\cdot)$ strips off the subscript from a symbol returning the symbol back to its original alphabet. There is an edge in $E^{(\mathrm{N})}$ from $v = \ell r$ to $v' = \ell' r'$ if

[N1'] $v' \in V_0^{(\mathrm{N})}$; or

[6] Under this new definition, a grain of length 2 is represented by a pair $(e-1, e)$ in contrast with our earlier notation for grains of length 2 (as defined in Section I), where a grain was represented only by the index of its ending position e.

[7] In bit-patterned media recording, nested grain errors might occur. The restriction on grain nesting proves helpful in the construction of the respective graph $\mathcal{G}^{(\mathrm{N})}$, because otherwise we would have to encapsulate the hierarchy of the nested grains in the states of the graph which would quickly render the computation impractical, as then the number of states would grow exponentially with g.

[N2'] for some $s \in \langle g{-}1\rangle$, either $\ell = \ell' \in \Sigma$ and $rr' \in \mathtt{v}_s(\Sigma)\mathtt{v}_{s+1}(\Sigma)$, or $\ell\ell' \in \mathtt{v}_s(\Sigma)\mathtt{v}_{s+1}(\Sigma)$ and $r = r' \in \Sigma$; or

[N3'] for some $s \in \langle g\rangle \setminus \{0\}$, either $\ell = r' \in \Sigma$, $\ell' r \in \mathtt{v}_1(\Sigma)\mathtt{v}_s(\Sigma)$ and $\phi(\ell') = \phi(r)$, or $\ell' = r \in \Sigma$, $\ell r' \in \mathtt{v}_s(\Sigma)\mathtt{v}_1(\Sigma)$ and $\phi(\ell) = \phi(r')$.

We redefine the component $\nu(\cdot)$ of the vector function $f^{(\mathrm{N})}$ as follows, while leaving the component $\chi(\cdot)$ intact (compare with (10)):

$$\nu(e) = \begin{cases} 2 & \text{if } e \text{ satisfies Condition [N2'] where} \\ & \text{either } \ell' \in \mathtt{v}_1(\Sigma) \text{ or } r' \in \mathtt{v}_1(\Sigma) \\ 2 & \text{if } e \text{ satisfies Condition [N3']} \\ 1 & \text{if } e \text{ satisfies Condition [N2'] where,} \\ & \text{for } s \geq 2, \text{ either } \ell' \in \mathtt{v}_s(\Sigma) \text{ or } r' \in \mathtt{v}_s(\Sigma) \\ 0 & \text{otherwise} \end{cases} .$$

We switch again to the wide-sense notion of confusability, namely, given a positive integer t, we will call two words $\boldsymbol{x}, \boldsymbol{y}$ *t-confusable in the wide sense* if there exist grain patterns $\mathcal{S} = \mathcal{S}(g)$ and $\mathcal{S}' = \mathcal{S}'(g)$ such that

$$\sum_{(b,e)\in\mathcal{S}} (e{-}b{+}1) + \sum_{(b,e)\in\mathcal{S}'} (e{-}b{+}1) \leq 4t$$

and $\sigma_{\mathcal{S}}(\boldsymbol{x}) = \sigma_{\mathcal{S}'}(\boldsymbol{y})$. It can be verified that the counterparts of Lemmas 2.10, 2.11, Proposition 2.13, and Theorem 2.1 hold also with this generalization. The reduced adjacency matrix $\mathcal{A}_{\mathcal{G}}^{(\mathrm{N})}$ is of order $g \times g$ and is obtained after merging the states of $V_s^{(\mathrm{N})}$ into one superstate s for every $s \in \langle g\rangle$:

$$\left[\mathcal{A}_{\mathcal{G}}^{(\mathrm{N})}(z,h)\right]_{s,s'} = \begin{cases} 1{+}h^2 & \text{if } (s,s') = (0,0) \\ 2h & \text{if } s \neq 0 \text{ and } s' = 0 \\ 2hz^2 & \text{if } (s,s') = (0,1) \\ h^2 z^2 & \text{if } s \neq 0 \text{ and } s' = 1 \\ z & \text{if } s \notin \{0, g{-}1\} \text{ and } s' = s{+}1 \\ 0 & \text{otherwise} \end{cases} .$$

Example 2.17: For $g = 4$,

$$\mathcal{A}_{\mathcal{G}}^{(\mathrm{N})}(z,h) = \begin{array}{c|cccc} & 0 & 1 & 2 & 3 \\ \hline 0 & 1{+}h^2 & 2hz^2 & 0 & 0 \\ 1 & 2h & h^2z^2 & z & 0 \\ 2 & 2h & h^2z^2 & 0 & z \\ 3 & 2h & h^2z^2 & 0 & 0 \end{array} .$$

□

Similarly to Theorem 2.1, one can obtain lower bounds $\varrho_2^{(j)}(\tau, g)$ on the rate of $\lceil \tau n\rceil$-grain-correcting codes of length n when $j \in \{\mathrm{N}, \mathrm{O}\}$ and grains of length at most g are allowed. Next, we present an analysis for the case when $g \to \infty$. By an argument akin to the one made in Remark 2.14, for any pair of overlapping grain patterns of total length at most $4t$ that confuse two given t-cws binary words $(\boldsymbol{x}, \boldsymbol{y})$, there exists a pair of nonoverlapping grain patterns of total length at most $4t$ that confuse $\boldsymbol{x}$ and $\boldsymbol{y}$; therefore, for any integer $g \geq 2$, we have $\varrho_2^{(\mathrm{N})}(\tau, g) = \varrho_2^{(\mathrm{O})}(\tau, g)$. For the simplicity of computations, we keep assuming that $j = \mathrm{N}$.

The characteristic polynomial $Q^{(\mathrm{N})}(\zeta) = Q^{(\mathrm{N})}(\zeta; z, h)$ of $\mathcal{A}_{\mathcal{G}}^{(\mathrm{N})}(z,h)$ can be verified to be

$$Q^{(\mathrm{N})}(\zeta) = \zeta^{g-2}\Big[\zeta^2 - (1{+}h^2{+}h^2z^2)\zeta + h^2z^2(h^2{-}3) - (1 - (z/\zeta)^{g-2})\frac{z^3h^2(\zeta + 3 - h^2)}{\zeta - z}\Big].$$

For[8] $g \to \infty$, $Q^{(\mathrm{N})}(\zeta)$ converges to

$$\frac{\zeta^{g-1}}{\zeta{-}z}\left(\zeta^2{-}(1{+}z{+}h^2{+}h^2z^2)\zeta + h^2z^2(h^2{-}3) + z(1+h^2)\right),$$

the largest root of which, $\lambda(\mathcal{A}_{\mathcal{G}}^{(\mathrm{N})})$, equals

$$\lambda(\mathcal{A}_{\mathcal{G}}^{(\mathrm{N})}) = \frac{1}{2}\Big(1 + z + h^2 + h^2z^2 + \sqrt{(1 + z + h^2 + h^2z^2)^2 - 4(h^2z^2(h^2 - 3) + z(1 + h^2))}\Big) . \tag{18}$$

Using (18) to find the expression for $K^{(\mathrm{N})}(\tau, p)$, defined in (14), and plugging this expression into Theorem 2.1 yield a lower bound

$$\xi(\tau) = \lim_{g\to\infty} \varrho_2^{(\mathrm{N})}(\tau, g)$$

on the rate of $\lceil\tau n\rceil$-grain-correcting codes of length n when $g \to \infty$ (and therefore $n \to \infty$) and overlaps are disallowed. Figure 2 depicts $\xi(\tau)$ along with $\varrho_2^{(\mathrm{N})}(\tau) = \varrho_2^{(\mathrm{N})}(\tau, 2)$ for $\tau \in [0, 0.0566]$ and the lower bound 0.5 attained by Construction 4.2 (to be presented later on) for $\tau \in [0.0566, 0.25]$. For comparison, we present the Gilbert–Varshamov bounds $\varrho_2^{(\mathrm{GV})}(\tau) = 1 - H_2(2\tau)$ (corresponding to ordinary binary $\lceil\tau n\rceil$-error-correcting codes) and $\varrho_2^{(\mathrm{GV})}(2\tau) = 1 - H_2(4\tau)$ (corresponding to binary $\lceil 2\tau n\rceil$-error-correcting codes); readily, $\xi(\tau)$ and $\varrho_2^{(\mathrm{N})}(\tau)$ improve on $\varrho_2^{(\mathrm{GV})}(2\tau)$ and $\varrho_2^{(\mathrm{GV})}(\tau)$, respectively, on the entire interval[9] $[0, 0.25]$.

Remark 2.18: We mention that it is also possible to obtain Gilbert–Varshamov-like bounds in the general case when the values of q and g are arbitrary positive integers greater than 1. Both reduced adjacency matrices $\mathcal{A}_{\mathcal{G}}^{(\mathrm{N})}$ and $\mathcal{A}_{\mathcal{G}}^{(\mathrm{O})}$ are then of order $\binom{g+1}{2} \times \binom{g+1}{2}$ when $q \geq 3$. □

III. Upper bounds

In this section, we restrict the discussion to $q = g = 2$ and compute upper bounds on the size $M_2^{(j)}(n, t)$ of t-grain-correcting codes of length n for $j \in \{\mathrm{N}, \mathrm{O}\}$. Given any word $\boldsymbol{x} = (x_i)_{i\in\langle n\rangle}$ in Σ^n and a positive integer t, let $\mathcal{B}_t^{(\mathrm{N})}(\boldsymbol{x})$ and $\mathcal{B}_t^{(\mathrm{O})}(\boldsymbol{x})$ be the sets of all words $\boldsymbol{y} \in \Sigma^n$ for which there exists a grain pattern $\mathcal{S}$ of size $|\mathcal{S}| \leq t$ such that $\sigma_{\mathcal{S}}(\boldsymbol{x}) = \boldsymbol{y}$ when overlaps are disallowed and allowed, respectively. Since a grain ending at location e alters $\boldsymbol{x}$ only when $x_{e-1} \neq x_e$, we can assume without loss of generality (w.l.o.g.) that the grain pattern $\mathcal{S}$ consists only of grains at such locations. In

[8] We are abusing notation here, implying by $g \to \infty$ that $n \to \infty$, while keeping the ratio between the sum of grain lengths and the word length to be at most 2τ.

[9] When the grain lengths are allowed to extend up to g, then, by our definition of a t-grain-correcting code, such a code should be able to correct as many as $2t(1{-}1/g) = 2\tau n(1{-}1/g)$ actual errors. Hence the comparison with $\varrho_2^{(\mathrm{GV})}(2\tau)$ when $g \to \infty$.

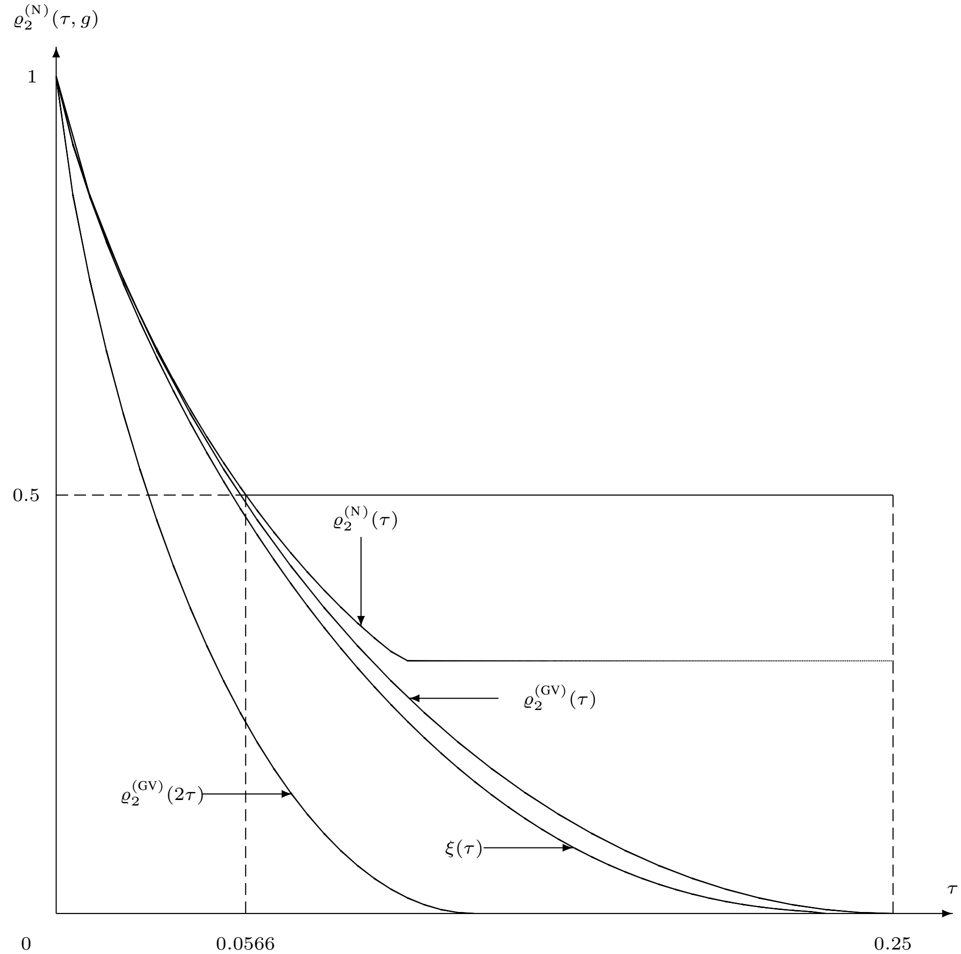

Fig. 2. Functions $\varrho_2^{(\mathrm{N})}(\tau)$ and $\xi(\tau)$.

particular, $|\mathcal{S}| \le \mathsf{r}(\boldsymbol{x}){-}1$ where $\mathsf{r}(\boldsymbol{x})$ is the number of runs[10] in $\boldsymbol{x}$.

Suppose now that for any $\boldsymbol{x} \in \Sigma^n$ we have a lower bound $\psi^{(j)}(r) = \psi_{n,t}^{(j)}(r)$ on $|\mathcal{B}_t^{(j)}(\boldsymbol{x})|$ that depends (other than on n and t) only on the number of runs $r = \mathsf{r}(\boldsymbol{x})$ in $\boldsymbol{x}$ and that is nondecreasing as a function of r for $j \in \{\mathrm{N}, \mathrm{O}\}$. Let $\mathsf{N}(r)$ be the number of words $\boldsymbol{x} \in \Sigma^n$ with r runs (i.e., $\mathsf{N}(r) = 2\binom{n-1}{r-1}$), and, for $j \in \{\mathrm{N}, \mathrm{O}\}$, let $\mathcal{U}^{(j)}$ be a largest set of words in Σ^n such that

$$\sum_{\boldsymbol{x} \in \mathcal{U}^{(j)}} \psi^{(j)}(\mathsf{r}(\boldsymbol{x})) \le 2^n$$

(if we list the words $\boldsymbol{x}$ in Σ^n according to increasing values of $\psi^{(j)}(\mathsf{r}(\boldsymbol{x}))$, then $\mathcal{U}^{(j)}$ can be assumed to consist of the first $\left|\mathcal{U}^{(j)}\right|$ words in this list). Since $\psi^{(j)}(r)$ is nondecreasing in r, we can group words with the same number of runs till the largest integer w such that

$$\sum_{r=1}^{w} \mathsf{N}(r)\psi^{(j)}(r) \le 2^n, \tag{19}$$

and thus,

$$\left|\mathcal{U}^{(j)}\right| = \sum_{r=1}^{w} \mathsf{N}(r) + \left\lfloor \frac{2^n - \sum_{r=1}^{w} \mathsf{N}(r)\psi^{(j)}(r)}{\psi^{(j)}(w{+}1)} \right\rfloor. \tag{20}$$

[10]By a run we mean a consecutive subword $x_i\ x_{i+1}\ \ldots\ x_{i'}$ of $\boldsymbol{x}$ such that $x_i = x_{i+1} = \ldots = x_{i'}$, where $x_{i-1} \ne x_i$ (if $i > 0$) and $x_{i'} \ne x_{i'+1}$ (if $i' < n{-}1$).

Now, let $\mathcal{C}^{(\mathrm{N})}$ and $\mathcal{C}^{(\mathrm{O})}$ be binary t-grain-correcting codes of length n (when overlaps are disallowed and allowed, respectively). By a sphere-packing argument, we have, for $j \in \{\mathrm{N}, \mathrm{O}\}$,

$$\sum_{\boldsymbol{c} \in \mathcal{C}^{(j)}} \psi^{(j)}(\mathsf{r}(\boldsymbol{c})) \le \sum_{\boldsymbol{c} \in \mathcal{C}^{(j)}} |\mathcal{B}_t^{(j)}(\boldsymbol{c})| \le 2^n .$$

It follows from the definition of $\mathcal{U}^{(j)}$ that $\left|\mathcal{C}^{(j)}\right| \le \left|\mathcal{U}^{(j)}\right|$ for $j \in \{\mathrm{N}, \mathrm{O}\}$. Hence, from (20) we get, for $j \in \{\mathrm{N}, \mathrm{O}\}$,

$$\left|\mathcal{C}^{(j)}\right| \le \sum_{r=1}^{w} \mathsf{N}(r) + \left\lfloor \frac{2^n - \sum_{r=1}^{w} \mathsf{N}(r)\psi^{(j)}(r)}{\psi^{(j)}(w{+}1)} \right\rfloor ,$$

where w is the largest integer that satisfies (19).

When overlaps are disallowed, we can bound $|\mathcal{B}_t^{(\mathrm{N})}(\boldsymbol{c})|$ with $\mathsf{r}(\boldsymbol{c}) = r$ from below using

$$\psi_{n,t}^{(\mathrm{N})}(r) = \sum_{s=0}^{\min\left\{t, \left\lfloor \frac{r}{2} \right\rfloor\right\}} \binom{r{-}s}{s} , \tag{21}$$

which is the number of ways of choosing up to t non-consecutive transitions out of the $r{-}1$ available transitions between adjacent runs. When overlaps are allowed, we are able to calculate the size of $\mathcal{B}_t^{(\mathrm{O})}(\boldsymbol{c})$ with $\mathsf{r}(\boldsymbol{c}) = r$ precisely, namely,

$$\psi_{n,t}^{(\mathrm{O})}(r) = \sum_{s=0}^{\min\{t, r-1\}} \binom{r{-}1}{s} . \tag{22}$$

Both (21) and (22) are clearly nondecreasing functions in r.

The next theorem summarizes the above discussion and, in fact, reformulates the sphere-packing bound that Abdel-Ghaffar and Weber [1, Th. 5] first used in a different context (see also [16, Sec. 7.3]).

Theorem 3.1: Let $\mathcal{C}^{(\mathrm{N})}$ and $\mathcal{C}^{(\mathrm{O})}$ be binary t-grain-correcting codes of length n (when overlaps are disallowed and allowed, respectively). Then, for $j \in \{\mathrm{N}, \mathrm{O}\}$, one has $\left|\mathcal{C}^{(j)}\right| \le \Delta^{(j)}(n,t)$, where

$$\Delta^{(j)}(n,t) = 2\sum_{r=1}^{w}\binom{n-1}{r-1} + \left\lfloor \frac{2^n - 2\sum_{r=1}^{w}\binom{n-1}{r-1}\psi_{n,t}^{(j)}(r)}{\psi_{n,t}^{(j)}(w+1)} \right\rfloor \tag{23}$$

and w is the largest integer such that

$$\textstyle\sum_{r=1}^{w}\binom{n-1}{r-1}\psi_{n,t}^{(j)}(r) \le 2^{n-1} . \tag{24}$$

The formulas for $\psi_{n,t}^{(\mathrm{N})}(r)$ and $\psi_{n,t}^{(\mathrm{O})}(r)$ are given in (21) and (22), respectively.

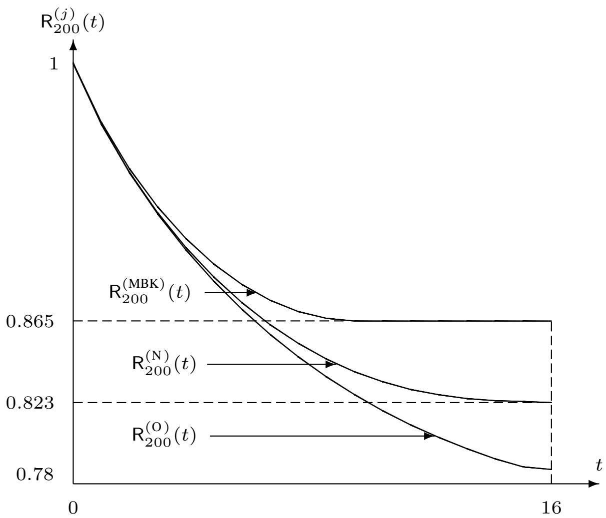

Fig. 3. Functions $\mathsf{R}_{200}^{(\mathrm{N})}(t)$, $\mathsf{R}_{200}^{(\mathrm{O})}(t)$, and $\mathsf{R}_{200}^{(\mathrm{MBK})}(t)$.

For $j \in \{\mathrm{N}, \mathrm{O}\}$, let $\mathsf{R}_n^{(j)}(t) = \log_2(\Delta^{(j)}(n,t))/n$. Figure 3 depicts the functions $t \mapsto \mathsf{R}_n^{(j)}(t)$ for $n = 200$ calculated at $t \in \{1, 2, \ldots, 16\}$.

Mazumdar *et al.* [17, Th. 3] obtained an upper bound on $M_2^{(\mathrm{N})}(n,t)$ using a similar technique by considering a t-grain-correcting code $\mathcal{C}$ of length n (when overlaps are disallowed) and partitioning it into two subcodes

$$\mathcal{C}_1 = \left\{ \boldsymbol{c} \in \mathcal{C} : |\mathsf{r}(\boldsymbol{c}) - n/2| \le \sqrt{nt\log_2 n} \right\} \text{ and } \mathcal{C}_2 = \mathcal{C} \backslash \mathcal{C}_1 .$$

The sizes of $\mathcal{B}_t^{(\mathrm{N})}(\boldsymbol{c})$ for $\boldsymbol{c} \in \mathcal{C}_1$ and $\boldsymbol{c} \in \mathcal{C}_2$ were then bounded from below by $\psi_t^{(\mathrm{N})}(\beta)$ and 1, respectively, where $\beta = \beta(n,t)$ is taken as $n/2 - \left\lfloor\sqrt{nt\log_2 n}\right\rfloor$. The obtained upper bound on $M_2^{(\mathrm{N})}(n,t)$ is

$$\Delta^{(\mathrm{MBK})}(n,t) = \frac{2^n t!}{(\beta - 1 - 3(t-1))^t} + 4\sum_{i=0}^{\beta}\binom{n-1}{i} . \tag{25}$$

For specific values of n and t, this bound can be optimized by taking β to minimize the right-hand side of (25). Let

$$\mathsf{R}_n^{(\mathrm{MBK})}(t) = \frac{1}{n}\log_2\left(\min_{3t-2<\beta<n/2}\Delta^{(\mathrm{MBK})}(n,t)\right).$$

The function $t \mapsto \mathsf{R}_{200}^{(\mathrm{MBK})}(t)$ is depicted in Figure 3 as well. It can be seen that the functions $\mathsf{R}_{200}^{(j)}(t)$ for $j \in \{\mathrm{N}, \mathrm{O}\}$ improve on $\mathsf{R}_{200}^{(\mathrm{MBK})}(t)$ for all $1 \le t \le 16$.

The limit of $\mathsf{R}_n^{(\mathrm{N})}(\lceil \tau n \rceil)$ as $n \to \infty$ is identical to that of $\mathsf{R}_n^{(\mathrm{MBK})}(\lceil \tau n \rceil)$ and is based on a lower bound on the asymptotic growth rate $\Psi(\tau, z)$ of $\psi_{n,\lceil \tau n \rceil}^{(\mathrm{N})}(\lceil zn \rceil)$ for fixed τ and z. Indeed, in the asymptotic analysis (for $n \to \infty$ and fixed τ and z), the largest integer $w = \lceil zn \rceil$ for which (24) holds, translates into the smallest positive solution $z = z^\dagger$ of the equation $H_2(z) + \Psi(\tau, z) = 1$. This implies that the asymptotic growth rate of the last summand of the right-hand side of (23),

$$\left\lfloor \frac{2^n - 2\sum_{r=1}^{w}\binom{n-1}{r-1}\psi_{n,t}^{(j)}(r)}{\psi_{n,t}^{(j)}(w+1)} \right\rfloor,$$

is at most $1 - \Psi(\tau, z^\dagger) = H_2(z^\dagger)$, which is exactly the growth rate of the sum $2\sum_{r=1}^{w}\binom{n-1}{r-1}$. In other words, the asymptotic growth rate of $\Delta^{(\mathrm{N})}(n,t)$ is $H_2(z^\dagger)$ where $z^\dagger$ is the smallest positive solution of $H_2(z) + \Psi(\tau, z) = 1$. If, like Mazumdar *et al.*, we bound $\Psi(\tau, z)$ from below by $\frac{1}{2}z \cdot H_2(2\tau/z)$, we will obtain the following upper bound on the rate of binary $\lceil \tau n \rceil$-grain-correcting codes (when overlaps are disallowed) for $\tau \le 0.0706$:

$$R_2^{(\mathrm{N})}(\tau) \le \rho^{(\mathrm{MBK})}(\tau) \triangleq H_2(z^\dagger) ,$$

where $z^\dagger$ is the smallest positive solution of

$$H_2(z) + \tfrac{1}{2}z \cdot H_2(2\tau/z) = 1,$$

and this is exactly how [17, Prop. 4] is formulated. However, if we use $\psi_{n,t}^{(\mathrm{N})}(r)$ from (21) as a lower bound on $|\mathcal{B}_t^{(\mathrm{N})}(\boldsymbol{c})|$, a better lower bound on $\Psi(\tau, z)$ can be obtained resulting in a better upper bound on $R_2^{(\mathrm{N})}(\tau)$ as stated in the following theorem, which essentially follows the proof of Mazumdar *et al.* in [17, Prop. 4].

Theorem 3.2: Let $\tau^{(\mathrm{N})}$ (≈ 0.070958) be the smallest positive solution of $H_2\left(\frac{5+\sqrt{5}}{2}\tau\right) + \frac{3+\sqrt{5}}{2}H_2\left(\frac{3-\sqrt{5}}{2}\right)\tau = 1$ (as an equation in τ). Then for $\tau \in [0, \tau^{(\mathrm{N})}]$,

$$R_2^{(\mathrm{N})}(\tau) \le \rho^{(\mathrm{N})}(\tau) \triangleq H_2(z^*),$$

where z^* is the smallest positive solution of

$$H_2(z) + (z - \tau)H_2\left(\frac{\tau}{z - \tau}\right) = 1$$

(as an equation in z).

Remark 3.3: Let $\tau^{(\mathrm{O})}$ (≈ 0.113546) be the smallest positive solution of $H_2(2\tau) + 2\tau = 1$. Similarly to Theorem 3.2 and using (22), one can bound $R_2^{(\mathrm{O})}(\tau)$ for $\tau \in [0, \tau^{(\mathrm{O})}]$ from above by $\rho^{(\mathrm{O})}(\tau) \triangleq H_2(z^\star)$, where $z^\star$ is the smallest positive solution of $H_2(z) + z \cdot H_2(\tau/z) = 1$. □

Figure 4 depicts the upper bounds $\rho^{(\mathrm{N})}(\tau)$ and $\rho^{(\mathrm{O})}(\tau)$ on the rate of $\lceil \tau n \rceil$-grain-correcting codes for $n \to \infty$ guaranteed by Theorem 3.2 and Remark 3.3, respectively (for $\tau > \tau^{(j)}$, the upper bounds become $\rho^{(j)}(\tau) = \rho^{(j)}(\tau^{(j)})$). For comparison, we present the upper bound $\rho^{(\mathrm{MBK})}(\tau)$ stated in [17, Prop. 4] and the lower bound $\varrho_2^{(\mathrm{N})}(\tau)$ found in Section II-B. One can observe the visible improvement of $\rho^{(\mathrm{N})}(\tau)$ over

$\rho^{(\text{MBK})}(\tau)$ as well as the proximity of $\rho^{(\text{N})}(\tau)$ and $\rho^{(\text{O})}(\tau)$. Recently, Kashyap and Zémor obtained in [9, Sec. 4] another new upper bound on the rate of binary $\lceil \tau n \rceil$-grain-correcting codes (when overlaps are disallowed), using information-theoretic arguments. Their bound, denoted in Figure 4 by $\rho^{(\text{KZ})}(\tau)$, improves on $\rho^{(\text{N})}(\tau)$ for $\tau \geq 0.02614$.

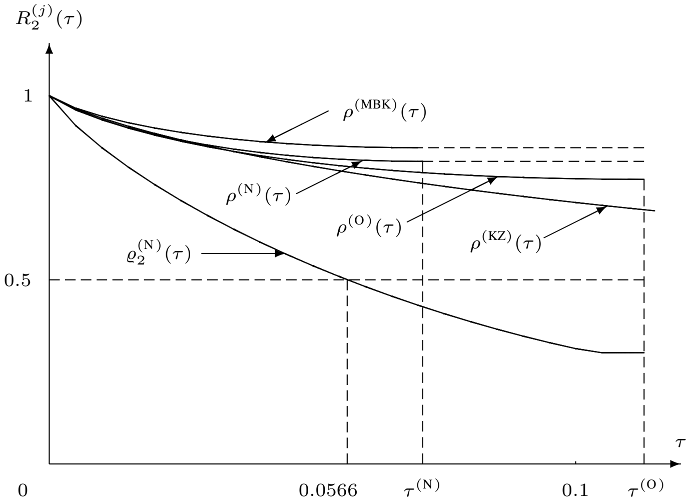

Fig. 4. Upper bounds $\rho^{(\text{N})}(\tau)$ and $\rho^{(\text{O})}(\tau)$ of Theorem 3.2 and Remark 3.3 along with $\rho^{(\text{MBK})}(\tau)$ and the lower bound $\varrho_2^{(\text{N})}(\tau)$.

IV. Constructions of grain-correcting codes

In this section, we present constructions of binary ($q = 2$) t-grain-correcting codes of length n for specific values of n and t, assuming a grain length $g = 2$. We also show the optimality and the uniqueness of some of these codes.

The following lemma will be useful later on in determining whether a pair of words is non-confusable (regardless of whether overlaps are allowed or not); in fact, the lemma applies also to wide-sense confusability, as this term was defined in Section II-B (see the discussion preceding Lemma 2.10).

Lemma 4.1: Let $\boldsymbol{x} = (x_i)_{i \in \langle n \rangle}$ and $\boldsymbol{x}' = (x'_i)_{i \in \langle n \rangle}$ be words in $\langle 2 \rangle^n$ such that $x_0 = x'_0$. Then $\boldsymbol{x}$ and $\boldsymbol{x}'$ are non-confusable if and only if there exists an index $e \in \langle n{-}1 \rangle$ such that

$$x_e = x_{e+1} \neq x'_e = x'_{e+1} \qquad (26)$$

(i.e., either $x_e x_{e+1} = 00$ and $x'_e x'_{e+1} = 11$, or vice versa).

Proof: The "if" part of the lemma is immediate. Turning to the "only if" part, assume that there does not exist an index $e \in \langle n{-}1 \rangle$ for which (26) holds. In what follows, we will show by induction on $i \in \langle n \rangle$ that there necessarily exists a pair of nonoverlapping patterns $\mathcal{S}$ and $\mathcal{S}'$ such that $\sigma_{\mathcal{S}}(\boldsymbol{x})$ and $\sigma_{\mathcal{S}'}(\boldsymbol{x}')$ are identical on their first $i{+}1$ positions. The induction basis is immediate as $x_0 = x'_0$ and we can take $\mathcal{S} = \mathcal{S}' = \emptyset$.

Assume now by induction that for $i \in \langle n{-}1 \rangle$, there exists a pair of nonoverlapping grain patterns $\mathcal{S}_i, \mathcal{S}'_i \subseteq \langle i{+}1 \rangle$ such that $\sigma_{\mathcal{S}_i}(\boldsymbol{x})$ and $\sigma_{\mathcal{S}'_i}(\boldsymbol{x}')$ are identical on their first $i{+}1$ positions, and $\mathcal{S}_i \cap \mathcal{S}'_i = \emptyset$ (i.e., no grain from $\mathcal{S}_i$ ends at the location where a grain from $\mathcal{S}'_i$ ends). We prove the inductive claim for $i{+}1$. If $x_{i+1} = x'_{i+1}$ then $\mathcal{S}_{i+1} = \mathcal{S}_i$ and $\mathcal{S}'_{i+1} = \mathcal{S}'_i$; otherwise, due to (26) not being satisfied, either [C1] $x_i = x'_i$ or [C2] $x_i = x'_{i+1} \neq x_{i+1} = x'_i$.

[C1] Notice that when $x_i = x'_i$, replacing $\mathcal{S}_i$ and $\mathcal{S}'_i$ with $\mathcal{S}_i \setminus \{i\}$ and $\mathcal{S}'_i \setminus \{i\}$, respectively, produce grain patterns that satisfy the induction hypothesis too, hence for $x_i = x'_i$, we may assume that $i \notin \mathcal{S}_i \cup \mathcal{S}'_i$. W.l.o.g., assume that $x_i = x'_i = x_{i+1} \neq x'_{i+1}$; in this case, $\mathcal{S}_{i+1} = \mathcal{S}_i$ and $\mathcal{S}'_{i+1} = \mathcal{S}'_i \cup \{i{+}1\}$ satisfy the inductive claim for $i{+}1$.

[C2] In this case, w.l.o.g., $i \notin \mathcal{S}_i$, and the grain patterns $\mathcal{S}_{i+1} = \mathcal{S}_i \cup \{i{+}1\}$ and $\mathcal{S}'_{i+1} = \mathcal{S}'_i$ satisfy the inductive claim for $i{+}1$.

For $i = n{-}1$, we obtain nonoverlapping grain patterns $\mathcal{S} = \mathcal{S}_{n-1}$ and $\mathcal{S}' = \mathcal{S}'_{n-1}$ that make $\boldsymbol{x}$ and $\boldsymbol{x}'$ confusable. □

Mazumdar *et al.* presented in [17, Sec. 2] the following simple construction for ∞-grain-correcting codes when overlaps are disallowed.

Construction 4.2: For any even positive n, the binary code

$$\mathcal{C}_n = \{\boldsymbol{c} = (c_i)_{i \in \langle n \rangle} : c_{2s} = c_{2s+1} \text{ for any } s \in \langle n/2 \rangle\}$$

is an ∞-grain-correcting code of length n and size $2^{n/2}$. For any odd positive n, the code $\mathcal{C}_n = (0\,\mathcal{C}_{n-1}) \cup (1\,\mathcal{C}_{n-1})$ is a binary ∞-grain-correcting code of length n and size $2^{(n+1)/2}$.

It is easy to see that the code $\mathcal{C}_n$ from Construction 4.2 is ∞-grain-correcting even when overlaps are allowed; therefore,

$$M_2^{(\text{N})}(n,\infty) = M_2^{(\text{N})}(n,\lfloor n/2 \rfloor) \geq M_2^{(\text{O})}(n,\lfloor n/2 \rfloor) \geq M_2^{(\text{O})}(n,\infty) \geq 2^{\lceil n/2 \rceil},$$

where $M_2^{(\text{N})}(n,\infty)$ and $M_2^{(\text{O})}(n,\infty)$ denote the largest size of any binary ∞-grain-correcting code of length n when overlaps are disallowed and allowed, respectively. Conversely, from [17, Prop. 1] it follows that $M_2^{(\text{N})}(n,\lfloor n/2 \rfloor) \leq 2^{\lceil n/2 \rceil}$, thereby implying the following result.

Theorem 4.3: For any positive integer n,

$$M_2^{(\text{N})}(n,\infty) = M_2^{(\text{N})}(n,\lfloor n/2 \rfloor) = M_2^{(\text{O})}(n,\lfloor n/2 \rfloor) = M_2^{(\text{O})}(n,\infty) = 2^{\lceil n/2 \rceil}.$$

It turns out that Construction 4.2 is the only way to construct binary ∞-grain-correcting codes of odd length n and size $2^{(n+1)/2}$.

Theorem 4.4: Let n be an odd positive integer. The binary ∞-grain-correcting code of length n and size $2^{(n+1)/2}$ is unique (whether overlaps are allowed or not).

Proof: It suffices to show uniqueness for the case where overlaps are disallowed, and our proof will be by induction on n. For $n = 1$, there is clearly only one binary ∞-grain-correcting code of size 2, which is $\langle 2 \rangle$.

Let now $\mathcal{C}$ be a binary ∞-grain-correcting code of odd length n and size $2^{(n+1)/2}$, and let $\mathcal{C}_{0*} \subseteq \mathcal{C}$ be the set of all codewords of $\mathcal{C}$ ending in either 00 or 01. Any two distinct codewords of $\mathcal{C}_{0*}$ are non-confusable, thus their prefixes of length $n{-}2$ are non-confusable as well. Therefore, the punctured code

$$\mathcal{C}'_{0*} = \{\boldsymbol{c} : \boldsymbol{c}00 \in \mathcal{C}_{0*} \text{ or } \boldsymbol{c}01 \in \mathcal{C}_{0*}\},$$

of length $n{-}2$, is ∞-grain-correcting. Likewise, the similarly defined punctured code $\mathcal{C}'_{1*}$ is ∞-grain-correcting as well. Hence, by Theorem 4.3 we have that $\max\{|\mathcal{C}'_{0*}|, |\mathcal{C}'_{1*}|\} \leq$

$2^{(n-1)/2}$. On the other hand, $|\mathcal{C}_{0*}|+|\mathcal{C}_{1*}| = |\mathcal{C}'_{0*}|+|\mathcal{C}'_{1*}| = |\mathcal{C}| = 2^{(n+1)/2}$, so $|\mathcal{C}'_{0*}| = |\mathcal{C}'_{1*}| = 2^{(n-1)/2}$. Applying the induction hypothesis to $\mathcal{C}'_{0*}$ and $\mathcal{C}'_{1*}$ yields that $\mathcal{C}'_{0*} = \mathcal{C}'_{1*}$.

By Lemma 4.1, the only way a codeword of $\mathcal{C}'_{0*} = \mathcal{C}'_{1*}$ can be a prefix of two distinct (non-confusable) codewords in $\mathcal{C}$ is when their suffixes are 00 and 11. This implies the uniqueness of the code $\mathcal{C}$ as well. The unique ∞-grain-correcting code of length n and size $2^{(n+1)/2}$ is in fact obtained from Construction 4.2. □

Remark 4.5: We point out that despite the fact that the induction step in the proof of Theorem 4.4 also holds when $\mathcal{C}$ is a binary ∞-grain-correcting code of even length n and size $2^{n/2}$, the proof cannot be generalized to include the even values of n as well, as it is impossible to find a basis for such an induction. Specifically, it is readily seen that there exist four different binary ∞-grain-correcting codes of length 2, namely,

$$\{00, 11\}, \{00, 10\}, \{01, 10\} \text{ and } \{01, 11\} . \tag{27}$$

For even $n \geq 4$, there exist at least four different constructions of ∞-grain-correcting codes of size $2^{n/2}$, obtained by prepending the prefixes in (27) to all the codewords of the code $\mathcal{C}_{n-2}$ obtained by Construction 4.2. Thus, for even n, a largest construction of ∞-grain-correcting codes is not unique. □

Construction 4.2 trivially yields $((n-3)/2)$-grain-correcting codes of odd length n and size $2^{(n+1)/2}$. We prove next that this size is optimal for $t = (n-3)/2$.

Theorem 4.6: Let $n \geq 5$ be an odd integer. Then

$$M_2^{(\mathrm{N})}(n, (n-3)/2) = M_2^{(\mathrm{O})}(n, (n-3)/2) = 2^{(n+1)/2}.$$

Proof: Construction 4.2 implies the lower bound

$$M_2^{(\mathrm{O})}(n, (n-3)/2) \geq 2^{(n+1)/2} ,$$

hence it remains to prove that

$$M_2^{(\mathrm{N})}(n, (n-3)/2) \leq 2^{(n+1)/2} .$$

The rest of the proof is similar to that of Theorem 4.4. The induction basis ($M_2^{(\mathrm{N})}(5, 1) = 8$) can be verified by a computer-based exhaustive search[11] (also see Table II).

Let now $\mathcal{C}$ be a binary $((n-3)/2)$-grain-correcting code (when overlaps are disallowed) of odd length $n \geq 7$, and let $\mathcal{C}'_{0*}$ and $\mathcal{C}'_{1*}$ be defined as in the proof of Theorem 4.4. These are $((n-5)/2)$-grain-correcting codes of length $n-2$ such that $|\mathcal{C}'_{0*}|+|\mathcal{C}'_{1*}| = |\mathcal{C}|$. By the induction hypothesis,

$$|\mathcal{C}| = |\mathcal{C}'_{0*}|+|\mathcal{C}'_{1*}| \leq 2 \cdot 2^{(n-1)/2} = 2^{(n+1)/2} .$$

□

Remark 4.7: Notice that the code $\mathcal{C}_n$ of Construction 4.2 is also ∞-grain-correcting under the criterion of wide-sense confusability. Hence Theorems 4.3 and 4.6 hold in this wide sense as well. □

[11]W.l.o.g., we can assume that a largest 1-grain-correcting code $\mathcal{C}$ of length 5 is closed under complementation (i.e., if $\boldsymbol{c} \in \mathcal{C}$, then the word obtained from $\boldsymbol{c}$ by changing each 0 to a 1 and each 1 to a 0 is also in $\mathcal{C}$). Thus, it suffices to verify that among any five words of length 5 starting with a 0 there is at least one 1-confusable pair. On the other hand, the code $\{00000, 00011, 00110, 01100\}$ is clearly 1-grain-correcting.

Turning to binary $(n/2-1)$-grain-correcting codes of even length n, we have the following result (for the proof, see Appendix B).

Theorem 4.8: Let $n \geq 4$ be an even integer. Then

$$M_2^{(\mathrm{N})}(n, n/2-1) = M_2^{(\mathrm{O})}(n, n/2-1) = 2^{n/2}+2.$$

The value of $M_2^{(\mathrm{N})}(n, n/2-1) = M_2^{(\mathrm{O})}(n, n/2-1)$ is realized by the augmentation of the code $\mathcal{C}_n$ in Construction 4.2 with the words[12] $(0110)^{n/4}$ and $(1001)^{n/4}$ when $n \equiv 0 \pmod 4$, or with the words $(0110)^{(n-2)/4}01$ and $(1001)^{(n-2)/4}10$ when $n \equiv 2 \pmod 4$ (see Appendix B).

Let $M_2^{(\mathrm{CWS})}(n, t)$ denote the size of a largest t-grain-correcting code of length n when assuming wide-sense confusability (notice that by Remark 2.14, it does not matter here whether overlaps are allowed or not). The next theorem shows that when n is even and $t = n/2-1$, the value of $M_2^{(\mathrm{CWS})}(n, t)$ is strictly smaller than $M_2^{(\mathrm{O})}(n, t)$.

Theorem 4.9: Let $n \geq 4$ be an even integer. Then

$$M_2^{(\mathrm{CWS})}(n, n/2-1) = 2^{n/2}.$$

Proof: Let $\mathcal{C}$ be a largest binary $(n/2-1)$-grain-correcting code of length n; w.l.o.g., we can assume that $\mathcal{C}$ is closed under complementation. The Hamming distance between two words $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ in $0\,\langle 2\rangle^{n-1}$ that are $n/2$-cws (that is, confusable) yet not $(n/2-1)$-cws, has to be[13] $n-1$, which, by Lemma 4.1, means that

$$\boldsymbol{c}_1 = (01)^{n/2} \quad \text{and} \quad \boldsymbol{c}_2 = 00(10)^{n/2-1} .$$

Therefore, if $\{\boldsymbol{c}_1, \boldsymbol{c}_2\} \not\subseteq \mathcal{C}$, then $\mathcal{C} \cap 0\,\langle 2\rangle^{n-1}$ — and, by closure under complementation, $\mathcal{C}$ itself — is an ∞-grain-correcting code (in the wide sense), and the result is implied by Theorem 4.3 and Remark 4.7.

Suppose now that $\{\boldsymbol{c}_1, \boldsymbol{c}_2\} \subseteq \mathcal{C}$. Every word $\boldsymbol{x} \in 0\,\langle 2\rangle^{n-1} \setminus \{\boldsymbol{c}_2\}$ is $(n/2-1)$-cws with $\boldsymbol{c}_1$, since the Hamming distance between $\boldsymbol{x} \in 0\,\langle 2\rangle^{n-1} \setminus \{\boldsymbol{c}_2\}$ and $\boldsymbol{c}_1$ is at most $n-2$ and, in addition, by Lemma 4.1, $\boldsymbol{c}_1$ is confusable with any word in $0\,\langle 2\rangle^{n-1}$. We then have $\mathcal{C} = \{\boldsymbol{c}_1, \boldsymbol{c}_2, \overline{\boldsymbol{c}}_1, \overline{\boldsymbol{c}}_2\}$, i.e., $|\mathcal{C}| = 4 \leq 2^{n/2}$. □

Using an inductive argument similar to that of Theorem 4.4, one can also prove that for $n \geq 5$, the binary $((n-3)/2)$-grain-correcting (in the wide sense) code of length n is unique.

An interesting (yet not provably optimal) construction of binary 1-grain-correcting codes (bearing some resemblance to the single asymmetric-error-correcting codes by Kim and Freiman [10]) can be obtained by the augmentation of a Hamming code with a subset of $\mathcal{C}_n$. We state this result in the following proposition.

[12]Recall that for a word $\boldsymbol{x}$, the notation $\boldsymbol{x}^s$ stands for s repetitions of $\boldsymbol{x}$.

[13]Since $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ are confusable, there exist grain patterns $\mathcal{S}_1$ and $\mathcal{S}_2$ such that $\sigma_{\mathcal{S}_1}(\boldsymbol{c}_1) = \sigma_{\mathcal{S}_2}(\boldsymbol{c}_2)$. W.l.o.g., we can assume that $\mathcal{S}_1 \cap \mathcal{S}_2 = \emptyset$, because had there been a location $e \in \mathcal{S}_1 \cap \mathcal{S}_2$, we could have obtained smaller grain patterns without that location $\mathcal{S}'_1 = \mathcal{S}_1 \setminus \{e\}$ and $\mathcal{S}'_2 = \mathcal{S}_2 \setminus \{e\}$ such that $\sigma_{\mathcal{S}'_1}(\boldsymbol{c}_1) = \sigma_{\mathcal{S}'_2}(\boldsymbol{c}_2)$. This implies that $|\mathcal{S}_1| + |\mathcal{S}_2|$ is at most the Hamming distance between $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$, which is at most $n-1$, since $\boldsymbol{c}_1, \boldsymbol{c}_2 \in 0\,\langle 2\rangle^{n-1}$. On the other hand, since $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ are not $(n/2-1)$-cws, one has $|\mathcal{S}_1| + |\mathcal{S}_2| > 2(n/2-1) = n-2$.

Proposition 4.10: Let $m \geq 2$ be an integer and let $n = 2^m - 1$. Then

$$M_2^{(\mathrm{N})}(n,1) = M_2^{(\mathrm{O})}(n,1) \geq 2^{n-m} + 2^{(n-1)/2}.$$

Proof: Consider a Hamming code $\mathcal{C}$ of length n with the parity-check matrix whose columns range over all the nonzero vectors in $\langle 2 \rangle^m$ in lexicographic order. Let

$$\mathcal{C}' = \left\{ \boldsymbol{x}' = 0(x_i' x_i')_{i \in \langle (n-1)/2 \rangle} : \mathsf{w}(\boldsymbol{x}') \in 4\mathbb{Z}+2 \right\}$$

and

$$\mathcal{C}'' = \left\{ \boldsymbol{x}'' = 1(x_i'' x_i'')_{i \in \langle (n-1)/2 \rangle} : \mathsf{w}(\boldsymbol{x}'') \in 4\mathbb{Z}+1 \right\} ,$$

where $\mathsf{w}(\boldsymbol{x})$ is the Hamming weight of the word $\boldsymbol{x}$. Denote $\mathcal{C}^* = \mathcal{C}' \cup \mathcal{C}''$. Any codeword $\boldsymbol{c} \in \mathcal{C}$ is at Hamming distance 1, 2, or 3 and higher from a word $\boldsymbol{x} \in \mathcal{C}^*$. The only codeword of $\mathcal{C}$ at distance 1 from $\boldsymbol{x}$ is $\boldsymbol{x} + 10^{n-1}$ (with addition taken componentwise modulo 2), but it is non-confusable with $\boldsymbol{x}$. Codewords of $\mathcal{C}$ at distance 2 from $\boldsymbol{x}$ differ from $\boldsymbol{x}$ on coordinates $1+2i$, $2+2i$ for some $i \in \langle (n-1)/2 \rangle$, yet this makes codewords of $\mathcal{C}$ at distance 2 non-confusable with $\boldsymbol{x}$. Codewords of $\mathcal{C}$ at distance 3 from $\boldsymbol{x}$ are not 1-confusable with $\boldsymbol{x}$ merely because 1-confusable words are at Hamming distance 2 (at most) from one another. Moreover, the code $\mathcal{C}^* \subseteq \mathcal{C}_n$ is ∞-grain-correcting therefore $\mathcal{C} \cup \mathcal{C}^*$ is a 1-grain-correcting code of size $2^{n-m} + 2^{(n-1)/2}$. □

Remark 4.11: Notice that the code $\mathcal{C} \cup \mathcal{C}^*$ in the proof of Proposition 4.10 is also 1-grain-correcting in the wide sense. □

Tables II and III contain the values of $M_2^{(\mathrm{N})}(n,t)$ and $M_2^{(\mathrm{CWS})}(n,t)$, respectively, for small n and t obtained using computer search (backtracking-based searching for a maximum independent set in a confusability graph [17, Sec. 3-A]). Values marked in bold are guaranteed by Theorems 4.3, 4.6, 4.8, or 4.9; values marked in italics are attained by unique codes due to Theorem 4.4 (and variations thereof). One can also observe that for $(n,t) = (7,1)$, the construction in Proposition 4.10 gives a code of size 24 which is close to the optimum $M_2^{(\mathrm{N})}(7,1) = 26$. Under the wide-sense confusability criterion, the very same code (by Remark 4.11) is a largest 1-grain-correcting code of length 7. It turns out that for all pairs (n,t) for which $M_2^{(\mathrm{N})}(n,t)$ is listed in Table II, we also have $M_2^{(\mathrm{O})}(n,t) = M_2^{(\mathrm{N})}(n,t)$; with the exception of $(n,t) = (8,2)$, this phenomenon can be explained by Theorems 4.3, 4.6, or 4.8, or by the basic observation that for $t = 1$ there can be no overlaps.

TABLE II
SIZES $M_2^{(\mathrm{N})}(n,t)$ OF LARGEST t-GRAIN-CORRECTING CODES OF LENGTH n.

t \ n	2	3	4	5	6	7	8	9
1	**2**	***4***	**6**	**8**	16	26	44	
2			**4**	***8***	**10**	**16**	22	
3					**8**	***16***	**18**	**32**

TABLE III
SIZES $M_2^{(\mathrm{CWS})}(n,t)$ OF LARGEST t-GRAIN-CORRECTING CODES OF LENGTH n, ASSUMING WIDE-SENSE CONFUSABILITY.

t \ n	2	3	4	5	6	7	8	9
1	**2**	***4***	**4**	***8***	12	24	32	
2			**4**	***8***	**8**	***16***	16	32
3					**8**	***16***	**16**	***32***

V. Error detection

Two words $\boldsymbol{x}, \boldsymbol{y} \in \Sigma^n$ are *t-similar* if there exists a grain pattern $\mathcal{S}$ of size at most t for which either $\sigma_{\mathcal{S}}(\boldsymbol{x}) = \boldsymbol{y}$ or $\sigma_{\mathcal{S}}(\boldsymbol{y}) = \boldsymbol{x}$. A code $\mathcal{C}$ of length n over Σ is called *t-grain-detecting* if no two distinct codewords in $\mathcal{C}$ are t-similar. In what follows, we will show the existence of an ∞-grain-detecting code $\mathcal{C}$ of length n over $\Sigma = \langle q \rangle$ with redundancy $n - \log_q |\mathcal{C}| \leq 1.5 \log_q(n) + O(1)$, for every $q \geq 2$ (here $O(1)$ stands for an absolute constant, independent of n and q).

Let n be a positive integer and define

$$\alpha(n) = \begin{cases} \left\lceil \frac{n+1}{2} \right\rceil & \text{if overlaps are disallowed} \\ n & \text{if overlaps are allowed} \end{cases} .$$

For $\boldsymbol{x} \in \Sigma^n$, let $\mathsf{s}(\boldsymbol{x})$ denote the sum of the indices of the starting positions of runs of $\boldsymbol{x}$. Let $\mathcal{F}$ denote the set of all binary words whose number of runs is either $\lfloor n(q-1)/q \rfloor$ or $\lfloor n(q-1)/q \rfloor + 1$ and which end with a run of length at least 2; partition $\mathcal{F}$ into blocks according to the value of $\mathsf{s}(\cdot)$ modulo $\alpha(n)$. By the pigeonhole principle, there has to be a partition block $\mathcal{C}_{\mathcal{F}} \subseteq \mathcal{F}$, of size at least $|\mathcal{F}| / \alpha(n)$ with the property that the value of $\mathsf{s}(\boldsymbol{x})$ modulo $\alpha(n)$ is the same for all $\boldsymbol{x} \in \mathcal{C}_{\mathcal{F}}$. Denote this common value by $\mathsf{s}_{\mathcal{F}}$.

Proposition 5.1: The code $\mathcal{C}_{\mathcal{F}}$ is an ∞-grain-detecting code with redundancy

$$n - \log_q |\mathcal{C}_{\mathcal{F}}| \leq 1.5 \log_q n + O\Big(\frac{1}{n}\Big)$$

(either when overlaps are allowed or not).

Proof: Let $m = \lfloor n(q-1)/q \rfloor$. A grain pattern $\mathcal{S}$ applied to a word $\boldsymbol{x} \in \mathcal{F}$ produces a word $\boldsymbol{y} = \sigma_{\mathcal{S}}(\boldsymbol{x})$ with a number of runs $\mathsf{r}(\boldsymbol{y})$ which is either equal to $\mathsf{r}(\boldsymbol{x})$ or is less than $\mathsf{r}(\boldsymbol{x})$ by at least 2. If the number of runs decreases by 2 (or more), then

$$\mathsf{r}(\boldsymbol{y}) \leq m - 1 < m \leq \mathsf{r}(\boldsymbol{x}),$$

and the error is detected. In particular, we will be able to detect such an error when words of $\mathcal{C}_{\mathcal{F}} \subseteq \mathcal{F}$ are transmitted.

Now, when the transmitted word $\boldsymbol{x}$ is from $\mathcal{C}_{\mathcal{F}}$ and $\mathsf{r}(\boldsymbol{y}) = \mathsf{r}(\boldsymbol{x})$, we can detect the inflicted errors by comparing $\mathsf{s}(\boldsymbol{y})$ with $\mathsf{s}_{\mathcal{F}}$. Specifically, since the maximal size[14] of $\mathcal{S}$ is $\alpha(n) - 1$, and any single grain increases the value of $\mathsf{s}(\boldsymbol{x})$ by 1, the maximal difference between $\mathsf{s}(\boldsymbol{y})$ and $\mathsf{s}(\boldsymbol{x})$ is $\alpha(n) - 1$, hence $\boldsymbol{y}$ and $\boldsymbol{x}$ are in different partition blocks of $\mathcal{F}$, *viz.*, $\boldsymbol{y} \notin \mathcal{C}_{\mathcal{F}}$.

[14]The maximal size of $\mathcal{S}$ when overlaps are allowed can be bounded from above by $\lfloor n(q-1)/q \rfloor$, rather than by $n-1$, but this will have no effect on the asymptotic analysis we are about to do.

The size of $\mathcal{C}_{\mathcal{F}}$ is at least

$$\frac{q}{\alpha(n)}\left(\binom{n-2}{m-1}(q-1)^{m-1}+\binom{n-2}{m}(q-1)^{m}\right)$$
$$\geq \frac{1}{\alpha(n)}\binom{n}{m}(q-1)^{m-1}$$
$$\geq \frac{1}{3(q-1)}\cdot\frac{1}{\alpha(n)\sqrt{n}}\cdot\frac{n^n(q-1)^m}{m^m(n-m)^{n-m}}$$
$$\geq \frac{1}{3(q-1)}\cdot\frac{1}{n\sqrt{n}}\,q^{nH_q\left(\frac{m}{n}\right)}$$
$$\geq \frac{1}{3(q-1)}\cdot\frac{1}{n\sqrt{n}}\,q^{nH_q\left(\frac{q-1}{q}-\frac{1}{n}\right)}$$
$$\geq \frac{1}{3(q-1)}\cdot\frac{1}{n\sqrt{n}}\,q^{n-O\left(\frac{1}{n}\right)}\ ,$$

where the second inequality follows from the known bounds on factorials [6, Sec. 2.9]:

$$\sqrt{2\pi}n^{n+1/2}\mathrm{e}^{-n+1/(12n+1)}<n!<\sqrt{2\pi}n^{n+1/2}\mathrm{e}^{-n+1/(12n)}\ ,$$

where e is the base of natural logarithms. Therefore, the redundancy of $\mathcal{C}_{\mathcal{F}}$ satisfies

$$n-\log_q|\mathcal{C}_{\mathcal{F}}|\leq 1.5\log_q n+O\Big(\frac{1}{n}\Big)\ .$$

□

An immediate corollary of Proposition 5.1 is that the rate of $\mathcal{C}_{\mathcal{F}}$ approaches 1 as $n\to\infty$, regardless of the number of grain errors to be detected.

APPENDICES

A. PROOF OF LEMMA 2.10

We prove the lemma for $j=\mathrm{N}$; the proof for $j=\mathrm{O}$ is similar. Let $\boldsymbol{x}=(x_i)_{i\in\langle n\rangle},\boldsymbol{y}=(y_i)_{i\in\langle n\rangle}\in\Sigma^n$ be a pair of t-cws words from $\mathcal{W}_t^{(\mathrm{N})}$. Then there exist grain patterns $\mathcal{S},\mathcal{S}'\subset\langle n\rangle\setminus\{0\}$ such that $\sigma_{\mathcal{S}}(\boldsymbol{x})=\sigma_{\mathcal{S}'}(\boldsymbol{y})$ and $|\mathcal{S}|+|\mathcal{S}'|$ $(\leq 2t)$ is the minimal possible sum of sizes of grain patterns that make $\boldsymbol{x}$ and $\boldsymbol{y}$ confusable. Construct the path $\gamma=(v_i=\ell_i r_i)_{i\in\langle n\rangle}$ corresponding to $(\boldsymbol{x},\boldsymbol{y})$ as follows: for each i,

$$\ell_i=\begin{cases}\underline{x}_i\ (\in\underline{\Sigma}) & \text{if } i\in\mathcal{S}\\ x_i\ (\in\Sigma) & \text{otherwise}\end{cases}$$

and

$$r_i=\begin{cases}\underline{y}_i\ (\in\underline{\Sigma}) & \text{if } i\in\mathcal{S}'\\ y_i\ (\in\Sigma) & \text{otherwise}\end{cases}\ .$$

Clearly, $|\mathsf{L}(\gamma)|=|\mathcal{S}|$, $|\mathsf{R}(\gamma)|=|\mathcal{S}'|$, thus $|\mathsf{L}(\gamma)|+|\mathsf{R}(\gamma)|\leq 2t$. Also, by construction, $x_i=\phi(\ell_i)$ and $y_i=\phi(r_i)$ for all $i\in\langle n\rangle$. Next, we verify that γ is indeed a path in $\mathcal{G}^{(\mathrm{N})}$. Every v_i constructed this way is indeed a state in $V^{(\mathrm{N})}$:

- It cannot be of the form $\ell_i r_i\in\Sigma\Sigma$ where $\ell_i\neq r_i$ because then $\sigma_{\mathcal{S}}(\boldsymbol{x})\neq\sigma_{\mathcal{S}'}(\boldsymbol{y})$.
- It cannot be of the form $\ell_i r_i\in\Sigma\underline{\Sigma}$ (or $\in\underline{\Sigma}\Sigma$) where $\phi(\ell_i)=\phi(r_i)$ because then either $\sigma_{\mathcal{S}}(\boldsymbol{x})\neq\sigma_{\mathcal{S}'}(\boldsymbol{y})$ or grain i is redundant in $\mathcal{S}'$ (or in $\mathcal{S}$), contradicting the minimality of $|\mathcal{S}|+|\mathcal{S}'|$.
- It cannot be of the form $\ell_i r_i\in\underline{\Sigma}\underline{\Sigma}$ where $\phi(\ell_i)=\phi(r_i)$ because then $\sigma_{\mathcal{S}\setminus\{i\}}(\boldsymbol{x})=\sigma_{\mathcal{S}'\setminus\{i\}}(\boldsymbol{y})$ contradicting the minimality of $|\mathcal{S}|+|\mathcal{S}'|$.

To verify that γ is indeed a path in $\mathcal{G}^{(\mathrm{N})}$, it is left to show that there are edges between the constructed $v_i=\ell_i r_i$ and $v_{i+1}=\ell_{i+1}r_{i+1}$ for any $i\in\langle n-1\rangle$. Indeed, if $v_{i+1}\in V_0$ then $\ell_{i+1}=r_{i+1}$ and by Condition [N1], $(v_i,v_{i+1})\in E^{(\mathrm{N})}$. If $v_{i+1}\in V_1$ and $v_{i+1}\in\Sigma\underline{\Sigma}$ (the case when $v_{i+1}\in\underline{\Sigma}\Sigma$ is similar), then $\phi(r_i)=\ell_{i+1}$ because otherwise $\sigma_{\mathcal{S}}(\boldsymbol{x})\neq\sigma_{\mathcal{S}'}(\boldsymbol{y})$. Hence v_i can be only of the following forms:

- $v_i\in\Sigma^2$ where $\ell_i{=}r_i{=}\ell_{i+1}$. This corresponds to Condition [N2].
- $v_i\in\underline{\Sigma}\Sigma$ where $r_i{=}\ell_{i+1}$. This corresponds to Condition [N3].

If $v_{i+1}\in V_2$ then $v_i\in V_0$ because otherwise $\sigma_{\mathcal{S}}(\boldsymbol{x})\neq\sigma_{\mathcal{S}'}(\boldsymbol{y})$. Moreover, $\ell_i=r_i\notin\{\phi(\ell_{i+1}),\phi(r_{i+1})\}$ since otherwise grain $i{+}1$ is redundant in $\mathcal{S}$ or in $\mathcal{S}'$. This corresponds to Condition [N4]. Finally, since $x_0=y_0$ (otherwise $\boldsymbol{x}$ and $\boldsymbol{y}$ are non-confusable), one has $v_0\in V_0$. From the above discussion we conclude that $\gamma\in\Pi_t^{(\mathrm{N})}$.

To prove that the above mapping from ordered pairs in $\mathcal{W}_t^{(\mathrm{N})}$ to paths in $\Pi_t^{(\mathrm{N})}$ is one-to-one, it remains to show that the above construction creates different paths for two different ordered pairs of t-cws words $(\boldsymbol{x}=(x_i)_{i\in\langle n\rangle},\boldsymbol{y}=(y_i)_{i\in\langle n\rangle})\neq(\boldsymbol{x'}=(x'_i)_{i\in\langle n\rangle},\boldsymbol{y'}=(y'_i)_{i\in\langle n\rangle})$. W.l.o.g., assume that there exists $s\in\langle n\rangle\setminus\{0\}$ such that $x_s\neq x'_s$; then state v_s in the respective path γ in $\mathcal{G}^{(\mathrm{N})}$ that was constructed from $(\boldsymbol{x},\boldsymbol{y})$ is different from state v'_s in the path γ' in $\mathcal{G}^{(\mathrm{N})}$ constructed from $(\boldsymbol{x}',\boldsymbol{y}')$. Therefore $\gamma\neq\gamma'$. □

B. PROOF OF THEOREM 4.8

It is easy to verify that the augmentation of $\mathcal{C}_n$ with

$$0110\ \ 0110\ \ 0110\ \ 0110\ldots$$

and its binary complement

$$1001\ \ 1001\ \ 1001\ \ 1001\ldots$$

results in an $(n/2{-}1)$-grain-correcting code (regardless of whether overlaps are allowed or not), implying the lower bound

$$M_2^{(\mathrm{N})}(n,n/2{-}1)\geq M_2^{(\mathrm{O})}(n,n/2{-}1)\geq 2^{n/2}{+}2.$$

Next, we prove the upper bound $M_2^{(\mathrm{N})}(n,n/2{-}1)\leq 2^{n/2}{+}2$.

Let $\mathcal{C}$ be a binary $(n/2{-}1)$-grain-correcting code of length n. For a word $\boldsymbol{x}=(x_i)_{i\in\langle n/2\rangle}\in\langle 2\rangle^{n/2}$, define $\mathcal{C}(\boldsymbol{x})$ to be the subcode of $\mathcal{C}$ with codewords containing $\boldsymbol{x}$ as a substring on the even-indexed positions, namely,

$$\mathcal{C}(\boldsymbol{x})=\left\{\boldsymbol{c}=(c_i)_{i\in\langle n\rangle}\in\mathcal{C}:\text{for all } i\in\langle n/2\rangle, c_{2i}=x_i\right\}.$$

By Lemma 4.1, all codewords of $\mathcal{C}(\boldsymbol{x})$ are pairwise confusable. Therefore, $|\mathcal{C}(\boldsymbol{x})|\leq 2$, as otherwise there would have existed two distinct (confusable) words in $\mathcal{C}(\boldsymbol{x})$ at Hamming distance less than $n/2$ apart, thus confusable by grain patterns $\mathcal{S}$ and $\mathcal{S}'$ such that $|\mathcal{S}|+|\mathcal{S}'|\leq n/2{-}1$. This, in turn, implies that $|\mathcal{S}|,|\mathcal{S}'|\leq n/2{-}1$, *viz.*, these two words in $\mathcal{C}(\boldsymbol{x})$ are $(n/2{-}1)$-confusable, which contradicts the fact that $\mathcal{C}$

is an $(n/2-1)$-grain-correcting code. We will call a word $\boldsymbol{x} \in \langle 2\rangle^{n/2}$ a 0-*profile*, a 1-*profile*, or a 2-*profile* if $|\mathcal{C}(\boldsymbol{x})|$ is 0, 1 or 2, respectively. For $m = 0, 1, 2$, denote the set of m-profiles by $\mathcal{P}_m$ (clearly, $|\mathcal{P}_2| = 2^{n/2} - |\mathcal{P}_0| - |\mathcal{P}_1|$). We will shortly demonstrate a one-to-one mapping from $\mathcal{P}_2 \setminus \{010101\ldots, 101010\ldots\}$ to $\mathcal{P}_0$. This, in turn, will imply that

$$|\mathcal{P}_0| \geq |\mathcal{P}_2| - 2 = 2^{n/2} - |\mathcal{P}_0| - |\mathcal{P}_1| - 2,$$

or, in other words, $2\,|\mathcal{P}_0| + |\mathcal{P}_1| \geq 2^{n/2} - 2$, which, combined with the fact that

$$\begin{aligned}|\mathcal{C}| &= |\mathcal{P}_1| + 2\,|\mathcal{P}_2| = |\mathcal{P}_1| + 2(2^{n/2} - \mathcal{P}_0 - \mathcal{P}_1) \\ &= 2^{n/2+1} - 2\,|\mathcal{P}_0| - |\mathcal{P}_1|\,,\end{aligned}$$

yields

$$|\mathcal{C}| \leq 2^{n/2+1} - (2^{n/2} - 2) = 2^{n/2} + 2\,.$$

Define a mapping $\eta : \mathcal{P}_2 \setminus \{010101\ldots, 101010\ldots\} \to \mathcal{P}_0$ in the following way. Let $\boldsymbol{x} = (x_i)_{i\in\langle n/2\rangle} \in \mathcal{P}_2 \setminus \{010101\ldots, 101010\ldots\}$ and let $j \in \langle n/2\rangle \setminus \{0\}$ be the smallest index for which $x_{j-1} = x_j$. Now, let $\eta(\boldsymbol{x})$ be defined as the word $\boldsymbol{y} = (y_i)_{i\in\langle n/2\rangle}$ such that for any $i \in \langle n/2\rangle$, one has

$$y_i = \begin{cases} x_i & i \neq j \\ \overline{x}_i & i = j \end{cases},$$

where $\overline{x}_i$ is the binary complement of x_i. Since $\boldsymbol{x}$ is a 2-profile, we have

$$\mathcal{C}(\boldsymbol{x}) = \left\{ \boldsymbol{c} = (x_i x_i)_{i\in\langle n/2\rangle}, \boldsymbol{c}^* = (x_i \overline{x}_i)_{i\in\langle n/2\rangle} \right\}.$$

The word $\boldsymbol{y}$ has to be in $\mathcal{P}_0$, because had there existed a word $\boldsymbol{c}' = (c_s)_{s\in\langle n\rangle} \in \mathcal{C}(\boldsymbol{y})$, it would have been either $(n/2-1)$-confusable with $\boldsymbol{c}^*$ (when $c_{2j-1} = \overline{x}_j$ or $c_{2j+1} = \overline{x}_j$), or $(n/2-1)$-confusable with $\boldsymbol{c}$ (when $c_{2j-1} = c_{2j+1} = x_j$). It is left to prove that η is one-to-one, namely, to refute the existence of another 2-profile $\boldsymbol{z} = (z_i)_{i\in\langle n/2\rangle} \in \mathcal{P}_2 \setminus \{010101\ldots, 101010\ldots, \boldsymbol{x}\}$ such that $\eta(\boldsymbol{x}) = \eta(\boldsymbol{z}) = \boldsymbol{y}$.

Suppose to the contrary that such a word $\boldsymbol{z}$ exists, and let $k \in \langle n/2\rangle \setminus \{0\}$ be the smallest index such that $z_{k-1} = z_k$; clearly, $j \neq k$. Since $\eta(\boldsymbol{z}) = \boldsymbol{y}$, then, for any $i \in \langle n/2\rangle$,

$$y_i = \begin{cases} z_i & i \neq k \\ \overline{z}_i & i = k \end{cases}.$$

In other words, $x_i = z_i$ when $i \notin \{j, k\}$, and $x_i = \overline{z}_i$ otherwise. Since $\boldsymbol{z}$ is a 2-profile, we have

$$\mathcal{C}(\boldsymbol{z}) = \left\{ (z_i z_i)_{i\in\langle n/2\rangle}, \widetilde{\boldsymbol{c}} = (z_i \overline{z}_i)_{i\in\langle n/2\rangle} \right\}.$$

It turns out that the words $\boldsymbol{c}^*$ and $\widetilde{\boldsymbol{c}}$ are 2-confusable by grain patterns without overlaps in each one of the possible cases:

- $k = j+1$. In this case, $\boldsymbol{c}^*$ and $\widetilde{\boldsymbol{c}}$ are confusable by the respective grain patterns $\{2j, 2j+2\}$ and $\{2j+1, 2j+3\}$. By definition, these grain patterns have no overlaps.
- $k = j-1$. In this case, $\boldsymbol{c}^*$ and $\widetilde{\boldsymbol{c}}$ are confusable by the respective grain patterns $\{2j-1, 2j+1\}$ and $\{2j-2, 2j\}$. By definition, these grain patterns have no overlaps.
- $|k-j| \geq 2$. In this case, $\boldsymbol{c}^*$ and $\widetilde{\boldsymbol{c}}$ are confusable by the respective grain patterns $\{2j, 2k+1\}$ and $\{2j+1, 2k\}$. Since $|k-j| \geq 2$, we have $|2j-(2k+1)| \geq 2$ and $|(2j+1)-2k| \geq 2$, therefore these grain patterns have no overlaps.

Notice that since $j, k \neq 0$ and $j \neq k$, either j or k has to be at least 2, which means $n \geq 6$. At any rate, $\boldsymbol{c}^*$ and $\widetilde{\boldsymbol{c}}$ cannot both be in a code correcting $(n/2-1) \geq 2$ grain errors that do not overlap, implying, in turn, that η is one-to-one. □

ACKNOWLEDGEMENT

We would like to thank the anonymous referees for their comments and suggestions. In particular, one of the referees suggested a better lower bound on $|\mathcal{B}_t^{(\mathrm{N})}(\boldsymbol{c})|$, which now appears in (21).

REFERENCES

[1] K.A.S. Abdel-Ghaffar, J.H. Weber, “Bounds and constructions for run-length limited error-control block codes,” *IEEE Trans. Inform. Theory,* vol. 37, pp. 789–800, 1991.

[2] T.M. Cover, J.A. Thomas, *Elements of Information Theory,* Wiley, New York, NY, 1991.

[3] I. Csiszár, T.M. Cover, B.-S. Choi, “Conditional limit theorems under Markov conditioning,” *IEEE Trans. Inform. Theory,* vol. 33, pp. 788–801, 1987.

[4] F.R. Gantmacher, *Matrix Theory, Volume II,* Chelsea Publishing Company, New York, 1960.

[5] S. Greaves, Y. Kanai, H. Muraoka, “Shingled magnetic recording on bit patterned media,” *IEEE Trans. Magn.,* vol. 46, pp. 1460–1463, 2010.

[6] W. Feller, *An Introduction to Probability Theory and Its Applications, Volume 1, 3rd Edition,* Wiley, New York, 1968.

[7] A.R. Iyengar, P.H. Siegel, J.K. Wolf, “Write channel model for bit-patterned media recording,” *IEEE Trans. Magn.,* vol. 47, pp. 35–45, 2011.

[8] A. Jiang, H. Zhou, Z. Wang, J. Bruck, “Patterned cells for phase change memories,” *Proc. IEEE Symp. Inform. Theory (ISIT 2011),* St. Petersburg, Russia, 2011, pp. 2333–2337.

[9] N. Kashyap, G. Zémor, “Upper bounds on the size of grain-correcting codes,” *Proc. IEEE Symp. Inform. Theory (ISIT 2013),* Istanbul, Turkey, 2013.

[10] W.H. Kim, C.V. Freiman, “IRE Trans. Inform. Theory,” vol. 5(2), pp. 62–66, 1959.

[11] V.D. Kolesnik, V.Yu. Krachkovsky, “Generating functions and lower bounds on rates for limiting error-correcting codes,” *IEEE Trans. Inform. Theory,* vol. 37, pp. 778–788, 1991.

[12] V.D. Kolesnik, V.Yu. Krachkovsky, “Lower bounds on achievable rates for limited bitshift correcting codes,” *IEEE Trans. Inform. Theory,* vol. 40, pp. 1443–1458, 1994.

[13] A.V. Kuznetsov, A.J.H. Vinck, “A coding scheme for single peak-shift correction in (d, k)-constrained channels,” *IEEE Trans. Inform. Theory,* vol. 39, pp. 1444–1450, 1993.

[14] D.G. Luenberger, *Introduction to Linear and Nonlinear Programming,* Addison-Wesley, Reading, MA, 1973.

[15] B.H. Marcus, R.M. Roth, “Improved Gilbert–Varshamov bound for constrained systems,” *IEEE Trans. Inform. Theory,* vol. 38, pp. 1213–1221, 1992.

[16] B.H. Marcus, R.M. Roth, P.H. Siegel, “Constrained systems and coding for recording channels,” in *Handbook of Coding Theory,* V.S. Pless and W.C. Huffman (Editors), Elsevier Scientific Publishers, Amsterdam, Netherlands, 1998.

[17] A. Mazumdar, A. Barg, N. Kashyap, “Coding for high-density recording on a 1-D granular magnetic medium,” *IEEE Trans. Inform. Theory,* vol. 57, pp. 7403–7417, 2011.

[18] W. Parry, S. Tuncel, “Classification Problems in Ergodic Theory,” *London Math. Soc. Lecture Note Series,* vol. 67, Cambridge University Press, 1982.

[19] T. Rockafellar, *Convex Analysis,* Princeton University Press, 1970.

[20] S. Shamai, E. Zehavi, “Bounds on the capacity of the bit-shift magnetic recording channel,” *IEEE Trans. Inform. Theory,* vol. 37, pp. 863–871, 1991.

[21] L.M.G.M. Tolhuizen, “The generalized Gilbert–Varshamov bound is implied by Turán’s theorem,” *IEEE Trans. Inform. Theory,* vol. 43, pp. 1605–1606, 1997.

[22] M.A. Tsfasman, S.G. Vlădut, T. Zink, “Modular curves, Shimura curves, and Goppa codes, better than Varshamov–Gilbert bound,” *Math. Nachr.,* vol. 109, pp. 21–28, 1982.
[23] R. Wood, M. Williams, A. Kavčić, J. Miles, “The feasibility of magnetic recording at 10 terabits per square inch on conventional media,” *IEEE Trans. Magn.,* vol. 45, pp. 917–923, 2009.

Artyom Sharov was born in Perm, U.S.S.R., in 1982. He received the B.Sc. degree in computer science, and the M.Sc. degree in computer science from Technion—Israel Institute of Technology, Haifa, Israel, in 2004, and 2010, respectively. He is currently pursuing the Ph.D. degree in computer science at Technion. His research interests include constrained coding and coding theory.

Ron M. Roth (M’88–SM’97–F’03) 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, and during 1996–97, 2004–05, and 2011–2012 he was on sabbatical leave at Hewlett–Packard Laboratories, Palo Alto, 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 the theory of complexity.