New Bounds and Constructions for Granular Media Coding

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

***Abstract*—Improved lower and upper bounds on the size and the rate of grain-correcting codes are presented. The lower bound is Gilbert–Varshamov-like combined with a construction by Gabrys *et al.*; it improves on the previously best known lower bounds on the asymptotic rate of $\lceil \tau n \rceil$-grain-correcting codes of length n on the interval $[0, 0.0668]$. One of the two newly presented upper bounds improves on the best known upper bounds on the asymptotic rate of $\lceil \tau n \rceil$-grain-correcting codes of length n on the interval $\tau \in (0, \frac{1}{8}]$ and meets the lower bound of $\frac{1}{2}$ for $\tau \geq \frac{1}{8}$. Moreover, in a nonasymptotic regime, both upper bounds improve on the previously best known results on the largest size of t-grain-correcting codes of length n, for certain values of n and t. Constructions of 1-grain-correcting codes based on a partitioning technique are presented for lengths up to 18. Finally, a lower bound of $\frac{1}{2} \log_2 n$ on the minimum redundancy of ∞-grain-detecting codes of length n is presented.**

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

I. Introduction

In a paper by Wood *et al.* [31], a certain improvement to the write and readback mechanisms of magnetic recording media was proposed, allowing for a higher storage density due to the ability of magnetizing areas commensurate with the dimensions of basic units forming the media called *grains*. Due to the higher writing density, one physical grain can be shared among several adjacent logical cells into which the media were partitioned, thereby introducing a new type of nonoverlapping smearing error to the information stored on these media. After the publication of [31], the granular media have been studied in several papers [11], [13], [14], [22], [27]. Mazumdar *et al.* [22] described a one-dimensional model of the errors occurring in these media restricting the grains to be only of lengths 1 and 2, and gave the first constructions and bounds on the size of codes that correct these so-called *grain errors*. In our earlier work [27], with a different yet conceptually similar application to shingled writing on bit-patterned media [12] in mind, we generalized the notion of grain errors to account for overlapping error patterns as well. Information-theoretic properties of the write channels representing the one-dimensional versions of both applications were studied by Iyengar *et al.* [13]. Kashyap and Zémor [14], as well as Fazeli *et al.* [8], [9], using a reduction to the problem of bounding the size of packings in hypergraphs (see [18]), presented the best known upper bounds on the size and rate of codes correcting grain errors for the nonoverlapping case, whereas Gabrys *et al.* [11] used that reduction to derive the best known upper bounds on the size of codes for the overlapping case. The best known lower bounds on the size and rate of those codes are due to our earlier work [27], where we present a general technique for improving Gilbert–Varshamov lower bounds (see the results of Kolesnik and Krachkovsky [17]). Several constructions of codes correcting a small number of grain errors were presented in [11].

Next, we formally present the model of grain errors. Let $\langle s \rangle$ denote the set $\{0, 1, \ldots, s{-}1\}$, for any positive integer s, and let $\Sigma_2 = \langle 2 \rangle$. A smearing error inflicted by a *grain (of length 2)*, ending at location e, to a word $\boldsymbol{x} = (x_i)_{i \in \langle n \rangle}$ of length n over Σ_2 results in an output word $\boldsymbol{y} = (y_i)_{i \in \langle n \rangle}$, wherein the value of y_e equals that of x_{e-1} instead of x_e. Given n consecutive positions on the medium (where words of length n over Σ_2 are to be written), define a *grain pattern* as a set $\mathcal{S} \subseteq \langle n \rangle \setminus \{0\}$ containing all the grain locations in these n positions. 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_2 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_2 in the following way: for any index $e \in \langle n \rangle \setminus \{0\}$,

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

If overlaps are disallowed in a grain pattern $\mathcal{S}$, for any two distinct grains $e, e' \in \mathcal{S}$, we require in addition that $|e' - e| > 1$.

Example 1.1: Let $n = 7$, $\boldsymbol{x} = 1010110$, $\mathcal{S} = \{1, 3, 5\}$, and $\mathcal{S}' = \{1, 2\}$. Then $\sigma_{\mathcal{S}}(\boldsymbol{x}) = 1111110$ and $\sigma_{\mathcal{S}'}(\boldsymbol{x}) = 1100110$. The grain pattern $\mathcal{S}$ is nonoverlapping, whereas $\mathcal{S}'$ has overlaps. □

For a positive integer t and $\boldsymbol{x}, \boldsymbol{x}' \in \Sigma_2^n$, the words $\boldsymbol{x}$ and $\boldsymbol{x}'$ are said to be as *t-confusable* if there exist grain patterns $\mathcal{S}, \mathcal{S}'$ of size at most t for which $\sigma_{\mathcal{S}}(\boldsymbol{x}) = \sigma_{\mathcal{S}'}(\boldsymbol{x}')$ (depending on the context, we may disallow overlaps in $\mathcal{S}$ and $\mathcal{S}'$). Words $\boldsymbol{x}$ and $\boldsymbol{x}'$ are *confusable* if they are t-confusable for some finite t; otherwise, they are said to be *non-confusable*. A binary code $\mathcal{C}$ of length n (namely, a nonempty subset of Σ_2^n) is

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 Grant No. 1092/12 from the Israel Science Foundation. Parts of this work were presented at the 51st Annual Allerton Conference on Communication, Control, and Computing, Monticello, Illinois, October 2013, and at the IEEE International Symposium on Information Theory, Honolulu, Hawaii, June–July 2014.

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. A code $\mathcal{C}$ will be called ∞-*grain-detecting* if for any codeword $\boldsymbol{x} \in \mathcal{C}$ and any nonempty grain pattern $\mathcal{S}$, one has $\sigma_{\mathcal{S}}(\boldsymbol{x}) \notin \mathcal{C}$.

Let $M(n,t)$ denote the largest size of any t-grain-correcting code of length n over Σ_2 with overlaps disallowed. For $\tau \in [0,1]$, define the (asymptotic) *rate* of $\lceil \tau n \rceil$-grain-correcting codes over Σ_2 (with overlaps disallowed) as

$$R(\tau) \triangleq \varlimsup_{n\to\infty} \frac{1}{n} \log_2 M(n, \lceil \tau n \rceil).$$

The rest of this paper is organized as follows. In Sections II and III, we present improved lower and upper bounds, respectively, on the size $M(n,t)$ and the rate $R(\tau)$ of $\lceil \tau n \rceil$-grain-correcting codes over Σ_2 (with overlaps disallowed). In Section IV, we show a construction of 1-grain-correcting codes of lengths up to 18 over Σ_2 using a well-known partitioning technique [1], [2, Sec. 6], [24]. Finally, Section V presents an upper bound on the size of ∞-grain-detecting codes of length n over Σ_2.

II. Lower bound on $R(\tau)$

In this section, we will develop Gilbert–Varshamov-like lower bounds on the size and the rate of grain-correcting codes. In the basis of our technique lies a variation of a construction by Gabrys *et al.* [11, Constr. 3] combined with our method from [27]; the latter technique, in turn, is based on the results of Kolesnik and Krachkovsky [17]. We start off by citing a modified version of [11, Constr. 3] next.

For simplicity of the presentation, we switch, as in our earlier work [27], to a different notion of confusability which we refer to as *wide-sense confusability*. Given a positive integer t, two words $\boldsymbol{x}$ and $\boldsymbol{x}'$ are t-*confusable in the wide sense* (t-*cws*, in short) if there exist grain patterns $\mathcal{S}$ and $\mathcal{S}'$ (with overlaps allowed) such that

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

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 (with overlaps allowed). Our lower bounds will, in fact, apply to t-grain-correcting codes in the wide sense and will imply the existence of t-grain-correcting codes in the ordinary sense of at least the same size. Note that for a pair of confusable words $\boldsymbol{x}$, $\boldsymbol{x}'$ over Σ_2, the minimal sum $|\mathcal{S}| + |\mathcal{S}'|$ of sizes of grain patterns $\mathcal{S}$ and $\mathcal{S}'$ that confuse these two words equals the Hamming distance $\mathsf{d}(\boldsymbol{x}, \boldsymbol{x}')$ between the words. This observation gives rise to an equivalent definition for wide-sense confusability, namely, two words $\boldsymbol{x}, \boldsymbol{x}' \in \Sigma_2^n$ are t-cws if they are confusable and $\mathsf{d}(\boldsymbol{x}, \boldsymbol{x}') \le 2t$.

Let $\Sigma_3 = \{\mathit{00}, \mathit{01}, \mathit{10}\}$ be a ternary alphabet[1] and let $\mathcal{E}^\star : \Sigma_2^2 \to \Sigma_3$ be a mapping which is defined for every binary pair $aa' \in \Sigma_2^2$ as follows:

$$\mathcal{E}^\star(aa') = \begin{cases} \mathit{00} & aa' \in \{00, 11\} \\ \mathit{01} & aa' = 01 \\ \mathit{10} & aa' = 10 \end{cases} .$$

Let n be a positive even integer denoting the length of our prospective code. We extend $\mathcal{E}^\star$ to a mapping $\mathcal{E} : \Sigma_2^n \to \Sigma_3^{n/2}$ which maps binary words $\boldsymbol{y} = (y_i y_{i+1})_{i \in \langle n/2 \rangle}$ of length n to ternary words of length $n/2$ as follows:

$$\mathcal{E}(\boldsymbol{y}) = (\mathcal{E}^\star(y_i y_{i+1}))_{i \in \langle n/2 \rangle}.$$

A *cloud* $\mathcal{E}^{-1}(\boldsymbol{x}) \subseteq \Sigma_2^n$ of a word $\boldsymbol{x} \in \Sigma_3^{n/2}$ is defined as the subset of all words in Σ_2^n whose image under $\mathcal{E}$ is $\boldsymbol{x}$, namely,

$$\mathcal{E}^{-1}(\boldsymbol{x}) = \{\boldsymbol{y} \in \Sigma_2^n : \mathcal{E}(\boldsymbol{y}) = \boldsymbol{x}\} .$$

Given words $\boldsymbol{x}, \boldsymbol{x}' \in \Sigma_3^{n/2}$, we will say that the clouds $\mathcal{E}^{-1}(\boldsymbol{x})$ and $\mathcal{E}^{-1}(\boldsymbol{x}')$ are t-*cws* if there exist binary words $\boldsymbol{y} \in \mathcal{E}^{-1}(x)$ and $\boldsymbol{y}' \in \mathcal{E}^{-1}(\boldsymbol{x}')$ of length n which are t-cws. Notice that any cloud is an ∞-grain-correcting code of length n (see [27, Lemma 4.1]), hence one way of constructing a t-grain-correcting code of length n is by joining clouds which are pairwise non-t-cws. The same basic idea underpins [11, Constr. 3]; our approach will differ by a more fine-grained analysis of the lower bound on the size of the obtained binary code of length n, by the virtue of [17].

Let $m \le n/2$ be a nonnegative integer. Let $\mathcal{C}_m(n)$ be a ternary code of length $n/2$ over Σ_3 such that the number of times that the symbol $\mathit{00} \in \Sigma_3$ appears in each codeword $\boldsymbol{c} = (c_i)_{i \in \langle n/2 \rangle} \in \mathcal{C}_m(n)$ is at least m. Define a binary code $\mathcal{C}$ of length n as the following set

$$\mathcal{C} = \bigcup_{\boldsymbol{c} \in \mathcal{C}_m(n)} \mathcal{E}^{-1}(\boldsymbol{c}). \tag{1}$$

The following lemma follows immediately; it summarizes how non-confusability of clouds in the code $\mathcal{C}_m(n)$ can be converted into grain-correcting capability of the code $\mathcal{C}$.

Lemma 2.1: Let $t \le n-1$ be a positive integer and let $\mathcal{C}_m(n)$ have the following property: for any two distinct codewords $\boldsymbol{c}, \boldsymbol{c}' \in \mathcal{C}_m(n)$, the clouds $\mathcal{E}^{-1}(\boldsymbol{c})$ and $\mathcal{E}^{-1}(\boldsymbol{c}')$ are not t-cws. Then the binary code $\mathcal{C}$ of length n defined in (1) is a t-grain-correcting code (in the wide sense) of size at least $2^m |\mathcal{C}_m(n)|$.

Let t be a positive integer. For a subset $\mathcal{J} \subseteq \Sigma_3^{n/2}$ and a word $\boldsymbol{x} \in \mathcal{J}$, let $\mathcal{R}_t(\boldsymbol{x}; \mathcal{J})$ be the set of all the words in $\mathcal{J}$ whose clouds are t-cws with $\mathcal{E}^{-1}(\boldsymbol{x})$. Let

$$\mathsf{W}_t(\mathcal{J}) = \sum_{\boldsymbol{x} \in \mathcal{J}} |\mathcal{R}_t(\boldsymbol{x}; \mathcal{J})|$$

be the number of ordered pairs of t-cws clouds of words in $\mathcal{J}$. In other words,

$$\mathsf{W}_t(\mathcal{J}) = \Big| \Big\{ (\boldsymbol{x}, \boldsymbol{x}') \in \mathcal{J} \times \mathcal{J} : \exists (\boldsymbol{y}, \boldsymbol{y}') \in \mathcal{E}^{-1}(\boldsymbol{x}) \times \mathcal{E}^{-1}(\boldsymbol{x}'), \text{ such that } \boldsymbol{y} \text{ and } \boldsymbol{y}' \text{ are } t\text{-cws} \Big\} \Big|.$$

Reformulating [17, Lemma 3] for grain-correcting codes, we are able to establish the following lower bound on $M(n,t)$.

Lemma 2.2: Let n, t be positive integers, $t \le n-1$, and let m be a nonnegative integer, $m \le n/2$. Let $\mathcal{J} = \mathcal{J}_m(n) \subseteq$

[1] We will use italics to denote the three (two-bit) symbols of Σ_3 to distinguish them better from the elements of Σ_2.

$\Sigma_3^{n/2}$ be a set of ternary words such that the number of times that the symbol *00* appears in each word of $\mathcal{J}$ is at least m. Then

$$M(n,t) \geq \frac{2^m \left|\mathcal{J}\right|^2}{4\mathsf{W}_t(\mathcal{J})}\,. \tag{2}$$

Proof: Combine [27, Lemma 2.5] and Lemma 2.1. □

Assuming we are able to compute the size of a selected subset $\mathcal{J}$, to obtain a lower bound on the rate of grain-correcting codes it remains to estimate the asymptotic growth rate of $\mathsf{W}_t(\mathcal{J})$, to which we will devote the rest of this section. The subsets $\mathcal{J}$ that we examine are certain sets of ternary words with prescribed empirical distribution of transitions from one symbol to another (such that the number of occurrences of *00* in each word is approximately m), and our method of assessing the asymptotic growth rate of $\mathsf{W}_t(\mathcal{J})$ relies on counting a certain type of cycle in a specifically designed finite directed graph.

For a word $\boldsymbol{x} = (x_i)_{i\in\langle n/2\rangle} \in \Sigma_3^{n/2}$ and symbols $a, a' \in \Sigma_3$, let

$$\kappa(\boldsymbol{x}; a, a') = \{i \in \langle n/2{-}1\rangle : (x_i, x_{i+1}) = (a, a')\}$$

count the number of transitions from the symbol a to the symbol a' in $\boldsymbol{x}$. Let p_0, p_1 be positive real numbers that satisfy $p_0{+}p_1 < 1$. For $\epsilon > 0$ and $n \geq \frac{16}{5\epsilon}$, let $\mathcal{J}_{p_0,p_1,\epsilon}(n)$ be the set of all the words $\boldsymbol{x}$ in $\Sigma_3^{n/2}$ such that for any $a, a' \in \Sigma_3$,

$$\left|\frac{\kappa(\boldsymbol{x}; a, a')}{\frac{1}{2}n{-}1} - \mu_{a,a'}\right| \leq \frac{\epsilon}{8}, \tag{3}$$

where

$$\mu_{00,00} = p_0, \tag{4}$$

$$\mu_{a,a'} = \frac{p_1}{4} \text{ for any } a, a' \in \{01, 10\} \text{ and} \tag{5}$$

$$\mu_{00,a} = \mu_{a,00} = \frac{1{-}p_0{-}p_1}{4} \text{ for any } a \in \{01, 10\}\,. \tag{6}$$

Let $z \in (0,1]$ and $h, m \in (0,\infty)$ be indeterminates and define the following matrix, whose rows and columns are indexed by $\langle 4\rangle$:

$$\mathcal{A}_{\mathcal{G}}(z,h,m) = \begin{pmatrix} h^2 & 4hz & 2 & 2z^2 \\ h & (2+2hm)z & 2m & 2mz^2 \\ 1 & 4mz & 2m^2 & 2m^2z^2 \\ 1 & 4mz & 2m^2 & m^2z^2 \end{pmatrix}. \tag{7}$$

The following lemma states an upper bound on the asymptotic growth rate of $\mathsf{W}_t(\mathcal{J})$ for $t = \lceil 2\tau(n/2{-}1)\rceil$ and $\mathcal{J} = \mathcal{J}_{p_0,p_1,\epsilon}(n)$ where $\tau \in (0,1)$ is a prescribed number of errors per symbol. The proof is similar to [27] and is included in Appendix A for completeness.

Lemma 2.3: Let $\tau \in (0,1)$, and let p_0, p_1 be positive real numbers that satisfy $p_0{+}p_1 < 1$. Then

$$\overline{\lim_{\epsilon\to 0}}\,\overline{\lim_{n\to\infty}}\,\frac{2}{n}\log_2 \mathsf{W}_{\lceil 2\tau(n/2-1)\rceil}(\mathcal{J}_{p_0,p_1,\epsilon}(n)) \leq K(\tau, p_0, p_1), \tag{8}$$

where

$$K(\tau,p_0,p_1) = \inf_{z\in(0,1],\ h,m\in(0,\infty)} \Big\{\log_2 \lambda(\mathcal{A}_{\mathcal{G}}(z,h,m)) - 4\tau\log_2 z - 2p_0\log_2 h - 2p_1\log_2 m\Big\}. \tag{9}$$

□

For positive real p_0, p_1 such that $p_0{+}p_1 < 1$, let $\mathsf{M} = \mathsf{M}(p_0, p_1)$ be the stationary Markov chain with the states *00*, *01*, *10* and the (non-conditional) probabilities of transitions as shown in Figure 1. Since the stationary probabilities of the states *00*, *01*, *10* are $\frac{1+p_0-p_1}{2}$, $\frac{1-p_0+p_1}{4}$ and $\frac{1-p_0+p_1}{4}$, respectively, the binary entropy $\mathtt{h}(\mathsf{M})$ of the Markov chain M is

$$\begin{aligned}\mathtt{h}(\mathsf{M}) = {} & \frac{1{-}p_0{+}p_1}{2}\log_2\frac{1{-}p_0{+}p_1}{4} \\ & + \frac{1{+}p_0{-}p_1}{2}\log_2\frac{1{+}p_0{-}p_1}{2} \\ & - p_0\log_2 p_0 - p_1\log_2\tfrac{p_1}{4} - (1{-}p_0{-}p_1)\log_2\tfrac{1-p_0-p_1}{4}.\end{aligned} \tag{10}$$

Next is our main theorem of this section.

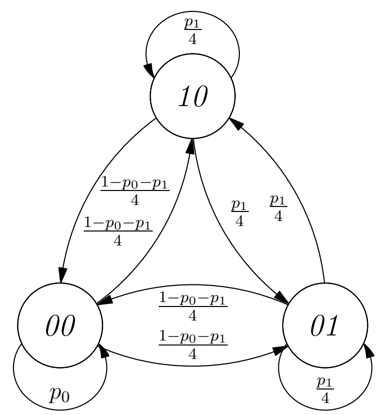

Fig. 1. Stationary Markov chain from the proof of Theorem 2.4.

Theorem 2.4: Let $\tau \in (0,1)$. Then

$$R(\tau) \geq \varrho(\tau) = \sup_{p_0,p_1}\left\{\frac{1{+}p_0{-}p_1}{4} + \mathtt{h}(\mathsf{M}) - \frac{1}{2}K(\tau,p_0,p_1)\right\},$$

where p_0, p_1 range over positive real numbers that satisfy $p_0{+}p_1 < 1$, and $K(\tau,p_0,p_1)$ and $\mathtt{h}(\mathsf{M})$ are defined in (9) and (10), respectively.

Proof: The words of $\mathcal{J}_{p_0,p_1,\epsilon}(n)$, defined in (3)–(6), are the typical sequences of the stationary Markov chain which appears in Figure 1. It is known (generalization of the asymptotic equipartition property [4, Th. 3.1.1] for Markovian sources) that the asymptotic growth rate of $\mathcal{J}_{p_0,p_1,\epsilon}(n)$ as n goes to infinity and as ϵ goes to 0 is

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{2}{n}\log_2\left|\mathcal{J}_{p_0,p_1,\epsilon}(n)\right| = \mathtt{h}(\mathsf{M}), \tag{11}$$

Moreover, the number of times that the symbol *00* appears in each word $\boldsymbol{x}$ of $\mathcal{J}_{p_0,p_1,\epsilon}(n)$ is at least

$$\begin{aligned}\sum_{a\in\Sigma_3}\kappa(\boldsymbol{x}; a, 00) &\geq \Big(\sum_{a\in\Sigma_3}\mu_{a,00} - \frac{3\epsilon}{8}\Big)\left(\frac{n}{2}{-}1\right) \\ &\geq \left(\frac{1{+}p_0{-}p_1}{2} - \epsilon\right)\frac{n}{2}.\end{aligned} \tag{12}$$

Therefore, for every positive real p_0 and p_1 that satisfy

$p_0{+}p_1 < 1$, we have

$$R(\tau) \overset{(2),\,(12)}{\geq} \varlimsup_{\epsilon\to 0} \varlimsup_{n\to\infty} \frac{1}{n}\Bigg(\left(\frac{1{+}p_0{-}p_1}{2}{-}\epsilon\right)\frac{n}{2} + 2\log_2|\mathcal{J}_{p_0,p_1,\epsilon}(n)| - \log_2 \mathsf{W}_{\lceil 2\tau(n/2-1)\rceil}(\mathcal{J}_{p_0,p_1,\epsilon}(n))\Bigg)$$
$$\overset{(8),\,(11)}{\geq} \frac{1{+}p_0{-}p_1}{4} + \mathtt{h}(\mathsf{M}) - \frac{1}{2}K(\tau,p_0,p_1).$$

□

Figure 2 shows the function $\tau \mapsto \varrho(\tau)$ depicted alongside the function $\tau \mapsto \varrho_{\mathrm{N}}(\tau)$ from [27, Th. 2.1] and the lower bound of 0.5 on the rate attained by a simple construction from [22, Sec. 2]. A visible improvement over the last two lower bounds can be observed on the interval $[0, \tau^{\star}{=}0.0668]$. Notice that due to the relationship between wide-sense confusability and ordinary confusability (with overlaps allowed), $\varrho(\tau)$ is a lower bound on the rate of $\lceil \tau n\rceil$-grain-correcting codes with overlaps, which, in turn, is a lower bound on $R(\tau)$.

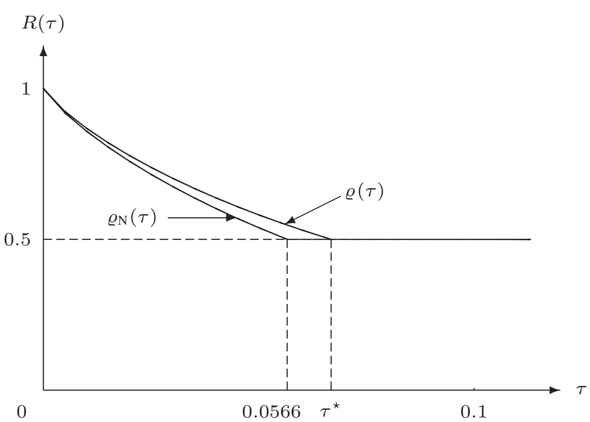

Fig. 2. Lower bound $\varrho(\tau)$ along with $\varrho_{\mathrm{N}}(\tau)$ from [27, Th. 2.1].

III. Upper bounds on $M(n,t)$ and $R(\tau)$

In this section, we will establish new upper bounds on $M(n,t)$ using two different approaches. The first approach is based on a connection that we establish between grain-correcting codes and codes correcting asymmetric errors, and will also bring about a new upper bound on $R(\tau)$. The second approach is based on the iterative procedure due to Cullina and Kiyavash [6]. We present the two new upper bounds on $M(n,t)$ in Section III-A and the upper bound on $R(\tau)$ in Section III-B.

A. Upper bounds on $M(n,t)$

For positive integers n and t, define $M_Z(n,t)$ to be the size of the largest code of length n correcting t asymmetric errors $1 \to 0$ [3]. We start off by establishing a correspondence between $M(n,t)$ and $M_Z(n,t)$, giving rise to the first upper bound on $M(n,t)$.

Theorem 3.1: Let n be a positive integer and $t \leq n/2$ be an integer. Then

$$M(n,t) \leq 2^{\lceil n/2\rceil} \cdot M_Z(\lfloor n/2\rfloor, t)\ .$$

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

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

Next, we note that a grain ending at position e can introduce an error to a word $\boldsymbol{c} = (c_i)_{i\in\langle n\rangle}$ only if $c_{e-1} \oplus c_e = 1$, where $\oplus$ is the addition modulo 2, and that the value of $c_{e-1} \oplus c_e$ changes to a 0, as a consequence. Therefore, if we restrict the grain patterns to the subset $\{e \in \langle n\rangle : e \text{ is odd}\}$, the code $\mathcal{C}(\boldsymbol{x})$ can be transformed into the following code $\mathcal{C}^{\oplus}(\boldsymbol{x})$ of length $\lfloor n/2\rfloor$ and of size $|\mathcal{C}(\boldsymbol{x})|$ correcting t asymmetric errors $1 \to 0$:

$$\mathcal{C}^{\oplus}(\boldsymbol{x}) = \left\{\boldsymbol{y}{=}(c_{2i}{\oplus}c_{2i+1})_{i\in\langle\lfloor n/2\rfloor\rangle} : \boldsymbol{c} = (c_i)_{i\in\langle n\rangle} \in \mathcal{C}(\boldsymbol{x})\right\}.$$

This implies that

$$M(n,t) = |\mathcal{C}| = \sum_{\boldsymbol{x}\in\Sigma_2^{\lceil n/2\rceil}} |\mathcal{C}(\boldsymbol{x})| = \sum_{\boldsymbol{x}\in\Sigma_2^{\lceil n/2\rceil}} \left|\mathcal{C}^{\oplus}(\boldsymbol{x})\right| \leq 2^{\lceil n/2\rceil} M_Z(\lfloor n/2\rfloor, t).$$

□

We now turn to developing the second upper bound on $M(n,t)$. To this end, we first recast the problem of finding the largest t-grain-correcting code of length n as a problem of finding a largest matching in a hypergraph, as defined next (see also [8][9][11][14][18]).

Let n, t be positive integers, let $\boldsymbol{x} \in \Sigma_2^n$ be a word and let $\Phi_t(\boldsymbol{x})$ be the set of all binary words of length n which are obtained by applying a grain pattern (with overlaps disallowed) of size at most t to $\boldsymbol{x}$, namely,

$$\Phi_t(\boldsymbol{x}) = \{\sigma_{\mathcal{S}}(\boldsymbol{x}) : |\mathcal{S}| \leq t\}\ .$$

Define a hypergraph $\mathcal{H} = \mathcal{H}_{n,t} = (V,E)$ where $V = \Sigma_2^n$ denotes the set of hypergraph states and $E = \{\Phi_t(\boldsymbol{x}) : \boldsymbol{x} \in \Sigma_2^n\}$ denotes the set of hyperedges in $\mathcal{H}$, that is, $|V| = |E| = 2^n$. Let $B_{\mathcal{H}}$ be the $2^n \times 2^n$ vertex-hyperedge incidence matrix of $\mathcal{H}$, namely, for a state $v \in V$ and a hyperedge $e \in E$,

$$(B_{\mathcal{H}})_{v,e} = \begin{cases} 1 & v \in e \\ 0 & \text{otherwise} \end{cases}.$$

Define the *matching number*

$$\mathsf{m}(\mathcal{H}) = \max\left\{\mathbf{1}^\top \boldsymbol{z} : \boldsymbol{z} \in \Sigma_2^{|V|}, B_{\mathcal{H}}\boldsymbol{z} \leq \mathbf{1}\right\}\ ,$$

where $\mathbf{1}$ is the all-ones column vector of length $|V| = 2^n$, all operations are over the reals, and the vector inequality holds component-wise; namely, $\mathsf{m}(\mathcal{H})$ is the largest size of a matching (i.e., a pairwise disjoint set of hyperedges) in $\mathcal{H}$, and so, $\mathsf{m}(\mathcal{H}) = M(n,t)$. The matching number $\mathsf{m}(\mathcal{H})$ is clearly bounded from above by the following *fractional matching number*

$$\max\left\{\mathbf{1}^\top \boldsymbol{z} : \boldsymbol{z} \in (\mathbb{R}^+ \cup \{0\})^{|V|}, B_{\mathcal{H}}\boldsymbol{z} \leq \mathbf{1}\right\}\ ,$$

which, by the strong linear programming duality, equals

$$\min\left\{\mathbf{1}^\top \boldsymbol{w} : \boldsymbol{w} \in (\mathbb{R}^+ \cup \{0\})^{|V|}, B_{\mathcal{H}}^\top \boldsymbol{w} \geq \mathbf{1}\right\}.$$

Thus, the sum of entries $\mathbf{1}^\top \boldsymbol{w}$ of any $\boldsymbol{w} \in (\mathbb{R}^+ \cup \{0\})^{|V|}$ such that $B_{\mathcal{H}}^\top \boldsymbol{w} \geq \mathbf{1}$ is a (generalized sphere-packing) upper bound on $M(n,t)$. The following lemma by Cullina and Kiyavash [6, Lemma 2], if applied iteratively, demonstrates how to take such a vector $\boldsymbol{w}$ and obtain an improvement on the upper bound of $\mathbf{1}^\top \boldsymbol{w}$.

For a real vector $\boldsymbol{z} = (z_{\boldsymbol{x}})_{\boldsymbol{x} \in \Sigma_2^n}$ and a word $\boldsymbol{y} \in (\Sigma_2)^{2^n}$, let

$$S_{\boldsymbol{z}}(\boldsymbol{y}) \triangleq \left(B_{\mathcal{H}}^\top \boldsymbol{z}\right)_{\boldsymbol{y}}$$

denote the $\boldsymbol{y}$-th component of the vector $B_{\mathcal{H}}^\top \boldsymbol{z}$.

Lemma 3.2: Let $\boldsymbol{w} \in (\mathbb{R}^+ \cup \{0\})^{|V|}$ be a vector such that $B_{\mathcal{H}}^\top \boldsymbol{w} \geq \mathbf{1}$. Let $\boldsymbol{w}' = (w'_{\boldsymbol{x}})_{\boldsymbol{x} \in \Sigma_2^n}$ be a vector of length 2^n, defined by

$$w'_{\boldsymbol{x}} = \frac{w_{\boldsymbol{x}}}{\min_{\boldsymbol{y}: \boldsymbol{x} \in \Phi_t(\boldsymbol{y})} \{S_{\boldsymbol{w}}(\boldsymbol{y})\}}, \quad \boldsymbol{x} \in \Sigma_2^n . \tag{13}$$

Then

$$B_{\mathcal{H}}^\top \boldsymbol{w}' \geq \mathbf{1} \tag{14}$$

and

$$\mathbf{1}^\top \boldsymbol{w}' \leq \mathbf{1}^\top \boldsymbol{w} . \tag{15}$$

Proof: To prove (14), we notice that

$$\begin{aligned} S_{\boldsymbol{w}'}(\boldsymbol{x}) &= \sum_{\boldsymbol{y} \in \Phi_t(\boldsymbol{x})} \frac{w_{\boldsymbol{y}}}{\min_{\boldsymbol{z}: \boldsymbol{y} \in \Phi_t(\boldsymbol{z})} \{S_{\boldsymbol{w}}(\boldsymbol{z})\}} \\ &\geq \sum_{\boldsymbol{y} \in \Phi_t(\boldsymbol{x})} \frac{w_{\boldsymbol{y}}}{S_{\boldsymbol{w}}(\boldsymbol{x})} = \frac{\sum_{\boldsymbol{y} \in \Phi_t(\boldsymbol{x})} w_{\boldsymbol{y}}}{S_{\boldsymbol{w}}(\boldsymbol{x})} = \frac{S_{\boldsymbol{w}}(\boldsymbol{x})}{S_{\boldsymbol{w}}(\boldsymbol{x})} = 1 \end{aligned}$$

for any $\boldsymbol{x} \in \Sigma_2^n$. To prove (15), we use the fact that $S_{\boldsymbol{w}}(\boldsymbol{y}) \geq 1$ for any $\boldsymbol{y} \in \Sigma_2^n$, hence

$$\mathbf{1}^\top \boldsymbol{w}' = \sum_{\boldsymbol{x} \in \Sigma_2^n} \frac{w_{\boldsymbol{x}}}{\min_{\boldsymbol{y}: \boldsymbol{x} \in \Phi_t(\boldsymbol{y})} \{S_{\boldsymbol{w}}(\boldsymbol{y})\}} \leq \sum_{\boldsymbol{x} \in \Sigma_2^n} \frac{w_{\boldsymbol{x}}}{1} = \mathbf{1}^\top \boldsymbol{w} .$$

□

Remark 3.3: Since the proof of (14) in Lemma 3.2 does not depend on the validity of the condition $B_{\mathcal{H}}^\top \boldsymbol{w} \geq \mathbf{1}$, we can first apply (13) to an arbitrary vector $\boldsymbol{w} \in (\mathbb{R}^+ \cup \{0\})^{|V|}$ (say, $\boldsymbol{w} = \mathbf{1}$) to obtain the vector $\boldsymbol{w}'$ yielding a first upper bound of $\mathbf{1}^\top \boldsymbol{w}'$ on $M(n,t)$, and then continue applying Lemma 3.2 repeatedly to (the descendants of) $\boldsymbol{w}'$ to obtain ever-improving upper bounds on $M(n,t)$. □

Using Theorem 3.1 along with the best known bounds[2] on $M_Z(\lfloor n/2 \rfloor, t)$ from [30, Table 10], as well as the iterative process described by Lemma 3.2 along with Remark 3.3, results in improvements on the best known upper bounds on $M(n,t)$, as shown in Table I. This table contains the best known upper bounds (with the corresponding best known lower bounds in parenthesis) on $M(n,t)$ for small values of n and t. Therein, the best upper bounds due to Theorem 3.1 or Lemma 3.2 are marked in bold, whereas the best upper bounds due to linear optimization of $\mathsf{m}(\mathcal{H})$ are typeset in medium font (with the exception of the best upper bound of 88 on $M(9,1)$ which was obtained by doubling $M(8,1) = 44$). The best lower bounds on $M(n,1)$ due to [11, Constr. 1] are marked with daggers, the best lower bounds on $M(n,2)$ and $M(n,3)$ due to [11, Ex. 4] (or variations thereof) are marked with diamonds, and the lower bounds typeset in medium font are derived from Table IV below (based on our construction in Section IV), [27, Table 2], and variations thereof. Tight upper bounds, obtained using exhaustive computer search, are marked in italics.

[2]Our definition of $M_Z(n,t)$ is equivalent to $Z(n,t+1)$ in [30].

TABLE I
BOUNDS ON THE SIZES $M(n,t)$ OF THE LARGEST KNOWN t-GRAIN-CORRECTING CODES OF LENGTH n.

t \ n	2	3	4	5	6	7	8	9	10	11	12	13
1	*2*	*4*	*6*	*8*	*16*	*26*	*44*	88(72)	172(112)	316(210†)	588(372)	1098(702†)
2			*4*	*8*	*10*	*16*	*22*	*32*	**64**(44)	106(68$^\diamond$)	182(88)	312(136$^\diamond$)
3					*8*	*16*	*18*	*32*	*38*	64(64)	96(76)	152(128)

t \ n	14	15	16	17	18
1	2054(1272)	**3930**(2400†)	**7396**(4522)	**13974**(8428)	**26488**(15348)
2	**512**(176)	**1024**(312$^\diamond$)	**1792**(418$^\diamond$)	**3264**(836$^\diamond$)	**5810**(1318$^\diamond$)
3	242(152)	**490**(260$^\diamond$)	**802**(304)	**1316**(520$^\diamond$)	**2048**(608)

B. Upper bound on $R(\tau)$

For a positive integer n, define the *asymmetric distance* $\Delta(\boldsymbol{c}, \boldsymbol{c}')$ between two words $\boldsymbol{c} = (c_i)_{i \in \langle n \rangle}$ and $\boldsymbol{c}' = (c'_i)_{i \in \langle n \rangle}$ over Σ_2 as

$$\Delta(\boldsymbol{c}, \boldsymbol{c}') \triangleq \max\left\{\Delta^\star(\boldsymbol{c}, \boldsymbol{c}'), \Delta^\star(\boldsymbol{c}', \boldsymbol{c})\right\} ,$$

where

$$\Delta^\star(\boldsymbol{c}, \boldsymbol{c}') = |\{i \in \langle n \rangle : c_i = 0, c'_i = 1\}| ,$$

and the *minimum asymmetric distance* of a code $\mathcal{C} \subseteq \Sigma_2^n$ as

$$\Delta(\mathcal{C}) \triangleq \min_{\boldsymbol{c}, \boldsymbol{c}' \in \mathcal{C}: \boldsymbol{c} \neq \boldsymbol{c}'} \{\Delta(\boldsymbol{c}, \boldsymbol{c}')\} .$$

Let $\mathsf{d}(\boldsymbol{c}, \boldsymbol{c}')$ denote the *Hamming distance* between two words $\boldsymbol{c}, \boldsymbol{c}' \in \Sigma_2^n$ and let $\mathsf{d}(\mathcal{C})$ denote the *minimum Hamming distance* of the code $\mathcal{C} \subseteq \Sigma_2^n$. Let $M_H(n,t)$ denote the size of a largest code of length n over Σ_2 correcting t (Hamming) errors and let

$$R_H(\tau) = \varlimsup_{n \to \infty} \frac{1}{n} \log_2 M_H(n, \lceil \tau n \rceil)$$

denote the (asymptotic) *rate* of codes of length n over Σ_2 correcting $\lceil \tau n \rceil$ (Hamming) errors. For a real $p \in [0,1]$, let

$$\mathsf{H}(p) \triangleq -p \log_2 p - (1-p) \log_2 (1-p)$$

denote the binary entropy of p. In the following theorem, which is the main result of this section, we prove a new upper bound on $R(\tau)$.

Theorem 3.4: Let $\tau \in [0, \frac{1}{8}]$. Then

$$R(\tau) \leq \frac{1}{2}(1 + R_H(2\tau)) . \tag{16}$$

In particular,

$$R(\tau) \leq \rho(\tau) \triangleq \frac{1}{2}\Big(1 + \min_{0 < x \leq 1 - 8\tau} \{b(x)\}\Big) , \tag{17}$$

where

$$b(x) = 1 + g(x^2) - g(x^2 + 8\tau x + 8\tau)$$

and

$$g(x) = \mathsf{H}\left(0.5(1 - \sqrt{1-x})\right) .$$

Proof: Let n be a positive integer and let $\mathcal{C}$ be a code of length n correcting $\lceil \tau n \rceil$ asymmetric errors of size $M_Z(n, \lceil \tau n \rceil)$. Its asymmetric distance $\Delta(\mathcal{C})$ is therefore at least $\lceil \tau n \rceil + 1$ (see [7, Th. 1]). By an averaging argument, there exists a constant-weight subcode $\mathcal{C}(w)$ of $\mathcal{C}$ whose codewords are of Hamming weight $w \in \langle n \rangle \setminus \{0\}$, whose size is at least $(|\mathcal{C}| - 2)/(n-1)$, and whose asymmetric distance is clearly at least $\lceil \tau n \rceil + 1$. Since $\mathsf{d}(\boldsymbol{c}, \boldsymbol{c}') = 2\Delta(\boldsymbol{c}, \boldsymbol{c}')$ for any two codewords $\boldsymbol{c}, \boldsymbol{c}' \in \mathcal{C}(w)$ (see [15, Sec. 2]), one has $\mathsf{d}(\mathcal{C}(w)) \geq 2(\lceil \tau n \rceil + 1)$, therefore $\mathcal{C}(w)$ can correct at least $\lceil \tau n \rceil$ (Hamming) errors.

The above discussion[3] implies

$$\begin{aligned} M_Z(n, \lceil \tau n \rceil) = |\mathcal{C}| &\leq (n-1)\, |\mathcal{C}(w)| + 2 \\ &\leq (n-1) M_H(n, \lceil \tau n \rceil) + 2, \end{aligned}$$

which, combined with the result of Theorem 3.1, yields

$$M(n, \lceil \tau n \rceil) \leq 2^{\lceil n/2 \rceil} \cdot \Big((\lfloor n/2 \rfloor - 1) \cdot M_H(\lfloor n/2 \rfloor, \lceil \tau n \rceil) + 2 \Big). \tag{18}$$

Asymptotically, the inequality (18) implies (16). Finally, to obtain the upper bound (17), we use the second MRRW upper bound [20, Ch. 17, Th. 37] on $R_H(2\tau)$. □

Figure 3 depicts the upper bound $\rho(\tau)$ of Theorem 3.4 along with two previously best known upper bounds $\rho_1(\tau)$ (see [14, Th. 6.1]) and $\rho_2(\tau)$ (see [26, Th. 3.3]) obtained using information-theoretic and sphere-packing arguments, respectively. The best known lower bound $\varrho(\tau)$ of Theorem 2.4 is plotted therein for comparison, along with the (ordinary) Gilbert–Varshamov bound

$$\varrho_4(\tau) \triangleq 1 - \mathsf{H}(2\tau) \; .$$

In addition, the dotted curve presents the Gilbert–Varshamov lower bound

$$\varrho_5(\tau) \triangleq 1 - \frac{1}{2} \mathsf{H}(4\tau)$$

on the rate of the largest $\lceil \tau n \rceil$-grain-correcting codes of length n when the grain patterns are restricted to the subset $\{e \in \langle n \rangle : e \text{ is odd}\}$. The upper bound $\rho(\tau)$ improves on $\rho_1(\tau)$ and on $\rho_2(\tau)$ on the entire interval $(0, \frac{1}{8}]$, and at $\tau = \frac{1}{8}$, it coincides with the lower bound of $\frac{1}{2}$ on $R(\tau)$ obtained by a simple construction from [22, Sec. 2]. The upper bound $\rho(\tau)$ also improves on the entire interval $(0, \frac{1}{8}]$ on the upper bound $\rho_3(\tau)$ derived from [11, Th. 1] on the rate of $\lceil \tau n \rceil$-grain-correcting codes of length n when overlaps are allowed.

The fact that the new upper bound $\rho(\tau)$ meets the lower bound of $\frac{1}{2}$ at $\tau = \frac{1}{8}$ implies a very slow decrease in the size $M(n, \lceil \tau n \rceil)$ of a largest $\lceil \tau n \rceil$-grain-correcting code of length n when τ runs from $\frac{1}{8}$ to $\frac{1}{2}$, which we demonstrate next for $\tau \geq \frac{1}{4}$. Let t be a positive integer and let $n = 4t$. Since a largest code of length $n/2 = 2t$ correcting $t = n/4$ asymmetric errors is of size[4] 2, by Theorem 3.1, the size

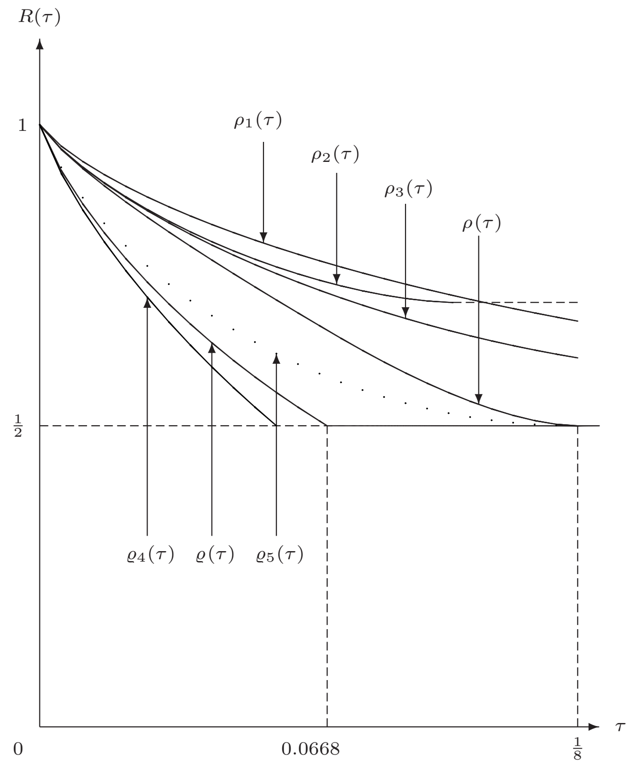

Fig. 3. Upper bound $\rho(\tau)$ along with upper bounds $\rho_1(\tau)$, $\rho_2(\tau)$ and $\rho_3(\tau)$ and lower bounds $\varrho(\tau)$, $\varrho_4(\tau)$, $\varrho_5(\tau)$.

$M(n, n/4)$ of a largest $n/4$-grain-correcting code of length n is at most $2^{n/2+1}$. As, due to [22, Prop. 1], $M(n, n/2) = 2^{n/2}$, when t runs from $\frac{n}{4}$ to $\frac{n}{2}$, the largest code size $M(n, t)$ decreases only by at most a factor of 2.

IV. Constructions of 1-grain-correcting codes

In this section, we present a construction of 1-grain-correcting codes based on the well-known partitioning technique (Al-Bassam *et al.* [1] used it to construct asymmetric single-error-correcting codes; also see [2] and [24]). To this end, we will need a somewhat stronger definition of confusability. Two binary words $\boldsymbol{x}, \boldsymbol{x}' \in \Sigma_2^n$ of length n will be referred to as 1-*strongly-confusable* (in short, 1-*sc*) if either $0\boldsymbol{x} \in \Sigma_2^{n+1}$ and $0\boldsymbol{x}' \in \Sigma_2^{n+1}$ are 1-confusable or $1\boldsymbol{x} \in \Sigma_2^{n+1}$ and $1\boldsymbol{x}' \in \Sigma_2^{n+1}$ are 1-confusable (in the ordinary sense). A code of length n will be referred to as 1-*grain-correcting in the strong sense* if its codewords are pairwise not 1-sc.

INPUT: graph $\mathcal{G} = (V, E)$;
$\boldsymbol{\pi} \leftarrow \emptyset$; // $\boldsymbol{\pi}$ is a partition of V
while $V \neq \emptyset$ **do** {
  $\mathcal{B} \leftarrow$ MAXIMUMCLIQUE($\mathcal{G}$);
  $\mathcal{G} \leftarrow \mathcal{G} \setminus \mathcal{B}$; // remove all the states of $\mathcal{B}$ from V and // all the edges connected to $\mathcal{B}$ from E
  ADD($\boldsymbol{\pi}, \mathcal{B}$); // append $\mathcal{B}$ to the end of the list $\boldsymbol{\pi}$
}

Fig. 4. Greedy procedure to obtain a partition $\boldsymbol{\pi}$ of the graph $\mathcal{G}$.

Example 4.1: The words $\boldsymbol{x} = 0001 \in \Sigma_2^4$ and $\boldsymbol{x}' = 1011 \in \Sigma_2^4$ are ∞-confusable in the ordinary sense, but 1-sc because $1\boldsymbol{x} \in \Sigma_2^5$ and $1\boldsymbol{x}' \in \Sigma_2^5$ are 1-confusable in the

[3] The result we developed in this discussion is a particular case of a more general claim obtained by substituting $\tau_1 = \tau_2 = \delta_1 = \delta_2 = \delta_3 = 0$ in [29, Th. 4].

[4] One such code is $\{0^{2t}, 1^{2t}\}$. Conversely, any binary code $\mathcal{C}$ of length $2t$ correcting t asymmetric errors must have at most one codeword of weight less than t, by the definition of the asymmetric distance and [7, Th. 1]. Since the supports of any two distinct words $\boldsymbol{x}, \boldsymbol{x}'$ of length $2t$ and weight at least $t+1$ have a nonempty intersection, we have $0 < \Delta(\boldsymbol{x}, \boldsymbol{x}') \leq t$, thus only one word of weight at least $t+1$ can be in $\mathcal{C}$, implying $|\mathcal{C}| \leq 2$.

TABLE II
SIZES OF PARTITIONS $\mathcal{J}_i$ OF Σ_2^b FOR VARIOUS VALUES OF b.

b	$\mathcal{J}_0$	$\mathcal{J}_1$	$\mathcal{J}_2$	$\mathcal{J}_3$	$\mathcal{J}_4$	$\mathcal{J}_5$	$\mathcal{J}_6$	$\mathcal{J}_7$	$\mathcal{J}_8$	$\mathcal{J}_9$	$\mathcal{J}_{10}$
1	1	1									
2	2	1	1								
3	2	2	2	2							
4	4	4	3	3	2						
5	7	6	6	5	5	2	1				
6	12	10	10	10	9	8	5				
7	22	19	19	18	17	15	11	6	1		
8	35	35	33	32	30	27	23	20	15	5	1

TABLE III
SIZES OF PARTITIONS $\mathcal{K}_i$ OF $\Sigma_{2,\text{EVEN}}^q$ FOR VARIOUS VALUES OF q.

q	$\mathcal{K}_0$	$\mathcal{K}_1$	$\mathcal{K}_2$	$\mathcal{K}_3$	$\mathcal{K}_4$	$\mathcal{K}_5$	$\mathcal{K}_6$	$\mathcal{K}_7$	$\mathcal{K}_8$	$\mathcal{K}_9$
1	1									
2	2									
3	2	1	1							
4	4	2	2							
5	5	4	3	3	1					
6	8	6	6	6	4	2				
7	15	12	11	10	7	6	3			
8	24	20	20	19	16	15	10	4		
9	40	38	36	34	31	27	22	17	10	1
9	39	38	37	34	32	28	23	16	6	3

ordinary sense by the grain patterns $\mathcal{S} = \{1\}$ and $\mathcal{S}' = \{3\}$ applied to $1\boldsymbol{x}$ and $1\boldsymbol{x}'$, respectively. □

Let b, $b^\star$ be positive integers and, for $i \in \langle b^\star \rangle$, let the sets $\mathcal{J}_i$ be 1-grain-correcting codes in the strong sense that form a partition of Σ_2^b (viz., $\mathcal{J}_i \cap \mathcal{J}_{i'} = \emptyset$ for any distinct $i, i' \in \langle b^\star \rangle$, and $\bigcup_{i \in \langle b^\star \rangle} \mathcal{J}_i = \Sigma_2^b$). Let q, $q^\star$ be positive integers and, for $i \in \langle q^\star \rangle$, let the sets $\mathcal{K}_i$ be 1-grain-correcting codes in the strong sense that form a partition of the set

$$\Sigma_{2,\text{even}}^q = \{\boldsymbol{x} \in \Sigma_2^q : \mathsf{w}(\boldsymbol{x}) \text{ is even}\}$$

(of even-weighted binary words of length q). Finally, let

$$\mathcal{C} = \bigcup_{i \in \langle \min\{b^\star, q^\star\} \rangle} \Sigma_2 \times \mathcal{J}_i \times \mathcal{K}_i \tag{19}$$

be the union of Cartesian products $\Sigma_2 \times \mathcal{J}_i \times \mathcal{K}_i$ for $i \in \langle \min\{b^\star, q^\star\} \rangle$. The code $\mathcal{C}$ is of length $n = b+q+1$ and of size

$$|\mathcal{C}| = 2 \cdot \sum_{i \in \langle \min\{b^\star, q^\star\} \rangle} |\mathcal{J}_i| \cdot |\mathcal{K}_i|. \tag{20}$$

The following theorem states that the code $\mathcal{C}$ is a 1-grain-correcting code.

Theorem 4.2: The code $\mathcal{C}$ of length $n = b+q+1$ defined as in (19) is a 1-grain-correcting code.

Proof: Let $\boldsymbol{c} = (z\ \boldsymbol{x}\ \boldsymbol{y})$, $\boldsymbol{c}' = (z\ \boldsymbol{x}'\ \boldsymbol{y}')$ be two distinct codewords in $\mathcal{C}$ such that $z \in \Sigma_2$, $\boldsymbol{x} = (x_j)_{j \in \langle b \rangle} \in \mathcal{J}_i$, $\boldsymbol{x}' = (x'_j)_{j \in \langle b \rangle} \in \mathcal{J}_{i'}$, $\boldsymbol{y} \in \mathcal{K}_i$, and $\boldsymbol{y}' \in \mathcal{K}_{i'}$, for $i, i' \in \langle \min\{b^\star, q^\star\} \rangle$, and w.l.o.g. assume that $z = 0$.

TABLE IV
SIZES OF THE 1-GRAIN-CORRECTING CODES DUE TO THEOREM 4.2.

n	10	11	12	13	14	15	16	17	18
Size	**112**	206	**372**	686	**1272**	2384	**4522**	**8428**	**15348**

- If $i = i'$ and $\boldsymbol{x} \neq \boldsymbol{x}'$, then $\boldsymbol{c}$ and $\boldsymbol{c}'$ are not 1-confusable because their prefixes of length $b+1$, $0\boldsymbol{x}$ and $0\boldsymbol{x}'$, are not 1-confusable (due to the fact that $\mathcal{J}_i$ is a 1-grain-correcting code in the strong sense).
- If $i = i'$ and $\boldsymbol{y} \neq \boldsymbol{y}'$, then $\boldsymbol{c}$ and $\boldsymbol{c}'$ are not 1-confusable because their suffixes of length $q+1$, $x_{b-1}\boldsymbol{y}$ and $x'_{b-1}\boldsymbol{y}'$, are not 1-confusable (due to the fact that $\mathcal{K}_i$ is a 1-grain-correcting code in the strong sense).
- Finally, if $i \neq i'$, then necessarily $\boldsymbol{x} \neq \boldsymbol{x}'$ and $\boldsymbol{y} \neq \boldsymbol{y}'$, therefore

$$\mathsf{d}(\boldsymbol{x}, \boldsymbol{x}') \geq 1 \text{ and } \mathsf{d}(\boldsymbol{y}, \boldsymbol{y}') \geq 2,$$

implying that $\mathsf{d}(\boldsymbol{c}, \boldsymbol{c}') \geq 3$ which makes it impossible for $\boldsymbol{c}$ and $\boldsymbol{c}'$ to be 1-confusable (as 1-confusable words must be at Hamming distance at most 2 from one another). □

It remains to show how to obtain the partitions $\mathcal{J}_i$ for $i \in \langle b^\star \rangle$ and $\mathcal{K}_i$ for $i \in \langle q^\star \rangle$ with the desired properties. Define a *non-confusability graph* $\mathcal{G}_n(\mathcal{Z}) = (\mathcal{Z}, E)$, where $\mathcal{Z} \subseteq \Sigma_2^n$, as an undirected graph whose states are all the words in $\mathcal{Z}$ and whose set of edges E contains an edge between $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{Z}$ if and only if the words $\boldsymbol{x}$ and $\boldsymbol{x}'$ are <u>not</u> 1-sc (contrast with the definition of the confusability graph in [22, Sec. 3-C]). Figure 4 shows a greedy procedure which we used to find partitions $\mathcal{J}_i$ for $i \in \langle b^\star \rangle$ and $\mathcal{K}_i$ for $i \in \langle q^\star \rangle$. The sets $\mathcal{J}_i$ are obtained by applying the procedure to $\mathcal{G} = \mathcal{G}_b(\Sigma_2^b)$ for $i \in \langle b^\star \rangle$ and then the sets $\mathcal{K}_i$ are obtained by applying the procedure to $\mathcal{G} = \mathcal{G}_q(\Sigma_{2,\text{even}}^q)$ for $i \in \langle q^\star \rangle$ (for the obtained results, see Tables II and III).

Using Tables II and III, we are able to construct codes $\mathcal{C}$ of length $n = b+q+1$ in the fashion of (19) as illustrated in the following example.

Example 4.3: Let $b = 6$ and $q = 7$. Then the size of the code $\mathcal{C}$ of length $b+q+1 = 14$ obtained in the vein of (19) is

$$\begin{aligned}|\mathcal{C}| &= 2 \cdot (12{\cdot}15+10{\cdot}12+10{\cdot}11+10{\cdot}10+9{\cdot}7+8{\cdot}6+5{\cdot}3) \\ &= 2 \cdot 636 = 1272.\end{aligned}$$

This code is currently the largest known 1-grain-correcting codes of length 14 (see Table IV). □

Table IV lists the sizes of the 1-grain-correcting codes we were able to obtain due to Theorem 4.2 for $n \in \{10, 11, \ldots, 18\}$. Marked in bold are the sizes of the largest known 1-grain-correcting codes (also to be found in the first row of Table I).

V. GRAIN DETECTION

In [27, Prop. 5.1], we have proved the existence of ∞-grain-detecting codes $\mathcal{C}$ (that is, codes capable of detecting any number of grain errors) of length n over Σ_2 with *redundancy*

$$n - \log_2 |\mathcal{C}| \leq 1.5 \log_2 n + O\left(\tfrac{1}{n}\right)$$

for the overlapping and nonoverlapping scenarios. Employing arguments similar to those used in the proof of Theorem 3.1, we conclude that the size of a largest ∞-grain-detecting code of length n over Σ_2 is bounded from above by $2^{\lceil n/2 \rceil}$ times the size of a largest code of length $\lfloor n/2 \rfloor$ over Σ_2 capable of detecting any number of asymmetric errors, which is known to be $\binom{\lfloor n/2 \rfloor}{\lfloor n/4 \rfloor}$ [28]. Altogether, this implies a lower bound of $\frac{1}{2}\log_2 n + O(1)$ on the *minimum redundancy*

$$r_n \triangleq n - \max_{\substack{\mathcal{C}\subseteq\Sigma_2^n \text{ is an} \\ \infty\text{-grain-detecting code}}} \{\log_2 |\mathcal{C}|\}$$

of ∞-grain-detecting codes of length n when overlaps are allowed or disallowed.

For the overlapping scenario, the upper bound on the size of a largest ∞-grain-detecting code of length n over Σ_2 can be improved by a constant factor (namely, by an additive constant term in the redundancy). In what follows, we will show how to obtain such an upper bound; the proof technique is inspired by the Christmas tree pattern [16, Sec. 7.2.1.6] of arranging 2^n binary strings into chains of subsets.

Define the following (partial) order relation $\preceq$ between two words $\boldsymbol{x}$ and $\boldsymbol{y}$ of the same length over Σ_2: a word $\boldsymbol{x}$ is *dominated* by a word $\boldsymbol{y}$, $\boldsymbol{x} \preceq \boldsymbol{y}$, if there exists a grain pattern $\mathcal{S}$ such that $\sigma_{\mathcal{S}}(\boldsymbol{y}) = \boldsymbol{x}$. Our construction will be iterative where at each step $\ell = 1, 2, 3, \ldots$ we will create s_ℓ new sets $\mathsf{C}_{\ell;j}$ of words of length ℓ for $j \in \langle \mathsf{s}_\ell \rangle$ out of $\mathsf{s}_{\ell-1}$ sets $\mathsf{C}_{\ell-1;j}$ of words of length $\ell-1$ for $j \in \langle \mathsf{s}_{\ell-1} \rangle$. Each set $\mathsf{C}_{\ell;j}$ will be shown (in Theorem 5.4) to be totally ordered with respect to $\preceq$, and the "biggest" and "smallest" words in $\mathsf{C}_{\ell;j}$ will be denoted by $\mathsf{F}(\mathsf{C}_{\ell;j})$ and $\mathsf{f}(\mathsf{C}_{\ell;j})$, respectively. The value of $2\mathsf{s}_n$ will then determine an improved upper bound on the size of a largest ∞-grain-detecting code of length n over Σ_2 when overlaps are allowed, as will be explained in Appendix B.

Construction 5.1: **Basis** ($\ell = 1$). Let $\mathsf{C}_{1;0} = \{0\}$.
Step ($\ell \geq 2$). For $j \in \langle \mathsf{s}_{\ell-1} \rangle$, from a set $\mathsf{C}_{\ell-1;j}$ of size 1, we derive a new set

(C1) $\mathsf{C}_{\ell-1;j} \times \Sigma_2$.

From a set $\mathsf{C}_{\ell-1;j}$ of size at least 2 whose words all end with $a \in \Sigma_2$, we derive two new sets

(C2) $(\mathsf{C}_{\ell-1;j} \times \{\overline{a}\}) \cup \{\mathsf{f}(\mathsf{C}_{\ell-1;j})a\}$,
(C3) $(\mathsf{C}_{\ell-1;j} \times \{a\}) \setminus \{\mathsf{f}(\mathsf{C}_{\ell-1;j})a\}$,

where $\overline{a}$ denotes the *binary complement* of the symbol $a \in \Sigma_2$.

In sets $\mathsf{C}_{\ell-1;j}$ of size at least 2 whose words not all end with the same symbol, this construction will guarantee the existence of only one word $\boldsymbol{c} \in \mathsf{C}_{\ell-1;j}$ whose last symbol $\overline{a}$ differs from that of $\mathsf{F}(\mathsf{C}_{\ell-1;j})$. For a set $\mathsf{C}_{\ell-1;j}$ of this kind, we derive two new sets

(C4) $(\mathsf{C}_{\ell-1;j} \times \{\overline{a}\}) \cup \{\boldsymbol{c}a\}$,
(C5) $(\mathsf{C}_{\ell-1;j} \times \{a\}) \setminus \{\boldsymbol{c}a\}$. □

Remark 5.2: Notice that in cases (C3) and (C5) of Construction 5.1, we create new sets whose words all end with the same symbol, whereas in cases (C1), (C2) and (C4), the newly created sets $\mathsf{C}_{\ell;j}$ include only one word whose last symbol differs from that of $\mathsf{F}(\mathsf{C}_{\ell;j})$. Therefore these are the only two types of sets with which Construction 5.1 operates. □

Example 5.3: The first four rounds of Construction 5.1 yield $\mathsf{C}_{1;0} = \{0\}$, $\mathsf{C}_{2;0} = \{00, 01\}$, $\mathsf{C}_{3;0} = \{000, 001, 010\}$, $\mathsf{C}_{3;1} = \{011\}$, $\mathsf{C}_{4;0} = \{0001, 0011, 0010, 0101\}$, $\mathsf{C}_{4;1} = \{0000, 0100\}$, $\mathsf{C}_{4;2} = \{0111, 0110\}$. □

We have reached the main theorem of this section (with proof in Appendix B).

Theorem 5.4: For any positive integer n and any $j \in \langle \mathsf{s}_n \rangle$, the set $\mathsf{C}_{n;j}$ is totally ordered with respect to $\preceq$. Moreover, $\mathsf{s}_n = \binom{n-1}{\lfloor (n-1)/2 \rfloor}$ for any positive integer n, thereby implying the upper bound of $2\binom{n-1}{\lfloor (n-1)/2 \rfloor}$ on the size of a largest ∞-grain-detecting code of length n over Σ_2 (with overlaps allowed). □

Since $\lim_{n\to\infty} 2^{\lceil n/2 \rceil}\binom{\lfloor n/2 \rfloor}{\lfloor n/4 \rfloor}/2\binom{n-1}{\lfloor (n-1)/2 \rfloor} = \sqrt{2}$, for large values of n, the upper bound on the size of ∞-grain-detecting codes of length n over Σ_2 (with overlaps allowed) due to Theorem 5.4 is $\approx\sqrt{2}$ times smaller than the upper bound $2^{\lceil n/2 \rceil}\binom{\lfloor n/2 \rfloor}{\lfloor n/4 \rfloor}$ on the size of ∞-grain-detecting code of length n over Σ_2 that can be obtained from Theorem 3.1 (see the discussion at the beginning of this section).

TABLE V
SIZES OF LARGEST t-GRAIN-DETECTING CODES OF LENGTH n WHEN OVERLAPS ARE DISALLOWED.

t \ n	2	3	4	5	6	7	8
1	2	4	8	16	32	64	128
2			8	10	18	34	58
3					18	32	56
4							56

TABLE VI
SIZES OF LARGEST t-GRAIN-DETECTING CODES OF LENGTH n WHEN OVERLAPS ARE ALLOWED.

t \ n	2	3	4	5	6	7	8
1	2	4	8	16	32	64	128
2		4	6	10	18	30	52
3			6	8	12	22	42
4				8	12	20	32
5					12	20	32

Tables V and VI list the sizes of the largest t-grain-detecting codes of length n over Σ_2 when overlaps are disallowed and allowed, respectively, for small values of n and t, found using a computer search.[5] It can be seen that already for length $n = 5$, there is a gap between the upper bound of $2\binom{4}{2} = 12$ on the size of ∞-grain-detecting codes of length 5 when overlaps are allowed due to Construction 5.1 and the size 8 of a largest ∞-grain-detecting code. However, using *ad hoc* arguments, it is still possible to partition the 16 words in $0\Sigma_2^4$ into the four sets

$$\mathsf{C}_{5;0} = \{00000, 00100, 01000, 01001\},$$
$$\mathsf{C}_{5;1} = \{00001, 00011, 00010, 00101\},$$
$$\mathsf{C}_{5;2} = \{00110, 01110, 01100, 01010\},$$
$$\mathsf{C}_{5;3} = \{00111, 01111, 01101, 01011\}$$

of size 4, which are totally ordered with respect to $\preceq$. This, in turn, results in a tight upper bound of 8 on the size of ∞-grain-detecting codes of length 5 when overlaps are allowed.

On the other hand, using a computer search, one can establish that for $n = 6$, the smallest number of totally ordered sets $\mathsf{C}_{6;j}$ required to partition $0\Sigma_2^5$ is 7, which results in the

[5] The entries for $t = 1$ in both tables follow from the simple observation that the Hamming distance between two distinct codewords that start with the same symbol in a binary 1-grain-detecting code must be at least 2 and that a binary parity code of any length is 1-grain-detecting.

upper bound of 14 on the size of a largest ∞-grain-detecting code of length 6 with overlaps; this bound is strictly greater than the size 12 of a largest such code. One such partition is given by

$$\begin{aligned}
\mathsf{C}_{6;0} &= \{000000, 000001, 000010, 000101, 001010\}\,,\\
\mathsf{C}_{6;1} &= \{000110, 000100, 001100, 001101, 001010\}\,,\\
\mathsf{C}_{6;2} &= \{000011, 000111, 001011, 010111, 010101\}\,,\\
\mathsf{C}_{6;3} &= \{001000, 011000, 010000, 010001, 010010\}\,,\\
\mathsf{C}_{6;4} &= \{001001, 011001, 011011, 010011\}\,,\\
\mathsf{C}_{6;5} &= \{001111, 011111, 011110, 011101\}\,,\\
\mathsf{C}_{6;6} &= \{001110, 011100, 011010, 010100\}\,.
\end{aligned}$$

Similar phenomena occur when overlaps are disallowed: for $n = 5$ it is possible to partition $0\Sigma_2^4$ into 5 totally ordered sets using *ad hoc* arguments, yet for $n = 6$ it is provably impossible to partition $0\Sigma_2^5$ into 9 totally ordered sets.

APPENDICES

A. PROOF OF LEMMA 2.3

Define the following "almost complete" directed graph $\mathcal{G} = (V, E)$. Its set of states is defined as $V = \Sigma_3^2$, whereas its set of edges is

$$E = (V \times V) \setminus \{(\mathit{01\ 10}, \mathit{10\ 01}), (\mathit{10\ 01}, \mathit{01\ 10})\}\,.$$

Define the subset of states

$$V_0 = \{\mathit{00\ 00}, \mathit{00\ 01}, \mathit{01\ 00}, \mathit{01\ 01}, \mathit{10\ 10}\}$$

as the set of *safe states*. Traversing a path $\gamma = (\ell_i r_i)_{i\in\langle n/2\rangle}$ of length[6] $\frac{1}{2}n-1$ in the graph $\mathcal{G}$ produces a pair of words $\boldsymbol{\ell} = (\ell_i)_{i\in\langle n/2\rangle}$ and $\boldsymbol{r} = (r_i)_{i\in\langle n/2\rangle}$ over Σ_3.

Remark A.1: Notice that the pair of edges

$$(\mathit{01\ 10}, \mathit{10\ 01}) \text{ and } (\mathit{10\ 01}, \mathit{01\ 10}),$$

which make up the difference between $\mathcal{G}$ and the complete directed graph on $|V| = 9$ states, correspond to the pair of subwords $\mathit{01\ 10}$ and $\mathit{10\ 01}$. For any pair of ternary words $\boldsymbol{x} = (x_i)_{i\in\langle n/2\rangle}, \boldsymbol{x}' = (x'_i)_{i\in\langle n/2\rangle} \in \Sigma_3^{n/2}$ and $j \in \langle n/2-1\rangle$ such that $x_j x_{j+1} = \mathit{01\ 10}$ and $x'_j x'_{j+1} = \mathit{10\ 01}$, one has that $\mathcal{E}^{-1}(\boldsymbol{x})$ and $\mathcal{E}^{-1}(\boldsymbol{x}')$ are non-confusable. Therefore, we did not include these two edges in our graph, as we are interested in counting finitely-cws pairs of clouds. □

The following lemma (with proof very similar to [27, App. A]) establishes a correspondence between pairs of clouds and paths in the graph $\mathcal{G}$.

Lemma A.2: Let t be a positive integer, $t \le n-1$. Let $\mathcal{W}_t$ denote the set of all (ordered) pairs $(\boldsymbol{x}, \boldsymbol{x}') \in \Sigma_3^{n/2} \times \Sigma_3^{n/2}$ of ternary words whose clouds are t-cws and let Π_t be the following set of paths (of length $\frac{1}{2}n-1$) in $\mathcal{G}$:

$$\Pi_t = \Big\{(\ell_i r_i)_{i\in\langle n/2\rangle} : (\ell_0 r_0) \in V_0,\ \mathsf{d}(\boldsymbol{\ell}, \boldsymbol{r}) \le 2t\Big\},$$

where $\boldsymbol{\ell} = (\ell_i)_{i\in\langle n/2\rangle}$ and $\boldsymbol{r} = (r_i)_{i\in\langle n/2\rangle}$, and $\mathsf{d}(\boldsymbol{\ell}, \boldsymbol{r})$ denotes the Hamming distance between $\boldsymbol{\ell}$ and $\boldsymbol{r}$ when $\boldsymbol{\ell}$ and $\boldsymbol{r}$ are viewed as binary words of length n. Then there exists a one-to-one mapping from $\mathcal{W}_t$ to Π_t that maps pairs of words $(\boldsymbol{x}, \boldsymbol{x}') = ((x_i)_{i\in\langle n/2\rangle}, (x'_i)_{i\in\langle n/2\rangle})$ whose clouds are t-cws to paths $(x_i x'_i)_{i\in\langle n/2\rangle}$. □

For an edge $e = (\ell r, \ell' r')$, define the function $\varphi : E \to \langle 3\rangle^3$ by $\varphi(e) = (\nu(e)\ \omega(e)\ \chi(e))$, where $\nu(e) = \mathsf{d}(\ell', r')$ (with ℓ' and r' being viewed as binary words of length 2),

$$\omega(e) = \begin{cases} 2 & \ell\ell' = \mathit{00\ 00} \text{ and } rr' = \mathit{00\ 00} \\ 1 & \ell\ell' = \mathit{00\ 00} \text{ and } rr' \ne \mathit{00\ 00} \text{ or} \\ & \ell\ell' \ne \mathit{00\ 00} \text{ and } rr' = \mathit{00\ 00} \\ 0 & \text{otherwise} \end{cases},$$

and

$$\chi(e) = \begin{cases} 2 & \ell\ell', rr' \in \{\mathit{01}, \mathit{10}\}^2 \\ 1 & \ell\ell' \in \{\mathit{01}, \mathit{10}\}^2 \text{ and } rr' \notin \{\mathit{01}, \mathit{10}\}^2 \text{ or} \\ & \ell\ell' \notin \{\mathit{01}, \mathit{10}\}^2 \text{ and } rr' \in \{\mathit{01}, \mathit{10}\}^2 \\ 0 & \text{otherwise} \end{cases}.$$

The function $\nu(e)$ counts the smallest number of (possibly, overlapping) grains making ℓ' and r' cws (when viewed as words of length 2 over Σ_2); the function $\omega(e)$ counts the number of transitions from $\mathit{00}$ to $\mathit{00}$ in the words $\ell\ell'$ and rr' (viewed as words of length 2 over Σ_3); the function $\chi(e)$ counts the number of transitions from either $\mathit{01}$ or $\mathit{10}$ to either $\mathit{01}$ or $\mathit{10}$ in the words $\ell\ell'$ and rr' (again, viewed as words of length 2 over Σ_3).

Define Γ as the set of all the cycles in $\mathcal{G}$ of length $\frac{1}{2}n$ that start and terminate in the same state of V_0. Now, set $\tau \in (0, 1)$, set $p_0, p_1 \in (0, 1)$ such that $p_0 + p_1 < 1$, let $\epsilon > 0$ and define

$$U_{\tau,p_0,p_1,\epsilon} = \{(u_1\ u_2\ u_3) : -\epsilon < u_1 < 4\tau + \epsilon,\ |u_2 - 2p_0|, |u_3 - 2p_1| < 2\epsilon\}\,,$$

and

$$\begin{aligned}\mathit{\Gamma}_{\tau,p_0,p_1,\epsilon} &= \mathit{\Gamma}_{\tau,p_0,p_1,\epsilon}(n) \\ &= \big\{\gamma = (v_i)_{i\in\langle n/2\rangle} \in \Gamma : \mathsf{E}_{\mathsf{P}_\gamma}\{\varphi\} \in U_{\tau,p_0,p_1,\epsilon}\big\}\,,\end{aligned}$$

where $\mathsf{E}_{\mathsf{P}_\gamma}\{\varphi\} = \sum_{e\in E} \mathsf{P}_\gamma(e)\varphi(e)$ is the expected value of φ with respect to the empirical probability distribution

$$\mathsf{P}_\gamma(e) = \frac{2}{n}\,|i \in \langle n/2\rangle : (v_i, v_{i+1}) = e|\,.$$

The set $\mathit{\Gamma}_{\tau,p_0,p_1,\epsilon}$ stands for all the cycles of length $\frac{1}{2}n$ in $\mathcal{G}$ starting in a safe state that represent pairs of ternary words $(\boldsymbol{x}, \boldsymbol{x}')$ of length $\frac{1}{2}n$ at Hamming distance at most $(4\tau+\epsilon)\frac{n}{2}$ from one another (when viewed as binary words of length n), whose total number of transitions from $\mathit{00}$ to $\mathit{00}$ is within $(p_0 \pm \epsilon)n$ and whose total number of transitions from either $\mathit{01}$ or $\mathit{10}$ to either $\mathit{01}$ or $\mathit{10}$ is within $(p_1 \pm \epsilon)n$. Also, for the same τ, p_0, p_1, ϵ, define

$$\begin{aligned}\mathit{\Pi}_{\tau,p_0,p_1,\epsilon} = \mathit{\Pi}_{\tau,p_0,p_1,\epsilon}(n) = \big\{\gamma \in \Pi_{\lceil 2\tau(n/2-1)\rceil} :\ & \\ \big|\mathsf{E}_{\mathsf{P}_\gamma}\{\omega\} - 2p_0\big| \le \epsilon, \big|\mathsf{E}_{\mathsf{P}_\gamma}\{\chi\} - 2p_1\big| \le \epsilon\big\}\,.&\end{aligned}$$

The set $\mathit{\Pi}_{\tau,p_0,p_1,\epsilon}$ contains paths of length $\frac{1}{2}n-1$ in $\mathcal{G}$ that represents pairs of ternary words $(\boldsymbol{x}, \boldsymbol{x}')$ of length n, whose clouds are $\lceil 2\tau(n/2-1)\rceil$-cws, whose total number of transitions from $\mathit{00}$ to $\mathit{00}$ is within $(2p_0 \pm \epsilon)(\frac{1}{2}n-1)$ and whose total

[6] The length of a path γ in the graph equals the number of edges along γ; since $\mathcal{G}$ has no parallel edges, we will specify paths in $\mathcal{G}$ as sequences of states.

number of transitions from either *01* or *10* to either *01* or *10* is within $(2p_1 \pm \epsilon)(\frac{1}{2}n - 1)$.

The following lemma claims that there exist at least as many cycles in $\Gamma_{\tau,p_0,p_1,\epsilon}$ as paths in $\Pi_{\tau,p_0,p_1,\epsilon}$ (for a similar proof, see [27, Lemma 2.11]).

Lemma A.3: Let $\tau \in (0,1)$, let $\epsilon > 0$, and let $p_0, p_1 \in (0,1)$ such that $p_0 + p_1 < 1$. Then, for $n \geq 4/\epsilon$,

$$|\Pi_{\tau,p_0,p_1,\epsilon}(n)| \leq |\Gamma_{\tau,p_0,p_1,\epsilon}(n)| \ .$$

□

In the proof of Lemma 2.3, we use special cases of [17, Lemma 2] and [17, Lemma 5] which we cite below (in Lemmas A.4 and A.5) and which will aid us in establishing the connection between the number of cycles $\Gamma_{\tau,p_0,p_1,\epsilon}$ and an optimization of a convex function subject to linear equality and inequality constraints (we also refer the reader to [5, Lemma 2], [19, pp. 312–316], [23, Ch. 2, Th. 25], and [25, Sec. 28]). In both lemmas, $\mathcal{M}_G(f;U)$ denotes the set of all stationary Markov chains M on a graph $G = (V_G, E_G)$ such that $\mathsf{E}_{\mathsf{M}}\{f\} \in U \subseteq \mathbb{R}^k$, for a positive integer k and a given function $f : E_G \to \mathbb{R}^k$.

Lemma A.4: Let $G = (V_G, E_G)$ be a primitive[7] directed graph and $f : E_G \to \mathbb{R}^k$ be a function. Let U be an open rectangular parallelepiped $\prod_{i\in\langle k\rangle}(\widetilde{s}_i, s_i)$ and let Γ_n denote the set of all cycles of length n in G. Then

$$\lim_{n\to\infty} \frac{1}{n}\log_2 \left|\left\{\gamma \in \Gamma_n : \mathsf{E}_{\mathsf{P}_\gamma}\{f\} \in U\right\}\right| = \sup_{\mathsf{M}\in\mathcal{M}_G(f;U)} \mathtt{h}(\mathsf{M}),$$

where

$$\mathtt{h}(\mathsf{M}) = -\sum_{\substack{v\in V_G:\\ \pi(v)>0}} \ \sum_{\substack{v':e=(v,v')\in E_G\\ \text{s.t. } \mathsf{M}(e)>0}} \mathsf{M}(e)\log_2\frac{\mathsf{M}(e)}{\pi(v)}$$

is the binary entropy of a stationary Markov chain M and

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

is the stationary probability to be in a state $v \in V_G$ along a random walk on G. □

Let k be a positive integer, $G = (V_G, E_G)$ be a directed graph, $\boldsymbol{z} = (z_i)_{i\in\langle k\rangle}$ be a vector of positive real indeterminates and $f = (f_i)_{i\in\langle k\rangle} : E_G \to \mathbb{R}^k$ be a function. Define the parametric matrix $\mathsf{A}_G(\boldsymbol{z})$ over $\mathbb{R}$ (with rows and columns indexed by the states of V_G) as

$$[\mathsf{A}_G(\boldsymbol{z})]_{v,v'\in V_G} = \begin{cases} \boldsymbol{z}^{f(e)} = \prod_{i\in\langle k\rangle} z_i^{f_i(e)} & \text{if } e=(v,v')\in E_G \\ 0 & \text{otherwise} \end{cases} . \tag{21}$$

Lemma A.5: 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

$$\sup_{\mathsf{M}\in\mathcal{M}_G(f;U):\ \mathsf{E}_P\{f'\}=\boldsymbol{p}} \mathtt{h}(\mathsf{M}) = \inf_{\boldsymbol{z},\boldsymbol{h}}\Big\{\log_2\lambda(\mathsf{A}_G(\boldsymbol{z},\boldsymbol{h})) - \sum_{i\in\langle k\rangle} s_i\log_2 z_i - \sum_{i\in\langle k'\rangle} p_i\log_2 h_i\Big\},$$

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 $(0,\infty)^{k'}$. □

Now we are in a position to prove Lemma 2.3.

Proof of Lemma 2.3: We will apply Lemmas A.4 and A.5 to our graph $\mathcal{G} = (V, E)$ with $f = \varphi = (\nu\ \omega\ \chi)$. Specifically, let $z \in (0,1]$ and $h, m \in (0,\infty)$ be indeterminates, and define the matrix $\mathsf{A}_{\mathcal{G}}(z,h,m)$ indexed by the states of $\mathcal{G}$ as a particular case of (21):

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

Apply Lemma A.4 to the case where $G = \mathcal{G}$, $U = U_{\tau,p_0,p_1,\epsilon}$, and $f = \varphi$, and combine it with the result of Lemma A.3 to obtain

$$\overline{\lim_{n\to\infty}}\ \frac{2}{n}\log_2|\Pi_{\tau,p_0,p_1,\epsilon}(n)| \leq \sup_{\mathsf{M}\in\mathcal{M}_{\mathcal{G}}(\varphi;U_{\tau,p_0,p_1,\epsilon})} \mathtt{h}(\mathsf{M}) \ .$$

By the continuity of the functions $\mathsf{M} \mapsto \mathsf{E}_{\mathsf{M}}(\varphi)$ and $\mathsf{M} \mapsto \mathtt{h}(\mathsf{M})$,

$$\overline{\lim_{\epsilon\to 0}}\ \overline{\lim_{n\to\infty}}\ \frac{2}{n}\log_2|\Pi_{\tau,p_0,p_1,\epsilon}(n)| \leq \sup_{\mathsf{M}\in\mathcal{M}_{\mathcal{G}}(\varphi;U_{\tau,p_0,p_1})} \mathtt{h}(\mathsf{M}) \ ,$$

where

$$U_{\tau,p_0,p_1} = \{(u\ 2p_0\ 2p_1) : u \in [0,4\tau]\} \ .$$

Applying Lemma A.5 with $G = \mathcal{G}$, $f = \nu$, $f' = (\omega\ \chi)$, $U = [0,4\tau]$, and $\boldsymbol{p} = (2p_0\ 2p_1)$ yields

$$\begin{aligned} &\overline{\lim_{\epsilon\to 0}}\ \overline{\lim_{n\to\infty}}\ \frac{2}{n}\log_2|\Pi_{\tau,p_0,p_1,\epsilon}(n)| \\ &\leq \inf_{\substack{z\in(0,1],\\ h,m\in(0,\infty)}} \Big\{\log_2\lambda(\mathsf{A}_{\mathcal{G}}(z,h,m)) \\ &\quad - 4\tau\log_2 z - 2p_0\log_2 h - 2p_1\log_2 m\Big\} \ . \end{aligned} \tag{22}$$

It follows from Lemma A.2 that

$$\begin{aligned} &\overline{\lim_{\epsilon\to 0}}\ \overline{\lim_{n\to\infty}}\ \frac{2}{n}\log_2\mathsf{W}_{\lceil 2\tau(n/2-1)\rceil}(\mathcal{J}_{p_0,p_1,\epsilon}(n)) \\ &\quad \leq \overline{\lim_{\epsilon\to 0}}\ \overline{\lim_{n\to\infty}}\ \frac{2}{n}\log_2|\Pi_{\tau,p_0,p_1,\epsilon}(n)|. \end{aligned} \tag{23}$$

Employing the Moore algorithm [21, Sec. 2.6], we merge the states of $\mathcal{G}$ to reduce the order of the matrix $\mathsf{A}_{\mathcal{G}}(z,h,m)$ all the while keeping its spectral radius intact. Specifically, the states in $\{$*00 01*, *10 00*, *00 10*, *01 00*$\}$ can be merged into superstate 1, the states in $\{$*01 01*, *10 10*$\}$ — into superstate 2, the states in $\{$*01 10*, *10 01*$\}$ — into superstate 3 (and state *00 00* can be renamed to superstate 0). The resulting reduced matrix $\mathcal{A}_{\mathcal{G}}(z,h,m)$ appears in (7). Plugging $\mathcal{A}_{\mathcal{G}}(z,h,m)$ instead of $\mathsf{A}_{\mathcal{G}}(z,h,m)$ in (22) and combining the obtained result with (23) yield (8). □

[7] A directed graph is *primitive* if it is strongly connected and the greatest common divisor of the lengths of its cycles is 1.

B. Proof of Theorem 5.4

The result of Theorem 5.4 follows by combining the results of Lemmas B.1, B.2, and Corollary B.3 below. Lemma B.1, coming next, demonstrates by induction on ℓ that each set $\mathsf{C}_{\ell;j}$ is totally ordered with respect to $\preceq$ which, in turn, justifies the use of the operators $\mathsf{f}(\cdot)$ and $\mathsf{F}(\cdot)$ in Construction 5.1.

Lemma B.1: For any positive integer ℓ and any $j \in \langle \mathsf{s}_\ell \rangle$, the set $\mathsf{C}_{\ell;j}$ is totally ordered with respect to $\preceq$.

Proof: Readily, the set $\mathsf{C}_{1;0} = \{0\}$ is totally ordered, which is the basis of our induction proof. As for the induction step, let us assume that each one of the sets $\mathsf{C}_{\ell-1;j}$ is totally ordered for every $j \in \langle \mathsf{s}_{\ell-1} \rangle$. To prove the statement of the lemma, it will suffice to take two words $\boldsymbol{x}, \boldsymbol{y} \in \mathsf{C}_{\ell-1;j}$ such that $\boldsymbol{x} \preceq \boldsymbol{y}$ and show the order between all the words in $\mathsf{C}_{\ell;j'}$ whose prefixes of length $\ell-1$ are $\boldsymbol{x}$ and $\boldsymbol{y}$, for each one of the cases (C1)–(C5) in Construction 5.1.

- (C1) In this case, $\boldsymbol{x} = \boldsymbol{y}$. When $\boldsymbol{x}$ ends with a 0, the order between $\boldsymbol{x}0$ and $\boldsymbol{x}1$ is $\boldsymbol{x}0 \preceq \boldsymbol{x}1$, whereas when $\boldsymbol{x}$ ends with a 1, the order is $\boldsymbol{x}1 \preceq \boldsymbol{x}0$.
- (C2) When $\boldsymbol{x} \neq \mathsf{f}(\mathsf{C}_{i-1;j})$, the order between $\boldsymbol{x}\bar{a}$ and $\boldsymbol{y}\bar{a}$ is $\boldsymbol{x}\bar{a} \preceq \boldsymbol{y}\bar{a}$; when $\boldsymbol{x} = \mathsf{f}(\mathsf{C}_{i-1;j})$, the order between $\boldsymbol{x}a$, $\boldsymbol{x}\bar{a}$, and $\boldsymbol{y}\bar{a}$ is $\boldsymbol{x}a \preceq \boldsymbol{x}\bar{a}$, $\boldsymbol{x}\bar{a} \preceq \boldsymbol{y}\bar{a}$, and $\boldsymbol{x}a \preceq \boldsymbol{y}\bar{a}$.
- (C3) The order between $\boldsymbol{x}a$ and $\boldsymbol{y}a$ is $\boldsymbol{x}a \preceq \boldsymbol{y}a$.
- (C4) When $\boldsymbol{x}, \boldsymbol{y} \neq \boldsymbol{c}$, the order between $\boldsymbol{x}\bar{a}$ and $\boldsymbol{y}\bar{a}$ is $\boldsymbol{x}\bar{a} \preceq \boldsymbol{y}\bar{a}$; when $\boldsymbol{x} = \boldsymbol{c}$, the order between $\boldsymbol{x}a$, $\boldsymbol{x}\bar{a}$, and $\boldsymbol{y}\bar{a}$ is $\boldsymbol{x}\bar{a} \preceq \boldsymbol{x}a$, $\boldsymbol{x}a \preceq \boldsymbol{y}\bar{a}$, and $\boldsymbol{x}\bar{a} \preceq \boldsymbol{y}\bar{a}$; when $\boldsymbol{y} = \boldsymbol{c}$, the order between $\boldsymbol{x}\bar{a}$, $\boldsymbol{y}a$, and $\boldsymbol{y}\bar{a}$ is $\boldsymbol{x}\bar{a} \preceq \boldsymbol{y}a$, $\boldsymbol{y}\bar{a} \preceq \boldsymbol{y}a$, and $\boldsymbol{x}\bar{a} \preceq \boldsymbol{y}\bar{a}$.
- (C5) The order between $\boldsymbol{x}a$ and $\boldsymbol{y}a$ is $\boldsymbol{x}a \preceq \boldsymbol{y}a$. □

In light of Lemma B.1 and by the simple observation that $\{\mathsf{C}_{n;j} : j \in \langle \mathsf{s}_n \rangle\}$ is a partition of $0\Sigma_2^{n-1}$ for a positive integer n, each set $\mathsf{C}_{n;j}$ for $j \in \langle \mathsf{s}_n \rangle$ can contribute at most one word to an ∞-grain-detecting code of length n. Therefore, by extending the above argument to all the words of length n that start with a 1, we obtain an upper bound of $2\mathsf{s}_n$ on the size of a largest ∞-grain-detecting code of length n. It is left to find the value of s_n; we will show that it equals the number $\binom{n-1}{\lfloor (n-1)/2 \rfloor}$ of walks of length $n-1$ on the square lattice from the origin $(0,0)$ by moving down or moving right, all the while staying on the points (x, y) satisfying $\mathsf{x}+\mathsf{y} \geq 0$ [10, Sec. 2]. This observation gives rise to the following lemma.

Lemma B.2: Let ℓ be a positive integer and let $\boldsymbol{x} = (x_i)_{i \in \langle \ell \rangle}$ be a word of length ℓ over Σ_2. For a positive integer $k \in \langle \ell \rangle \setminus \{0\}$, let

$$\begin{aligned}\mathsf{p}_k(\boldsymbol{x}) &= 2\,|\{s \in \langle k \rangle : x_s \neq x_{s+1}\}| - k \\ &= \ |\{s \in \langle k \rangle : x_s \neq x_{s+1}\}| - |\{s \in \langle k \rangle : x_s = x_{s+1}\}|\end{aligned}$$

be the difference between the number of symbol alternations and the number of symbol repetitions in the prefix of length $k+1$ of $\boldsymbol{x}$. Then for any $j \in \langle \mathsf{s}_\ell \rangle$, the only word $\boldsymbol{x}$ in $\mathsf{C}_{\ell;j}$ which satisfies $\mathsf{p}_k(\boldsymbol{x}) \geq 0$ for all $k \in \langle \ell \rangle \setminus \{0\}$ is $\mathsf{F}(\mathsf{C}_{\ell;j})$.

Proof: By induction on ℓ, one can readily see that for any positive integer $\ell \geq 2$ and any $j \in \langle \mathsf{s}_\ell \rangle$, one has

$$|\mathsf{C}_{\ell;j}| = \mathsf{p}_{\ell-1}(\mathsf{F}(\mathsf{C}_{\ell;j}))+1 \ . \qquad (24)$$

We will prove the claim of the lemma by induction on ℓ. Clearly, the claim holds for $\ell = 2$, namely, the only word $\boldsymbol{x}$ in $\mathsf{C}_{2;0}$ that satisfies $\mathsf{p}_1(\boldsymbol{x}) \geq 0$ is $\mathsf{F}(\mathsf{C}_{2;0}) = 01$. As for the induction step, let us assume that for $\ell \geq 3$, the only word $\boldsymbol{x}$ in each one of the sets $\mathsf{C}_{\ell-1;j}$ which satisfies $\mathsf{p}_k(\boldsymbol{x}) \geq 0$ for all $k \in \langle \ell-1 \rangle \setminus \{0\}$ is $\mathsf{F}(\mathsf{C}_{\ell-1;j})$. To prove the claim of the lemma, it will suffice to take $\boldsymbol{x} = \mathsf{F}(\mathsf{C}_{\ell-1;j})$ and, for each one of the cases (C1)–(C5), show that the word $\boldsymbol{y}$ in $\mathsf{C}_{\ell;j'}$, whose prefix of length $\ell-1$ is $\boldsymbol{x}$, satisfies $\mathsf{p}_k(\boldsymbol{y}) \geq 0$ for all $k \in \langle \ell \rangle \setminus \{0\}$ if and only if $\boldsymbol{y} = \mathsf{F}(\mathsf{C}_{\ell;j'})$.

- (C1) Without loss of generality, $\boldsymbol{x}$ ends with a 0 and $\boldsymbol{x}1 = \mathsf{F}(\mathsf{C}_{\ell;j'})$. Since $\mathsf{p}_{\ell-1}(\boldsymbol{x}1) = 1 + \mathsf{p}_{\ell-2}(\boldsymbol{x}) \geq 1$ by the induction hypothesis and $\mathsf{p}_k(\boldsymbol{x}1) = \mathsf{p}_k(\boldsymbol{x})$ for $k \in \langle \ell-1 \rangle \setminus \{0\}$, the word $\boldsymbol{x}1$ satisfies $\mathsf{p}_k(\boldsymbol{x}1) \geq 0$ for all $k \in \langle \ell \rangle \setminus \{0\}$. Moreover, by (24), $\mathsf{p}_{\ell-2}(\boldsymbol{x}) = 0$, therefore $\mathsf{p}_{\ell-1}(\boldsymbol{x}0) = -1$ implying that $\boldsymbol{x}1$ is the only word $\boldsymbol{y}$ in $\mathsf{C}_{\ell;j'}$ satisfying $\mathsf{p}_k(\boldsymbol{y}) \geq 0$ for all $k \in \langle \ell \rangle \setminus \{0\}$.
- (C2),(C4) In these cases, the only word in $\mathsf{C}_{\ell;j'}$ whose prefix is $\boldsymbol{x}$ is $\boldsymbol{x}\bar{a}$. Since $\boldsymbol{x}$ ends with a, by the induction hypothesis one has $\mathsf{p}_{\ell-1}(\boldsymbol{x}\bar{a}) = \mathsf{p}_{\ell-2}(\boldsymbol{x})+1 \geq 1$, so the only word $\boldsymbol{y}$ in $\mathsf{C}_{\ell;j'}$ satisfying $\mathsf{p}_k(\boldsymbol{y}) \geq 0$ for all $k \in \langle \ell \rangle \setminus \{0\}$ is $\boldsymbol{x}\bar{a}$.
- (C3),(C5) In these cases, the only word in $\mathsf{C}_{\ell;j'}$ whose prefix is $\boldsymbol{x}$ is $\boldsymbol{x}a$. Since $\boldsymbol{x}$ ends with a, by the induction hypothesis and by (24), one has $\mathsf{p}_{\ell-1}(\boldsymbol{x}a) = \mathsf{p}_{\ell-2}(\boldsymbol{x})-1 \geq 0$, hence the only word $\boldsymbol{y}$ in $\mathsf{C}_{\ell;j'}$ satisfying $\mathsf{p}_k(\boldsymbol{y}) \geq 0$ for all $k \in \langle \ell \rangle \setminus \{0\}$ is $\boldsymbol{x}a$.

Corollary B.3: Let ℓ be a nonnegative integer. Then

$$\mathsf{s}_\ell = \binom{\ell-1}{\lfloor (\ell-1)/2 \rfloor} .$$

Proof: Due to the result of Lemma B.2 and the observation that $\{\mathsf{C}_{\ell;j} : j \in \langle \mathsf{s}_\ell \rangle\}$ is a partition of $0\Sigma_2^{\ell-1}$, instead of counting different sets $\mathsf{C}_{\ell;j}$, we can count the number of "biggest" words $\boldsymbol{x} = (x_i)_{i \in \langle \ell \rangle} \in 0\Sigma_2^{\ell-1}$ which satisfy $|\mathsf{p}_k(\boldsymbol{x})| \geq 0$ for all $k \in \langle \ell \rangle \setminus \{0\}$. Now, there is a natural 1-to-1 correspondence between such words and walks of length $\ell-1$ on the square lattice from the origin $(0,0)$ by moving down or moving right, all the while staying on the points (x, y) satisfying $\mathsf{x}+\mathsf{y} \geq 0$, specifically, we move right at step k of that walk if $x_{k-1} \neq x_k$ and move down otherwise. The number of such walks is, in turn, $\binom{\ell-1}{\lfloor (\ell-1)/2 \rfloor}$. □

Acknowledgement

The authors would like to thank the anonymous referees for their comments and suggestions.

References

[1] S. Al-Bassam, R. Venkatesan, S. Al-Muhammadi, "New single asymmetric error-correcting codes," *IEEE Trans. Inform. Theory,* vol. 43, pp. 1619–1623, 1997.

[2] A.E. Brower, J.B. Shearer, N.J.A. Sloane, W.D. Smith, "A new table of constant weight codes," *IEEE Trans. Inform. Theory,* vol. 36, pp. 1334–1380, 1990.

[3] S.D. Constantin, T.R.N. Rao "On the theory of binary asymmetric error correcting codes," *Inform. Contr.,* vol. 40, pp. 20–36, 1979.

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

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

[6] D. Cullina, N. Kiyavash, "Generalized sphere-packing upper bounds on the size of codes for combinatorial channels," *Proc. IEEE Int'l Symp. Inform. Theory (ISIT 2014),* Honolulu, HI, 2014, pp. 1266–1270.

[7] P. Delsarte, P. Piret, "Bounds and constructions for binary asymmetric error-correcting codes," *IEEE Trans. Inform. Theory,* vol. 27, pp. 125–128, 1981.
[8] A. Fazeli, A. Vardy, E. Yaakobi, "Generalized sphere packing bound: basic principles," *Proc. IEEE Int'l Symp. Inform. Theory (ISIT 2014),* Honolulu, HI, 2014, pp. 1256–1260.
[9] A. Fazeli, A. Vardy, E. Yaakobi, "Generalized sphere packing bound: applications," *Proc. IEEE Int'l Symp. Inform. Theory (ISIT 2014),* Honolulu, HI, 2014, pp. 1261–1265.
[10] L. Ferrari, "Some combinatorics related to central binomial coefficients: Grand-Dyck paths, coloured noncrossing partitions and signed pattern avoiding permutations," *Graphs and Combinat.*, vol. 26, pp. 51–70, 2010.
[11] R. Gabrys, E. Yaakobi, L. Dolecek, "Correcting grain-errors in magnetic media," *IEEE Trans. Inform. Theory,* vol. 61, pp. 2256–2272, 2015.
[12] S. Greaves, Y. Kanai, H. Muraoka, "Shingled magnetic recording on bit patterned media," *IEEE Trans. Magn.,* vol. 46, pp. 1460–1463, 2010.
[13] 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.
[14] N. Kashyap, G. Zémor, "Upper bounds on the size of grain-correcting codes," *IEEE Trans. Inform. Theory,* vol. 60, pp. 4699–4709, 2014.
[15] T. Kløve, "Upper bounds on codes correcting asymmetric errors," *IEEE Trans. Inform. Theory,* vol. 27, pp. 128–131, 1981.
[16] D.E. Knuth, *The Art of Computer Programming, Vol. 4A, Combinatorial Algorithms, Part 1,* 1st edition, Addison–Wesley, Reading, MA, 2011.
[17] 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.
[18] A.A. Kulkarni, N. Kiyavash, "Nonasymptotic upper bounds for deletion correcting codes," *IEEE Trans. Inform. Theory,* vol. 59, pp. 5115–5130, 2013.
[19] D.G. Luenberger, *Introduction to Linear and Nonlinear Programming,* Addison-Wesley, Reading, MA, 1973.
[20] F.J. MacWilliams, N.J.A. Sloane, *The Theory of Error-Correcting Codes,* North-Holland Publishing Company, New York, 1977.
[21] 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.
[22] 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.
[23] W. Parry, S. Tuncel, "Classification Problems in Ergodic Theory," *London Math. Soc. Lecture Note Series,* vol. 67, Cambridge University Press, 1982.
[24] C.L.M. van Pul, T. Etzion, "New lower bounds for constant weight codes," *IEEE Trans. Inform. Theory,* vol. 35, pp. 1324–1329, 1989.
[25] T. Rockafellar, *Convex Analysis,* Princeton University Press, 1970.
[26] A. Sharov, R.M. Roth, "Improved bounds and constructions for granular media coding," *Proc. Allerton Conf. Commun., Control Comput.,* Allerton Retreat Center, Monticello, IL, 2013, pp. 637–644.
[27] A. Sharov, R.M. Roth, "Bounds and constructions for granular media coding," *IEEE Trans. Inform. Theory,* vol. 60, pp. 2010–2027, 2014.
[28] E. Sperner, "Ein satz über untermengen einer endlichen menge," *Mathematische Zeitschrift* (in German), vol. 27, pp. 544–548, 1928.
[29] J.H. Weber, "Asymptotic results on codes for symmetric, unidirectional, and asymmetric error control," *IEEE Trans. Inform. Theory,* vol. 40, pp. 2073–2075, 1994.
[30] J.H. Weber, C. de Vroedt, D.E. Boekee, "Bounds and constructions for binary codes of length less than 24 and asymmetric distance less than 6," *IEEE Trans. Inform. Theory,* vol. 34, pp. 1321–1331, 1988.
[31] 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. He was a recipient of the Best Student Paper Award at the 2014 IEEE International Symposium on Information Theory (ISIT).

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.