Upper Bounds on the Size of Grain-Correcting Codes

Navin Kashyap Gilles Zémor

Abstract—**In this paper, we re-visit the combinatorial error model of Mazumdar et al. that models errors in high-density magnetic recording caused by lack of knowledge of grain boundaries in the recording medium. We present new upper bounds on the cardinality/rate of binary block codes that correct errors within this model. All our bounds, except for one, are obtained using combinatorial arguments based on hypergraph fractional coverings. The exception is a bound derived via an information-theoretic argument. Our bounds significantly improve upon existing bounds from the prior literature.**

Index Terms—**fractional coverings, grain-correcting codes, high-density magnetic recording,**

I. Introduction

The combinatorial error model studied by Mazumdar et al. [6] is a highly simplified model of an error mechanism encountered in a magnetic recording medium at terabit-per-square-inch storage densites [5], [7], [10]. In this model, a one-dimensional track on a magnetic recording medium is divided into evenly spaced bit cells, each of which can store one bit of data. Bits are written sequentially into these bit cells. The sequence of bit cells has an underlying "grain" distribution, which may be described as follows: bit cells are grouped into non-overlapping blocks called *grains*, which may consist of up to b adjacent bit cells. We focus on the case $b = 2$, so that a grain can contain at most two bit cells. We define the *length* of a grain to be the number of bit cells it contains.

Each grain can store only one bit of information, i.e., all the bit cells within a grain carry the same bit value (0 or 1), which we call the *polarity* of the grain. We assume, following [6], that in the sequential write process, the first bit to be written into a grain sets the polarity of the grain, so that all the bit cells within this grain must retain this polarity.[1] This implies that any subsequent attempts at writing bits within this grain make no difference to the value actually stored in the bit cells in the grain. If the grain boundaries were known to the write head (encoder) and the read head (decoder), then the maximum storage capacity of one bit per grain can be achieved. However, in a more realistic scenario where the underlying grain distribution is fixed but *unknown*, the lack of knowledge of grain boundaries reduces the storage capacity.

Constructions and rate/cardinality bounds for codes that correct errors caused by an underlying grain distribution have been studied in the prior literature [3], [4], [6], [8], [9]. In this paper, we derive improved rate/cardinality bounds for such codes. More precisely, following [6], we define a *t-grain-correcting code* to be a code that can correct errors caused by a fixed but unknown underlying grain distribution consisting of at most t length-2 grains. We derive upper bounds on the cardinality of t-grain correcting codes of blocklength n. Furthermore, in the regime of t being a constant fraction of n, we present upper bounds on the rate achievable by such codes, asymptotically in n. Our upper bounds on cardinality are derived using the fractional covering technique previously used in a different context by Kulkarni and Kiyavash [2]. Standard asymptotic analysis allows us to translate these cardinality bounds to bounds on the maximum rate asymptotically achievable. We also use an alternative, information-theoretic approach to obtain a different upper bound on this rate. Comparing our bounds with those from the prior literature [6], [8], [9], we find that our bounds improve significantly upon the existing bounds, although the gap to the best known lower bounds remains large (except at some small blocklengths).

The paper is organized as follows. In Section II, we provide the necessary definitions and notation, and derive some basic results needed for the subsequent development. The fractional covering technique from [2] is described in Section III, and to illustrate the use of this technique, an upper bound on the cardinality of t-grain-correcting codes is proposed. A stronger upper bound, also based on the fractional covering method, is conjectured in Section IV. We can prove that the stronger bound indeed holds for $t = 1, 2, 3$; the details of the proof are in Appendix A. The fractional covering technique is also used in Section V to obtain an upper bound on the maximum rate asymptotically achievable by codes correcting a constant fraction of grain errors. An information-theoretic upper bound on the same quantity is derived in Section VI. We conclude in Section VII with some remarks concerning the bounds. The current state-of-the-art on upper bounds on the maximum rate asymptotically achievable, including the bounds derived in this paper, is summarized in Figures 2 and 3. Some of the technical proofs from Sections IV and V are given in appendices.

This work was supported in part by a research grant from the Department of Science & Technology, Government of India. This paper was presented in part at the 2013 IEEE International Symposium on Information Theory (ISIT 2013), Istanbul, Turkey, July 7–12, 2013.

N. Kashyap is with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore. Email: nkashyap@ece.iisc.ernet.in

G. Zémor is with the Institut de Mathématiques de Bordeaux, UMR 5251, Université de Bordeaux, France. Email: gilles.zemor@math.u-bordeaux.fr

[1]Considering the physics of the write process, it would make more sense to assume that the *last* bit to be written within a grain sets the polarity of the grain, thus overwriting all other bits previously stored in the bit cells comprising the grain. However, mathematically, this is completely equivalent to the polarity-set-by-first-bit model; see also [5, Section III-C].

II. Preliminaries

A. Definitions and Notation

Let $\Sigma = \{0, 1\}$, and for a positive integer n, let $[n]$ denote the set $\{1, 2, \ldots, n\}$. A track on the recording medium consists of n bit cells indexed by the integers in $[n]$. The bit cells on the track are grouped into non-overlapping grains of length at most two. A length-2 grain consists of bit cells with indices $j-1$ and j, for some $j \in [n]$; we denote such a grain by the pair $(j-1, j)$. Let $E \subseteq \{2, \ldots, n\}$ be the set of all indices j such that $(j-1, j)$ is a length-2 grain. Since grains cannot overlap[2], E contains no pair of consecutive integers.

A binary sequence $\mathbf{x} = (x_1, \ldots, x_n) \in \Sigma^n$ to be written on to the track can be affected by errors only at the indices $j \in E$. Indeed, what actually gets recorded on the track is the sequence $\mathbf{y} = (y_1, \ldots, y_n)$, where

$$y_j = \begin{cases} x_{j-1} & \text{if } j \in E \\ x_j & \text{otherwise.} \end{cases} \tag{1}$$

For example, if $\mathbf{x} = (000101011100010)$ and $E = \{2, 4, 7, 9, 14\}$, then $\mathbf{y} = (000001111100000)$.

Note that the set E completely specifies the positions and locations of all the grains (both length-1 and length-2) in the track. We will call this set the *grain pattern*. It is assumed that the grain pattern is unknown to both the write head and the read head. The effect of the grain pattern E on a binary sequence $\mathbf{x} \in \Sigma^n$ defines an operator $\phi_E : \Sigma^n \to \Sigma^n$, where $\mathbf{y} = \phi_E(\mathbf{x})$ is as specified by (1) above.

For integers $n \geq 1$ and $t \geq 0$, let $\mathcal{E}_{n,t}$ denote the set of all grain patterns E with $|E| \leq t$. In other words, $\mathcal{E}_{n,t}$ consists of all subsets $E \subseteq \{2, \ldots, n\}$ of cardinality at most t, such that E contains no pair of consecutive integers. For an $\mathbf{x} \in \Sigma^n$, we define

$$\Phi_t(\mathbf{x}) = \{\phi_E(\mathbf{x}) : E \in \mathcal{E}_{n,t}\}.$$

Thus, $\Phi_t(\mathbf{x})$ is the set of all possible sequences that can be obtained from $\mathbf{x}$ by the action of some grain pattern E with $|E| \leq t$. Two sequences $\mathbf{x}_1$ and $\mathbf{x}_2$ are *t-confusable* if $\Phi_t(\mathbf{x}_1) \cap \Phi_t(\mathbf{x}_2) \neq \emptyset$. A binary code $\mathcal{C}$ of length n is said to correct t grain errors, or be a *t-grain-correcting code*, if no two distinct vectors $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$ are t-confusable. Let $M(n, t)$ denote the maximum cardinality of a t-grain-correcting code of length n. Also, for $\tau \in [0, \frac{1}{2}]$, the maximum asymptotic rate of a $\lceil \tau n \rceil$-grain-correcting code is defined to be

$$R(\tau) = \limsup_{n \to \infty} \frac{1}{n} \log_2 M(n, \lceil \tau n \rceil). \tag{2}$$

We note here that in an optimal t-grain-correcting code of blocklength n, the codewords that start with a 0 and those that start with a 1 must be equal in number. In particular, $M(n, t)$ is always an even number. To see why this is the case, consider a t-grain-correcting code $\mathcal{C} = \mathcal{C}_0 \cup \mathcal{C}_1$, where, $\mathcal{C}_b$, $b \in \{0, 1\}$, is the subcode of $\mathcal{C}$ consisting of codewords that start with a b. Without loss of generality, assume that $|\mathcal{C}_0| > |\mathcal{C}_1|$. Delete $\mathcal{C}_1$ from $\mathcal{C}$, and replace it with $\overline{\mathcal{C}_0} = \{\overline{\mathbf{c}} : \mathbf{c} \in \mathcal{C}_0\}$, where $\overline{\mathbf{c}}$ is the binary sequence obtained by complementing each bit of $\mathbf{c}$. Since codewords beginning with a 0 are not confusable with those beginning with a 1, it follows that $\mathcal{C}_0 \cup \overline{\mathcal{C}_0}$ is also t-grain-correcting. Moreover, it should be clear that $|\mathcal{C}_0 \cup \overline{\mathcal{C}_0}| > |\mathcal{C}|$.

A grain pattern E changes a sequence $\mathbf{x}$ to a different sequence $\mathbf{y}$ iff $x_{j-1} \neq x_j$ for some $j \in E$, i.e., the length-2 grain $(j-1, j)$ straddles the boundary between two successive runs in $\mathbf{x}$. Here, a *run* is a maximal substring of consecutive identical symbols (0s or 1s) in $\mathbf{x}$. A run consisting of 0s (resp. 1s) is called a 0-*run* (resp. 1-*run*). The number of distinct runs in $\mathbf{x}$ is denoted by $r(\mathbf{x})$.

A convenient means of keeping track of run boundaries in $\mathbf{x}$ is via its *derivative sequence*, $\mathbf{x}'$: for $\mathbf{x} = (x_1, \ldots, x_n)$, the sequence $\mathbf{x}' = (x'_2, \ldots, x'_n)$ is defined by $x'_j = x_{j-1} \oplus x_j$, $j = 2, \ldots, n$, where $\oplus$ denotes modulo-2 addition. The 1s in $\mathbf{x}'$ identify the boundaries between successive runs in $\mathbf{x}$. Thus, $\omega(\mathbf{x}') = r(\mathbf{x}) - 1$, where $\omega(\cdot)$ denotes the Hamming weight of a binary sequence. We define the *support* of $\mathbf{x}'$ to be the set $\text{supp}(\mathbf{x}') = \{j : x'_j = 1\}$; note that $\text{supp}(\mathbf{x}') \subseteq \{2, \ldots, n\}$.

Example 1. *Taking* $\mathbf{x} = (000101011100010)$, *as in the example given after* (1), *we have* $\mathbf{x}' = (00111110010011)$, *with supp*$(\mathbf{x}') = \{4, 5, 6, 7, 8, 11, 14, 15\}$. *Observe that the grain pattern* $E = \{2, 4, 7, 9, 14\}$ *has the same effect on* $\mathbf{x}$ *as the grain pattern* $\hat{E} = \{4, 7, 14\}$, *which is the intersection of* E *with supp*$(\mathbf{x}')$. *Note also that the grain patterns having the maximum effect on* $\mathbf{x}$ *are* $\{4, 6, 8, 11, 14\}$ *and* $\{4, 6, 8, 11, 15\}$, *which are obtained by fitting into the support of* $\mathbf{x}'$ *the largest possible number of integers, no two of which are consecutive.*

B. Counting $|\Phi_t(\mathbf{x})|$

For $\mathbf{x} \in \Sigma^n$, the sequences $\mathbf{y} \in \Phi_t(\mathbf{x})$ are in one-to-one correspondence with the different ways of selecting at most t non-consecutive integers[3] from $\text{supp}(\mathbf{x}')$ to form a grain pattern $E \in \mathcal{E}_{n,t}$. Thus, $|\Phi_t(\mathbf{x})|$ counts the number of ways of forming such grain patterns.

Example 1 (cont'd)**.** *We saw above that there are two grain patterns of cardinality* 5 *that fit into supp*$(\mathbf{x}')$. *To systematically count the number of grain patterns of cardinality* 4 *that can fit into supp*$(\mathbf{x}')$, *we note that there are three* 1*-runs in* $\mathbf{x}'$, *which have supports* $S_1 = \{4, 5, 6, 7, 8\}$, $S_2 = \{11\}$ *and* $S_3 = \{14, 15\}$. *To form a grain pattern of cardinality* 4, *we can do one of the following: (a) pick* 3 *non-consecutive integers from* S_1 *and* 1 *integer from* S_2*; (b) pick* 3 *non-consecutive integers from* S_1 *and* 1 *integer from* S_3*; or (c) pick* 2 *non-consecutive integers from* S_1 *and* 1 *integer each from* S_2 *and* S_3. *There is* 1 *way of doing (a),* 2 *ways of doing (b), and* $6 \times 2 = 12$ *ways of doing (c). Thus, there are* $1 + 2 + 12 = 15$ *grain patterns of size* 4 *that are subsets of supp*$(\mathbf{x}')$.

In general, the count can be obtained as follows. Let $\ell_1, \ell_2, \ldots, \ell_m$ denote the lengths of the distinct 1-runs in $\mathbf{x}'$, and define the set

$$T = \Big\{(t_1, \ldots, t_m) \in \mathbb{Z}_+^m : \sum_{j=1}^{m} t_j \leq t\Big\}, \tag{3}$$

[2]An overlapping grains model has also been considered in the literature [3], [4], [8], [9].

[3]A sequence or set of non-consecutive integers is one that does not contain a pair of consecutive integers.

where $\mathbb{Z}_+$ denotes the set of non-negative integers. In the above expression, t_j represents the number of integers from the support of the jth 1-run that are to be included in a grain pattern E being formed. The number of distinct ways in which t_j non-consecutive integers can be chosen from the ℓ_j consecutive integers forming the support of the jth 1-run is, by an elementary counting argument, equal to $\binom{\ell_j - t_j + 1}{t_j}$. Thus,

$$|\Phi_t(\mathbf{x})| = \sum_{(t_1,\ldots,t_m)\in T} \prod_{j=1}^{m} \binom{\ell_j - t_j + 1}{t_j}. \quad (4)$$

Simplified expressions can be obtained for small values of t.

Proposition II.1. *For $\mathbf{x} \in \Sigma^n$, let $\omega = \omega(\mathbf{x}')$ denote the Hamming weight of the derivative sequence $\mathbf{x}'$. Also, let m be the number of 1-runs in $\mathbf{x}'$.*

(a) $|\Phi_1(\mathbf{x})| = 1 + \omega = r(\mathbf{x})$.
(b) $|\Phi_2(\mathbf{x})| = 1 + m + \binom{\omega}{2}$.
(c) $|\Phi_3(\mathbf{x})| = 1 + m_1 + m(\omega - 3) + \binom{\omega}{3} - \binom{\omega}{2} + 2\omega$, *where m_1 denotes the number of 1-runs of length 1 in $\mathbf{x}'$.*

Proof: (a) While the expression for $|\Phi_1(\mathbf{x})|$ can be directly obtained from (4), it is simpler to observe that the set $\Phi_1(\mathbf{x})$ consists of the sequence $\mathbf{x}$ itself, and the ω distinct sequences in the set $\{\phi_E(\mathbf{x}) : E = \{j\} \text{ for some } j \in \text{supp}(\mathbf{x}')\}$.

(b) For $t = 2$, it is easy to see that the expression in (4) simplifies to

$$|\Phi_2(\mathbf{x})| = 1 + \sum_{j=1}^{m} \ell_j + \sum_{j=1}^{m} \binom{\ell_j - 1}{2} + \sum_{(i,j):i<j} \ell_i \ell_j.$$

Writing $\sum_j \ell_j$ as $m + \sum_j (\ell_j - 1)$, and using the fact that $(\ell_j - 1) + \binom{\ell_j - 1}{2} = \binom{\ell_j}{2}$, we obtain

$$\begin{aligned}
|\Phi_2(\mathbf{x})| &= 1 + m + \sum_{j=1}^{m} \binom{\ell_j}{2} + \sum_{(i,j):i<j} \ell_i \ell_j \\
&= 1 + m + \frac{1}{2} \sum_{j=1}^{m} (\ell_j^2 - \ell_j) + \sum_{(i,j):i<j} \ell_i \ell_j \\
&= 1 + m + \frac{1}{2} \left[\sum_{j=1}^{m} \ell_j^2 + \sum_{(i,j):i\neq j} \ell_i \ell_j - \sum_{j=1}^{m} \ell_j \right] \\
&= 1 + m + \frac{1}{2} \left[\left(\sum_{j=1}^{m} \ell_j \right)^2 - \sum_{j=1}^{m} \ell_j \right] \\
&= 1 + m + \binom{\omega}{2},
\end{aligned}$$

the last equality being due to the fact that $\omega = \sum_{j=1}^{m} \ell_j$.

(c) For $t = 3$, the expression in (4) can be written as

$$\begin{aligned}
|\Phi_3(\mathbf{x})| = |\Phi_2(\mathbf{x})| &+ \sum_{(i,j,k):i<j<k} \ell_i \ell_j \ell_k \\
&+ \sum_{i=1}^{m} \binom{\ell_i - 1}{2} \left(\sum_{j:j\neq i} \ell_j \right) + \sum_{i=1}^{m} \binom{\ell_i - 2}{3}.
\end{aligned}$$

From here on, straightforward algebraic manipulations lead to the expression given in the statement of the proposition. We omit the details, noting only that m_1 enters the picture when we write the last term above as

$$\sum_{i=1}^{m} \frac{(\ell_i - 2)(\ell_i - 3)(\ell_i - 4)}{6} + \sum_{i:\ell_i=1} 1.$$

The extra term $\sum_{i:\ell_i=1} 1$, which equals m_1, accounts for the fact that the expansion of $\binom{\ell_i - 2}{3}$ as $\frac{(\ell_i-2)(\ell_i-3)(\ell_i-4)}{6}$ is invalid when $\ell_i = 1$; by convention, $\binom{a}{b} = 0$ when $a < 0$. ∎

We will also find the following simple lower bound on $|\Phi_t(\mathbf{x})|$, valid for any $t \geq 1$, to be useful.

Proposition II.2. *For $\mathbf{x} \in \Sigma^n$ and $t \geq 1$, we have*

$$|\Phi_t(\mathbf{x})| \geq \sum_{j=0}^{t} \binom{r(\mathbf{x}) - j}{j}.$$

Proof: Consider the number of different ways of choosing exactly j non-consecutive integers from $\text{supp}(\mathbf{x}')$. This number is smallest when $\text{supp}(\mathbf{x}')$ consists of consecutive integers, e.g., $\text{supp}(\mathbf{x}) = [r(\mathbf{x}) - 1]$. The number of different ways of choosing exactly j non-consecutive integers from $[r(\mathbf{x}) - 1]$ is, by an elementary counting argument, equal to $\binom{r(\mathbf{x})-j}{j}$. ∎

III. The Fractional Covering Method

In this section, we explore the applicability to grain-correcting codes of a technique used by Kulkarni and Kiyavash [2] to derive upper bounds on the cardinalities of deletion-correcting codes. This approach was independently used by Gabrys et al. [3], [4] to derive upper bounds on the size of codes within the overlapping grains model.

A *hypergraph* $\mathcal{H}$ is a pair $(V, \mathcal{X})$, where V is a finite set, called the *vertex set*, and $\mathcal{X}$ is a family of subsets of V. The members of $\mathcal{X}$ are called *hyperedges*. A *matching* of $\mathcal{H}$ is a pairwise disjoint collection of hyperedges. A *(vertex) covering* of $\mathcal{H}$ is a subset $T \subseteq V$ such that T meets every hyperedge of $\mathcal{H}$, i.e., $T \cap X \neq \emptyset$ for all $X \in \mathcal{X}$. The *matching number* $\nu(\mathcal{H})$ is the largest size of a matching of $\mathcal{H}$, while the *covering number*, $\tau(\mathcal{H})$, is the smallest size of a covering of $\mathcal{H}$.

The problems of computing the matching and covering numbers can be expressed as a dual pair of integer programs. This is done via the vertex-hyperedge incidence matrix, A, of $\mathcal{H}$, which is defined as follows. Let $v_1, v_2, \ldots, v_{|V|}$ and $X_1, X_2, \ldots, X_{|\mathcal{X}|}$ be a listing of the vertices and hyperedges, respectively, of $\mathcal{H}$. Then, $A = (A_{i,j})$ is the $|V| \times |\mathcal{X}|$ matrix with $0/1$ entries, with $A_{i,j} = 1$ iff $v_i \in X_j$. It is easy to verify that

$$\nu(\mathcal{H}) = \max\{\mathbf{1}^T \mathbf{z} : \mathbf{z} \in \{0,1\}^{|\mathcal{X}|},\ A\mathbf{z} \leq \mathbf{1}\} \quad (5)$$

and

$$\tau(\mathcal{H}) = \min\{\mathbf{1}^T \mathbf{w} : \mathbf{w} \in \{0,1\}^{|V|},\ A^T \mathbf{w} \geq \mathbf{1}\}, \quad (6)$$

where $\mathbf{1}$, $\mathbf{w}$ and $\mathbf{z}$ are column vectors, with $\mathbf{1}$ in particular denoting an all-ones vector. Note that the linear programming (LP) relaxations of (5),

$$\nu_f(\mathcal{H}) = \max\{\mathbf{1}^T \mathbf{z} : \mathbf{z} \geq \mathbf{0},\ A\mathbf{z} \leq \mathbf{1}\}, \quad (7)$$

and (6),

$$\tau_f(\mathcal{H}) = \min\{\mathbf{1}^T\mathbf{w} : \mathbf{w} \geq \mathbf{0},\ A^T\mathbf{w} \geq \mathbf{1}\}, \tag{8}$$

where $\mathbf{0}$ denotes an all-zeros vector, are duals of each other. By strong LP duality, we have $\nu_f(\mathcal{H}) = \tau_f(\mathcal{H})$, and hence,

$$\nu(\mathcal{H}) \leq \nu_f(\mathcal{H}) = \tau_f(\mathcal{H}) \leq \tau(\mathcal{H}). \tag{9}$$

The quantities $\nu_f(\mathcal{H})$ and $\tau_f(\mathcal{H})$ are called the *fractional matching number* and *fractional covering number*, respectively, of the hypergraph $\mathcal{H}$. Any non-negative vector $\mathbf{w}$ such that $A^T\mathbf{w} \geq 1$ is called a *fractional covering*[4] of $\mathcal{H}$. To put it in another way, a fractional covering is a function $\mathsf{w} : V \to \mathbb{R}_+$ such that $\sum_{v\in X} \mathsf{w}(v) \geq 1$ for all $X \in \mathcal{X}$. The *value* of a fractional covering w is defined to be $|\mathsf{w}| := \sum_{v\in V} \mathsf{w}(v)$. From the inequality $\nu(\mathcal{H}) \leq \tau_f(\mathcal{H})$ in (9), we see that $\nu(\mathcal{H}) \leq |\mathsf{w}|$ for any fractional covering w of $\mathcal{H}$. We use this to suggest an upper bound on the largest size, $M(n,t)$, of a t-grain-correcting code of blocklength n.

Consider the hypergraph $\mathcal{H}_{n,t} = (V, \mathcal{X})$, where $V = \Sigma^n$, and $\mathcal{X} = \{\Phi_t(\mathbf{x}) : \mathbf{x} \in \Sigma^n\}$. Note that $\nu(\mathcal{H}_{n,t}) = M(n,t)$; thus, fractional coverings of $\mathcal{H}_{n,t}$ yield upper bounds on $M(n,t)$. Bounding the size of packings in this way has been extensively used in combinatorics, see e.g. [1]. As a first attempt, taking inspiration from [2], we consider the function $\mathsf{w}_t : \Sigma^n \to \mathbb{R}_+$, defined for $\mathbf{x} \in \Sigma^n$ as

$$\mathsf{w}_t(\mathbf{x}) = \frac{1}{|\Phi_t(\mathbf{x})|}. \tag{10}$$

For $t = 1, 2, 3$, we can prove that w_t is a fractional covering of $\mathcal{H}_{n,t}$, and conjecture that this is in fact the case for all $t \geq 1$.

Conjecture III.1. *For all positive integers n and t, the function w_t defined in (10) is a fractional covering of $\mathcal{H}_{n,t}$, i.e., for all $\mathbf{x} \in \Sigma^n$,*

$$\sum_{\mathbf{y}\in\Phi_t(\mathbf{x})} \frac{1}{|\Phi_t(\mathbf{y})|} \geq 1. \tag{11}$$

Therefore,

$$M(n,t) \leq |\mathsf{w}_t| = \sum_{\mathbf{x}\in\Sigma^n} \frac{1}{|\Phi_t(\mathbf{x})|}. \tag{12}$$

We do not give a proof of the conjecture for $t = 1, 2, 3$ here, as these cases are implied by the stronger Theorem IV.1 of the next section. We instead suggest a possible approach to proving Conjecture III.1 for general t. To prove (11), it would be enough to show that for each $\mathbf{x} \in \Sigma^n$,

$$\sum_{\mathbf{y}\in\Phi_t(\mathbf{x})} [|\Phi_t(\mathbf{y})| - |\Phi_t(\mathbf{x})|] \leq 0.$$

Indeed, the above inequality is equivalent to showing that the arithmetic mean $\frac{1}{|\Phi_t(\mathbf{x})|}\sum_{\mathbf{y}\in\Phi_t(\mathbf{x})} |\Phi_t(\mathbf{y})|$ is at most $|\Phi_t(\mathbf{x})|$. If this is true, then by convexity of the function $f(x) = \frac{1}{x}$, we would have

$$\frac{1}{|\Phi_t(\mathbf{x})|}\sum_{\mathbf{y}\in\Phi_t(\mathbf{x})} \frac{1}{|\Phi_t(\mathbf{y})|} \geq \frac{1}{\frac{1}{|\Phi_t(\mathbf{x})|}\sum\limits_{\mathbf{y}\in\Phi_t(\mathbf{x})} |\Phi_t(\mathbf{y})|} \geq \frac{1}{|\Phi_t(\mathbf{x})|},$$

which is the desired inequality (11).

IV. A Stronger Upper Bound on $M(n,t)$

Our second attempt at finding a fractional covering for $\mathcal{H}_{n,t}$ involves the cardinality, $V(n,t)$, of a Hamming ball of radius t in Σ^n:

$$V(n,t) = \sum_{j=0}^{t} \binom{n}{j}.$$

Note that for any $\mathbf{x} \in \Sigma^n$, we have $|\Phi_t(\mathbf{x})| \leq V(\omega(\mathbf{x}'), t)$, where $\omega(\mathbf{x}')$ is the Hamming weight of the derivative sequence $\mathbf{x}'$. This is because $V(\omega(\mathbf{x}'), t)$ counts the number of ways that a pattern of up to t length-2 grains could affect $\mathbf{x}$ if the grains were not constrained to be non-overlapping.

We conjecture that the function $\widetilde{\mathsf{w}}_t : \Sigma^n \to \mathbb{R}_+$, defined by

$$\widetilde{\mathsf{w}}_t(\mathbf{x}) = \frac{1}{V(\omega(\mathbf{x}'), t)} \tag{13}$$

is a fractional covering of the hypergraph $\mathcal{H}_{n,t}$. Note that

$$|\widetilde{\mathsf{w}}_t| = \sum_{\mathbf{x}\in\Sigma^n} \frac{1}{V(\omega(\mathbf{x}'), t)} = 2\sum_{\omega=0}^{n-1} \binom{n-1}{\omega} \frac{1}{V(\omega, t)},$$

since $2\binom{n-1}{\omega}$ is the number of sequences $\mathbf{x} \in \Sigma^n$ whose derivative sequence $\mathbf{x}'$ has Hamming weight ω.

Conjecture IV.1. *For all positive integers n and t, and for all $\mathbf{x} \in \Sigma^n$, we have*

$$\sum_{\mathbf{y}\in\Phi_t(\mathbf{x})} \frac{1}{V(\omega(\mathbf{y}'), t)} \geq 1. \tag{14}$$

Therefore,

$$M(n,t) \leq 2\sum_{\omega=0}^{n-1} \binom{n-1}{\omega} \frac{1}{V(\omega, t)}. \tag{15}$$

Clearly, (15) is tighter than (12), since $|\Phi_t(\mathbf{x})| \leq V(\omega(\mathbf{x}'), t)$. Thus, Conjecture IV.1 implies Conjecture III.1. We can prove that these conjectures hold for $t = 1, 2, 3$.

Theorem IV.1. *For any positive integer n and $t = 1, 2, 3$, we have*

$$M(n,t) \leq 2\sum_{\omega=0}^{n-1} \binom{n-1}{\omega} \frac{1}{V(\omega, t)}.$$

The proof of the theorem is presented in Appendix A. We remark here that $\widetilde{\mathsf{w}}_t$ can easily be shown to be a fractional covering for the hypergraph analogous to $\mathcal{H}_{n,t}$ in the *overlapping* grains model. Thus, Conjecture IV.1 and, in particular, the inequality (15) hold for that model [4, Theorem 1].

[4] A fractional matching is correspondingly defined, but we will have no further use for this concept.

For $t = 1$, the summation in the above theorem can be expressed in closed form. Indeed, we have

$$\begin{aligned} 2\sum_{\omega=0}^{n-1}\binom{n-1}{\omega}\frac{1}{V(\omega,1)} &= 2\sum_{\omega=0}^{n-1}\binom{n-1}{\omega}\frac{1}{1+\omega} \\ &= 2\sum_{r=1}^{n}\binom{n-1}{r-1}\frac{1}{r} \\ &\overset{\text{(a)}}{=} 2\sum_{r=1}^{n}\frac{1}{n}\binom{n}{r} \\ &= \frac{2}{n}(2^n-1). \end{aligned}$$

Equality (a) above uses the identity $\frac{1}{r}\binom{n-1}{r-1} = \frac{1}{n}\binom{n}{r}$. Thus, we have the following corollary to Theorem IV.1.

Corollary IV.2. *$M(n,1) \le \frac{1}{n}(2^{n+1}-2)$ for all positive integers n.*

n	$t=1$	$t=2$	$t=3$	n	$t=1$	$t=2$	$t=3$
				11	372	168	126
2	2 (2)			12	682	280	198
3	4 (4)			13	1260	476	312
4	6 (6)	6 (4)		14	2340	814	496
5	12 (8)	10 (8)		15	4368	1406	800
6	20 (16)	14 (10)	14 (8)	16	8190	2448	1300
7	36 (26)	24 (16)	22 (16)	17	15420	4302	2132
8	62 (44)	38 (22)	34 (18)	18	29126	7612	3528
9	112	62	52 (32)	19	55188	13560	5892
10	204	102	80	20	104856	24306	9920

TABLE I
SOME NUMERICAL VALUES OF THE UPPER BOUND ON $M(n,t)$ OF THEOREM IV.1, ROUNDED DOWN TO THE NEAREST EVEN INTEGER. WITHIN PARENTHESES ARE THE CORRESPONDING LOWER BOUNDS FROM TABLE II OF [9].

Table I lists the numerical values of the bound in Theorem IV.1, evaluated for some small values of n and rounded down to the nearest even integer[5]. The corresponding best known lower bounds, taken from Table II of [9], are given in parentheses. Two other upper bounds on $M(n,t)$ exist in the prior literature, namely Corollary 6 of [6] and Theorem 3.1 of [9]. Numerical computations for $n \le 20$ show that our bounds are consistently better than the bounds obtained from [9, Theorem 3.1]. On the other hand, the bound of [6, Corollary 6] may be better than our bound for small values of n: for example, the bound in [6] yields $M(10,2) \le 92$. However, our bound is better for all n sufficiently large: for $t = 1$, our bound is better for all $n \ge 8$; for $t = 2$, our bound wins for $n \ge 12$; for $t = 3$, it wins for $n \ge 17$.

V. An Upper Bound on $R(\tau)$

Were they to be proved, Conjectures III.1 and IV.1 would yield upper bounds on the asymptotic rate $R(\tau)$, as defined in (2). Instead, a slightly different approach can be used to obtain a fractional covering that does result in a provable upper bound on $R(\tau)$.

[5]As explained in Section II-A, $M(n,t)$ is always an even number.

Suppose that for any fixed n, t, we could find a lower bound $\varphi_{n,t}(r)$ on $|\Phi_t(\mathbf{x})|$, $\mathbf{x} \in \Sigma^n$, that depends on $\mathbf{x}$ only through $r = r(\mathbf{x})$, the number of distinct runs in $\mathbf{x}$. Furthermore, suppose that the function $\varphi_{n,t}(r)$ is non-decreasing in r [9, Section III]. Then, it is straightforward to see that the function $\mathbf{x} \mapsto \frac{1}{\varphi_{n,t}(r(\mathbf{x}))}$ is a fractional covering of the hypergraph $\mathcal{H}_{n,t}$ for all positive integers n and t. Indeed, by Lemma A.1, we have

$$\sum_{\mathbf{y}\in\Phi_t(\mathbf{x})}\frac{1}{\varphi_{n,t}(r(\mathbf{y}))} \ge \sum_{\mathbf{y}\in\Phi_t(\mathbf{x})}\frac{1}{\varphi_{n,t}(r(\mathbf{x}))} = \frac{|\Phi_t(\mathbf{x})|}{\varphi_{n,t}(r(\mathbf{x}))} \ge 1.$$

Thus, for any such $\varphi_{n,t}$, we have

$$M(n,t) \le \sum_{\mathbf{x}\in\Sigma^n}\frac{1}{\varphi_{n,t}(r(\mathbf{x}))} = 2\sum_{r=1}^{n}\binom{n-1}{r-1}\frac{1}{\varphi_{n,t}(r)}. \tag{16}$$

Theorem V.1. *For all positive integers n and t, the upper bound* (16) *holds with*

$$\varphi_{n,t}(r) = \sum_{j=0}^{t}\binom{r-j}{j} \tag{17}$$

Proof: The expression on the right-hand side of (17) is clearly non-decreasing in r, and by Proposition II.2, $\varphi_{n,t}(r(\mathbf{x}))$ is a lower bound on $|\Phi_t(\mathbf{x})|$ for any $\mathbf{x} \in \Sigma^n$ and $t > 0$. ■

The bound of Theorem V.1 is weaker than that of Theorem IV.1 for $t = 1, 2, 3$. However, it has the advantage of being provably true for all values of n and t. It can therefore be used to derive an upper bound on $R(\tau)$ by studying the asymptotics of $\varphi_{n,t}(r)$ as $n \to \infty$, with $t = \lceil \tau n \rceil$ and $r = \lceil \rho n \rceil$ for fixed $\tau \in [0, \frac{1}{2}]$ and $\rho \in (0, 1]$. The following theorem is proved in Appendix B.

Theorem V.2. *Let $\phi = \frac{1+\sqrt{5}}{2}$ (the golden ratio), and define $\theta = \frac{1}{\sqrt{5}\phi(\phi+1)} = \frac{1}{5+2\sqrt{5}}$. For $\tau \in [0, \frac{1}{2}]$, we have*

$$R(\tau) \le \begin{cases} \displaystyle\max_{\sqrt{5}\phi\tau \le \rho \le 1}\left[\mathsf{h}(\rho) - (\rho-\tau)\,\mathsf{h}\left(\frac{\tau}{\rho-\tau}\right)\right] & \text{if } \tau < \theta \\ \log_2\phi & \text{if } \tau \ge \theta \end{cases}$$

Numerically, $\theta \approx 0.1056$, and $\log_2\phi \approx 0.6942$.

Figure 2 contains a plot of the above upper bound. The figure shows that this is the best known upper bound for values of τ up to about 0.1103, beyond which it is beaten by the bound of the next section.

VI. An Information-Theoretic Upper Bound on $R(\tau)$

In this section, we use an information-theoretic approach to derive an upper bound on $R(\tau)$. For every even n, by grouping together adjacent coordinates, we can view any code $\mathcal{C} \in \{0,1\}^n$ as a code of blocklength $n/2$ over the alphabet $\{00, 01, 10, 11\}$. Let us say that a binary n-tuple, alternatively an $n/2$-tuple over the quaternary alphabet, has *quaternary distribution* (or simply *distribution*) $(f_{00}, f_{11}, f_{01}, f_{10})$ if it has $f_{00}n/2$ symbols 00, $f_{11}n/2$ symbols 11, $f_{01}n/2$ symbols

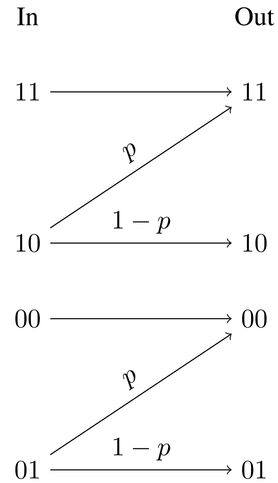

Fig. 1. A DMC whose effect can be mimicked by grain patterns

01 and $f_{10}n/2$ symbols 10. We will say that a code has *constant distribution* if each of its codewords has the same quaternary distribution $(f_{00}, f_{11}, f_{01}, f_{10})$. Our goal is to find upper bounds on the rate of $\lceil \tau n \rceil$-grain-correcting codes of constant distribution: since the number of possible quaternary distributions for a code of length n is $O(n^3)$, which is sub-exponential in n, the maximum of these upper bounds on constrained codes will yield an unconstrained upper bound.

Let us introduce the following notation:

$$R_f(\tau) = \limsup_{n\to\infty} \frac{1}{n} \log_2 M(n, f, \lceil \tau n \rceil)$$

where $M(n, f, t)$ denotes the maximum cardinality of a t-grain error correcting code of length n and constant quaternary distribution f.

Our strategy is the following: for any given distribution $f = (f_{00}, f_{11}, f_{01}, f_{10})$, we associate to it a discrete memoryless channel (DMC) with input and output alphabets $\{00, 01, 10, 11\}$ such that any infinite family of $\lceil \tau n \rceil$-grain-correcting codes of constant distribution f achieves vanishing error-probability when submitted through this channel. By a standard information-theoretic argument, this implies that the asymptotic rate R of any family of $\lceil \tau n \rceil$-grain-correcting codes of constant distribution f is bounded from above by half the mutual information between the channel input with probability distribution f and the channel output.

Consider the channel depicted in Figure 1. Let $\mathcal{C}$ be a member of a family of $\lceil \tau n \rceil$-grain-correcting codes of length n and constant distribution f. Suppose that

$$(f_{10} + f_{01})pn/2 \leq \tau n(1 - \varepsilon),$$

where p is the transition probability shown in Figure 1. When a binary n-tuple, equivalently a word of length $n/2$ over the alphabet $\{00, 01, 10, 11\}$, is transmitted over the channel, then with probability tending to 1 as n goes to infinity, the number of transitions $01 \to 00$ plus the number of transitions $10 \to 11$ is not more than $\lceil \tau n \rceil$. Since these transitions are of the kind caused by grain errors, if there are no more than $\lceil \tau n \rceil$ such transitions, then the errors they cause are correctable by any $\lceil \tau n \rceil$-grain-correcting code. Therefore, for any $\varepsilon > 0$, any family of $\lceil \tau n \rceil$-grain-correcting codes of constant distribution f can be transmitted over the above channel with vanishing error probability after decoding. By a continuity argument we conclude that:

$$R_f(\tau) \leq \frac{1}{2} I(X;Y) \tag{18}$$

where X is the channel input with probability distribution $p(X) = f$, and Y is the corresponding output of the channel with parameter

$$p = \frac{2\tau}{f_{10} + f_{01}}. \tag{19}$$

It remains to compute the mutual information $I(X;Y)$. Since $p \leq 1$, (19) implies that we can write

$$f_{10} + f_{10} = 2\tau + x \tag{20}$$

$$f_{00} + f_{11} = 1 - 2\tau - x \tag{21}$$

with x non-negative. Now, for every distribution satisfying (20) and (21) we have

$$H(Y \mid X) = (2\tau + x)\,\mathsf{h}\left(\frac{2\tau}{2\tau + x}\right),$$

where $\mathsf{h}(\cdot)$ is the binary entropy function defined by $\mathsf{h}(\xi) = -\xi \log_2 \xi - (1 - \xi)\log_2(1 - \xi)$, for $\xi \in [0, 1]$. This implies that $I(X;Y) = H(Y) - H(Y \mid X)$ is maximum under the constraints (20) and (21) when $H(Y)$ is maximum, i.e., under the distribution:

$$P(Y = 10) = P(Y = 01) = \frac{x}{2},$$
$$P(Y = 00) = P(Y = 11) = \frac{1 - x}{2}.$$

Therefore, we obtain

$$I(X;Y) \leq 1 + \mathsf{h}(x) - (2\tau + x)\,\mathsf{h}\left(\frac{2\tau}{2\tau + x}\right), \tag{22}$$

which together with (18) gives

$$R_f(\tau) \leq \frac{1}{2}\Bigg[1 + \mathsf{h}(f_{10} + f_{01} - 2\tau) - (f_{10} + f_{01})\,\mathsf{h}\left(\frac{2\tau}{f_{10} + f_{01}}\right)\Bigg].$$

The right hand side of (22) is maximized for $x = 1/2 - \tau$, thus yielding the unconstrained upper bound stated in the theorem below.

Theorem VI.1. *For $\tau \in [0, \frac{1}{2}]$, we have*

$$R(\tau) \leq \frac{1}{2}\left(1 + \mathsf{h}\left(\frac{1}{2} - \tau\right) - \left(\frac{1}{2} + \tau\right)\mathsf{h}\left(\frac{2\tau}{\frac{1}{2} + \tau}\right)\right).$$

The upper bounds of Theorems V.2 and VI.1 are plotted in Figure 2. For comparison, also plotted are upper and lower bounds from [6, Figure 1], and the upper bound of Sharov and Roth [9, Theorem 3.2], which were the best bounds in the prior literature. The bounds of Theorem V.2 and VI.1 improve upon the previous bounds. For $\tau \in [0, 0.1103]$, the bound of Theorem V.2 beats all the other upper bounds, and

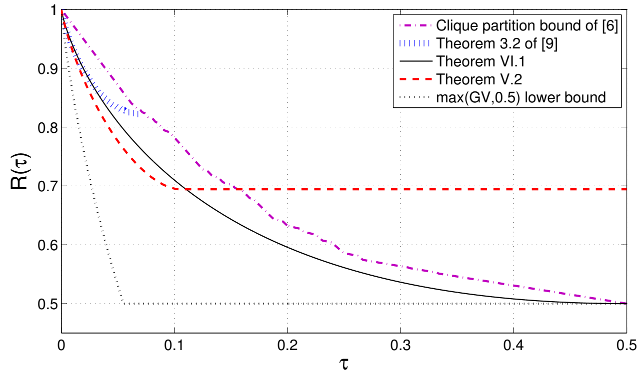

Fig. 2. The upper bounds of Theorems V.2 and VI.1, along with bounds from [6] and [9].

for $\tau \in [0.1104, 0.5]$, the bound of Theorem VI.1 takes over as the best upper bound. However, all upper bounds still remain far from the lower bound plotted. It should be pointed that a slightly better lower bound was found by Sharov and Roth [9, Theorem 2.1]. Unfortunately, the improvement is only minor: the lower bound of [9] remains above 0.5 only in the interval $[0, 0.0566]$, and in that interval, the improvement does not exceed 0.012 [9, Section II-C].

VII. Concluding Remarks

In this paper, we derived upper bounds on the maximum cardinality, $M(n,t)$, of a binary t-grain-correcting code of blocklength n, and also on the asymptotic rate $R(\tau)$. In nearly all cases, the gap between the upper bound and the best known lower bound remains significant. A natural question to ask is whether the putative upper bounds on $M(n,t)$ in Conjectures III.1 and IV.1 would yield a better bound on $R(\tau)$.

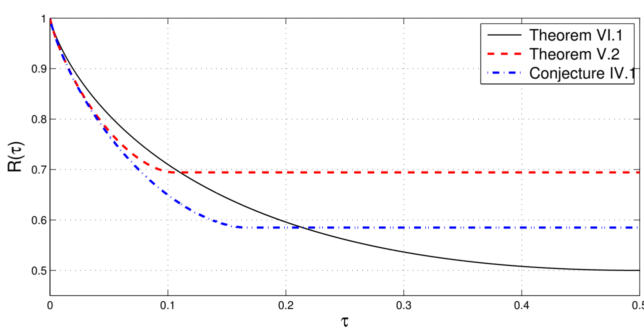

Fig. 3. The upper bounds of Theorems V.2 and VI.1 compared with the asymptotic bound obtained from Conjecture IV.1.

The bound in Conjecture IV.1 is the stronger of the two conjectured bounds, and its asymptotics are straightforward to analyze. Let $\overline{M}(n,t) = 2\sum_{\omega=0}^{n-1}\binom{n-1}{\omega}\frac{1}{V(\omega,t)}$ denote the upper bound in (15), and let $\overline{R}(\tau) = \lim_{n\to\infty}\frac{1}{n}\log_2 \overline{M}(n,n\tau)$. Conjecture IV.1 implies that $R(\tau) \le \overline{R}(\tau)$. By standard asymptotic analysis, we obtain

$$\overline{R}(\tau) = \max_{0\le\nu\le 1}[\mathsf{h}(\nu) - \eta(\nu)],$$

where $\eta(\nu)$ equals ν if $\nu \le 2\tau$, and equals $\nu\mathsf{h}(\tau/\nu)$ otherwise. Thus,

$$\overline{R}(\tau) = \max\left\{\max_{0\le\nu\le 2\tau}[\mathsf{h}(\nu)-\nu],\ \max_{2\tau\le\nu\le 1}[\mathsf{h}(\nu)-\nu\mathsf{h}(\tau/\nu)]\right\}.$$

Using elementary calculus to solve the two maximization problems within the braces in the above equation, and comparing the solutions (details of these calculations are omitted), we obtain the following:

$$\overline{R}(\tau) = \begin{cases}\mathsf{h}(\nu^*) - \nu^*\mathsf{h}(\tau/\nu^*) & \text{if } \tau < 1/6\\ \mathsf{h}(1/3) - 1/3 & \text{if } \tau \ge 1/6\end{cases} \tag{23}$$

where $\nu^* = \frac{1}{4}(\tau + 1 + \sqrt{\tau^2 - 6\tau + 1})$. Numerically, $\mathsf{h}(1/3) - 1/3 \approx 0.5850$. This bound is compared with the bounds of Theorem VI.1 in Figure 3. The plot shows that the conjectured upper bound (23) is (expectedly) better than the bound of Theorem V.2, and improves upon the bound of Theorem VI.1 for $\tau < 0.214$.

Appendix A

In this appendix, we give a proof of Theorem IV.1. The idea, of course, is to show that for $t = 1, 2, 3$, the function $\widetilde{\mathsf{w}}_t$ defined in (13) is a fractional covering of $\mathcal{H}_{n,t}$, i.e., that (14) holds. For this, we need to understand how, for $\mathbf{y} \in \Phi_t(\mathbf{x})$, the distribution of 1s changes in going from $\mathbf{x}'$ to $\mathbf{y}'$.

A. *Effect of Grains on the Derivative Sequence*

Recall that 1s in $\mathbf{x}'$ correspond to run boundaries in $\mathbf{x}$. We say that a (length-2) grain *acts on* a 1 in $\mathbf{x}'$ if it straddles the corresponding run boundary in $\mathbf{x}$. We need to distinguish between two types of 1s in the derivative sequence $\mathbf{x}'$. A *trailing* 1 is the last 1 in a 1-run, while a *non-trailing* 1 is any 1 that is not a trailing 1. Grains act on trailing 1s in a manner different from non-trailing 1s.

A segment of $\mathbf{x}'$ that contains a trailing 1 is of the form $*10*$, or $*1$ in case the trailing 1 is a suffix of $\mathbf{x}$. Up to complementation, the corresponding segment of $\mathbf{x}$ is of the form $*011*$ or $*01$. A grain acting on the trailing 1 in $\mathbf{x}'$ straddles the 01 run boundary in $\mathbf{x}$. In the sequence $\mathbf{y}$ obtained through the action of this grain, the segment under observation becomes $*001*$ or $*00$, and the corresponding segment of the derivative sequence $\mathbf{y}'$ is $*01*$ or $*0$.

On the other hand, a non-trailing 1 in $\mathbf{x}'$ belongs to a segment of the form $*11*$; the first 1 shown is the non-trailing 1 under consideration. Again, up to complementation, the corresponding segment in $\mathbf{x}$ is of the form $*010*$. A grain acting on the non-trailing 1 in $\mathbf{x}'$ straddles the 01 run boundary shown in $\mathbf{x}$. This grain causes the segment being observed to become $*000*$ in $\mathbf{y}$, and hence $*00*$ in $\mathbf{y}'$.

To summarize, the action of a grain on a trailing 1 converts a segment of the form $*10*$ or $*1$ in $\mathbf{x}'$ to $*01*$ or $*0$ in $\mathbf{y}'$, and a grain acting on a non-trailing 1 converts a segment of the form $*11*$ in $\mathbf{x}'$ to $*00*$ in $\mathbf{y}'$. It should be clear that the bits depicted by $*$s on either side of these segments remain unchanged by the action of the grain. Note, in particular, that a grain acting on a 1 in $\mathbf{x}'$ does not increase the Hamming weight of $\mathbf{x}'$. A grain acting on a trailing 1 either leaves the

Hamming weight of $\mathbf{x}'$ unchanged, or reduces it by 1; in the case of a non-trailing 1, the Hamming weight of $\mathbf{x}'$ is always reduced by 2.

Finally, when dealing with a grain pattern containing $t > 1$ length-2 grains, since the grains are non-overlapping, the actions of individual grains can be considered independently. Thus, the discussion above immediately implies the following useful fact.

Lemma A.1. *For any $\mathbf{y} \in \Phi_t(\mathbf{x})$, we have $\omega(\mathbf{x}') - 2t \leq \omega(\mathbf{y}') \leq \omega(\mathbf{x}')$, or equivalently, $r(\mathbf{x}) - 2t \leq r(\mathbf{y}) \leq r(\mathbf{x})$.*

B. Proof of Theorem IV.1 for $t = 1$

Note first that for any $\mathbf{y} \in \Phi_1(\mathbf{x})$, we have

$$V(\omega(\mathbf{y}'), 1) = 1 + \omega(\mathbf{y}') \stackrel{\text{(a)}}{\leq} 1 + \omega(\mathbf{x}') \stackrel{\text{(b)}}{=} |\Phi_1(\mathbf{x})|,$$

where (a) above is by Lemma A.1, and (b) is by Proposition II.1(a). Therefore,

$$\sum_{\mathbf{y}\in\Phi_1(\mathbf{x})} \frac{1}{V(\omega(\mathbf{y}'), 1)} \geq \sum_{\mathbf{y}\in\Phi_1(\mathbf{x})} \frac{1}{|\Phi_1(\mathbf{x})|} = 1,$$

which proves (14) for $t = 1$.

C. Proof of Theorem IV.1 for $t = 2$

Fix $\mathbf{x} \in \Sigma^n$, and let $\omega = \omega(\mathbf{x}')$. Lemma A.1 tells us that for any $\mathbf{y} \in \Phi_2(\mathbf{x})$, the Hamming weight of $\mathbf{y}'$ must lie between $\omega - 4$ and ω. For $j = 0, 1, 2, 3, 4$, let A_j be the number of sequences $\mathbf{y} \in \Phi_2(\mathbf{x})$ such that $\omega(\mathbf{y}') = \omega - j$.

Lemma A.2. *Let m denote the number of 1-runs in $\mathbf{x}'$, and let m_1 be the number of these that are of length 1. Then,*

(a) $A_0 + A_1 = 1 + m + \binom{m}{2}$;
(b) $A_2 + A_3 = (\omega - m)(m + 1) - (m - m_1)$;
(c) $A_4 = \binom{\omega-m}{2} - (\omega - m) + (m - m_1)$.

Proof: Let $\mathbf{y} = \phi_E(\mathbf{x})$ for some $E \in \mathcal{E}_{n,2}$. Write $\mathbf{x}' = (x_2', \ldots, x_n')$. Note that $\mathbf{x}'$ contains m trailing 1s and $\omega - m$ non-trailing 1s.

(a) We have $\omega(\mathbf{y}') = \omega$ or $\omega - 1$ iff each $j \in E$ acts upon a trailing 1 of $\mathbf{x}'$. Let $J = \{j \in \{2, \ldots, n\} : x_j' \text{ is a trailing } 1\}$ be the positions of the trailing 1s in $\mathbf{x}'$. Thus, $|J| = m$, and J does not contain consecutive integers. The sequence $\mathbf{y}$ is counted by $A_0 + A_1$ iff $E \subseteq J$. The number of such grain patterns E is precisely $1 + m + \binom{m}{2}$.

(b) We have $\omega(\mathbf{y}') = \omega - 2$ or $\omega - 3$ iff exactly one $j \in E$ acts upon a non-trailing 1 in $\mathbf{x}'$. Thus, for a grain pattern $E \in \mathcal{E}_{n,2}$ to contribute to $A_2 + A_3$, exactly one grain in the pattern must act on a non-trailing 1. The number of such grain patterns E with $|E| = 1$ is precisely $\omega - m$. It remains to count the number of grain patterns E of cardinality 2 that contribute to $A_2 + A_3$. Let $E = \{i, j\}$, where i and j are the grains acting on a trailing 1 and a non-trailing 1, respectively. If i acts on an "isolated" 1, i.e., a 1-run of length 1, then j can act on any of the $\omega - m$ non-trailing 1s. On the other hand, if i acts on a trailing 1 from a 1-run of length at least 2, then j can be any of the non-trailing 1s *except* for the 1 at position $i - 1$. It follows that the number of grain patterns of cardinality 2 contributing to $A_2 + A_3$ equals $m_1(\omega - m) + (m - m_1)(\omega - m - 1)$. Thus,

$$\begin{aligned} A_2 + A_3 &= (\omega - m) + m_1(\omega - m) \\ &\quad + (m - m_1)(\omega - m - 1) \\ &= (\omega - m)(m + 1) - (m - m_1). \end{aligned}$$

(c) This part follows from the fact that $A_4 = |\Phi_2(\mathbf{x})| - \sum_{j=0}^{3} A_j$, using the expression for $|\Phi_2(\mathbf{x})|$ given in Proposition II.1(b). ■

We are now ready to prove (14). For convenience, we use $V(a)$ to denote $1 + a + \binom{a}{2}$. We start with

$$\begin{aligned} \sum_{\mathbf{y}\in\Phi_2(\mathbf{x})} \frac{1}{V(\omega(\mathbf{y}'), 2)} &\geq \frac{A_0 + A_1}{V(\omega)} + \frac{A_2 + A_3}{V(\omega - 2)} + \frac{A_4}{V(\omega - 4)} \\ &= 1 - \frac{\frac{1}{2}(\omega - m)(\omega + m + 1)}{V(\omega)} \\ &\quad + \frac{A_2 + A_3}{V(\omega - 2)} + \frac{A_4}{V(\omega - 4)}. \end{aligned} \tag{24}$$

The equality above simply uses the fact that $V(\omega) - V(m) = \frac{1}{2}(\omega - m)(\omega + m + 1)$. Now, note that $\frac{A_2+A_3}{V(\omega-2)} + \frac{A_4}{V(\omega-4)}$ is lower bounded by

$$\frac{(\omega - m)(m + 1)}{V(\omega - 2)} + \frac{\binom{\omega-m}{2} - (\omega - m)}{V(\omega - 4)},$$

which equals

$$\frac{(\omega - m)(m + 1)}{V(\omega - 2)} + \frac{\frac{1}{2}(\omega - m)(\omega - m - 3)}{V(\omega - 4)}.$$

Therefore, carrying on from (24), we have

$$\begin{aligned} &\sum_{\mathbf{y}\in\Phi_2(\mathbf{x})} \frac{1}{V(\omega(\mathbf{y}'), 2)} \\ &\geq 1 + (\omega - m)\Bigg[\frac{m + 1}{V(\omega - 2)} + \frac{\frac{1}{2}(\omega - m - 3)}{V(\omega - 4)} \\ &\qquad - \frac{\frac{1}{2}(\omega + m + 1)}{V(\omega)}\Bigg] \qquad (25) \\ &\geq 1 + (\omega - m)\Bigg[\frac{m + 1 + \frac{1}{2}(\omega - m - 3)}{V(\omega - 2)} \\ &\qquad - \frac{\frac{1}{2}(\omega + m + 1)}{V(\omega)}\Bigg] \\ &= 1 + \frac{1}{2}(\omega - m)\left[\frac{\omega + m - 1}{V(\omega - 2)} - \frac{\omega + m + 1}{V(\omega)}\right] \\ &= 1 + \frac{\frac{1}{2}(\omega - m)}{V(\omega - 2)V(\omega)}\left[\omega^2 + 2m\omega - (m + 3)\right]. \end{aligned} \tag{26}$$

If $\omega = m$, then (25) proves (14). Else, if $\omega \geq m + 1$, then the term within square brackets in (26) can be further bounded as follows:

$$\begin{aligned} \omega^2 + 2m\omega - (m + 3) &\geq (m + 1)^2 + 2m(m + 1) - (m + 3) \\ &= 3m^2 + 3m - 2, \end{aligned}$$

which is positive for $m \geq 1$. Thus, again, we have (14), which completes the proof of the $t = 2$ case.

D. Proof of Theorem IV.1 for $t = 3$

The approach is the same as that for $t = 2$, but the computations are more cumbersome. So, let $\mathbf{x} \in \Sigma^n$ be fixed, and let $\omega = \omega(\mathbf{x}')$. The Hamming weight of $\mathbf{y}'$, for any $\mathbf{y} \in \Phi_3(\mathbf{x})$, lies between $\omega - 6$ and ω. For $j = 0, 1, \ldots, 6$, let B_j be the number of $\mathbf{y} \in \Phi_3(\mathbf{x})$ such that $\omega(\mathbf{y}') = \omega - j$.

Lemma A.3. *Let m denote the number of 1-runs in $\mathbf{x}'$, and let m_i, $i = 1, 2$, be the number of these that are of length i. Then,*

(a) $B_0 + B_1 = 1 + m + \binom{m}{2} + \binom{m}{3}$;
(b) $B_2 + B_3 = (\omega - m)\big(1 + m + \binom{m}{2}\big) - m(m - m_1)$;
(c) $B_4 + B_5 = (1 + m)\left[\binom{\omega-m}{2} - (\omega - m)\right] - (\omega - 2m - 3)(m - m_1) - m_2$;
(d) $B_6 = \binom{\omega-m}{3} - (\omega - m)(\omega - m + 1) + (\omega - m)(m - m_1) + 4(\omega - 2m + m_1) + m_2$.

Proof: Let $\mathbf{y} = \phi_E(\mathbf{x})$ for some $E \in \mathcal{E}_{n,3}$.

(a) This is proved by an easy extension of the proof of Lemma A.2(a).

(b) For a grain pattern $E \in \mathcal{E}_{n,2}$ to contribute to $B_2 + B_3$, exactly one grain in the pattern must act on a non-trailing 1. The number of such grain patterns E with $|E| \leq 2$ is equal to $(\omega - m)(m+1) - (m - m_1)$ by Lemma A.2(b). Extending the arguments in the proof of Lemma A.2(b), we determine that the number of grain patterns of cardinality 3 that contribute to $B_2 + B_3$ is equal to $\binom{m_1}{2}(\omega - m) + m_1(m - m_1)(\omega - m - 1) + \binom{m-m_1}{2}(\omega - m - 2)$. Thus,

$$\begin{aligned} B_2 + B_3 = \;& (\omega - m)(m + 1) - (m - m_1) \\ & + \binom{m_1}{2}(\omega - m) + m_1(m - m_1)(\omega - m - 1) \\ & + \binom{m - m_1}{2}(\omega - m - 2), \end{aligned}$$

which simplifies to $(\omega - m)\big(1 + m + \binom{m}{2}\big) - m(m - m_1)$.

(c) This part follows from the fact that $B_4 + B_5 = |\Phi_3(\mathbf{x})| - \sum_{j=0}^{3} B_j - B_6$, using the expression for $|\Phi_3(\mathbf{x})|$ given in Proposition II.1(c).

(d) B_6 equals the number of grain patterns $E \in \mathcal{E}_{n,3}$ with $|E| = 3$, in which all three grains act on non-trailing 1s of $\mathbf{x}'$. The sequence $\mathbf{x}'$ has $m - m_1$ 1-runs of length at least 2; let $\ell_1, \ldots, \ell_{m-m_1}$ denote the lengths of these runs. Then, for $i = 1, \ldots, m - m_1$, $\ell_i^- = \ell_i - 1$ denotes the number of non-trailing 1s in these runs. With this, we can write

$$\begin{aligned} B_6 = \sum_{(i,j,k): i<j<k} \ell_i^- \ell_j^- \ell_k^- + \sum_{i=1}^{m-m_1} \binom{\ell_i^- - 1}{2} \Bigg(\sum_{j: j \neq i} \ell_j^-\Bigg) \\ + \sum_{i=1}^{m-m_1} \binom{\ell_i^- - 2}{3}. \end{aligned}$$

From this, straightforward algebraic manipulations yield the expression in the statement of the lemma. The algebra here is analogous to that needed to prove Proposition II.1(c). ■

For convenience, we define $U(a)$ to be $1 + a + \binom{a}{2} + \binom{a}{3}$. We then have

$$\sum_{\mathbf{y} \in \Phi_2(\mathbf{x})} \frac{1}{V(\omega(\mathbf{y}'), 3)} \geq \sum_{j=0}^{2} \frac{B_{2j} + B_{2j+1}}{U(\omega - 2j)} + \frac{B_6}{U(\omega - 6)}. \quad (27)$$

The aim is to show, using Lemma A.3, that the right-hand side of the above inequality is at least 1. We dispose of an easy case first. If $\omega = m$, then note that we must have $m = m_1 = \omega$, and $m_2 = 0$. With this, Lemma A.3 yields $B_0 + B_1 = U(\omega)$, and $B_2 + B_3 = B_4 + B_5 = B_6 = 0$. Hence, the right-hand side of (27) simplifies to $\frac{U(\omega)}{U(\omega)} = 1$. This proves the desired inequality (14) when $\omega = m$.

Also, for small values of ω, it can be checked by direct computation using Lemma A.3 that the right-hand side of (27) is at least 1. We used a computer to check this for $\omega \leq 16$ and all valid choices of m, m_1 and m_2. Here, "valid" means that these quantities must be realizable as the number of 1-runs of the appropriate type in a binary sequence of Hamming weight ω. In particular, m, m_1 and m_2 must be non-negative integers constrained by the relations $m_1 + m_2 \leq m$ and $w \geq m_1 + 2m_2 + 3(m - m_1 - m_2)$.

Thus, we may henceforth assume that $1 \leq m \leq \omega - 1$ and $\omega \geq 17$.

We carry out some more simplifications. The idea is to justify ignoring the terms that involve m_1 and m_2 in the formulae stated in Lemma A.3. When we expand out $\frac{B_2+B_3}{U(\omega-2)} + \frac{B_4+B_5}{U(\omega-4)} + \frac{B_6}{U(\omega-6)}$ using Lemma A.3, we obtain an expression that includes the following terms:

$$\begin{aligned} -\frac{m(m - m_1)}{U(\omega - 2)} - \frac{(\omega - 2m - 3)(m - m_1) + m_2}{U(\omega - 4)} \\ + \frac{(\omega - m)(m - m_1) + 4(\omega - 2m + m_1) + m_2}{U(\omega - 6)}. \end{aligned}$$

Re-write this as

$$\begin{aligned} & m(m - m_1)\left[\frac{1}{U(\omega - 4)} - \frac{1}{U(\omega - 2)}\right] \\ & + [(\omega - m)(m - m_1) + m_2]\left[\frac{1}{U(\omega - 6)} - \frac{1}{U(\omega - 4)}\right] \\ & + \frac{3(m - m_1)}{U(\omega - 4)} + \frac{4(\omega - 2m + m_1)}{U(\omega - 6)}. \end{aligned}$$

The above expression is a sum of four terms, each of which is non-negative. (To see that the last term is non-negative, observe that $\omega \geq m_1 + 2(m - m_1) = 2m - m_1$.) Therefore, the sum $\frac{B_2+B_3}{U(\omega-2)} + \frac{B_4+B_5}{U(\omega-4)} + \frac{B_6}{U(\omega-6)}$ is at least

$$\begin{aligned} \frac{(\omega - m)(1 + m + \binom{m}{2})}{U(\omega - 2)} + \frac{(1 + m)\left[\binom{\omega-m}{2} - (\omega - m)\right]}{U(\omega - 4)} \\ + \frac{\binom{\omega-m}{3} - (\omega - m)(\omega - m + 1)}{U(\omega - 6)}, \end{aligned}$$

which can also be expressed as

$$\begin{aligned} (\omega - m)\Bigg[\frac{1 + m + \binom{m}{2}}{U(\omega - 2)} + \frac{\frac{1}{2}(1 + m)(\omega - m - 3)}{U(\omega - 4)} \\ + \frac{\frac{1}{6}\left[(\omega - m)^2 - 9(\omega - m) - 4\right]}{U(\omega - 6)}\Bigg]. \end{aligned} \quad (28)$$

Next, we write

$$\frac{B_0+B_1}{U(\omega)} = \frac{U(m)}{U(\omega)} = 1-\frac{U(\omega)-U(m)}{U(\omega)} = 1-\frac{\frac{1}{6}(\omega-m)(\omega^2+\omega m+m^2+5)}{U(\omega)}. \quad (29)$$

Putting (28) and (29) together, we find that the right-hand side of (27) is lower bounded by

$$1+(\omega-m)g_\omega(m), \quad (30)$$

where

$$g_\omega(m) = \frac{1+m+\binom{m}{2}}{U(\omega-2)}+\frac{\frac{1}{2}(1+m)(\omega-m-3)}{U(\omega-4)} + \frac{\frac{1}{6}\left[(\omega-m)^2-9(\omega-m)-4\right]}{U(\omega-6)} - \frac{\frac{1}{6}(\omega^2+\omega m+m^2+5)}{U(\omega)}.$$

For a fixed ω, consider g_ω as a function of m. Some tedious computations (some of which were performed with the aid of Maple) show the following:

- for $\omega \geq 6$, g_ω is a convex function, i.e., $g''_\omega(x) \geq 0$ for $1 \leq x \leq \omega$;
- for $\omega \geq 12$, $g'_\omega(\omega-1) \leq 0$;
- for $\omega \geq 13$, $g_\omega(\omega-1) \geq 0$.

From this, we obtain the fact that, as long as $\omega \geq 13$, we have $g_\omega(m) \geq 0$ for $1 \leq m \leq \omega-1$. Thus, for these values of ω and m, (30) yields that the right-hand side of (27) is lower bounded by 1. Recalling that we only needed to show this for $\omega \geq 17$, the proof of the $t=3$ case in Theorem IV.1 is complete.

Appendix B

We prove Theorem V.2 here. Throughout this appendix, we set $t=\lceil \tau n\rceil$ and $r=\lceil \rho n\rceil$ for some $\tau\in[0,\frac{1}{2}]$ and $\rho\in[0,1]$.

The asymptotics of $\varphi_{n,t}(r)$ is determined by the largest term within the summation in (17). Letting $\zeta_j=\binom{r-j}{j}$, it is easy to verify that the ratio ζ_{j-1}/ζ_j is at most 1 when $j\leq \frac{1}{10}(5r+7-\sqrt{5r^2+10r+9})$, and is strictly larger than 1 for $\frac{1}{10}(5r+7-\sqrt{5r^2+10r+9}) < j \leq r/2$; for $j>r/2$, we have $\zeta_j=0$. Therefore, setting $J=\lfloor \frac{1}{10}(5r+7-\sqrt{5r^2+10r+9})\rfloor$, we see that if $t<J$, then the dominant term in (17) is ζ_t; and if $t\geq J$, the dominant term is ζ_J. Passing to asymptotics, it follows that if we define $\alpha=\frac{5-\sqrt{5}}{10}$, then

$$\lim_{n\to\infty}\frac{1}{n}\log_2\varphi_{n,\lceil\tau n\rceil}(\lceil\rho n\rceil) = \begin{cases}(\rho-\tau)\mathsf{h}\left(\frac{\tau}{\rho-\tau}\right) & \text{if } \tau\leq\alpha\rho\\ \rho(1-\alpha)\mathsf{h}(\frac{\alpha}{1-\alpha}) & \text{if } \tau\geq\alpha\rho\end{cases} \quad (31)$$

We record in the following lemma some facts about the constant $\alpha=\frac{5-\sqrt{5}}{10}$ that will be useful in the sequel. They are proved by straightforward algebraic manipulations. For ease of verification, we give a proof of part (c) at the end of this appendix.

Lemma B.1. *Recall that $\phi=\frac{1+\sqrt{5}}{2}$ is the golden ratio.*
(a) $\alpha^{-1}=\sqrt{5}\phi$
(b) $\frac{\alpha}{1-\alpha}=\frac{1}{1+\phi}$
(c) $(1-\alpha)\mathsf{h}(\frac{\alpha}{1-\alpha})=\log_2\phi$.

Resuming the proof of Theorem V.2, from (16), we obtain

$$R(\tau)\leq\max_{0\leq\rho\leq1}\left[\mathsf{h}(\rho)-\lim_{n\to\infty}\frac{1}{n}\varphi_{n,\lceil\tau n\rceil}(\lceil\rho n\rceil)\right].$$

Hence, using (31) and Lemma B.1, we have

$$R(\tau)\leq\max\{A(\tau),B(\tau)\}, \quad (32)$$

where

$$A(\tau)=\max_{0\leq\rho\leq\min\{\alpha^{-1}\tau,1\}}[\mathsf{h}(\rho)-\rho\log_2\phi], \quad (33)$$

and

$$B(\tau)=\max_{\alpha^{-1}\tau\leq\rho\leq1}\left[\mathsf{h}(\rho)-(\rho-\tau)\mathsf{h}\left(\frac{\tau}{\rho-\tau}\right)\right] \quad (34)$$

For convenience, we define $B(\tau)=0$ if $\alpha^{-1}\tau>1$. Note that the term within square brackets in (34) reduces to $\mathsf{h}(\alpha^{-1}\tau)-\alpha^{-1}\tau\log_2\phi$ if we set $\rho=\alpha^{-1}\tau$; therefore, $\mathsf{h}(\alpha^{-1}\tau)-\alpha^{-1}\tau\log_2\phi\leq B(\tau)$.

Now, using elementary calculus to solve the maximization problem in (33), we obtain

$$A(\tau)=\begin{cases}\mathsf{h}(\alpha^{-1}\tau)-\alpha^{-1}\tau\log_2\phi & \text{if } \alpha^{-1}\tau\leq\frac{1}{1+\phi}\\ \mathsf{h}\left(\frac{1}{1+\phi}\right)-\frac{1}{1+\phi}\cdot\log_2\phi & \text{if } \alpha^{-1}\tau\geq\frac{1}{1+\phi}\end{cases}$$

Somewhat miraculously, the expression $\mathsf{h}(\frac{1}{1+\phi})-\frac{1}{1+\phi}\cdot\log_2\phi$ simplifies to $\log_2\phi$ using parts (b) and (c) of Lemma B.1: replace $\frac{1}{1+\phi}$ and $\log_2\phi$ by $\frac{\alpha}{1-\alpha}$ and $(1-\alpha)\mathsf{h}(\frac{\alpha}{1-\alpha})$, respectively, and simplify. Thus, we have

$$A(\tau)=\begin{cases}\mathsf{h}(\alpha^{-1}\tau)-\alpha^{-1}\tau\log_2\phi & \text{if } \alpha^{-1}\tau\leq\frac{1}{1+\phi}\\ \log_2\phi & \text{if } \alpha^{-1}\tau\geq\frac{1}{1+\phi}\end{cases} \quad (35)$$

As a result, when $\alpha^{-1}\tau\leq\frac{1}{1+\phi}$, we have $A(\tau)=\mathsf{h}(\alpha^{-1}\tau)-\alpha^{-1}\tau\log_2\phi\leq B(\tau)$. Thus, (32) reduces to $R(\tau)\leq B(\tau)$, which proves one half of Theorem V.2.

To complete the proof of the theorem, we must show that when $\alpha^{-1}\tau\geq\frac{1}{1+\phi}$, we have $A(\tau)\geq B(\tau)$. This would then imply that $\max\{A(\tau),B(\tau)\}=A(\tau)=\log_2\phi$ by (35). The above clearly holds when $\alpha^{-1}\tau>1$, since $B(\tau)=0$ in this case; so we henceforth assume $1\geq\alpha^{-1}\tau\geq\frac{1}{1+\phi}$.

We will show that the maximum in the definition of $B(\tau)$ is achieved at $\rho=\alpha^{-1}\tau$. With this, $B(\tau)=\mathsf{h}(\alpha^{-1}\tau)-\alpha^{-1}\tau\log_2\phi\leq\max_{0\leq\rho\leq\alpha^{-1}\tau}[\mathsf{h}(\rho)-\rho\log_2\phi]=A(\tau)$.

Define $f_\tau(\rho)=\mathsf{h}(\rho)-(\rho-\tau)\mathsf{h}(\frac{\tau}{\rho-\tau})$, so that $B(\tau)=\max_{\alpha^{-1}\tau\leq\rho\leq1}f_\tau(\rho)$. We want to show that, under the assumption $1\geq\alpha^{-1}\tau\geq\frac{1}{1+\phi}$, the function $f_\tau(\rho)$ is monotonically decreasing in the range $\alpha^{-1}\tau\leq\rho\leq1$. We accomplish this by showing that $f'_\tau(\alpha^{-1}\tau)\leq0$, and $f''_\tau(\rho)<0$ for $\alpha^{-1}\tau\leq\rho\leq1$. Here, all derivatives are with respect to the variable ρ.

$\underline{f'_\tau(\alpha^{-1}\tau)\leq0}$: Computing the derivative $f'_\tau(\rho)$ by direct differentiation, then plugging in $\rho=\alpha^{-1}\tau$ and simplifying using Lemma B.1, we obtain

$$f'_\tau(\alpha^{-1}\tau)=\log_2\frac{1-\alpha^{-1}\tau}{\alpha^{-1}\tau}-\log_2\phi=g'(\alpha^{-1}\tau),$$

where g is the function defined by $g(x) = \mathsf{h}(x) - (\log_2 \phi)x$. Observe that $g(x)$ is strictly concave on $[0, 1]$, and attains its unique maximum at $x = \frac{1}{1+\phi}$. Hence, for $x \geq \frac{1}{1+\phi}$, $g'(x) \leq 0$. In particular, $g'(\alpha^{-1}\tau) \leq 0$.

$f''_\tau(\rho) < 0$ for $\alpha^{-1}\tau \leq \rho \leq 1$: Routine differentiation yields

$$f''_\tau(\rho) = -\frac{1}{1-\rho} - \frac{1}{\rho} + \frac{\tau}{(\rho-\tau)(\rho-2\tau)}.$$

For $\tau \leq \alpha\rho$, we have

$$\begin{aligned} \frac{\tau}{(\rho-\tau)(\rho-2\tau)} &\leq \frac{\alpha\rho}{(\rho-\alpha\rho)(\rho-2\alpha\rho)} \\ &= \frac{\alpha}{(1-\alpha)(1-2\alpha)} \cdot \frac{1}{\rho} \\ &= \frac{5-\sqrt{5}}{1+\sqrt{5}} \cdot \frac{1}{\rho}, \end{aligned}$$

which is strictly less than $\frac{1}{\rho}$. Hence, $f''_\tau(\rho) < -\frac{1}{1-\rho} < 0$.

This completes the proof of Theorem V.2, modulo the promised proof of Lemma B.1(c).

Proof of Lemma B.1(c): We first write $(1-\alpha)\mathsf{h}\left(\frac{\alpha}{1-\alpha}\right)$ as

$$-\alpha \log_2(\alpha(1-\alpha)) - (1-2\alpha)\log_2(1-2\alpha) + \log_2(1-\alpha).$$

Using $\alpha(1-\alpha) = \frac{1}{5}$ and $1 - 2\alpha = \frac{1}{\sqrt{5}}$, the above simplifies to

$$\frac{1}{2}\log_2 5 + \log_2(1-\alpha) \; = \; \log_2[\sqrt{5}(1-\alpha)].$$

It is easy to verify that $\sqrt{5}(1-\alpha) = \phi$. □

Acknowledgements

The authors thank Artyom Sharov and Ronny Roth for suggesting the approach described in Section V. The authors are also grateful to the organisers of the 2012 “Trends in Coding Theory” meeting in Monte Verita, Ascona, Switzerland, from which this research emerged.

References

[1] C. Berge, “Packing Problems and Hypergraph Theory: A Survey,” *Annals of Discrete Mathematics*, vol. 4, pp. 3–37, 1979.

[2] A.A. Kulkarni and N. Kiyavash, “Non-asymptotic upper bounds for deletion correcting codes,” arXiv:1211.3128, Nov. 2012.

[3] R. Gabrys, E. Yaakobi and L. Dolecek, “Correcting grain-errors in magnetic media,” *Proc. 2013 IEEE Int. Symp. Inform. Theory (ISIT 2013)*, Istanbul, Turkey, July 7–12, 2013, pp. 689–693.

[4] R. Gabrys, E. Yaakobi and L. Dolecek, “Correcting grain-errors in magnetic media,” submitted to *IEEE Trans. Inform. Theory*, 2013. Available: http://arxiv.org/abs/1307.7087

[5] A. R. Iyengar, P. H. Siegel, and J. K. Wolf, “Write channel model for bit-patterned media recording,” *IEEE Trans. Magn.*, vol. 47, no. 1, pp. 35–45, Jan. 2011.

[6] A. Mazumdar, A. Barg and N. Kashyap, “Coding for high-density recording on a 1-d granular magnetic medium,” *IEEE Trans. Inform. Theory*, vol. 57, no. 11, pp. 7403–7417, Nov. 2011.

[7] L. Pan, W.E. Ryan, R. Wood and B. Vasic, “Coding and detection for rectangular grain models,” *IEEE Trans. Magn.*, vol. 47, no. 6, pp. 1705–1711, June 2011.

[8] A. Sharov and R.M. Roth, “Bounds and constructions for granular media coding,” *Proc. 2011 IEEE Int. Symp. Inform. Theory (ISIT 2011)*, Saint Petersburg, Russia, July 31 – Aug. 5, 2011, pp. 2304–2308.

[9] A. Sharov and R.M. Roth, “Bounds and constructions for granular media coding,” *IEEE Trans. Inform. Theory*, vol. 60, no. 4, pp. 2010–2027, April 2014.

[10] R. Wood, M. Williams, A. Kavcic and J. Miles, “The feasibility of magnetic recording at 10 Terabits per square inch on conventional media,” *IEEE Trans. Magn.*, vol. 45, no. 2, pp. 917–923, Feb. 2009.