

Two-Dimensional Weight-Constrained Codes through Enumeration Bounds

ERIK ORDENTLICH*

iCompression, Inc.

2500 Walsh Avenue

Santa Clara, CA 95051, USA

eor@alum.mit.edu

RON M. ROTH†

Computer Science Department

Technion

Haifa 32000, Israel

ronny@cs.technion.ac.il

February 17, 2000

Abstract

For a rational $\alpha \in (0, 1)$, let $\mathcal{A}_{n \times m, \alpha}$ be the set of binary $n \times m$ arrays in which each row has Hamming weight αm and each column has Hamming weight αn , where αm and αn are integers. (The special case of two-dimensional balanced arrays corresponds to $\alpha = 1/2$ and even values for n and m .) The redundancy of $\mathcal{A}_{n \times m, \alpha}$ is defined by $\rho_{n \times m, \alpha} = nmH(\alpha) - \log_2 |\mathcal{A}_{n \times m, \alpha}|$, where $H(x) = -x \log_2 x - (1-x) \log_2 (1-x)$. Bounds on $\rho_{n \times m, \alpha}$ are obtained in terms of the redundancies of the sets $\mathcal{A}_{\ell, \alpha}$ of all binary ℓ -vectors with Hamming weight $\alpha \ell$, $\ell \in \{n, m\}$. Specifically, it is shown that

$$\rho_{n \times m, \alpha} \leq n\rho_{m, \alpha} + m\rho_{n, \alpha},$$

where $\rho_{\ell, \alpha} = \ell H(\alpha) - \log_2 |\mathcal{A}_{\ell, \alpha}|$, and that this bound is tight up to an additive term $O(n + \log m)$. A polynomial-time coding algorithm is presented that maps unconstrained input sequences into $\mathcal{A}_{n \times m, \alpha}$ at a rate $H(\alpha) - (\rho_{m, \alpha}/m) - (\rho_{n, \alpha}/n)$.

Keywords: Two-dimensional coding; Weight-constrained codes; Balanced codes; DC-free codes; Enumerative coding.

*Work done at Hewlett-Packard Laboratories, 1501 Page Mill Road, Palo Alto, CA 94304, USA.

†Work done in part while on sabbatical leave at Hewlett-Packard Laboratories, 1501 Page Mill Road, Palo Alto, CA 94304, USA.

1 Introduction

In currently-available magnetic and optical memory devices, data is recorded along tracks, thus treating the recording device as one-dimensional. Recent proposals for the design of optical storage—in particular holographic memory—try to take advantage of the fact that the recording device is two-dimensional (or even three-dimensional), thereby increasing the recording density [4], [10], [15]. The new approach, however, introduces new types of constraints on the data—those constraints now become multi-dimensional in nature, rather than one-dimensional. The specific constraints to be used in the recently suggested recording techniques are yet to be crystallized. Nevertheless, experiments reported on holographic memory, and experience gathered in other existing optical devices, suggest that 0's and 1's in the recorded data need to be balanced within certain areas or patterns.

This work is motivated primarily by the coding problem of two-dimensional balanced binary $n \times m$ arrays, in which each row, and respectively each column, has the same number of 0's and 1's (n and m are even). We study this problem here as part of the following more general setting: we will consider the enumeration and coding problem of binary $n \times m$ arrays in which all rows have the same Hamming weight, and so do all the columns. We will refer to such arrays as *two-dimensional weight-constrained arrays*.

By a binary n -vector we refer to a vector in \mathbb{R}^n with entries restricted to $\{0, 1\}$. Let α be a rational in the open interval $(0, 1)$ and let n be a positive integer such that αn is an integer. A binary n -vector \mathbf{v} is called α -*weighted* if the Hamming weight of \mathbf{v} is αn . We denote the set of all α -weighted binary n -vectors by $\mathcal{A}_{n,\alpha}$; clearly,

$$|\mathcal{A}_{n,\alpha}| = \binom{n}{\alpha n}.$$

Let $H : [0, 1] \rightarrow [0, 1]$ be the entropy function

$$H(x) = -x \log_2 x - (1-x) \log_2 (1-x) \tag{1}$$

and define

$$\lambda_{n,\alpha} = 2^{-nH(\alpha)} \binom{n}{\alpha n}.$$

The (*information*) *redundancy* of a subset $C \subseteq \mathcal{A}_{n,\alpha}$ is defined as $nH(\alpha) - \log_2 |C|$. In particular, the redundancy of $\mathcal{A}_{n,\alpha}$, denoted $\rho_{n,\alpha}$, is given by

$$\rho_{n,\alpha} = n - \log_2 |\mathcal{A}_{n,\alpha}| = -\log_2 \lambda_{n,\alpha}. \tag{2}$$

It is known that

$$\frac{1}{\sqrt{8n\alpha(1-\alpha)}} \leq \lambda_{n,\alpha} \leq \frac{1}{\sqrt{2\pi n\alpha(1-\alpha)}} \leq 1 \tag{3}$$

(see [9, p. 309]), where the rightmost inequality follows from $1/n \leq \alpha \leq 1 - (1/n)$. So, $\rho_{n,\alpha} \geq (1/2) \log_2(2\pi n\alpha(1-\alpha)) \geq 0$. The reference $nH(\alpha)$ in the definition of redundancy will turn out to be convenient for the purpose of this paper. Note, however, that when $\alpha = 1/2$ this definition coincides with the usual one used in coding theory.

Much is known about codes that map unconstrained input sequences to one-dimensional α -weighted binary n -vectors. The most studied case is $\alpha = 1/2$, and the respective codes go by the names DC-free or zero-disparity codes [11]. There is a known efficient encoding algorithm due to Knuth [6] which maps, in a one-to-one manner, unconstrained binary words of length $n - \log_2 n - O(\log \log n)$ into a subset of $\mathcal{A}_{n,1/2}$; the redundancy of (the range of) such a coding scheme is therefore $\log_2 n + O(\log \log n)$. By ‘efficient’ we refer to the time (and space) complexity of the encoding, which amounts in Knuth’s algorithm to $O(n)$ increments/decrements of a $\lceil \log_2 n \rceil$ -bit counter (memory trade-offs allow to reduce the redundancy to $\lceil \log_2 n \rceil$). Improvements to Knuth’s algorithm have been obtained by Al-Bassam and Bose [1] and Tallini, Capocelli, and Bose [14].

The method of enumerative coding [5], [11], when applied to $\mathcal{A}_{n,\alpha}$, allows to map unconstrained binary words of length $\lfloor nH(\alpha) - \rho_{n,\alpha} \rfloor$ into $\mathcal{A}_{n,\alpha}$. Enumerative coding results in a polynomial-time algorithm; yet, for the case $\alpha = 1/2$, the overall complexity of enumerative coding is still higher than that of Knuth’s algorithm.

Less is known about the redundancy in the two-dimensional case. By a binary $n \times m$ array we mean an $n \times m$ array whose columns are binary n -vectors. Let α be a rational in $(0, 1)$ and let n and m be positive integers such that αn and αm are integers. A binary $n \times m$ array Γ is called α -weighted if each row and column in Γ is α -weighted. We denote by $\mathcal{A}_{n \times m, \alpha}$ the set (or the code) of all α -weighted $n \times m$ arrays. The (information) redundancy of a subset $C \subseteq \mathcal{A}_{n \times m, \alpha}$ is defined by $nmH(\alpha) - \log_2 |C|$, and the redundancy of $\mathcal{A}_{n \times m, \alpha}$, denoted $\rho_{n \times m, \alpha}$, is given by

$$\rho_{n \times m, \alpha} = nmH(\alpha) - \log_2 |\mathcal{A}_{n \times m, \alpha}|.$$

Observe that since each row (say) in an α -weighted array is α -weighted, we have

$$\rho_{n \times m, \alpha} \geq n\rho_{m, \alpha} \geq 0. \tag{4}$$

Estimates on $|\mathcal{A}_{n \times m, \alpha}|$ exist in the literature for the case where α goes to zero (or one) as n and m go to infinity. See [3, p. 48] and the references therein (e.g., [2]).

For the case $\alpha = 1/2$, an efficient coding algorithm into a subset of $\mathcal{A}_{n \times m, 1/2}$ is presented in [13] that has redundancy $n \log_2 m + m \log_2 n + O(n + m \log \log n)$. In its simpler version, the algorithm in [13] balances the rows using one of the algorithms in [1], [6], or [14]; by trading those algorithms with the (more computationally complex)

enumerative coding of $\mathcal{A}_{m,1/2}$, the redundancy can be reduced to $\frac{1}{2}(n \log_2 m) + m \log_2 n + O(n + m \log \log n)$.

In Section 2 we prove the upper bound

$$\rho_{n \times m, \alpha} \leq n \rho_{m, \alpha} + m \rho_{n, \alpha} , \quad (5)$$

and in Section 3 we show that the bound is tight up to an additive term $O(n + \log m)$ whenever $\alpha(1-\alpha)m$ is at least some absolute constant. The bound (5) implies that requiring α -weighted rows in a binary array does not “interfere” with requiring α -weighted columns. Note, however, that those requirements are not independent: for instance, if all n rows in a binary $n \times m$ array are α -weighted, and $m-1$ of the columns are α -weighted as well, then so must be the remaining column.

From (3), (4), and (5) it follows that

$$0 \leq \rho_{n \times m, \alpha} \leq \frac{1}{2} \left(n \log_2(2m) + m \log_2(2n) \right) .$$

Therefore, for every fixed rational $\alpha \in (0, 1)$ we have

$$\lim_{n, m \rightarrow \infty} \frac{\log_2 |\mathcal{A}_{n \times m, \alpha}|}{nmH(\alpha)} = 1 ,$$

where the limit is taken over integers n and m such that αn and αm are integers.

In Section 4, we present a fixed-rate lossless coding scheme into a subset of $\mathcal{A}_{n \times m, \alpha}$ with redundancy at most $n \rho_{m, \alpha} + m \rho_{n, \alpha}$; this corresponds to an encoding rate of at least $H(\alpha) - (\rho_{m, \alpha}/m) - (\rho_{n, \alpha}/n)$. In particular, for $\alpha = 1/2$, the redundancy of our encoder is smaller than the redundancies of the encoders in [13]. The new coding scheme is based on a modification of the enumerative coding technique and can be implemented by a polynomial-time algorithm (yet, for $\alpha = 1/2$, the complexity of our algorithm will be higher than the algorithms in [13]).

2 Upper bound on the redundancy of $\mathcal{A}_{n \times m, \alpha}$

In this section, we prove the following upper bound on $\rho_{n \times m, \alpha}$.

Proposition 2.1 *Let α be a rational in $(0, 1)$ and let n and m be positive integers such that αn and αm are integers. Then,*

$$\rho_{n \times m, \alpha} \leq n \rho_{m, \alpha} + m \rho_{n, \alpha} .$$

Proposition 2.1 is a direct corollary of the following lower bound on the size of $\mathcal{A}_{n \times m, \alpha}$, combined with (2).

Proposition 2.2 *Let α be a rational in $(0, 1)$ and let n and m be positive integers such that αn and αm are integers. Then,*

$$|\mathcal{A}_{n \times m, \alpha}| \geq 2^{nmH(\alpha)} \lambda_{m, \alpha}^n \lambda_{n, \alpha}^m .$$

Our proof of Proposition 2.2 is based on the following idea. Regard the entries of an $n \times m$ binary array as independent Bernoulli random variables taking on $\{0, 1\}$, where each entry equals 1 with probability α . Then, as we show, the event that all the rows are α -weighted and the event that all the columns are α -weighted are positively related. In other words, conditioning on all the rows being α -weighted increases the likelihood that all the columns are α -weighted as well.

A key ingredient in the proof is the next lemma. Denote the set $\{1, 2, \dots, n\}$ by $\langle n \rangle$.

Lemma 2.3 *Let X_1, \dots, X_n be independent Bernoulli random variables taking on $\{0, 1\}$ with probabilities $\text{Prob}\{X_i = 1\} = p_i$, $i \in \langle n \rangle$, and suppose that $\sum_{i=1}^n p_i = \alpha n$, where $\alpha \in (0, 1)$ and αn is an integer. Then,*

$$\text{Prob} \left\{ \sum_{i=1}^n X_i = \alpha n \right\} \geq \lambda_{n, \alpha} ,$$

with equality holding if and only if $p_i = \alpha$ for all i .

Lemma 2.3 follows from a result due to Hoeffding [8]. For the sake of completeness, we provide a proof of Lemma 2.3 in Appendix A (see Proposition A.1 therein). The proof we present is different and simpler than the one in [8], as Lemma 2.3, which is what we need here, is less general than Hoeffding's result.

We introduce some notations that will be used hereafter in this work.

Let Γ be a binary $n \times m$ array. The *row type* of Γ is an integer n -vector $\mathbf{w} = (w_1, \dots, w_n)$ where w_i is the sum of the entries of the i th row of Γ .

For an integer n -vector $\mathbf{w} = (w_1, \dots, w_n)$, define $\mathcal{R}_m(\mathbf{w})$ to be the set of all binary $n \times m$ arrays whose row type is \mathbf{w} . Clearly,

$$|\mathcal{R}_m(\mathbf{w})| = \prod_{i=1}^n \binom{m}{w_i}$$

(we define $\binom{m}{w} = 0$ if $w < 0$ or $w > m$).

Let α be a rational in $(0, 1)$ and let n be a positive integer such that αn is an integer. For an integer n -vector \mathbf{w} , denote by $\mathcal{U}_{m,\alpha}(\mathbf{w})$ the set of all arrays in $\mathcal{R}_m(\mathbf{w})$ whose columns are α -weighted. If m is a positive integer such that αm is an integer, then $\mathcal{A}_{n \times m, \alpha} = \mathcal{U}_{m,\alpha}(\alpha m \cdot \mathbf{1}_n)$, where $\mathbf{1}_n$ denotes the all-one vector in \mathbb{R}^n .

For a real vector \mathbf{y} , we denote by $\|\mathbf{y}\| = \|\mathbf{y}\|_1$ the sum of the absolute values of the entries of \mathbf{y} and by $\|\mathbf{y}\|_\infty$ the largest absolute value of any entry of \mathbf{y} .

Proposition 2.2 is a special case of the following lemma.

Lemma 2.4 *Let α be a rational in $(0, 1)$ and let n be a positive integer such that αn is an integer. Also, let m be a positive integer and \mathbf{w} be an integer n -vector with $\|\mathbf{w}\| = \alpha n m$. Then,*

$$|\mathcal{U}_{m,\alpha}(\mathbf{w})| \geq \lambda_{n,\alpha}^m \cdot |\mathcal{R}_m(\mathbf{w})| .$$

Proof. Let P be the measure on binary $n \times m$ arrays where the entries in an array Γ are independent Bernoulli trials with probability α of taking the value 1. Consider the conditional measure $Q(\Gamma) = P(\Gamma | \Gamma \in \mathcal{R}_m(\mathbf{w}))$. Since all the elements in $\mathcal{R}_m(\mathbf{w})$ have the same number of 1's, the measure Q is uniform on $\mathcal{R}_m(\mathbf{w})$; namely,

$$Q(\Gamma) = \begin{cases} |\mathcal{R}_m(\mathbf{w})|^{-1} & \text{if } \Gamma \in \mathcal{R}_m(\mathbf{w}) \\ 0 & \text{otherwise} \end{cases} . \quad (6)$$

In order to prove the lemma, it suffices to show that, with respect to the measure Q ,

$$\text{Prob}_Q \{ \Gamma \in \mathcal{U}_{m,\alpha}(\mathbf{w}) \} \geq \lambda_{n,\alpha}^m . \quad (7)$$

(Note that the right-hand side of (7) is the probability, with respect to the (unconditional) measure P , that all the columns are α -weighted. The inequality (7) is what we referred to earlier in saying that conditioning on $\Gamma \in \mathcal{R}_m(\mathbf{w})$ —in particular, conditioning on all the rows being α -weighted—increases the likelihood that all the columns are α -weighted.)

For $j \in \langle m \rangle$, let \mathbf{c}_j denote the j th column of the random array $\Gamma \in \mathcal{R}_m(\mathbf{w})$, and let \mathcal{B}_j denote the event that \mathbf{c}_j is α -weighted. The first (and main) step in our proof is showing that for $j \in \langle m \rangle$,

$$\text{Prob}_Q \left\{ \mathcal{B}_j \mid \mathbf{c}_1 = \mathbf{v}_1, \dots, \mathbf{c}_{j-1} = \mathbf{v}_{j-1} \right\} \geq \lambda_{n,\alpha} , \quad (8)$$

where $(\mathbf{v}_1, \dots, \mathbf{v}_{j-1})$ is any $(j-1)$ -tuple of α -weighted vectors in $\mathcal{A}_{n,\alpha}$ for which we have $\text{Prob}_Q \{ \mathbf{c}_1 = \mathbf{v}_1, \dots, \mathbf{c}_{j-1} = \mathbf{v}_{j-1} \} > 0$. The left-hand side of (8) is the conditional probability, implied by the measure Q , that \mathbf{c}_j is α -weighted, given that columns \mathbf{c}_1 through

\mathbf{c}_{j-1} are equal respectively to the α -weighted vectors $\mathbf{v}_1, \dots, \mathbf{v}_{j-1}$, while the right-hand side is the corresponding probability under the measure P . For $j = 1$, the inequality (8) becomes

$$\text{Prob}_Q \{ \mathcal{B}_1 \} \geq \lambda_{n,\alpha} . \quad (9)$$

To prove (8), we assume that the first $j-1$ columns of Γ are equal to $\mathbf{v}_1, \dots, \mathbf{v}_{j-1}$, and we let m_i be the number of 1's in the first $j-1$ positions of the i th row of Γ (with $m_i = 0$ if $j = 1$); note that the condition $\text{Prob}_Q \{ \mathbf{c}_1 = \mathbf{v}_1, \dots, \mathbf{c}_{j-1} = \mathbf{v}_{j-1} \} > 0$ implies that $m_i \leq w_i$ for all $i \in \langle n \rangle$. It is easy to see that

$$\text{Prob}_Q \{ \mathcal{B}_j \mid \mathbf{c}_1 = \mathbf{v}_1, \dots, \mathbf{c}_{j-1} = \mathbf{v}_{j-1} \} = \text{Prob} \left\{ \sum_{i=1}^n X_i = \alpha n \right\} , \quad (10)$$

where the X_i are independent Bernoulli random variables taking on $\{0, 1\}$ with probabilities $\text{Prob} \{ X_i = 1 \} = p_i = (w_i - m_i) / (m - j + 1)$. Note further that since \mathbf{v}_ℓ is α -weighted for every $\ell \in \langle j-1 \rangle$, then $\sum_{i=1}^n m_i = \alpha n (j-1)$. Recalling that $\|\mathbf{w}\| = \alpha n m$ we thus have,

$$\begin{aligned} \sum_{i=1}^n p_i &= \frac{1}{m-j+1} \sum_{i=1}^n (w_i - m_i) = \frac{\alpha n m}{m-j+1} - \frac{1}{m-j+1} \sum_{i=1}^n m_i \\ &= \frac{\alpha n m}{m-j+1} - \frac{\alpha n (j-1)}{m-j+1} \\ &= \alpha n . \end{aligned}$$

Incorporating this and (10) into Lemma 2.3 yields (8).

Having established (8), we next compute the probability that the first j columns of Γ are α -weighted:

$$\begin{aligned} &\text{Prob}_Q \{ \mathcal{B}_1, \dots, \mathcal{B}_{j-1}, \mathcal{B}_j \} \\ &= \sum_{(\mathbf{v}_1, \dots, \mathbf{v}_{j-1}) \in \mathcal{A}_{n,\alpha}^{j-1}} \text{Prob}_Q \{ \mathcal{B}_j, \mathbf{c}_1 = \mathbf{v}_1, \dots, \mathbf{c}_{j-1} = \mathbf{v}_{j-1} \} \\ &= \sum_{(\mathbf{v}_1, \dots, \mathbf{v}_{j-1}) \in \mathcal{A}_{n,\alpha}^{j-1}} \text{Prob}_Q \{ \mathcal{B}_j \mid \mathbf{c}_1 = \mathbf{v}_1, \dots, \mathbf{c}_{j-1} = \mathbf{v}_{j-1} \} \cdot \text{Prob}_Q \{ \mathbf{c}_1 = \mathbf{v}_1, \dots, \mathbf{c}_{j-1} = \mathbf{v}_{j-1} \} \\ &\stackrel{(8)}{\geq} \lambda_{n,\alpha} \cdot \sum_{(\mathbf{v}_1, \dots, \mathbf{v}_{j-1}) \in \mathcal{A}_{n,\alpha}^{j-1}} \text{Prob}_Q \{ \mathbf{c}_1 = \mathbf{v}_1, \dots, \mathbf{c}_{j-1} = \mathbf{v}_{j-1} \} \\ &= \lambda_{n,\alpha} \cdot \text{Prob}_Q \{ \mathcal{B}_1, \dots, \mathcal{B}_{j-1} \} \end{aligned}$$

(the summations are over all $(j-1)$ -tuples of α -weighted vectors in $\mathcal{A}_{n,\alpha}$). By induction on j we thus obtain

$$\text{Prob}_Q \{ \mathcal{B}_1, \dots, \mathcal{B}_j \} \geq \lambda_{n,\alpha}^j , \quad j \in \langle m \rangle .$$

In particular,

$$\text{Prob}_Q \{\Gamma \in \mathcal{U}_{m,\alpha}(\mathbf{w})\} = \text{Prob}_Q \{\mathcal{B}_1, \dots, \mathcal{B}_m\} \geq \lambda_{n,\alpha}^m,$$

thereby yielding (7). □

3 Lower bound on the redundancy of $\mathcal{A}_{n \times m, \alpha}$

In this section, we prove the following lower bound on $\rho_{n \times m, \alpha}$.

Proposition 3.1 *Let α be a rational in $(0, 1)$ and let n and m be positive integers such that αn and αm are integers. Then,*

$$\rho_{n \times m, \alpha} \geq n\rho_{m, \alpha} + m\rho_{n, \alpha} - O(n + \log m)$$

whenever $\alpha(1-\alpha)m \geq K$ for some absolute constant K .

Observe that there is asymmetry between n and m in the bound of Proposition 3.1, so transposition of the arrays may yield a better bound. Note, however, that the presentation of the bounds here is not suitable for specific values of n and m , since we will not be explicit in the constant multipliers of the $O(\cdot)$ expressions; yet, we point out that those multipliers do *not* depend on α .

Throughout this section, we fix α to be a rational in $(0, 1)$ and let n and m be positive integers such that αn and αm are integers. We denote by $\delta_{m,\alpha}$ the value $\lfloor 2\sqrt{\alpha(1-\alpha)m} \rfloor$ and define $\mathcal{D}_{n \times m, \alpha}$ as the set of all integer n -vectors \mathbf{w} such that $\|\mathbf{w}\| = \alpha mn$ and $\|\mathbf{w} - \alpha m \cdot \mathbf{1}_n\|_\infty \leq \delta_{m,\alpha}$. It can be readily verified that $\delta_{m,\alpha} \leq \min\{\alpha m, (1-\alpha)m\}$; hence, every vector $\mathbf{w} \in \mathcal{D}_{n \times m, \alpha}$ is nonnegative and $\|\mathbf{w}\|_\infty \leq m$. Let \mathbf{w}_{\min} be a vector $\mathbf{w} \in \mathcal{D}_{n \times m, \alpha}$ for which $|\mathcal{U}_{m,\alpha}(\mathbf{w})|$ is minimal, and define $\tau_{n \times m, \alpha}$ by

$$\tau_{n \times m, \alpha} = nmH(\alpha) - \log_2 |\mathcal{U}_{m,\alpha}(\mathbf{w}_{\min})|.$$

The proof of Proposition 3.1 will be carried out through a sequence of lemmas. The first two lemmas lead to a lower bound on $\tau_{n \times m, \alpha}$, and the remaining lemmas provide a lower bound on $\rho_{n \times m, \alpha}$ in terms of $\tau_{n \times m, \alpha}$.

Lemma 3.2

$$|\mathcal{D}_{n \times m, \alpha}| \geq \frac{(2\delta_{m,\alpha} + 1)^{n-1}}{n-1}.$$

Proof. Let $\mathcal{X}_{(n-1) \times m, \alpha}$ denote the set of all integer $(n-1)$ -vectors $\mathbf{v} = (v_1, \dots, v_{n-1})$ such that $\|\mathbf{v} - \alpha m \cdot \mathbf{1}_{n-1}\|_\infty \leq \delta_{m, \alpha}$. For such a vector \mathbf{v} and an index $i \in \langle n-1 \rangle$, let \mathbf{v}_i denote the vector $(2\alpha m - v_1, \dots, 2\alpha m - v_i, v_{i+1}, \dots, v_{n-1})$; namely, \mathbf{v}_i is obtained from \mathbf{v} by changing the first i entries into the respective entries in $2\alpha m \cdot \mathbf{1}_{n-1} - \mathbf{v}$. Generalizing the balancing technique of Knuth in [6], it can be shown that for every $\mathbf{v} \in \mathcal{X}_{(n-1) \times m, \alpha}$ there is at least one index $i \in \langle n-1 \rangle$ such that $|\|\mathbf{v}_i\| - \alpha m(n-1)| \leq \delta_{m, \alpha}$. Let $i(\mathbf{v})$ denote the first such index i and let $\mathbf{w}(\mathbf{v})$ be the n -vector obtained by appending $\alpha mn - \|\mathbf{v}_{i(\mathbf{v})}\|$ as an n th entry to $\mathbf{v}_{i(\mathbf{v})}$. The mapping

$$\mathbf{v} \mapsto \mathbf{w}(\mathbf{v})$$

sends $\mathcal{X}_{(n-1) \times m, \alpha}$ to a subset of $\mathcal{D}_{n \times m, \alpha}$. Furthermore, each element of $\mathcal{D}_{n \times m, \alpha}$ has at most $n-1$ pre-images in $\mathcal{X}_{(n-1) \times m, \alpha}$. Hence, $|\mathcal{D}_{n \times m, \alpha}| \geq |\mathcal{X}_{(n-1) \times m, \alpha}| / (n-1) = (2\delta_{m, \alpha} + 1)^{n-1} / (n-1)$. \square

Lemma 3.3

$$\begin{aligned} \tau_{n \times m, \alpha} &\geq m\rho_{n, \alpha} + (n-1) \log_2(2\delta_{m, \alpha} + 1) - \log_2(n-1) \\ &= n\rho_{m, \alpha} + m\rho_{n, \alpha} - O(\log m + \log n). \end{aligned}$$

Proof. The set of all binary $n \times m$ arrays whose columns are α -weighted can be written as $\bigcup_{\mathbf{w}} \mathcal{U}_{m, \alpha}(\mathbf{w})$, where the union is taken over all integer n -vectors \mathbf{w} . Now, $\mathcal{U}_{m, \alpha}(\mathbf{w})$ is nonempty only when $\|\mathbf{w}\| = \alpha mn$, and $\mathcal{U}_{m, \alpha}(\mathbf{w})$ and $\mathcal{U}_{m, \alpha}(\mathbf{w}')$ are disjoint when $\mathbf{w} \neq \mathbf{w}'$. So,

$$\sum_{\mathbf{w}: \|\mathbf{w}\| = \alpha nm} |\mathcal{U}_{m, \alpha}(\mathbf{w})| = \left| \bigcup_{\mathbf{w}} \mathcal{U}_{m, \alpha}(\mathbf{w}) \right| = |\mathcal{A}_{n, \alpha}|^m. \quad (11)$$

On the other hand,

$$|\mathcal{U}_{m, \alpha}(\mathbf{w}_{\min})| \leq \frac{1}{|\mathcal{D}_{n \times m, \alpha}|} \sum_{\mathbf{w} \in \mathcal{D}_{n \times m, \alpha}} |\mathcal{U}_{m, \alpha}(\mathbf{w})| \leq \frac{1}{|\mathcal{D}_{n \times m, \alpha}|} \sum_{\mathbf{w}: \|\mathbf{w}\| = \alpha nm} |\mathcal{U}_{m, \alpha}(\mathbf{w})|. \quad (12)$$

Combining (11) and (12) yields

$$|\mathcal{U}_{m, \alpha}(\mathbf{w}_{\min})| \leq \frac{|\mathcal{A}_{n, \alpha}|^m}{|\mathcal{D}_{n \times m, \alpha}|},$$

and by taking logarithms we obtain

$$\tau_{n \times m, \alpha} \geq m\rho_{n, \alpha} + \log_2 |\mathcal{D}_{n \times m, \alpha}|.$$

The result now follows from Lemma 3.2, (2), and (3). \square

Let $\mathbf{w} = (w_1, \dots, w_n)$ and $\mathbf{w}' = (w'_1, \dots, w'_n)$ be two vectors in $\mathcal{D}_{n \times m, \alpha}$. We say that $(\mathbf{w}, \mathbf{w}')$ is an *incremental pair* if the following conditions hold:

1. There are indexes $i, \ell \in \langle n \rangle$ such that $w'_i + 1 = w_i \leq \alpha m \leq w_\ell = w'_\ell - 1$.
2. $w_j = w'_j$ for all $j \in \langle n \rangle \setminus \{i, \ell\}$.

The next lemma is proved in Appendix B.

Lemma 3.4 *Let $(\mathbf{w}, \mathbf{w}')$ be an incremental pair. Then*

$$\frac{|\mathcal{U}_{m,\alpha}(\mathbf{w}')|}{|\mathcal{U}_{m,\alpha}(\mathbf{w})|} \geq 1 - O(1/\delta_{m,\alpha}) .$$

Lemma 3.5

$$\rho_{n \times m, \alpha} = \tau_{n \times m, \alpha} - O(n)$$

whenever $\alpha(1-\alpha)m \geq K$ for some absolute constant K .

Proof. Let $\mathbf{w}_0, \mathbf{w}_1, \dots, \mathbf{w}_h$ be a shortest sequence of distinct elements of $\mathcal{D}_{n \times m, \alpha}$ such that $\mathbf{w}_0 = \alpha m \cdot \mathbf{1}_n$, $\mathbf{w}_h = \mathbf{w}_{\min}$, and $(\mathbf{w}_{j-1}, \mathbf{w}_j)$ is an incremental pair for every $j \in \langle h \rangle$. One can easily see that h is bounded from above by $\delta_{m,\alpha}n/2$. Hence, by Lemma 3.4 we have

$$\frac{|\mathcal{U}_{m,\alpha}(\mathbf{w}_{\min})|}{|\mathcal{U}_{m,\alpha}(\alpha m \cdot \mathbf{1}_n)|} = \prod_{j=1}^h \frac{|\mathcal{U}_{m,\alpha}(\mathbf{w}_j)|}{|\mathcal{U}_{m,\alpha}(\mathbf{w}_{j-1})|} \geq (1 - O(1/\delta_{m,\alpha}))^{\delta_{m,\alpha}n/2} = \exp(-O(n)) ,$$

where we select the constant K so that the term $1 - O(1/\delta_{m,\alpha})$ is at least $1/2$ (say). Taking logarithms, we obtain the desired result. \square

Proof of Proposition 3.1. Combine Lemmas 3.3 and 3.5. \square

4 Coding scheme

In this section, we describe an encoding scheme for $\mathcal{A}_{n \times m, \alpha}$ with redundancy $n\rho_{m,\alpha} + m\rho_{n,\alpha}$, thus attaining the bound of Proposition 2.1; note that this redundancy corresponds to a rate $H(\alpha) - (\rho_{m,\alpha}/m) - (\rho_{n,\alpha}/n)$. Specifically, our encoder, which we denote by $\mathcal{E}_{n \times m, \alpha}$, maps any integer u in the range $0 \leq u < 2^{nmH(\alpha)}\lambda_{m,\alpha}^n\lambda_{n,\alpha}^m$ to elements of $\mathcal{A}_{n \times m, \alpha}$ in a one-to-one manner. In case the input is allowed to take the value of any binary k -vector for a given k , then k can be any positive integer in the range $k \leq nmH(\alpha) - (n\rho_{m,\alpha} + m\rho_{n,\alpha})$. In particular, by (3) we can take

$$k = \left\lfloor nmH(\alpha) - \frac{1}{2} \left(n \log_2(8\alpha(1-\alpha)m) + m \log_2(8\alpha(1-\alpha)n) \right) \right\rfloor .$$

The mapping $\mathcal{E}_{n \times m, \alpha}$ is obtained through a modification of the enumerative coding technique [5], [11, p. 117]. Specifically, our coding scheme is an extension of classical enumerative coding in the following two aspects:

1. Our coding scheme effectively applies two levels of enumerative coding: a ‘coarse’ phase, in which each column is regarded as a super-symbol, and a ‘fine’ phase, where the bit contents of each super-symbol is determined.
2. The application of (proper) enumerative coding would essentially require computing the numbers of arrays with certain row types (those row types depending on the specific encoded array). Here we compute *lower bounds* on those numbers instead, since we do not know how to obtain the exact numbers. We point out that the idea of using enumerative coding with estimates rather than exact numbers has been recently applied by Immink in [12] to speed up enumerative coding for one-dimensional constrained words.

The encoder $\mathcal{E}_{n \times m, \alpha}$ will be described through a recursive procedure, $\mathcal{E}_m(\mathbf{w})$, where m is a positive integers and \mathbf{w} is an integer n -vector such that $\|\mathbf{w}\| = \alpha m$. The input to $\mathcal{E}_m(\mathbf{w})$ is an integer u in the range $0 \leq u < \lambda_{n, \alpha}^m |\mathcal{R}_m(\mathbf{w})|$, and the output is an array $\Gamma \in \mathcal{U}_{m, \alpha}(\mathbf{w})$. Thus, we regard the procedure $\mathcal{E}_m(\mathbf{w})$ as a mapping $\mathcal{E}_m(\mathbf{w}) : u \mapsto \Gamma$ from the integer set $\{u : 0 \leq u < \lambda_{n, \alpha}^m |\mathcal{R}_m(\mathbf{w})|\}$ into $\mathcal{U}_{m, \alpha}(\mathbf{w})$. For encoding purposes, we will need this mapping to be one-to-one.

The encoder $\mathcal{E}_{n \times m, \alpha}$ is a special case of $\mathcal{E}_m(\mathbf{w})$ where $\mathbf{w} = \alpha m \cdot \mathbf{1}_n$.

Hereafter we assume the standard lexicographic ordering on $\{0, 1\}^n$. Specifically, let $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n)$ be two binary n -vectors. Then $\mathbf{x} < \mathbf{y}$ if there is an index i such that $x_\ell = y_\ell$ for $1 \leq \ell < i$ and $x_i < y_i$. Also, for every binary n -vector \mathbf{y} , we define the expression

$$\sigma(\mathbf{y}) = \lambda_{n, \alpha}^{m-1} \sum_{\mathbf{v} \in \mathcal{A}_{n, \alpha} : \mathbf{v} < \mathbf{y}} |\mathcal{R}_{m-1}(\mathbf{w} - \mathbf{v})| \quad (13)$$

(the values of m and \mathbf{w} will be clear from the context). A sum over an empty set in (13) is defined as zero. Also, we define $\mathcal{R}_0(\mathbf{0}) = \{\mathbf{0}\}$ and $\mathcal{R}_0(\mathbf{w}) = \emptyset$ if $\mathbf{w} \neq \mathbf{0}$.

The procedure $\mathcal{E}_m(\mathbf{w})$ appears in Figure 1.

To show that $\mathcal{E}_m(\mathbf{w})$ is well-defined, we need to prove that the value of u' computed in Step 3 is valid for Step 4; namely, the value of u' is in the range $0 \leq u' < \lambda_{n, \alpha}^{m-1} |\mathcal{R}_{m-1}(\mathbf{w} - \mathbf{c})|$. Suppose first that the vector \mathbf{c} found in Step 1 is not the maximal vector in $\mathcal{A}_{n, \alpha}$ (that maximal vector is $(1, 1, \dots, 1, 0, 0, \dots, 0)$). By the way \mathbf{c} is defined

1. Let $\mathbf{c} \leftarrow \max\{\mathbf{y} \in \mathcal{A}_{n,\alpha} : \sigma(\mathbf{y}) \leq u\}$, where $\sigma(\mathbf{y})$ is given by (13).
2. If $m = 1$ then let $\Gamma \leftarrow \mathbf{c}$ and return.
3. Let $u' \leftarrow u - \lceil \sigma(\mathbf{c}) \rceil$.
4. Apply $\mathcal{E}_{m-1}(\mathbf{w}-\mathbf{c})$ recursively on the input u' to produce the output $n \times (m-1)$ array Γ' .
5. Let $\Gamma \leftarrow (\mathbf{c} \Gamma')$ and return.

Figure 1: Procedure $\mathcal{E}_m(\mathbf{w})$.

we have

$$\sigma(\mathbf{c}) = \lambda_{n,\alpha}^{m-1} \sum_{\mathbf{v} \in \mathcal{A}_{n,\alpha} : \mathbf{v} < \mathbf{c}} |\mathcal{R}_{m-1}(\mathbf{w}-\mathbf{v})| \leq u < \lambda_{n,\alpha}^{m-1} \sum_{\mathbf{v} \in \mathcal{A}_{n,\alpha} : \mathbf{v} \leq \mathbf{c}} |\mathcal{R}_{m-1}(\mathbf{w}-\mathbf{v})|. \quad (14)$$

Since u is an integer, then $u \geq \sigma(\mathbf{c})$ implies $u \geq \lceil \sigma(\mathbf{c}) \rceil$. Combining this with (14) yields

$$0 \leq u' = u - \lceil \sigma(\mathbf{c}) \rceil \leq u - \sigma(\mathbf{c}) < \lambda_{n,\alpha}^{m-1} |\mathcal{R}_{m-1}(\mathbf{w}-\mathbf{c})|,$$

namely, u' is in the required range.

We now show that (14) holds also when the computed \mathbf{c} is the maximal vector in $\mathcal{A}_{n,\alpha}$. Indeed, this follows from the next result, together with the inequality $u < \lambda_{n,\alpha}^m |\mathcal{R}_m(\mathbf{w})|$.

Lemma 4.1

$$\lambda_{n,\alpha}^m |\mathcal{R}_m(\mathbf{w})| \leq \lambda_{n,\alpha}^{m-1} \sum_{\mathbf{v} \in \mathcal{A}_{n,\alpha}} |\mathcal{R}_{m-1}(\mathbf{w}-\mathbf{v})|.$$

Proof. Recall the measure Q given by (6) in the proof of Lemma 2.4. Also, as in that proof, let \mathcal{B}_1 stand for the event that the first column of a random array in $\mathcal{R}_m(\mathbf{w})$ is α -weighted. We have

$$\text{Prob}_Q \{\mathcal{B}_1\} \geq \lambda_{n,\alpha} \quad (15)$$

(see (9)). On the other hand, we also have

$$\text{Prob}_Q \{\mathcal{B}_1\} = \sum_{\mathbf{v} \in \mathcal{A}_{n,\alpha}} \sum_{\Gamma' \in \mathcal{R}_{m-1}(\mathbf{w}-\mathbf{v})} Q((\mathbf{v} \Gamma')) = \frac{\sum_{\mathbf{v} \in \mathcal{A}_{n,\alpha}} |\mathcal{R}_{m-1}(\mathbf{w}-\mathbf{v})|}{|\mathcal{R}_m(\mathbf{w})|}.$$

Combining this with (15), we obtain

$$\lambda_{n,\alpha} |\mathcal{R}_m(\mathbf{w})| \leq \sum_{\mathbf{v} \in \mathcal{A}_{n,\alpha}} |\mathcal{R}_{m-1}(\mathbf{w}-\mathbf{v})|, \quad (16)$$

which implies the desired result. Note that (16) holds also when $m = 1$: if \mathbf{w} is a binary vector, then the left-hand side of (16) equals $\lambda_{n,\alpha}$ and the right-hand side equals 1; otherwise, both sides of (16) are zero. \square

Observe that the recursion in Step 4 is applied with a row type, $\mathbf{w}-\mathbf{c}$, which satisfies $\|\mathbf{w}-\mathbf{c}\| = \alpha n(m-1)$. Furthermore, since $0 \leq u' < \lambda_{n,\alpha}^{m-1} |\mathcal{R}_{m-1}(\mathbf{w}-\mathbf{c})|$, the set $\mathcal{R}_{m-1}(\mathbf{w}-\mathbf{c})$ must be nonempty; that is, $\mathbf{w}-\mathbf{c}$ is a nonnegative vector.

The following lemma establishes the correctness of $\mathcal{E}_m(\mathbf{w})$.

Lemma 4.2 *The mapping $\mathcal{E}_m(\mathbf{w}) : u \mapsto \Gamma$ is one-to-one from the integer set $\{u : 0 \leq u < \lambda_{n,\alpha}^m |\mathcal{R}_m(\mathbf{w})|\}$ into $\mathcal{U}_{m,\alpha}(\mathbf{w})$.*

Proof. The proof is by induction on m . The case $m = 1$ is simple: recalling that $\|\mathbf{w}\| = \alpha n m = \alpha n$, we have $|\mathcal{R}_1(\mathbf{w})| = 1$ if \mathbf{w} is binary and $|\mathcal{R}_1(\mathbf{w})| = 0$ otherwise. In case \mathbf{w} is binary, we have $\sigma(\mathbf{y}) = 0$ if $\mathbf{y} \leq \mathbf{w}$ and $\sigma(\mathbf{y}) = 1$ otherwise. Therefore, the output Γ will be $\mathbf{w} \in \mathcal{U}_{1,\alpha}(\mathbf{w})$ and the mapping $\mathcal{E}_1(\mathbf{w}) : u \mapsto \Gamma$ is trivially one-to-one from $\{0\}$ into $\mathcal{U}_{1,\alpha}(\mathbf{w})$.

As for the induction step, suppose that the lemma holds for $m-1$. By the induction hypothesis, the $n \times (m-1)$ array Γ' that is computed in Step 4 is in $\mathcal{U}_{m-1,\alpha}(\mathbf{w}-\mathbf{c})$. Hence, the $n \times m$ array $\Gamma = (\mathbf{c} \Gamma')$ is in $\mathcal{U}_{m,\alpha}(\mathbf{w})$. That the mapping $\mathcal{E}_m(\mathbf{w}) : u \mapsto \Gamma$ is one-to-one follows from the fact that two distinct inputs u result in distinct vectors \mathbf{c} in Step 1 or in distinct values of u' in Step 3 (or both); by the induction hypothesis, distinct values of u' result in distinct arrays Γ' in Step 4. \square

We remark that $\mathcal{E}_m(\mathbf{w})$ could be regarded as an application of the (proper) enumerative coding technique if we had equality in (16). Still, the inequality in (16) suffices for the algorithm to work.

When viewing the procedure $\mathcal{E}_m(\mathbf{w})$ as an (extension of) enumerative coding, each column \mathbf{c} in Step 1 plays the role of one super-symbol. To obtain a polynomial-time implementation of the procedure $\mathcal{E}_m(\mathbf{w})$, we need to show how we can compute \mathbf{c} efficiently in Step 1. The computation of \mathbf{c} will be, in fact, a second, finer, phase of enumerative coding.

Let \mathbf{z} be a binary t -vector, $t \leq n$, and consider the set $\mathcal{A}_{n,\alpha}(\mathbf{z})$ of all vectors in $\mathcal{A}_{n,\alpha}$ whose t -prefix is \mathbf{z} . Namely,

$$\mathcal{A}_{n,\alpha}(\mathbf{z}) = \left\{ (\mathbf{z} \mathbf{v}') : \mathbf{v}' \in \{0, 1\}^{n-t}, \|\mathbf{v}'\| = \alpha n - \|\mathbf{z}\| \right\}.$$

Define the sum

$$\mu(\mathbf{z}) = \sum_{\mathbf{v} \in \mathcal{A}_{n,\alpha}(\mathbf{z})} |\mathcal{R}_{m-1}(\mathbf{w}-\mathbf{v})|$$

($\mu(\mathbf{z})$ depends on m and \mathbf{w}). Suppose we have an efficient way to compute $\mu(\mathbf{z})$ for any given t -vector \mathbf{z} . We show how we can find the vector \mathbf{c} in Step 1, bit by bit. Our computation makes use of the following lemma, which follows from Proposition 1 in [5].

Lemma 4.3 *For every binary n -vector $\mathbf{y} = (y_1, \dots, y_n)$,*

$$\sigma(\mathbf{y}) = \lambda_{n,\alpha}^{m-1} \cdot \sum_{1 \leq i \leq n : y_i = 1} \mu(y_1, \dots, y_{i-1}, 0).$$

Suppose that the entries c_1, \dots, c_{t-1} in \mathbf{c} have already been determined for some $1 \leq t \leq n$, and let $\mathbf{c}^{(t)}$ be the vector $(c_1, \dots, c_{t-1}, 1, 0, 0, \dots, 0)$. It is easy to see that c_t should be set to 1 if and only if $u \geq \sigma(\mathbf{c}^{(t)})$. Since we can apply Lemma 4.3 to compute $\sigma(\mathbf{c}^{(t)})$, it follows that an efficient procedure for computing $\mu(\mathbf{z})$ implies an efficient way for computing the bits of \mathbf{c} .

We next describe how we can compute $\mu(\mathbf{z})$ for $\mathbf{z} = (z_1, \dots, z_t) \in \{0, 1\}^t$. Let Q be as in (6) and let $\mathcal{B}_1(\mathbf{z})$ denote the event that the first column of a random array in $\mathcal{R}_m(\mathbf{w})$ is in $\mathcal{A}_{n,\alpha}(\mathbf{z})$. One can easily verify that

$$\text{Prob}_Q \{\mathcal{B}_1(\mathbf{z})\} = \text{Prob} \{(X_1, \dots, X_n) \in \mathcal{A}_{n,\alpha}(\mathbf{z})\},$$

where the X_i are independent Bernoulli random variables taking on $\{0, 1\}$ with probabilities $\text{Prob} \{X_i = 1\} = p_i = w_i/m$. That is,

$$\begin{aligned} \text{Prob}_Q \{\mathcal{B}_1(\mathbf{z})\} &= \sum_{(v_1, \dots, v_n) \in \mathcal{A}_{n,\alpha}(\mathbf{z})} \prod_{i=1}^n p_i^{v_i} (1-p_i)^{1-v_i} \\ &= S_{\alpha n - \|\mathbf{z}\|}(p_{t+1}, \dots, p_n) \cdot \prod_{i=1}^t p_i^{z_i} (1-p_i)^{1-z_i}, \end{aligned} \quad (17)$$

where $S_k(p_{t+1}, \dots, p_n)$ is the coefficient of x^k in the polynomial

$$S(x; p_{t+1}, \dots, p_n) = \prod_{i=t+1}^n (1-p_i + p_i x). \quad (18)$$

On the other hand,

$$\text{Prob}_Q \{\mathcal{B}_1(\mathbf{z})\} = \sum_{\mathbf{v} \in \mathcal{A}_{n,\alpha}(\mathbf{z})} \sum_{\Gamma' \in \mathcal{R}_{m-1}(\mathbf{w}-\mathbf{v})} Q((\mathbf{v}, \Gamma')) = \frac{\mu(\mathbf{z})}{|\mathcal{R}_m(\mathbf{w})|}.$$

Combining this with (17) we obtain the following formula for computing $\mu(\mathbf{z})$:

$$\mu(\mathbf{z}) = |\mathcal{R}_m(\mathbf{w})| \cdot S_{\alpha n - \|\mathbf{z}\|}(p_{t+1}, \dots, p_n) \cdot \prod_{i=1}^t p_i^{z_i} (1-p_i)^{1-z_i} . \quad (19)$$

(When substituting $p_i = w_i/m$ we can cancel out fractions in (19) to make the computation over the integers.)

The procedure in Figure 2 summarizes the computation of \mathbf{c} in Step 1 of $\mathcal{E}_m(\mathbf{w})$; it can be verified that the procedure can be implemented in polynomial time.

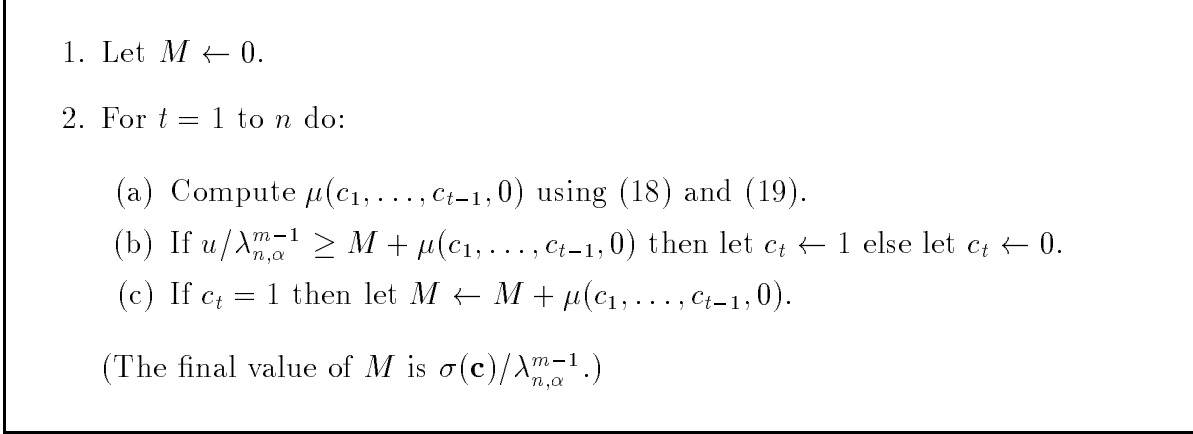


Figure 2: Detailed description of Step 1 in $\mathcal{E}_m(\mathbf{w})$.

Decoding of an array Γ that was produced by $\mathcal{E}_m(\mathbf{w})$ is simple: write $\Gamma = (\mathbf{c} \Gamma')$ and apply the decoder of $\mathcal{E}_{m-1}(\mathbf{w} - \mathbf{c})$ to Γ' to produce the respective input u' . Compute $\sigma(\mathbf{c})$ using Lemma 4.3 and, finally, let $u \leftarrow u' + \lceil \sigma(\mathbf{c}) \rceil$.

A Appendix

For a vector $\mathbf{p} = (p_1, \dots, p_n) \in \mathbb{R}^n$ and an integer k , $0 \leq k \leq n$, define

$$S_k(\mathbf{p}) = \sum_{I \subseteq \langle n \rangle : |I|=k} \prod_{i \in I} p_i \prod_{i \in \langle n \rangle \setminus I} (1-p_i) ,$$

and $S_k(\mathbf{p}) = 0$ if $k < 0$ or $k > n$. The closed unit n -dimensional real hyper-cube $[0, 1] \times [0, 1] \times \dots \times [0, 1]$ will be denoted by $[0, 1]^n$, and the respective open hyper-cube

will be denoted by $(0, 1)^n$. We also define $\mathcal{C}_k^{(n)} = \{\mathbf{p} \in [0, 1]^n : \|\mathbf{p}\| = k\}$ where (as before) $\|\mathbf{p}\| = \sum_{i=1}^n p_i$.

The quantity $S_k(\mathbf{p})$ equals $\text{Prob}\{\sum_{i=1}^n X_i = k\}$, where X_1, \dots, X_n are independent Bernoulli random variables taking on $\{0, 1\}$ with $\text{Prob}\{X_i = 1\} = p_i$. In particular, $S_k(\mathbf{p}) \leq 1$ for every $\mathbf{p} \in [0, 1]^n$.

Recalling the definition of the entropy function in (1), we prove the following result, which is a restatement of Lemma 2.3.

Proposition A.1 *Let k and n be integers, $0 \leq k \leq n$. For every $\mathbf{p} \in \mathcal{C}_k^{(n)}$,*

$$S_k(\mathbf{p}) \geq \binom{n}{k} \left(\frac{k}{n}\right)^k \left(\frac{n-k}{n}\right)^{n-k} = \binom{n}{k} \cdot 2^{-nH(k/n)},$$

with equality holding if and only if $\mathbf{p} = (k/n) \cdot \mathbf{1}_n$.

Lemma A.2 *For $0 \leq k \leq n$, let $\frac{\partial}{\partial p_i} S_k(\mathbf{p})$ be the partial derivative with respect to p_i of the mapping $(p_1, \dots, p_n) \mapsto S_k(p_1, \dots, p_n)$ when defined over \mathbb{R}^n . Then,*

$$\sum_{i=1}^n p_i(1-p_i) \frac{\partial S_k(\mathbf{p})}{\partial p_i} = (k - \|\mathbf{p}\|) S_k(\mathbf{p}).$$

Proof. For a vector $\mathbf{p} = (p_1, \dots, p_n) \in \mathbb{R}^n$, define the generating polynomial $S(x; \mathbf{p})$ in the indeterminate x by

$$S(x; \mathbf{p}) = \sum_{k=0}^n S_k(\mathbf{p}) \cdot x^k.$$

The generating polynomial can also be written as

$$S(x; \mathbf{p}) = \prod_{i=1}^n (1 - p_i + p_i x).$$

Taking partial derivatives of $S(x; \mathbf{p})$ with respect to p_i yields

$$\frac{\partial}{\partial p_i} S(x; \mathbf{p}) = (x-1) \prod_{\substack{1 \leq j \leq n \\ j \neq i}} (1 - p_j + p_j x) = \frac{(x-1)S(x; \mathbf{p})}{1 - p_i + p_i x}, \quad i \in \langle n \rangle. \quad (20)$$

Multiplying (20) by $p_i(1-p_i)$ and summing over i , we obtain

$$\sum_{i=1}^n p_i(1-p_i) \frac{\partial}{\partial p_i} S(x; \mathbf{p}) = \sum_{i=1}^n \frac{p_i(1-p_i)(x-1)S(x; \mathbf{p})}{1 - p_i + p_i x} \quad (21)$$

$$\begin{aligned}
&= \sum_{i=1}^n \frac{(p_i x - p_i(1-p_i + p_i x))S(x; \mathbf{p})}{1-p_i + p_i x} \\
&= x \cdot \sum_{i=1}^n \frac{p_i S(x; \mathbf{p})}{1-p_i + p_i x} - \sum_{i=1}^n p_i S(x; \mathbf{p}) \\
&= x \frac{\partial}{\partial x} S(x; \mathbf{p}) - \|\mathbf{p}\| S(x; \mathbf{p}) .
\end{aligned} \tag{22}$$

The lemma follows by equating the coefficient of x^k in (22) to its counterpart in the left-hand side of (21). \square

Lemma A.3 [7, p. 52] *For $r \in \langle n \rangle$ and any vector $\mathbf{p} \in (0, 1)^n$,*

$$(S_{r-1}(\mathbf{p}))^2 > S_{r-2}(\mathbf{p}) \cdot S_r(\mathbf{p}) .$$

Proof of Proposition A.1. The cases $k \in \{0, n\}$ are trivial since $|\mathcal{C}_0^{(n)}| = |\mathcal{C}_n^{(n)}| = 1$. Therefore, we assume from now on that $0 < k < n$. The set $\mathcal{C}_k^{(n)}$ is compact; so, the mapping $\mathbf{p} \mapsto S_k(\mathbf{p})$ over $\mathcal{C}_k^{(n)}$ attains a minimum (with value less than 1) at some point $\mathbf{q} = (q_1, \dots, q_n) \in \mathcal{C}_k^{(n)}$. Without loss of generality we can assume that $q_i \in (0, 1)$ for $i \in \langle m \rangle$ and $q_i \in \{0, 1\}$ for $i \in \langle n \rangle \setminus \langle m \rangle$; note that $m > 0$ (or else $S_k(\mathbf{q})$ would be 1). We denote by \mathbf{q}' and \mathbf{q}'' the vectors (q_1, \dots, q_m) and $(q_{m+1}, q_{m+2}, \dots, q_n)$, respectively.

Define the mapping $\Psi_k : \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$\Psi_k(p_1, \dots, p_n) = S_k(k - \sum_{j=2}^n p_j, p_2, p_3, \dots, p_n)$$

(Ψ_k does not depend on p_1 ; nevertheless, for the sake of having simpler notations we inserted p_1 as a redundant variable). For every real \mathbf{p} on the line $\|\mathbf{p}\| = k$ we have

$$\frac{\partial \Psi_k(\mathbf{p})}{\partial p_i} = \frac{\partial S_k(\mathbf{p})}{\partial p_i} - \frac{\partial S_k(\mathbf{p})}{\partial p_1}, \quad i \in \langle n \rangle . \tag{23}$$

If we fix $p_i = q_i$, $i \in \langle n \rangle \setminus \langle m \rangle$, then the mapping $\mathbf{p}' \mapsto \Psi_k(\mathbf{p}', \mathbf{q}'')$, over \mathbb{R}^m , must have a local minimum at $\mathbf{p}' = \mathbf{q}'$. Hence, by (23) we have

$$\frac{\partial S_k(\mathbf{p})}{\partial p_i} \Big|_{\mathbf{p}=\mathbf{q}} = \frac{\partial S_k(\mathbf{p})}{\partial p_1} \Big|_{\mathbf{p}=\mathbf{q}}, \quad i \in \langle m \rangle . \tag{24}$$

This, together with Lemma A.2, yields

$$\frac{\partial S_k(\mathbf{p})}{\partial p_1} \Big|_{\mathbf{p}=\mathbf{q}} \cdot \sum_{i=1}^m q_i(1-q_i) = \sum_{i=1}^n q_i(1-q_i) \frac{\partial S_k(\mathbf{p})}{\partial p_i} \Big|_{\mathbf{p}=\mathbf{q}} = (k - \|\mathbf{q}\|)S_k(\mathbf{q}) = 0 .$$

Since $\sum_{i=1}^m q_i(1-q_i) \neq 0$ we thus have

$$\frac{\partial S_k(\mathbf{p})}{\partial p_i} \Big|_{\mathbf{p}=\mathbf{q}} = \frac{\partial S_k(\mathbf{p})}{\partial p_1} \Big|_{\mathbf{p}=\mathbf{q}} = 0, \quad i \in \langle m \rangle. \quad (25)$$

We show next that $m = n$ by proving that $q_n \notin \{0, 1\}$. For a vector $\mathbf{p} = (p_1, \dots, p_n) \in \mathbb{R}^n$ and integers ℓ and i , $1 < i \leq n$, let $S_{\ell;i} = S_{\ell;i}(\mathbf{p})$ denote the expression $S_{\ell}(p_2, p_3, \dots, p_{i-1}, p_{i+1}, \dots, p_n)$. We define $S_{0;2} = 1$ if $n = 2$, and let $S_{\ell;i} = 0$ if $\ell > n-2$ or $\ell < 0$. Note that $S_{\ell;i}$ does not depend on p_1 or p_i . We have

$$S_k(\mathbf{p}) = p_1 p_i S_{k-2;i} + (p_1(1-p_i) + p_i(1-p_1)) S_{k-1;i} + (1-p_1)(1-p_i) S_{k;i}.$$

Taking partial derivatives with respect to p_1 and p_i , we obtain

$$\frac{\partial S_k(\mathbf{p})}{\partial p_1} = p_i S_{k-2;i} + (1-2p_i) S_{k-1;i} + (p_i-1) S_{k;i} \quad (26)$$

$$\frac{\partial S_k(\mathbf{p})}{\partial p_i} = p_1 S_{k-2;i} + (1-2p_1) S_{k-1;i} + (p_1-1) S_{k;i}. \quad (27)$$

Now, suppose to the contrary that $q_n = 0$ and compute $S_{k-2;n}$, $S_{k-1;n}$, and $S_{k;n}$ for $\mathbf{p} = \mathbf{q}$. By (25) and (26) we obtain

$$\frac{\partial S_k(\mathbf{p})}{\partial p_1} \Big|_{\mathbf{p}=\mathbf{q}} = S_{k-1;n} - S_{k;n} = 0 \quad (28)$$

i.e., $S_{k-1;n} = S_{k;n}$. Also, the partial derivative $\frac{\partial}{\partial p_n} \Psi_k(\mathbf{p})$ at $\mathbf{p} = \mathbf{q}$ must be nonnegative, or else we could increase q_n to some small $\epsilon > 0$ (and decrease q_1 by ϵ) to obtain a vector $\mathbf{q}_\epsilon \in \mathcal{C}_k^{(n)}$ such that $S_k(\mathbf{q}_\epsilon) < S_k(\mathbf{q})$, thereby contradicting the minimality of \mathbf{q} . Hence, by (23), (27), and (28) we have

$$\frac{\partial \Psi_k(\mathbf{p})}{\partial p_n} \Big|_{\mathbf{p}=\mathbf{q}} = \frac{\partial S_k(\mathbf{p})}{\partial p_n} \Big|_{\mathbf{p}=\mathbf{q}} = q_1 (S_{k-2;n} - S_{k-1;n}) \geq 0.$$

So, $S_{k-2;n} \geq S_{k-1;n} = S_{k;n} \geq 0$, or

$$S_{k-1;n}^2 \leq S_{k-2;n} \cdot S_{k;n}. \quad (29)$$

On the other hand, observe that $S_{\ell;n} = S_{\ell-\|\mathbf{q}''\|}(q_2, q_3, \dots, q_m)$ for every integer ℓ . Noting that $0 < \|\mathbf{q}'\| = k - \|\mathbf{q}''\| < m$ and that $(q_2, q_3, \dots, q_m) \in (0, 1)^{m-1}$, we can apply Lemma A.3 to the vector (q_2, q_3, \dots, q_m) with $r = k - \|\mathbf{q}''\|$ to obtain

$$S_{k-1;n}^2 > S_{k-2;n} \cdot S_{k;n},$$

thus contradicting (29). Hence, we cannot have $q_n = 0$.

A similar contradiction results if we assume that $q_n = 1$ (in this case, the partial derivative $\frac{\partial}{\partial p_n} \Psi_k(\mathbf{p})$ at $\mathbf{p} = \mathbf{q}$ must be nonpositive). Thus, we must have $m = n$, and \mathbf{q} is therefore a local minimum of $\mathbf{p} \mapsto \Psi_k(\mathbf{p})$. By (25), (26), and (27) it follows that the vector $(S_{k-2;i}, S_{k-1;i}, S_{k;i})^T$, when computed for $\mathbf{p} = \mathbf{q}$, belongs to the right null space of the array

$$A(q_1, q_i) = \begin{pmatrix} q_i & 1-2q_i & q_i-1 \\ q_1 & 1-2q_1 & q_1-1 \end{pmatrix}.$$

On the other hand, the vector $(1, 1, 1)^T$ is also in the right null space of $A(q_1, q_i)$. However, Lemma A.3, when applied to the vector $(q_2, q_3, \dots, q_{i-1}, q_{i+1}, \dots, q_n) \in (0, 1)^{n-2}$ with $r = k$, implies that the vectors $(S_{k-2;i}, S_{k-1;i}, S_{k;i})$ and $(1, 1, 1)$ are linearly independent. Therefore, the rank of $A(q_1, q_i)$ is less than 2, which is possible only when $q_1 = q_i$. Since i is any index between 2 and n , it follows that all the entries of \mathbf{q} are equal. And, since $\|\mathbf{q}\| = k$, we must have $q_i = k/n$ for all $i \in \langle n \rangle$. Finally, by symmetry it follows that $\mathbf{q} = (k/n) \cdot \mathbf{1}_n$ indeed satisfies (25). \square

It is worthwhile pointing out that the mappings $\mathbf{p} \mapsto S_k(\mathbf{p})$ are generally not U-convex over $\mathcal{C}_k^{(n)}$. For example, let $\mathbf{p}_1 = (.1, .1, .9, .9)$, $\mathbf{p}_2 = (0, .2, .9, .9)$, and $\mathbf{p}_3 = (.2, 0, .9, .9)$. Then $\mathbf{p}_1 = (\mathbf{p}_2 + \mathbf{p}_3)/2$, yet $S_2(\mathbf{p}_1) > (S_2(\mathbf{p}_2) + S_2(\mathbf{p}_3))/2 = S(\mathbf{p}_2)$.

B Appendix

We provide here the proof of Lemma 3.4. For a nonnegative integer vector $\mathbf{v} = (v_1, v_2)$ with $v_1 \leq v_2$, denote by $\mathcal{R}_m(\mathbf{v}, r)$ the set of all pairs of binary m -vectors $(\mathbf{y}_1, \mathbf{y}_2)$ such that $(\|\mathbf{y}_1\|, \|\mathbf{y}_2\|) = \mathbf{v}$ and $\|\mathbf{y}_1 - \mathbf{y}_2\| = 2r + v_2 - v_1$; note that $\|\mathbf{y}_1 - \mathbf{y}_2\|$ is the number of positions on which \mathbf{y}_1 and \mathbf{y}_2 differ. We have

$$|\mathcal{R}_m(\mathbf{v}, r)| = \binom{m}{v_2} \binom{v_2}{v_1 - r} \binom{m - v_2}{r}. \quad (30)$$

In particular, $\mathcal{R}_m(\mathbf{v}, r)$ is nonempty if and only if $0 \leq r \leq \min\{v_1, m - v_2\}$.

Lemma B.1 *Let $\mathbf{v} = (v_1, v_2)$ be an integer vector such that $\alpha m - \delta_{m,\alpha} \leq v_1 \leq \alpha m \leq v_2 \leq \alpha m + \delta_{m,\alpha}$. Then*

$$\frac{\sum_{r \leq \delta_{m,\alpha}^2/17} |\mathcal{R}_m(\mathbf{v}, r)|}{|\mathcal{R}_m(\mathbf{v})|} \leq 2^{-c \cdot \delta_{m,\alpha}^2}$$

for some constant $c > 0$.

Proof. Without loss of generality we assume that $\alpha \leq 1/2$. Write $\Delta = v_2 - v_1$ and $T = \lfloor \delta_{m,\alpha}^2/17 \rfloor$. We have,

$$\sum_{r \leq T} |\mathcal{R}_m(\mathbf{v}, r)| \leq \binom{m}{v_2} \sum_{k \leq 2T + \Delta} \binom{m}{k},$$

where the right-hand side of the inequality is an upper bound on the number of pairs of binary m -vectors $(\mathbf{y}_1, \mathbf{y}_2)$ such that $\|\mathbf{y}_2\| = v_2$ and $\|\mathbf{y}_1 - \mathbf{y}_2\| \leq 2T + \Delta$. Recalling that $|\mathcal{R}_m(\mathbf{v})| = \binom{m}{v_1} \binom{m}{v_2}$, we thus obtain

$$\frac{\sum_{r \leq T} |\mathcal{R}_m(\mathbf{v}, r)|}{|\mathcal{R}_m(\mathbf{v})|} \leq \frac{\sum_{k \leq 2T + \Delta} \binom{m}{k}}{\binom{m}{v_1}}.$$

We now combine this inequality with the lower bound

$$\binom{m}{v_1} \geq \frac{2^{mH(v_1/m)}}{\sqrt{8v_1(1-v_1/m)}} \geq \frac{2^{mH(v_1/m)}}{\sqrt{2m}} = 2^{mH(v_1/m) - o(m)}$$

and the following upper bound that holds for $2T + \Delta \leq m/2$,

$$\sum_{k \leq 2T + \Delta} \binom{m}{k} \leq 2^{mH((2T + \Delta)/m)}$$

(see [9, p. 310]). This yields

$$\frac{\sum_{r \leq T} |\mathcal{R}_m(\mathbf{v}, r)|}{|\mathcal{R}_m(\mathbf{v})|} \leq 2^{m(H(2(T + \delta_{m,\alpha})/m) - H(\alpha - \delta_{m,\alpha}/m)) + o(m)} \leq 2^{m(H(\alpha/2) - H(3\alpha/4)) + o(m)},$$

where the last inequality holds whenever $\delta_{m,\alpha}$ is sufficiently large so that

$$\lfloor \delta_{m,\alpha}^2/17 \rfloor + \delta_{m,\alpha} \leq \delta_{m,\alpha}^2/16$$

(recall that $\delta_{m,\alpha}^2/16$, in turn, is bounded from above by $\alpha m/4$). The result now follows by observing that when $\alpha \leq 1/2$, the difference $H(3\alpha/4) - H(\alpha/2)$ is bounded from below by $c' \cdot \alpha$ for some constant $c' > 0$. \square

For a k -vector \mathbf{y} and a nonempty subset B of $\langle k \rangle$, we denote by $(\mathbf{y})_B$ the subvector of \mathbf{y} indexed by B .

Let $\mathbf{w} = (w_1, \dots, w_n) \in \mathcal{D}_{n \times m, \alpha}$ and suppose that $w_1 \leq \alpha m \leq w_2$ (in fact, there are always $i, \ell \in \langle n \rangle$ such that $w_i \leq \alpha m \leq w_\ell$; we assume here that $i = 1$ and $\ell = 2$). We will use the notation $\mathbf{w}_{\langle 2 \rangle}$ for the vector $(\mathbf{w})_{\langle 2 \rangle} = (w_1, w_2)$. Also, the rows of an $n \times m$ array

Γ will be denoted by $[\Gamma]_1, \dots, [\Gamma]_n$. Let $(\mathbf{y}_1, \mathbf{y}_2)$ be a pair in $\mathcal{R}_m(\mathbf{w}_{\langle 2 \rangle}, r)$ and consider the set

$$\mathcal{U}_{m,\alpha}(\mathbf{w}; \mathbf{y}_1, \mathbf{y}_2) = \{\Gamma \in \mathcal{U}_{m,\alpha}(\mathbf{w}) : ([\Gamma]_1, [\Gamma]_2) = (\mathbf{y}_1, \mathbf{y}_2)\} .$$

The set of all arrays $\Gamma \in \mathcal{U}_{m,\alpha}(\mathbf{w})$ with $\|[\Gamma]_1 - [\Gamma]_2\| = 2r + w_2 - w_1$ is invariant under a fixed permutation on the columns of its elements. Therefore, the size of $\mathcal{U}_{m,\alpha}(\mathbf{w}; \mathbf{y}_1, \mathbf{y}_2)$ depends on r , but not on the particular choice of $(\mathbf{y}_1, \mathbf{y}_2) \in \mathcal{R}_m(\mathbf{w}_{\langle 2 \rangle}, r)$. We denote that size by $V_{m,\alpha}(\mathbf{w}, r)$ and prove the following result.

Lemma B.2 *Let $\mathbf{w} = (w_1, \dots, w_n) \in \mathcal{D}_{n \times m, \alpha}$ with $w_1 \leq \alpha m \leq w_2$. Then $V_{m,\alpha}(\mathbf{w}, r)$ is nondecreasing for values of r in the range*

$$0 \leq r \leq \min\{w_1, m - w_2\} . \quad (31)$$

Proof. Assume that both r and $r+1$ lie in the range (31). Let $(\mathbf{y}_1, \mathbf{y}_2) \in \mathcal{R}_m(\mathbf{w}_{\langle 2 \rangle}, r)$ be such that $(\mathbf{y}_1)_{\langle 2 \rangle} = (\mathbf{y}_2)_{\langle 2 \rangle} = (0, 1)$; the existence of such a pair follows from the assumption that $r+1$ satisfies (31). Let \mathbf{y}'_2 be the binary m -vector obtained from \mathbf{y}_2 by flipping the bits indexed by $\langle 2 \rangle$; that is, $(\mathbf{y}'_2)_{\langle 2 \rangle} = (1, 0)$ and $(\mathbf{y}'_2)_{\langle m \rangle \setminus \langle 2 \rangle} = (\mathbf{y}_2)_{\langle m \rangle \setminus \langle 2 \rangle}$. Clearly, the pair $(\mathbf{y}_1, \mathbf{y}'_2)$ is in $\mathcal{R}_m(\mathbf{w}_{\langle 2 \rangle}, r+1)$ and $V_{m,\alpha}(\mathbf{w}, r+1) = |\mathcal{U}_{m,\alpha}(\mathbf{w}; \mathbf{y}_1, \mathbf{y}'_2)|$.

Define a mapping

$$\eta : \mathcal{U}_{m,\alpha}(\mathbf{w}; \mathbf{y}_1, \mathbf{y}_2) \rightarrow \mathcal{U}_{m,\alpha}(\mathbf{w}; \mathbf{y}_1, \mathbf{y}'_2)$$

where $\Gamma' = \eta(\Gamma)$ is an $n \times m$ array obtained as follows:

1. $[\Gamma']_2 = \mathbf{y}'_2$.
2. Let U be the set of row indexes $b \in \langle n \rangle \setminus \langle 2 \rangle$ for which $([\Gamma]_b)_{\langle 2 \rangle} \in \{(0, 1), (1, 0)\}$. Denoting the first column of Γ by \mathbf{c}_1 and the first two columns of Γ' by \mathbf{c}'_1 and \mathbf{c}'_2 ,

$$(\mathbf{c}'_1)_U = \mathbf{1}_{|U|} - (\mathbf{c}'_2)_U = \chi((\mathbf{c}_1)_U) ,$$

where χ is a particular 1-1 mapping from the set of all binary $|U|$ -vectors \mathbf{y} with $\|\mathbf{y}\| = (|U|/2) + 1$ into the set $\mathcal{A}_{|U|, 1/2}$ (one can verify that $\|(\mathbf{c}_1)_U\| = (|U|/2) + 1$).

3. The remaining entries of Γ' (including the row $[\Gamma']_1$) are the same as in Γ .

It is easy to check that η is 1-1 and that Γ' is in $\mathcal{U}_{m,\alpha}(\mathbf{w}; \mathbf{y}_1, \mathbf{y}'_2)$. Hence,

$$V_{m,\alpha}(\mathbf{w}, r) = |\mathcal{U}_{m,\alpha}(\mathbf{w}; \mathbf{y}_1, \mathbf{y}_2)| \leq |\mathcal{U}_{m,\alpha}(\mathbf{w}; \mathbf{y}_1, \mathbf{y}'_2)| = V_{m,\alpha}(\mathbf{w}, r+1) ,$$

as desired. □

Proof of Lemma 3.4 Assume without loss of generality that \mathbf{w} and \mathbf{w}' are such that $w'_1 + 1 = w_1 \leq \alpha m \leq w_2 = w'_2 - 1$. Write $\Delta = w_2 - w_1$, and let $\mathcal{U}_{m,\alpha}(\mathbf{w}, r)$ be the set of all arrays in $\mathcal{U}_{m,\alpha}(\mathbf{w})$ with $||[\Gamma]_1 - [\Gamma]_2|| = 2r + \Delta$. Then

$$|\mathcal{U}_{m,\alpha}(\mathbf{w}, r)| = |\mathcal{R}_m(\mathbf{w}_{\langle 2 \rangle}, r)| \cdot V_{m,\alpha}(\mathbf{w}, r).$$

Observing that $V_{m,\alpha}(\mathbf{w}, r) = V_{m,\alpha}(\mathbf{w}', r-1)$ and using (30), we obtain

$$\frac{|\mathcal{U}_{m,\alpha}(\mathbf{w}', r-1)|}{|\mathcal{U}_{m,\alpha}(\mathbf{w}, r)|} = \frac{r}{\Delta + r + 1}.$$

Letting $L = \min\{w_1, m - w_2\}$, we have

$$\begin{aligned} \frac{|\mathcal{U}_{m,\alpha}(\mathbf{w}')|}{|\mathcal{U}_{m,\alpha}(\mathbf{w})|} &= \frac{\sum_{r=1}^L |\mathcal{U}_{m,\alpha}(\mathbf{w}', r-1)|}{\sum_{r=0}^L |\mathcal{U}_{m,\alpha}(\mathbf{w}, r)|} \\ &= \frac{\sum_{r=0}^L (r/(\Delta + r + 1)) |\mathcal{U}_{m,\alpha}(\mathbf{w}, r)|}{\sum_{r=0}^L |\mathcal{U}_{m,\alpha}(\mathbf{w}, r)|} \\ &\geq \frac{\sum_{r=0}^L (r/(\Delta + r + 1)) |\mathcal{R}_m(\mathbf{w}_{\langle 2 \rangle}, r)|}{\sum_{r=0}^L |\mathcal{R}_m(\mathbf{w}_{\langle 2 \rangle}, r)|}, \end{aligned}$$

where the last step follows from the monotonicity of $|V_{m,\alpha}(\mathbf{w}, r)|$ as stated in Lemma B.2. Now, letting $T = \lfloor \delta_{m,\alpha}^2/17 \rfloor + 1$, we obtain

$$\begin{aligned} \frac{\sum_{r=0}^L (r/(\Delta + r + 1)) |\mathcal{R}_m(\mathbf{w}_{\langle 2 \rangle}, r)|}{\sum_{r=0}^L |\mathcal{R}_m(\mathbf{w}_{\langle 2 \rangle}, r)|} &\geq \frac{\sum_{r=T}^L (r/(\Delta + r + 1)) |\mathcal{R}_m(\mathbf{w}_{\langle 2 \rangle}, r)|}{\sum_{r=0}^L |\mathcal{R}_m(\mathbf{w}_{\langle 2 \rangle}, r)|} \\ &\geq \frac{T}{\Delta + T + 1} \cdot \frac{\sum_{r=T}^L |\mathcal{R}_m(\mathbf{w}_{\langle 2 \rangle}, r)|}{\sum_{r=0}^L |\mathcal{R}_m(\mathbf{w}_{\langle 2 \rangle}, r)|} \\ &= (1 - O(1/\delta_{m,\alpha})) \left(1 - 2^{-c \cdot \delta_{m,\alpha}^2}\right), \end{aligned}$$

where the last step follows from Lemma B.1. The result now follows. \square

References

- [1] S. AL-BASSAM, B. BOSE, *On balanced codes*, *IEEE Trans. Inform. Theory*, 36 (1990), 406–408.
- [2] A. BÉKÉSSY, P. BÉKÉSSY, J. KOMLÓS, *Asymptotic enumeration of regular matrices*, *Studia Sci. Math. Hungar.*, 7 (1972), 343–353.

- [3] B. BOLLOBÁS, *Random Graphs*, Academic Press, London, 1985.
- [4] D. BRADY, D. PSALTIS, *Control of volume holograms*, *J. Opt. Soc. Am. A*, 9 (1992), 1167–1182.
- [5] T.M. COVER, *Enumerative source encoding*, *IEEE Trans. Inform. Theory*, IT-19 (1973), 73–77.
- [6] D.E. KNUTH, *Efficient balanced codes*, *IEEE Trans. Inform. Theory*, 32 (1986), 51–53.
- [7] G.H. HARDY, J.E. LITTLEWOOD, G. PÓLYA, *Inequalities*, Cambridge University Press, Cambridge, 1952.
- [8] W. HOEFFDING, *On the distribution of the number of successes in independent trials*, *Ann. Math. Statist.*, 27 (1956), 713–721.
- [9] F.J. MACWILLIAMS, N.J.A. SLOANE, *The Theory of Error-Correcting Codes*, North Holland, Amsterdam, 1977.
- [10] D. PSALTIS, M.A. NEIFELD, A. YAMAMURA, S. KOBAYASHI, *Optical memory disks in optical information processing*, *Applied Optics*, 29 (1990), 2038–2057.
- [11] K.A. SCHOUHAMER IMMINK, *Coding Techniques for Digital Recorders*, Prentice Hall, New York, 1991.
- [12] K.A. SCHOUHAMER IMMINK, *A practical method for approaching the channel capacity of constrained channels*, *IEEE Trans. Inform. Theory*, 43 (1997), 1389–1399.
- [13] R. TALYANSKY, T. ETZION, R.M. ROTH, *Efficient code constructions for certain two-dimensional constraints*, *IEEE Trans. Inform. Theory*, 45 (1999), 794–799.
- [14] L.G. TALLINI, R.M. CAPOCELLI, B. BOSE, *Design of some new balanced codes*, *IEEE Trans. Inform. Theory*, 42 (1996), 790–802.
- [15] W. WEEKS IV, R.E. BLAHUT, *The capacity and coding gain of certain checkerboard codes*, *IEEE Trans. Inform. Theory*, 44 (1998), 1193–1203.