

Improved Lower Bounds on Capacities of Symmetric 2-Dimensional Constraints using Rayleigh Quotients

Erez Loidor and Brian Marcus

March 5, 2009

Abstract

A method for computing lower bounds on capacities of 2-dimensional constraints having a symmetric presentation in either the horizontal or the vertical direction is presented. The method is a generalization of the method of Calkin and Wilf (*SIAM J. Discrete Math.*, 1998). Previous best lower bounds on capacities of certain constraints are improved using the method. It is also shown how this method, as well as their method for computing upper bounds on the capacity, can be applied to constraints which are not of finite-type. Additionally, capacities of 2 families of multi-dimensional constraints are given exactly.

Index-terms: Channel capacity, constrained-coding, min-max principle.

1 Introduction.

Fix an alphabet Σ and let \mathcal{G} be a directed graph whose edges are labeled with symbols in Σ . Each path in \mathcal{G} corresponds to a finite word obtained by reading the labels of the edges of the path in sequence. The path is said to generate the corresponding word, and the set of words generated by all finite paths in the graph is called a *1-dimensional constrained system* or a *1-dimensional constraint*. Such a graph is called a *presentation* of the constraint. We say that a word satisfies the constraint if it belongs to the constrained system. One-dimensional constraints have found widespread applications in digital storage systems, where they are used to model the set of sequences that can be written reliably to a medium. A central example is the binary runlength-limited constraint, denoted $\text{RLL}(d, k)$ for nonnegative integers $0 \leq d \leq k$, consisting of all binary sequences in which the number of '0's between consecutive '1's is at least d , and each runlength of '0's has length at most k . Another 1-dimensional constraint, often used in practice, is the bounded-charge constraint, denoted $\text{CHG}(b)$, for some positive integer b ; it consists of all words $w_1 w_2 \dots w_\ell$, where $\ell = 0, 1, 2, \dots$ and each w_i is either $+1$ or -1 , such that for all $1 \leq i \leq j \leq \ell$, $|\sum_{k=i}^j w_k| \leq b$. Other examples of 1-dimensional constraints are the *EVEN* and *ODD* constraints, which contain all finite binary sequences in which the number of '0's between consecutive '1's is even and odd, respectively. Presentations for these constraints are given in Figure 1.

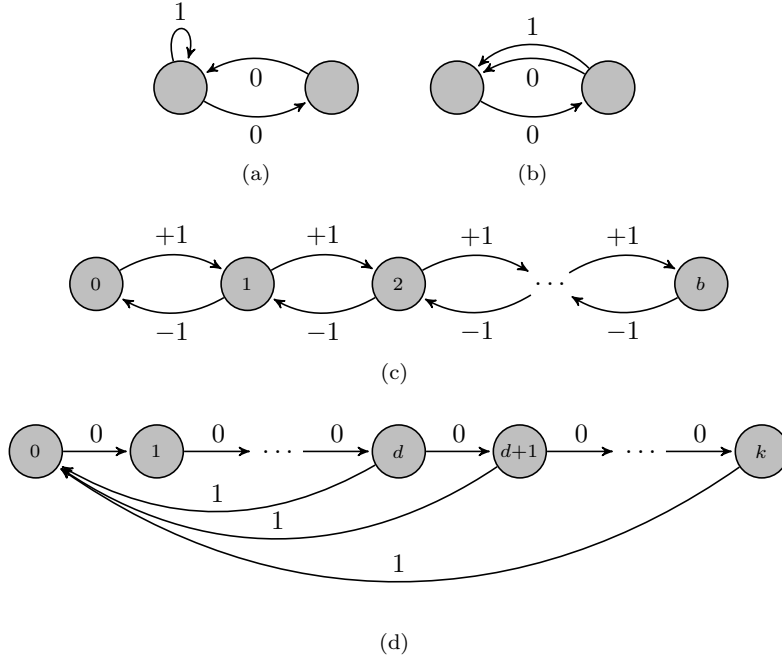


Figure 1: Presentations of 1-dimensional constraints: (a) EVEN ;(b) ODD; (c) CHG(b); (d) RLL(d, k).

A 1-dimensional constraint over an alphabet Σ is said to have memory m , for some positive integer m , if for every word w of more than m letters over Σ , in which every subword of $m+1$ consecutive letters satisfies S , it holds that w satisfies S as well, and m is the smallest integer for which this holds. A 1-dimensional constraint with finite-memory is called a *finite-type* constraint. Among our examples, RLL(d, k) is a finite-type constraint with memory k , whereas EVEN, ODD and CHG(b) for $b \geq 2$ are not finite-type constraints.

In this work, we consider multidimensional constraints of dimension \mathfrak{D} for some positive integer \mathfrak{D} . These are sets of finite-size \mathfrak{D} -dimensional arrays with entries over some finite alphabet specified by \mathfrak{D} edge-labeled directed graphs. In Section 2 we give a precise definition of what we mean by \mathfrak{D} -dimensional constraints. Here we just mention that they are closed under taking subarrays, meaning that if an array belongs to the constraint then any of its \mathfrak{D} -dimensional subarrays consisting of “adjacent” entries also belongs to the constraint. As before we say that an array satisfies the constraint if it belongs to it. Examples of multidimensional constraints can be obtained by generalizing 1-dimensional constraints, defining the constraint to consist of all arrays satisfying a given 1-dimensional constraint S on every “row” in every direction along an “axis” of the array. We denote such a \mathfrak{D} -dimensional constraint by $S^{\otimes \mathfrak{D}}$. We will almost exclusively be concerned with 2 dimensional constraints, namely $\mathfrak{D} = 2$. In this case $S^{\otimes 2}$ is the set of all 2-dimensional arrays where each row and each column satisfy S . A well-known 2-dimensional constraint studied in statistical mechanics is the so called “hard-square” constraint. It consists of all finite-size binary

arrays which do not contain 2 adjacent ‘1’s either horizontally or vertically. Two variations of this constraint are the isolated ‘1’s or “non-attacking-kings” constraint, denoted NAK, and the “read-write-isolated-memory” constraint, denoted RWIM. The former consists of all finite-size binary arrays in which there are no two adjacent ‘1’s either horizontally, vertically, or diagonally, and the latter consists of all finite-size binary arrays in which there are no two adjacent ‘1’s either horizontally, or diagonally. Like their 1-dimensional counterparts, 2-dimensional constraints play a role in storage systems, where with recent developments, information is written in a true 2-dimensional fashion rather than using essentially 1-dimensional tracks. The RWIM constraint is used to model sequences of states of a binary linear memory in which no two adjacent entries may contain a ‘1’, and in every update, no two adjacent entries are both changed. See [2] and [4] for more details.

Let S now be a \mathfrak{D} -dimensional constraint over an alphabet Σ . For a \mathfrak{D} -tuple $\mathbf{m} = (m_1, m_2, \dots, m_{\mathfrak{D}})$ of positive integers, $S_{\mathbf{m}}$ or $S_{m_1 \times \dots \times m_{\mathfrak{D}}}$ denotes the set of all $m_1 \times m_2 \times \dots \times m_{\mathfrak{D}}$ arrays in S , and $\langle \mathbf{m} \rangle$ denotes the product of the entries of \mathbf{m} . We say that a sequence $\mathbf{m}_i = (m_1^{(i)}, \dots, m_{\mathfrak{D}}^{(i)})$ diverges to infinity, denoted $\mathbf{m}_i \rightarrow \infty$, if $(m_j^{(i)})_{i=1}^{\infty}$ does for each j . The capacity of S is defined by

$$\text{cap}(S) = \lim_{i \rightarrow \infty} \frac{\log |S_{\mathbf{m}_i}|}{\langle \mathbf{m}_i \rangle}, \quad (1)$$

where $\mathbf{m}_i \rightarrow \infty$, $|\cdot|$ denotes cardinality, and $\log = \log_2$. The above limit exists and is independent of the choice of $(\mathbf{m}_i)_{i=1}^{\infty}$, for any set S of finite-size \mathfrak{D} -dimensional arrays over Σ which is closed under taking subarrays. This follows from subadditivity arguments (see [11] for a proof for $\mathfrak{D} = 2$, which can be generalized to higher dimensions). In fact

$$\text{cap}(S) = \inf_{\mathbf{m}} \frac{\log_2 |S_{\mathbf{m}}|}{\langle \mathbf{m} \rangle}. \quad (2)$$

While a closed form formula for the capacity of 1-dimensional constraints is known (up to finding the largest root of a polynomial), no such formula is known for constraints in higher dimensions, and currently there are only a few multidimensional constraints for which the capacity is known exactly and is nonzero (a highly non-trivial example can be found in [10]).

Let S be a 2-dimensional constraint over Σ , and m be a positive integer. The *horizontal* (resp. *vertical*) *strip of height* (resp. *width*) m of S , denoted $\mathcal{H}_m(S)$ (resp. $\mathcal{V}_m(S)$) is the subset of S given by

$$\mathcal{H}_m(S) = \bigcup_n S_{m \times n} \quad (\text{resp. } \mathcal{V}_m(S) = \bigcup_n S_{n \times m}).$$

We show in Section 2 that $\mathcal{H}_m(S)$ and $\mathcal{V}_m(S)$ are 1-dimensional constraints over Σ^m .

A method for computing very good lower and upper bounds on the capacity of the hard-square constraint is given in [1] (see also [15]). Their method can be shown to work on 2-dimensional constraints for which every horizontal or every vertical strip has memory 1 and is “symmetric”; that is, it is closed under reversing the order of symbols in words. The main contributions of our work are:

1. We establish a generalization of the method of [1] that gives improved lower bounds on capacities of 2-dimensional constraints, for instance for NAK and RWIM.
2. We show how this generalization as well as the original method for obtaining upper-bounds may be applied to a larger class of 2-dimensional constraints that includes constraints in which the vertical and horizontal strips are not necessarily finite-type. We illustrate this by computing lower and upper bounds on the capacities of the $\text{CHG}(3)^{\otimes 2}$ and $\text{EVEN}^{\otimes 2}$ constraints.
3. We show that $\text{cap}(\text{CHG}(2)^{\otimes \mathfrak{D}}) = 2^{-\mathfrak{D}}$ and $\text{cap}(\text{ODD}^{\otimes \mathfrak{D}}) = 1/2$, for all positive integers \mathfrak{D} .

Previous work involving applications of the method of [1] and generalizations include [5], [6], [16], and [18].

2 Framework.

In this section, we define the framework that we use in the rest of the paper. We deal with a directed graph $G = (V, E)$, sometimes simply called a graph, with vertices V and edges E . For $e \in E$ we denote by $\sigma_G(e)$ and $\tau_G(e)$ the initial and terminal vertices of e in G , respectively. We shall omit the subscript G from σ_G and τ_G when the graph is clear from the context. A *path of length ℓ* in G is a sequence of ℓ edges $(e_i)_{i=1}^{\ell} \subseteq E$, where for $i = 1, 2, \dots, \ell-1$, $\tau(e_i) = \sigma(e_{i+1})$. The path starts at the vertex $\sigma(e_1)$ and ends at the vertex $\tau(e_{\ell})$. A cycle in G is a path that starts and ends at the same vertex. Fix a finite alphabet Σ . A directed *labeled graph* \mathcal{G} with labels in Σ is a pair $\mathcal{G} = (G, \mathcal{L})$, where $G = (V, E)$ is a directed graph, and $\mathcal{L} : E \rightarrow \Sigma$ is a labeling of the edges of G with symbols of Σ . The paths and cycles of \mathcal{G} are inherited from G and we will sometime use $\sigma_{\mathcal{G}}$ and $\tau_{\mathcal{G}}$ to denote σ_G and τ_G respectively.

For a labeled graph $\mathcal{G} = ((V, E), \mathcal{L})$ with $\mathcal{L} : E \rightarrow \Sigma$ and a path $(e_i)_{i=1}^{\ell}$ of \mathcal{G} , we say the path *generates* the word $\mathcal{L}(e_1)\mathcal{L}(e_2)\dots\mathcal{L}(e_{\ell})$ in Σ^* . The graph \mathcal{G} is called *lossless* if for any two vertices u and v of \mathcal{G} , all paths starting at u and terminating at v generate distinct words. The graph \mathcal{G} is called *deterministic* if there are no two distinct edges with the same initial vertex and the same label. Every 1-dimensional constraint S has a deterministic, and therefore lossless, presentation [14].

We introduce two 1-dimensional constraints defined by general directed graphs. Let $G=(V, E)$ be a directed graph, the *edge constraint defined by G* , denoted $\mathsf{X}(G)$, is the 1-dimensional constraint over the alphabet E , presented by $\mathcal{G} = (G, I_E)$ where I_E is the identity map on E . For a graph $G=(V, E)$ with no parallel edges, the *vertex-constraint defined by G* , denoted $\widehat{\mathsf{X}}(G)$, is the set

$$\{(v_i)_{i=1}^{\ell} \subseteq V : \ell=0,1,2,\dots, \text{ and for } 1 \leq i < \ell, \exists e \in E \text{ s.t. } \sigma(e)=v_i, \tau(e)=v_{i+1}\}.$$

It's not hard to verify that vertex-constraints and edge-constraints are 1-dimensional constraints with memory 1. In fact, the vertex constraints are precisely the finite-type constraints with memory 1. And it can be shown that edge constraints are characterized as follows. The *follower set* of a symbol a in a constraint S is defined to be $\{b : ab \in S\}$; edge constraints are precisely the

constraints with memory 1 such that any two follower sets are either disjoint or identical [13, exercise 2.3.4].

We consider multidimensional arrays of dimension \mathfrak{D} —a positive integer. For an integer ℓ denote by $[\ell]$ the set $\{0, 1, \dots, \ell-1\}$. For a \mathfrak{D} -tuple $\mathbf{m} = (m_1, \dots, m_{\mathfrak{D}})$ of nonnegative integers and a finite set A , we call an $m_1 \times m_2 \times \dots \times m_{\mathfrak{D}}$ \mathfrak{D} -dimensional array with entries in A , a \mathfrak{D} -dimensional array of size \mathbf{m} over A . We shall index the entries of such an array by $[m_1] \times [m_2] \times \dots \times [m_{\mathfrak{D}}]$. Let $A^{\mathbf{m}}$ or $A^{m_1 \times \dots \times m_{\mathfrak{D}}}$ denote the set of all \mathfrak{D} -dimensional arrays of size \mathbf{m} over A . We define $A^{*\dots*} = A^{*\mathfrak{D}}$, where the number of $*$'s in the superscript is \mathfrak{D} , by

$$A^{*\mathfrak{D}} = \bigcup_{\mathbf{m}} \Sigma^{\mathbf{m}},$$

as the set of all finite-size \mathfrak{D} -dimensional arrays with entries in A . Let $\Gamma \in A^{*\mathfrak{D}}$ be such an array. Given an integer $1 \leq i \leq \mathfrak{D}$, a *row in direction i* of Γ is a sequence of entries of Γ of the form $(\Gamma_{(k_1, k_2, \dots, k_{i-1}, j, k_{i+1}, \dots, k_{\mathfrak{D}})})_{j=0}^{m_i-1}$ for some integers $k_l \in [m_l]$; $1 \leq l \leq \mathfrak{D}$, $l \neq i$. In this paper, for $\mathfrak{D} = 2$, we use the convention that direction 1 is the vertical direction and direction 2 is the horizontal; thus the columns of a 2-dimensional array are its rows in direction 1, and its “traditional rows” are its rows in direction 2. Let B be a finite set and $\mathcal{L} : A \rightarrow B$ be a mapping. We extend \mathcal{L} to a mapping $\mathcal{L} : A^{*\mathfrak{D}} \rightarrow B^{*\mathfrak{D}}$ as follows. For a \mathfrak{D} -dimensional array $\Gamma \in A^{\mathbf{m}}$, $\mathcal{L}(\Gamma)$ is the array in $B^{\mathbf{m}}$ obtained by applying \mathcal{L} to each entry of Γ , that is

$$(\mathcal{L}(\Gamma))_{\mathbf{j}} = \mathcal{L}(\Gamma_{\mathbf{j}}) \quad , \quad \mathbf{j} \in [m_1] \times \dots \times [m_{\mathfrak{D}}].$$

Additionally, for a subset $S \subseteq A^{*\mathfrak{D}}$ we define $\mathcal{L}(S) = \{\mathcal{L}(\Gamma) : \Gamma \in S\}$.

We generalize the definition of constrained system to \mathfrak{D} dimensions. Let $\bar{\mathcal{G}} = (\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_{\mathfrak{D}})$, be a \mathfrak{D} -tuple of labeled graphs with the same set of edges E and the same labeling $\mathcal{L} : E \rightarrow \Sigma$. The edge e has \mathfrak{D} pairs of initial and terminal vertices $(\sigma_{\mathcal{G}_i}(e), \tau_{\mathcal{G}_i}(e))$ —one for each graph \mathcal{G}_i in $\bar{\mathcal{G}}$. We say that an array $\Gamma \in \Sigma^{*\mathfrak{D}}$ of size \mathbf{m} is *generated* by $\bar{\mathcal{G}}$ if there exists an array $\Gamma' \in E^{*\mathfrak{D}}$ of size \mathbf{m} , such that for $i = 1, 2, \dots, \mathfrak{D}$, every row in direction i of Γ' is a path in \mathcal{G}_i , and $\mathcal{L}(\Gamma') = \Gamma$. We call the set of all arrays $\Gamma \in \Sigma^{*\mathfrak{D}}$ generated by $\bar{\mathcal{G}}$, the *\mathfrak{D} -dimensional constrained system* or the *\mathfrak{D} -dimensional constraint* presented by $\bar{\mathcal{G}}$, and denote it by $X(\bar{\mathcal{G}})$. We say that $\bar{\mathcal{G}}$ is a *presentation* for $X(\bar{\mathcal{G}})$.

In [7], 2-dimensional constrained systems are defined by vertex-labeled graphs, with a common set of vertices and a common labelling on the vertices. It can be shown that their definition (generalized to higher dimensions) is equivalent to ours. We find it more convenient to use our definition, since, just as in one dimension, it permits use of parallel edges and often enables a smaller presentation of a given constraint.

Figure 2 shows presentations for the NAK and RWIM constraints defined in Section 1. In these presentations \mathcal{G}_1 and \mathcal{G}_2 describe the vertical and horizontal constraints on the edges, respectively. Each edge $\mathbf{e} = (\mathbf{e})_{i,j}$ is a 2×2 binary matrix of the form

$$\mathbf{e} = \begin{pmatrix} (\mathbf{e})_{(0,0)} & (\mathbf{e})_{(0,1)} \\ (\mathbf{e})_{(1,0)} & (\mathbf{e})_{(1,1)} \end{pmatrix},$$

and it is labeled by $(\mathbf{e})_{(1,1)}$, i.e., the labelling of an edge simply picks out the entry in the lower-right corner. For NAK, the edges $E = E_{\text{NAK}}$ are the 2×2

matrices with at most one 1; note that a rectangular array of any size satisfies NAK iff each of its 2×2 sub-arrays belongs to E ; similarly, for RWIM, the edges $E = E_{\text{RWIM}}$ are the elements of E_{NAK} together with

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix},$$

and again a rectangular array of any size satisfies RWIM iff each of its 2×2 sub-arrays belongs to E . In the figures, each edge is drawn twice—once in \mathcal{G}_1 and once in \mathcal{G}_2 —and the matrix identifying it is written next to it. The states are 1×2 blocks for \mathcal{G}_1 and 2×1 blocks for \mathcal{G}_2 ; for an edge \mathbf{e} ,

$$\sigma_{\mathcal{G}_1}(e) = (\mathbf{e})_{(0,0)}(\mathbf{e})_{(0,1)}, \text{ and } \tau_{\mathcal{G}_1}(e) = (\mathbf{e})_{(1,0)}(\mathbf{e})_{(1,1)},$$

and

$$\sigma_{\mathcal{G}_2}(e) = \begin{pmatrix} (\mathbf{e})_{(0,0)} \\ (\mathbf{e})_{(1,0)} \end{pmatrix}, \text{ and } \tau_{\mathcal{G}_2}(e) = \begin{pmatrix} (\mathbf{e})_{(0,1)} \\ (\mathbf{e})_{(1,1)} \end{pmatrix}.$$

It follows that, for both constraints, and any rectangular array $\Gamma' \in E^{m \times n}$ with each of its rows a path in \mathcal{G}_2 and each of its columns a path in \mathcal{G}_1 , it holds that

$$\Gamma'_{(i,j)} = \begin{pmatrix} \mathcal{L}(\Gamma'_{(i-1,j-1)}) & \mathcal{L}(\Gamma'_{(i-1,j)}) \\ \mathcal{L}(\Gamma'_{(i,j-1)}) & \mathcal{L}(\Gamma'_{(i,j)}) \end{pmatrix}.$$

for $i = 1, 2, \dots, m-1$, $j = 1, 2, \dots, n-1$. Therefore, the only 2×2 sub-arrays appearing in the array $\mathcal{L}(\Gamma')$ are elements of E , and it follows that $\mathcal{L}(\Gamma')$ satisfies the corresponding constraint. Similarly, it can be shown that any array satisfying the constraint can be generated by the presentation.

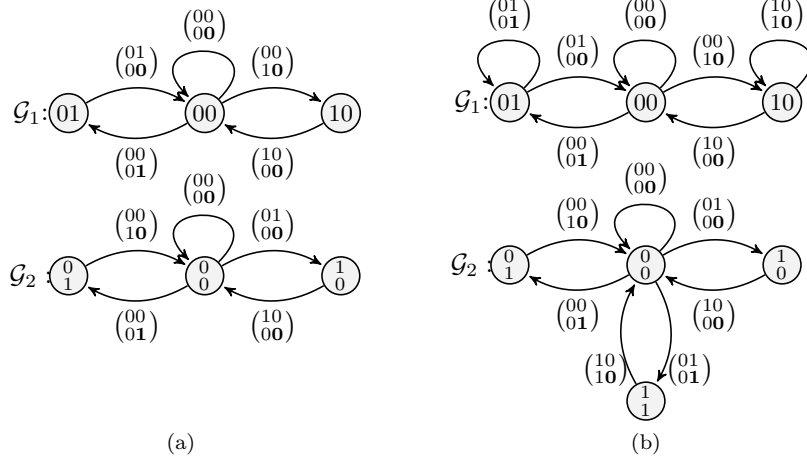


Figure 2: Presentations of 2-dimensional constraints: (a) NAK constraint; (b) RWIM constraint.

The *axial product* of \mathfrak{D} sets $L_1, \dots, L_{\mathfrak{D}} \subseteq \Sigma^*$, denoted $L_1 \otimes L_2 \otimes \dots \otimes L_{\mathfrak{D}} \subseteq \Sigma^{*\mathfrak{D}}$, is the set of all arrays $\Gamma \in \Sigma^{*\mathfrak{D}}$ such that for $i = 1, 2, \dots, \mathfrak{D}$ every row of Γ in direction i belongs to L_i . If $L_1 = L_2 = \dots = L_{\mathfrak{D}} = L$ we say that the

axial-product is *isotropic* and denote it by $L^{\otimes \mathfrak{D}}$. Given a presentation $\bar{\mathcal{G}} = ((G_1, \mathcal{L}), \dots, (G_{\mathfrak{D}}, \mathcal{L}))$ for a \mathfrak{D} -dimensional constraint S with a common set of edges E , the set $X(G_1) \otimes \dots \otimes X(G_{\mathfrak{D}}) \subseteq E^{*\mathfrak{D}}$ is a \mathfrak{D} -dimensional constraint represented by $((G_1, I_E), \dots, (G_{\mathfrak{D}}, I_E))$, where I_E is the identity map on E .

We say that $\bar{\mathcal{G}}$ is *capacity-preserving* if $\text{cap}(X(G_1) \otimes \dots \otimes X(G_{\mathfrak{D}})) = \text{cap}(S)$. For $\mathfrak{D} = 1$ any lossless, and therefore deterministic, presentation of S is capacity-preserving, since in a lossless graph with $|V|$ vertices there are at most $|V|^2$ paths generating any given word; for $\mathfrak{D} > 1$, the question whether every \mathfrak{D} -dimensional constraint has a capacity-preserving presentation is open. This is a major open problem in symbolic dynamics, although it is usually formulated in a slightly different manner; see [3], where it is shown that for every \mathfrak{D} -dimensional constraint S and $\epsilon > 0$, there is a presentation $\bar{\mathcal{G}} = ((G_1, \mathcal{L}), \dots, (G_{\mathfrak{D}}, \mathcal{L}))$ such that $\text{cap}(S) \leq \text{cap}(X(G_1) \otimes \dots \otimes X(G_{\mathfrak{D}})) < \text{cap}(S) + \epsilon$. We show in the next proposition that the answer to this question is positive, if S is an axial product of \mathfrak{D} one-dimensional constraints.

Proposition 1. *Let $S_1, S_2, \dots, S_{\mathfrak{D}} \subseteq \Sigma^*$ be \mathfrak{D} one-dimensional constraints over Σ , and let $S = S_1 \otimes \dots \otimes S_{\mathfrak{D}}$, then S is a \mathfrak{D} -dimensional constraint over Σ . Moreover, S has a capacity-preserving presentation.*

Remark . There are \mathfrak{D} -dimensional constraints which are not axial products of \mathfrak{D} one-dimensional constraints. For example for $\mathfrak{D}=2$, the NAK constraint defined in Section 1 is not an axial product.

Proof. Let $\mathcal{G}_{S_1} = ((V_1, E_1), \mathcal{L}_1), \mathcal{G}_{S_2} = ((V_2, E_2), \mathcal{L}_2), \dots, \mathcal{G}_{S_{\mathfrak{D}}} = ((V_{\mathfrak{D}}, E_{\mathfrak{D}}), \mathcal{L}_{\mathfrak{D}})$ be presentations of $S_1, \dots, S_{\mathfrak{D}}$, respectively. Define the \mathfrak{D} -tuple of labeled graphs $\bar{\mathcal{G}} = (\mathcal{G}_1, \dots, \mathcal{G}_{\mathfrak{D}})$, as follows. Let

$$E = \left\{ (e_1, \dots, e_{\mathfrak{D}}) \in \prod_{i=1}^{\mathfrak{D}} E_i : \mathcal{L}_1(e_1) = \mathcal{L}_2(e_2) = \dots = \mathcal{L}_{\mathfrak{D}}(e_{\mathfrak{D}}) \right\},$$

and let $\mathcal{L} : E \rightarrow \Sigma$ be given by

$$\mathcal{L}(e_1, \dots, e_{\mathfrak{D}}) = \mathcal{L}_1(e_1) ; (e_1, \dots, e_{\mathfrak{D}}) \in E.$$

For $i = 1, 2, \dots, \mathfrak{D}$, the graph \mathcal{G}_i is defined by $\mathcal{G}_i = (G_i, \mathcal{L})$ with $G_i = (V_i, E)$, where for $\mathbf{e} = (e_1, \dots, e_{\mathfrak{D}}) \in E$, $\sigma_{\mathcal{G}_i}(\mathbf{e}) = \sigma_{\mathcal{G}_{S_i}}(e_i)$ and $\tau_{\mathcal{G}_i}(\mathbf{e}) = \tau_{\mathcal{G}_{S_i}}(e_i)$. It's easy to verify that $\bar{\mathcal{G}}$ is a presentation of $S_1 \otimes S_2 \otimes \dots \otimes S_{\mathfrak{D}}$.

Assume now that every \mathcal{G}_{S_i} is lossless. We show that in this case $\bar{\mathcal{G}}$ is capacity preserving. Let $X = X(G_1) \otimes \dots \otimes X(G_{\mathfrak{D}})$, n be a positive integer, and let \mathbf{n} be the \mathfrak{D} -tuple with every entry equal to n . We extend the mapping \mathcal{L} to $\mathcal{L} : X_{\mathbf{n}} \rightarrow S_{\mathbf{n}}$ as described above. Now, fix an array $\Gamma \in S_{\mathbf{n}}$, and for $i = 1, 2, \dots, \mathfrak{D}$ let $\Gamma^{(i)} \in (E_i)^{\mathbf{n}}$ be an array such that every row in direction i , $(\Gamma_{\mathbf{j}_k}^{(i)})_{k=1}^n$ is a path in \mathcal{G}_{S_i} generating the corresponding row $(\Gamma_{\mathbf{j}_k})_{k=1}^n$ in Γ . Let $\Gamma' \in E^{\mathbf{n}}$ be the array with entries given by $\Gamma'_{\mathbf{j}} = (\Gamma_{\mathbf{j}}^{(1)}, \dots, \Gamma_{\mathbf{j}}^{(\mathfrak{D})})$, $\mathbf{j} \in [n]^{\mathfrak{D}}$. It follows from the construction of $\bar{\mathcal{G}}$ that $\Gamma' \in X_{\mathbf{n}}$, and that $\mathcal{L}(\Gamma') = \Gamma$. Moreover, any array $\Delta \in X_{\mathbf{n}}$ such that $\mathcal{L}(\Delta) = \Gamma$ can be constructed in this manner. Now, as each \mathcal{G}_{S_i} is lossless, there are at most $|V_i|^2$ possibilities of choosing each row in direction i of $\Gamma^{(i)}$, and as there are $n^{\mathfrak{D}-1}$ such rows, there are at most $|V_i|^{2n^{\mathfrak{D}-1}}$ possibilities of choosing each $\Gamma^{(i)}$. It follows that

$$|\mathcal{L}^{-1}(\{\Gamma\})| \leq \prod_{i=1}^{\mathfrak{D}} |V_i|^{2n^{\mathfrak{D}-1}},$$

where $\mathcal{L}^{-1}(\{\Gamma\}) = \{\Gamma' \in X_{\mathbf{n}} : \mathcal{L}(\Gamma') = \Gamma\}$. Summing the latter inequality over all $\Gamma \in \mathcal{S}_{\mathbf{n}}$, we obtain

$$|X_{\mathbf{n}}| \leq |\mathcal{S}_{\mathbf{n}}| \prod_{i=1}^{\mathfrak{D}} |V_i|^{2n^{\mathfrak{D}-1}}.$$

Taking the log, dividing through by $n^{\mathfrak{D}}$, and taking the limit as n approaches infinity, we have $\text{cap}(X) \leq \text{cap}(S)$. Clearly, $\text{cap}(X) \geq \text{cap}(S)$, since X is a presentation of S . The result follows. \blacksquare

A graph $G = (V, E)$ is *irreducible* if for any pair of vertices $u, v \in V$ there is a path in G starting at u and terminating at v ; it is *primitive* if it is irreducible and the gcd of the lengths of all cycles of G is 1. We denote by $A(G)$ the adjacency matrix of G : namely the $|V| \times |V|$ matrix where $(A(G))_{i,j}$ is the number of edges in G from i to j . We use $\mathbf{1}$ to denote the real vector each of whose entries is 1. For a nonnegative matrix A denote by $\lambda(A)$ its Perron eigenvalue, that is, its largest real eigenvalue. It is well-known that for a 1-dimensional constraint S presented by a lossless labeled graph $\mathcal{G} = (G, \mathcal{L})$, the capacity of S is $\log \lambda(A(G))$. In particular, for a graph G , it holds that $\text{cap}(X(G)) = \log \lambda(A(G))$. We say that a graph G is *symmetric* if $A(G)$ is symmetric. We say that a vertex of a graph is *isolated* if it has no outgoing nor incoming edges. We say that a vertex-constraint (resp. edge-constraint) is *symmetric* if it is defined by a symmetric graph. For a vertex-constraint, this definition is equivalent to requiring that the constraint is closed under reversal of the order of symbols in words. Note, that in a symmetric edge-constraint, up to removal of isolated vertices, the (unlabeled) graph defining the constraint is unique.

In this paper, we will mostly deal with 2-dimensional constraints. For a 2-dimensional array $\Gamma \in \Sigma^{m_1 \times m_2}$, for nonnegative integers m_1, m_2 , we denote by Γ^t its transpose, namely $(\Gamma^t)_{(i,j)} = (\Gamma)_{(j,i)}$ for all $(i,j) \in [m_1] \times [m_2]$. For a 2-dimensional constraint S over Σ , we use S^t to denote the set

$$S^t = \{\Gamma \in \Sigma^{**} : \Gamma^t \in S\},$$

Clearly S^t is a 2-dimensional constraint with $\text{cap}(S^t) = \text{cap}(S)$.

Let S be a 2-dimensional constraint over an alphabet A , and consider a horizontal (resp. vertical) strip $\mathcal{H}_m(S)$ (resp. $\mathcal{V}_m(S)$) of S for some positive integer m . We regard such a strip as a set of 1-dimensional words over Σ^m where each $m \times n$ (resp. $n \times m$) array in the strip is considered a word of length n over A^m . Below we show that the horizontal and vertical strips of S are 1-dimensional constraints over A^m . For this, we need the following definition. Let $G = (V, E)$ be a graph, and let m be a positive integer. Let $G^{\times m}$ be the graph given by $G^{\times m} = (V^m, E^m)$, where for each $\mathbf{e} = (e_1, \dots, e_m) \in E^m$, $\sigma_{G^{\times m}}(\mathbf{e}) = (\sigma_G(e_1), \dots, \sigma_G(e_m))$ and $\tau_{G^{\times m}}(\mathbf{e}) = (\tau_G(e_1), \dots, \tau_G(e_m))$. For a labeled graph $\mathcal{G} = (G, \mathcal{L})$ with $G = (V, E)$ and $\mathcal{L} : E \rightarrow \Sigma$, let $\mathcal{G}^{\times m}$ be the labeled graph defined by $\mathcal{G}^{\times m} = (G^{\times m}, \mathcal{L}^{\times m})$, where $\mathcal{L}^{\times m} : E^m \rightarrow \Sigma^m$ is given by

$$\mathcal{L}^{\times m}(e_1, \dots, e_m) = (\mathcal{L}(e_1), \dots, \mathcal{L}(e_m)) ; (e_1, \dots, e_m) \in E^m$$

We call $G^{\times m}$ (resp., $\mathcal{G}^{\times m}$) the *m*th tensor-power of G . (resp., \mathcal{G}). We can now state the following proposition.

Proposition 2. *Let S be a 2-dimensional constraint over Σ and let m be a positive integer. Then*

1. $\mathcal{H}_m(S)$ (resp. $\mathcal{V}_m(S)$) is a 1-dimensional constraint over Σ^m .
2. Let $S = T^{(\mathcal{V})} \otimes T^{(\mathcal{H})}$ for 1-dimensional constraints $T^{(\mathcal{V})}, T^{(\mathcal{H})}$ over Σ , presented by labeled graphs $\mathcal{G}^{(\mathcal{V})}, \mathcal{G}^{(\mathcal{H})}$, respectively. Then the 1-dimensional constraint $\mathcal{H}_m(S)$ is presented by the labeled graph $\mathcal{G}_m^{(\mathcal{H})}$ defined as the sub-graph of the labeled graph $(\mathcal{G}^{(\mathcal{H})})^{\times m}$ consisting of only those edges whose label (an m -letter word over Σ) satisfies $T^{(\mathcal{V})}$. An analogous statement holds for $\mathcal{V}_m(S)$, with respect to the graph $\mathcal{G}_m^{(\mathcal{V})}$ formed in a similar way from $(\mathcal{G}^{(\mathcal{V})})^{\times m}$.

Proof. It suffices to prove this only for horizontal strips $\mathcal{H}_m(S)$. We first prove part 2. It's easy to verify that the labeled graph $(\mathcal{G}^{(\mathcal{H})})^{\times m}$ presents the constraint over Σ^m , consisting of all $m \times n$ arrays of Σ^{**} such that every row satisfies $T^{(\mathcal{H})}$. It follows that the sub-graph $\mathcal{G}_m^{(\mathcal{H})}$, formed by removing all the edges of $(\mathcal{G}^{(\mathcal{H})})^{\times m}$ that are labeled with a word that does not satisfy $T^{(\mathcal{V})}$, presents the 1-dimensional constraint consisting of all $m \times n$ arrays with every row satisfying $T^{(\mathcal{H})}$ and every column satisfying $T^{(\mathcal{V})}$. This is precisely the constraint $\mathcal{H}_m(S)$.

We now prove part 1. Let the pair of labeled graphs $(\mathcal{G}^{(\mathcal{V})}, \mathcal{G}^{(\mathcal{H})})$ be a presentation of S , where $\mathcal{G}^{(\mathcal{V})} = ((V^{(\mathcal{V})}, E), \mathcal{L})$ and $\mathcal{G}^{(\mathcal{H})} = ((V^{(\mathcal{H})}, E), \mathcal{L})$. Define the edge-constraints $\mathcal{E}^{(\mathcal{V})} = \mathbf{X}(V^{(\mathcal{V})}, E)$ and $\mathcal{E}^{(\mathcal{H})} = \mathbf{X}(V^{(\mathcal{H})}, E)$. Since $(\mathcal{G}^{(\mathcal{V})}, \mathcal{G}^{(\mathcal{H})})$ is a presentation of S , we have $S = \mathcal{L}(\mathcal{E}^{(\mathcal{V})} \otimes \mathcal{E}^{(\mathcal{H})})$, and therefore $\mathcal{H}_m(S) = \mathcal{L}(\mathcal{H}_m(\mathcal{E}^{(\mathcal{V})} \otimes \mathcal{E}^{(\mathcal{H})}))$. By part 2, $\mathcal{H}_m(\mathcal{E}^{(\mathcal{V})} \otimes \mathcal{E}^{(\mathcal{H})})$ is a 1-dimensional constraint, presented by a labeled graph $\mathcal{G}_m^{(\mathcal{H})}$ with edges labeled by words in E^m . Replacing each such label $\mathbf{e} \in E^m$ in that graph with $\mathcal{L}(\mathbf{e})$ we clearly obtain a presentation of $\mathcal{L}(\mathcal{H}_m(\mathcal{E}^{(\mathcal{V})} \otimes \mathcal{E}^{(\mathcal{H})})) = \mathcal{H}_m(S)$. \blacksquare

We shall use the following notation for 2-dimensional arrays. Let A be a set and $\Gamma \in A^{s \times t}$. For integers $0 \leq s_1 \leq s_2 < s$ and $0 \leq t_1 \leq t_2 < t$, we denote by $\Gamma_{s_1:s_2, t_1:t_2}$ the sub-array:

$$(\Gamma_{s_1:s_2, t_1:t_2})_{i,j} = \Gamma_{s_1+i, t_1+j} \ ; \ (i, j) \in [s_2 - s_1 + 1] \times [t_2 - t_1 + 1],$$

and by $\Gamma_{s_1:s_2, *}$ (resp. $\Gamma_{*, t_1:t_2}$) the sub-array $\Gamma_{s_1:s_2, 0:t-1}$ (resp. $\Gamma_{0:s-1, t_1:t_2}$). We also abbreviate $x:x$ in the subscript by x . We shall use the same notation for one-dimensional vectors: for a vector $\mathbf{v} \in A^s$, $\mathbf{v}_{s:t}$ denotes the subvector

$$(\mathbf{v}_{s:t})_i = \mathbf{v}_{s+i} \ ; \ i \in [t-s+1].$$

3 Constraints with symmetric edge-constrained strips.

In this section we generalize the method presented in [1] to provide improved lower bounds on capacities of 2-dimensional constraints whose horizontal strips are symmetric edge-constraints.

Fix an alphabet Σ , and let S be a 2-dimensional constraint over Σ . We say that S has *horizontal edge-constrained-strips* if for every positive integer m , the constraint $\mathcal{H}_m(S)$ is an edge-constraint. If, in addition, every horizontal strip

is symmetric, we say that S has *symmetric horizontal edge-constrained strips*. Analogously, using $\mathcal{V}_m(S)$, we have the notions of a 2-dimensional constraint with vertical edge-constrained-strips and symmetric vertical edge-constrained-strips

Here, we consider constraints of the form $S = T \otimes \mathcal{E}$, where $\mathcal{E} = \mathcal{X}(G_{\mathcal{E}})$ is an edge-constraint defined by the graph $G_{\mathcal{E}} = (V_{\mathcal{E}}, E_{\mathcal{E}})$ and T is an arbitrary 1-dimensional constraint over Σ . Then \mathcal{E} is presented by $\mathcal{G}_{\mathcal{E}} = (G_{\mathcal{E}}, I_E)$ where I_E is the identity map on $E_{\mathcal{E}}$. Let m be a positive integer. By Proposition 2, part 2, $\mathcal{H}_m(S)$ is a 1-dimensional constraint presented by a subgraph $\mathcal{G}_m^{(\mathcal{H})} = (G_m^{(H)}, I_E^{\times m})$ of $\mathcal{G}_{\mathcal{E}}^{\times m}$. It follows that $\mathcal{H}_m(S) = \mathcal{X}(G_m^{(\mathcal{H})})$, and so S has horizontal edge-constrained strips. Henceforth, we further assume that it has symmetric horizontal edge-constrained strips; note that symmetry of the graph $G_{\mathcal{E}}$ is necessary but not sufficient for this assumption (see Proposition 3 below).

For a positive integer m , let $F_m = |(V_{\mathcal{E}})^m|$, and let H_m denote the $F_m \times F_m$ adjacency matrix of $G_m^{(\mathcal{H})}$. Since the limit in (1) exists independently of the choice of $(\mathbf{m}_i)_{i=1}^{\infty}$, and since $\lim_{n \rightarrow \infty} \log(|S_{m \times n}|)/n = \text{cap}(\mathcal{H}_m(S))$ for every positive integer m , we have

$$\begin{aligned} \text{cap}(S) &= \lim_{m, n \rightarrow \infty} \frac{\log |S_{m \times n}|}{mn} \\ &= \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{\log |S_{m \times n}|}{mn} \\ &= \lim_{m \rightarrow \infty} \frac{\text{cap}(\mathcal{H}_m(S))}{m} \\ &= \lim_{m \rightarrow \infty} \frac{\log \lambda(H_m)}{m}. \end{aligned} \quad (3)$$

For a matrix M , let M^t denote its transpose. Fix a positive integer m . Following [1], since H_m is real and symmetric, we obtain by the min-max principle [8]

$$\lambda(H_m^p) \geq \frac{\mathbf{y}_m^t H_m^p \mathbf{y}_m}{\mathbf{y}_m^t \mathbf{y}_m},$$

for any $F_m \times 1$ real vector $\mathbf{y}_m \neq \mathbf{0}$ and positive integer p . Choosing \mathbf{y}_m to be the vector $H_m^q \mathbf{x}_m$, for some positive integer q and $F_m \times 1$ real vector \mathbf{x}_m such that $\mathbf{y}_m \neq \mathbf{0}$, we have

$$\lambda(H_m^p) \geq \frac{\mathbf{x}_m^t H_m^{2q+p} \mathbf{x}_m}{\mathbf{x}_m^t H_m^{2q} \mathbf{x}_m}. \quad (4)$$

Thus by (3), it follows that

$$\text{cap}(S) \geq \frac{1}{p} \limsup_{m \rightarrow \infty} \frac{1}{m} \log \frac{\mathbf{x}_m^t H_m^{2q+p} \mathbf{x}_m}{\mathbf{x}_m^t H_m^{2q} \mathbf{x}_m}. \quad (5)$$

In [1], each \mathbf{x}_m is chosen to be the $F_m \times 1$ vector with every entry equal to 1. We obtain improved lower bounds in many cases by choosing other sequences of vectors, $(\mathbf{x}_m)_{m=1}^{\infty}$, as follows. We fix integers $\mu \geq 0$ and $\alpha \geq 1$, and let $\phi : (V_{\mathcal{E}})^{\mu+\alpha} \rightarrow [0, \infty)$ be a nonnegative function. Our method works for sequences $(\mathbf{x}_{m_k})_{k=1}^{\infty} = (\mathbf{z}_{m_k}^{\phi})_{k=1}^{\infty}$ where $m_k = \mu + k\alpha$ for positive integers k , and $\mathbf{z}_{m_k}^{\phi}$ is the $F_{m_k} \times 1$ nonnegative vector indexed by $(V_{\mathcal{E}})^{m_k}$:

$$(\mathbf{z}_{m_k}^{\phi})_{\mathbf{v}} = \prod_{i=0}^{k-1} \phi(\mathbf{v}_{i\alpha:i\alpha+\mu+\alpha-1}) \quad ; \quad \mathbf{v} \in (V_{\mathcal{E}})^{m_k}. \quad (6)$$

For such sequences and a fixed positive integer n , we will show that one can compute L_n , the growth rate of $\mathbf{x}_{m_k}^\dagger H_{m_k}^n \mathbf{x}_{m_k}$:

$$L_n = \lim_{k \rightarrow \infty} \frac{\log \mathbf{x}_{m_k}^\dagger H_{m_k}^n \mathbf{x}_{m_k}}{m_k},$$

and from (5) we obtain the lower bound $\text{cap}(S) \geq (L_{2q+p} - L_{2q})/p$.

Before doing this for general μ and α , it is instructive to look at the special case: $\mu = 0$ and $\alpha = 1$. In this case $m_k = k$, and

$$\left(\mathbf{z}_k^\phi \right)_{\mathbf{v}} = \prod_{i=0}^{k-1} \phi((\mathbf{v})_i) \quad ; \quad \mathbf{v} \in (V_{\mathcal{E}})^k.$$

Let $(\mathbf{x}_m)_{m=1}^\infty = (\mathbf{z}_m^\phi)_{m=1}^\infty$ and let n be a positive integer. For a word $w = w_1 \dots w_n \in \mathcal{E}$ define its weight, $\mathcal{W}_\phi(w)$, by $\mathcal{W}_\phi(w) = \phi(\sigma(w_1)) \phi(\tau(w_n))$, where w_1, w_n are regarded as edges in $G_{\mathcal{E}}$ and extend this to arrays $\Gamma \in S_{m \times n}$ by

$$\mathcal{W}_\phi(\Gamma) = \prod_{i=0}^{m-1} \mathcal{W}_\phi(\Gamma_{i,:}).$$

Observe that for an array $\Gamma \in S_{m \times n}$ that is a path in $G_m^{(H)}$ of length n starting at $\mathbf{v} \in (V_{\mathcal{E}})^m$ and ending at $\mathbf{u} \in (V_{\mathcal{E}})^m$, it holds that $\mathcal{W}_\phi(\Gamma) = (\mathbf{z}_m^\phi)_{\mathbf{v}} (\mathbf{z}_m^\phi)_{\mathbf{u}}$. It follows that

$$\mathbf{x}_m^\dagger H_m^n \mathbf{x}_m = \sum_{\Gamma \in S_{m \times n}} \mathcal{W}_\phi(\Gamma). \quad (7)$$

Now, pick a deterministic presentation, $\mathcal{G}_n^{(\mathcal{V})} = (V_n^{(\mathcal{V})}, E_n^{(\mathcal{V})}, \mathcal{L}_n^{(\mathcal{V})})$ of $\mathcal{V}_n(S)$, and let $\mathcal{W}_\phi : E \rightarrow [0, \infty)$ be the edge weighting defined by $\mathcal{W}_\phi(e) = \mathcal{W}_\phi(\mathcal{L}_n^{(\mathcal{V})}(e))$ for $e \in E_n^{(\mathcal{V})}$. Let $A(\mathcal{G}_n^{(\mathcal{V})}, \mathcal{W}_\phi)$ be the $|V_n^{(\mathcal{V})}| \times |V_n^{(\mathcal{V})}|$ weighted adjacency matrix of $\mathcal{G}_n^{(\mathcal{V})}$ with entries indexed by $V_n^{(\mathcal{V})} \times V_n^{(\mathcal{V})}$ and given by

$$\left(A(\mathcal{G}_n^{(\mathcal{V})}, \mathcal{W}_\phi) \right)_{i,j} = \sum_{\substack{e \in E_n^{(\mathcal{V})} \\ \sigma(e)=i, \tau(e)=j}} \mathcal{W}_\phi(e) \quad ; \quad i, j \in V_n^{(\mathcal{V})}.$$

Then

$$\mathbf{1}^\dagger A(\mathcal{G}_n^{(\mathcal{V})}, \mathcal{W}_\phi)^m \mathbf{1} = \sum_{\gamma} \mathcal{W}_\phi(\mathcal{L}_n^{(\mathcal{V})}(\gamma)),$$

where the sum is taken over all paths γ in $\mathcal{G}_n^{(\mathcal{V})}$ of length m and $\mathcal{L}_n^{(\mathcal{V})}(\gamma)$ denotes the array in $S_{m \times n}$ generated by γ . Since $\mathcal{G}_n^{(\mathcal{V})}$ is deterministic it follows that

$$\lim_{m \rightarrow \infty} \frac{\log \mathbf{1}^\dagger A(\mathcal{G}_n^{(\mathcal{V})}, \mathcal{W}_\phi)^m \mathbf{1}}{m} = \lim_{m \rightarrow \infty} \frac{\log \sum_{\Gamma \in S_{m \times n}} \mathcal{W}_\phi(\Gamma)}{m}.$$

By (7) the RHS is L_n and by Perron-Frobenius theory the LHS is $\log \lambda(A(\mathcal{G}_n^{(\mathcal{V})}, \mathcal{W}_\phi))$, and thus $L_n = \log \lambda(A(\mathcal{G}_n^{(\mathcal{V})}, \mathcal{W}_\phi))$.

For general μ, α , we proceed similarly. We pick a deterministic presentation $\mathcal{G}_n^{(\mathcal{V})} = (V_n^{(\mathcal{V})}, E_n^{(\mathcal{V})}, \mathcal{L}_n^{(\mathcal{V})})$ of $\mathcal{V}_n(S)$ and construct a labeled directed graph

$\mathcal{I} = \mathcal{I}(\mu, \alpha, n, \mathcal{G}_n^{(V)}, G_{\mathcal{E}}) = (V_{\mathcal{I}}, E_{\mathcal{I}}, \mathcal{L}_{\mathcal{I}})$, with nonnegative real weights on its edges given by $\mathcal{W}_{\phi} : E_{\mathcal{I}} \rightarrow [0, \infty)$. The graph \mathcal{I} and weight function \mathcal{W}_{ϕ} are defined as follows. The set of vertices $V_{\mathcal{I}}$ is given by

$$V_{\mathcal{I}} = \left\{ (\mathbf{f}, v, \mathbf{l}) : v \in V_n^{(V)}, \mathbf{f}, \mathbf{l} \in (V_{\mathcal{E}})^{\mu} \right\},$$

and the function $\mathcal{L}_{\mathcal{I}} : E_{\mathcal{I}} \rightarrow \Sigma^{\alpha \times n}$ labels each edge with an $\alpha \times n$ array over Σ . We specify the edges of \mathcal{I} by describing the outgoing edges of each of its vertices along with their weights. Let $\mathbf{v} = (\mathbf{f}, v, \mathbf{l}) \in V_{\mathcal{I}}$ be a vertex of \mathcal{I} . The set of outgoing edges of \mathbf{v} consists of exactly one edge for every path of length α in $\mathcal{G}_n^{(V)}$ starting at v . Let $\gamma = (e_i)_{i=0}^{\alpha-1} \subseteq E_n^{(V)}$ be such a path and let u be its terminating vertex. We regard the word generated by γ in $\mathcal{G}_n^{(V)}$ as an array $\Gamma \in \Sigma^{\alpha \times n}$ with entries given by $(\Gamma)_{i,j} = \left(\mathcal{L}_n^{(V)}(e_i) \right)_j$. Let $\mathbf{f} = (f_0, \dots, f_{\mu-1})$ and $\mathbf{l} = (l_0, \dots, l_{\mu-1})$ and for $i = \mu, \mu+1, \dots, \mu+\alpha-1$, define f_i to be $\sigma(\Gamma_{i-\mu,0})$ and l_i to be $\tau(\Gamma_{i-\mu, n-1})$, where $\Gamma_{i-\mu,0}$ and $\Gamma_{i-\mu, n-1}$ are regarded as edges in the graph $G_{\mathcal{E}}$. For such a path γ the corresponding outgoing edge $e \in E_{\mathcal{I}}$ of \mathbf{v} satisfies $\sigma(e) = \mathbf{v}$, $\mathcal{L}_{\mathcal{I}}(e) = \Gamma$, $\tau(e) = ((f_{\alpha}, f_{\alpha+1}, \dots, f_{\mu+\alpha-1}), u, (l_{\alpha}, l_{\alpha+1}, \dots, l_{\mu+\alpha-1}))$. The weight of e , $\mathcal{W}_{\phi}(e)$, is given by $\mathcal{W}_{\phi}(e) = \phi(f_0, \dots, f_{\mu+\alpha-1})\phi(l_0, \dots, l_{\mu+\alpha-1})$. We shall regard the label of a path $(e_i)_{i=0}^{\ell-1}$ in \mathcal{I} as the $\ell\alpha \times n$ array Γ over Σ resulting from concatenating the labels of the edges of γ in order in the vertical direction, namely $\Gamma_{i\alpha+k,j} = (\mathcal{L}_{\mathcal{I}}(e_i))_{k,j}$, for all $i \in [\ell]$, $k \in [\alpha]$ and $j \in [n]$. Finally, we define the weighted adjacency matrix of the labeled directed graph \mathcal{I} with weights given by \mathcal{W}_{ϕ} as the $|V_{\mathcal{I}}| \times |V_{\mathcal{I}}|$ nonnegative real matrix $A(\mathcal{I}, \mathcal{W}_{\phi})$ with entries indexed by $(V_{\mathcal{I}})^2$ and given by

$$(A(\mathcal{I}, \mathcal{W}_{\phi}))_{i,j} = \sum_{\substack{e \in E_{\mathcal{I}}: \\ \sigma(e)=i, \tau(e)=j}} \mathcal{W}_{\phi}(e) \quad ; \quad i, j \in V_{\mathcal{I}}.$$

The following lemma generalizes ideas in [1] and uses the weighted labeled graph \mathcal{I} to compute L_n , when $(\mathbf{x}_{m_k})_{k=1}^{\infty} = (\mathbf{z}_{m_k}^{\phi})_{k=1}^{\infty}$.

Lemma 1. For $(\mathbf{x}_{m_k})_{k=1}^{\infty} = (\mathbf{z}_{m_k}^{\phi})_{k=1}^{\infty}$, and $\mathcal{I} = \mathcal{I}(\mu, \alpha, n, \mathcal{G}_n^{(V)}, G_{\mathcal{E}})$,

$$\lim_{k \rightarrow \infty} \frac{\log \mathbf{x}_{m_k}^{\dagger} H_{m_k}^n \mathbf{x}_{m_k}}{m_k} = \frac{\log \lambda(A(\mathcal{I}, \mathcal{W}_{\phi}))}{\alpha}.$$

Proof. We shall show that there are positive real constants c, d such that for all positive integers k ,

$$c \cdot \mathbf{1}^{\dagger} (A(\mathcal{I}, \mathcal{W}_{\phi}))^{k + \lceil \mu/\alpha \rceil} \mathbf{1} \leq \mathbf{x}_{m_k}^{\dagger} H_{m_k}^n \mathbf{x}_{m_k} \leq d \cdot \mathbf{1}^{\dagger} (A(\mathcal{I}, \mathcal{W}_{\phi}))^k \mathbf{1}. \quad (8)$$

For a positive integer s and vector $\mathbf{e} = (e_0, \dots, e_{s-1})$ in $(E_{\mathcal{E}})^s$ denote by $\sigma(\mathbf{e}), \tau(\mathbf{e}) \in (V_{\mathcal{E}})^s$ the vectors with entries given by

$$\begin{aligned} (\sigma(\mathbf{e}))_i &= \sigma(e_i) \\ (\tau(\mathbf{e}))_j &= \tau(e_j) \end{aligned} \quad ; \quad i \in \{0, 1, \dots, s-1\}.$$

Now, fix a positive integer k . Let Γ be an array in $S_{m_k \times n}$. Recall that each entry of Γ is an edge in $G_{\mathcal{E}}$ and define the weight of Γ , denoted $\mathcal{W}_{\phi}(\Gamma)$, by

$$\mathcal{W}_{\phi}(\Gamma) = \prod_{i=0}^{k-1} \phi(\sigma(\Gamma_{i\alpha:i\alpha+\mu+\alpha-1,0}))\phi(\tau(\Gamma_{i\alpha:i\alpha+\mu+\alpha-1,n-1})).$$

For a positive integer ℓ , let \mathcal{P}_ℓ denote the set of paths of length ℓ in \mathcal{I} . We denote the label of a path $\gamma \in \mathcal{P}_\ell$ by $\mathcal{L}_\mathcal{I}(\gamma)$. It is easily verified that there exists a path in \mathcal{P}_ℓ with label $\Gamma \in \Sigma^{\ell\alpha \times n}$ if and only if there exists a path in $\mathcal{G}_n^{(\mathcal{V})}$ of length $\ell\alpha$ that generates Γ . As $\mathcal{G}_n^{(\mathcal{V})}$ is a presentation of \mathcal{V}_n , the set of labels of paths in \mathcal{P}_ℓ is $S_{\ell\alpha \times n}$.

For a finite path γ in \mathcal{I} , define its weight, denoted $\mathcal{W}_\phi(\gamma)$, as the product of the weights of the edges in the path. Recalling that the entries of $\mathbf{x}_{m_k} = \mathbf{z}_{m_k}^\phi$ are indexed by $(V_\mathcal{E})^{m_k}$, we observe that

$$\begin{aligned} \mathbf{x}_{m_k}^\dagger H_{m_k}^n \mathbf{x}_{m_k} &= \sum_{\Gamma \in S_{m_k \times n}} (\mathbf{x}_{m_k})_{\sigma(\Gamma_{*,0})} (\mathbf{x}_{m_k})_{\tau(\Gamma_{*,n-1})} \\ &= \sum_{\Gamma \in S_{m_k \times n}} \mathcal{W}_\phi(\Gamma). \end{aligned}$$

For an array $\Gamma \in S_{m_k \times n}$, we say that a path $\gamma \in \mathcal{P}_k$ *matches* Γ if it is labeled by the sub-array $\Gamma_{\mu:m_k-1,*}$ and starts at a vertex $(\mathbf{f}, v, \mathbf{l}) \in V_\mathcal{I}$ with $\mathbf{f} = \sigma(\Gamma_{0:\mu-1,0})$ and $\mathbf{l} = \tau(\Gamma_{0:\mu-1,n-1})$. It can be verified from the construction of \mathcal{I} that if γ matches Γ then $\mathcal{W}_\phi(\gamma) = \mathcal{W}_\phi(\Gamma)$.

Now, since $\mathcal{G}_n^{(\mathcal{V})}$ is a presentation of $\mathcal{V}_n(S)$, it follows from the construction of \mathcal{I} that every $\Gamma \in S_{m_k \times n}$ has a path in \mathcal{P}_k matching it. Conversely, since for a path $\gamma \in \mathcal{P}_k$ all arrays $\Gamma \in S_{m_k \times n}$ that it matches have the same sub-array $\Gamma_{\mu:m_k-1,*}$, it follows that there are at most $|\Sigma|^{\mu n}$ arrays in $S_{m_k \times n}$ that γ matches. Therefore,

$$\begin{aligned} \mathbf{x}_{m_k}^\dagger H_{m_k}^n \mathbf{x}_{m_k} &= \sum_{\Gamma \in S_{m_k \times n}} \mathcal{W}_\phi(\Gamma) \\ &\leq |\Sigma|^{\mu n} \sum_{\gamma \in \mathcal{P}_k} \mathcal{W}_\phi(\gamma) \\ &= |\Sigma|^{\mu n} \mathbf{1}^\dagger (A(\mathcal{I}, \mathcal{W}_\phi))^k \mathbf{1}. \end{aligned}$$

This shows the right inequality of (8), we now turn to the left. Set $k' = k + \lceil \mu/\alpha \rceil$, $s = \lceil \mu/\alpha \rceil \alpha - \mu$, and let $\psi : \mathcal{P}_{k'} \rightarrow S_{m_k \times n}$ be given by

$$\psi(\gamma) = (\mathcal{L}_\mathcal{I}(\gamma))_{s:k'\alpha-1,*} ; \gamma \in \mathcal{P}_{k'}.$$

For a path $\gamma \in \mathcal{P}_{k'}$, with $\gamma = (e_i)_{i=0}^{k'-1} \subseteq E_\mathcal{I}$, its weight satisfies

$$\mathcal{W}_\phi(\gamma) = \prod_{i=0}^{k'-1} \mathcal{W}_\phi(e_i) = \left(\prod_{i=0}^{\lceil \mu/\alpha \rceil - 1} \mathcal{W}_\phi(e_i) \right) \mathcal{W}_\phi(\psi(\gamma)) \leq \Phi^{2\lceil \mu/\alpha \rceil} \mathcal{W}_\phi(\psi(\gamma)),$$

where we take Φ to be a positive constant satisfying $\Phi \geq \max \{\phi(\mathbf{v}) : \mathbf{v} \in (V_\mathcal{E})^{\mu+\alpha}\}$. Now let Γ be an array in $S_{m_k \times n}$. Since $\mathcal{G}_n^{(\mathcal{V})}$ is deterministic, so is \mathcal{I} , and thus for every vertex $v \in V_\mathcal{I}$, the paths in $\mathcal{P}_{k'}$ starting at v are labeled distinctly. As all paths γ that map to Γ under ψ have the same sub-array $(\mathcal{L}_\mathcal{I}(\gamma))_{s:k'\alpha-1,*}$, it follows that there are at most $|\Sigma|^{sn}$ paths $\gamma \in \mathcal{P}_{k'}$ starting at v such that $\psi(\gamma) = \Gamma$. Consequently, there are at most $|V_\mathcal{I}| |\Sigma|^{sn}$ paths in $\mathcal{P}_{k'}$ that map to

Γ under ψ . Therefore,

$$\begin{aligned}
\mathbf{1}^\top (A(\mathcal{I}, \mathcal{W}_\phi))^{k+\lceil \mu/\alpha \rceil} \mathbf{1} &= \sum_{\gamma \in \mathcal{P}_{k'}} \mathcal{W}_\phi(\gamma) \\
&\leq \Phi^{2\lceil \mu/\alpha \rceil} \sum_{\gamma \in \mathcal{P}_{k'}} \mathcal{W}_\phi(\psi(\gamma)) \\
&\leq \Phi^{2\lceil \mu/\alpha \rceil} |V_{\mathcal{I}}| |\Sigma|^{sn} \sum_{\Gamma \in S_{m_k \times n}} \mathcal{W}_\phi(\Gamma) \\
&= \Phi^{2\lceil \mu/\alpha \rceil} |V_{\mathcal{I}}| |\Sigma|^{sn} \mathbf{x}_{m_k}^\top H_{m_k}^n \mathbf{x}_{m_k}.
\end{aligned}$$

Dividing both sides by $\Phi^{2\lceil \mu/\alpha \rceil} |V_{\mathcal{I}}| |\Sigma|^{sn}$ we obtain the left inequality of (8). The claim of the lemma now follows from Perron-Frobenius theory by taking the log of (8), dividing it by m_k and taking the limit as k approaches infinity. \blacksquare

We thus obtain the following lower bound on the capacity of a 2-dimensional constraint.

Theorem 1. *Let T, \mathcal{E} be 1-dimensional constraints over an alphabet Σ , with \mathcal{E} an edge constraint defined by a graph $G_{\mathcal{E}} = (V_{\mathcal{E}}, E_{\mathcal{E}})$. Set $S = T \otimes \mathcal{E}$ and suppose that S has symmetric horizontal edge-constrained strips. Let $\mu \geq 0$ and $\alpha, p, q > 0$ be integers and $\phi : (V_{\mathcal{E}})^{\mu+\alpha} \rightarrow [0, \infty)$ be a nonnegative real function. For a positive integer n let \mathcal{G}_n be a deterministic presentation of $\mathcal{V}_n(S)$ and set $A_{n,\phi} = A(\mathcal{I}(\mu, \alpha, n, \mathcal{G}_n, G_{\mathcal{E}}), \mathcal{W}_\phi)$. Then*

$$\text{cap}(S) \geq \frac{\log \lambda(A_{2q+p,\phi}) - \log \lambda(A_{2q,\phi})}{p\alpha}. \quad (9)$$

Remark 1. In addition to computing lower-bounds, [1] gives a method for computing upper bounds on the capacity of the hard-square constraint. It can be shown that this method can also be applied to all constraints of the form $T \otimes \mathcal{E}$, with \mathcal{E} an edge constraint, having symmetric horizontal edge-constrained strips.

Remark 2. Theorem 1 can be generalized to apply to 2-dimensional constraints having symmetric horizontal edge-constrained strips, which are not necessarily axial-products. Let S be such a constraint, and for every positive integer m , let $G_m^{(\mathcal{H})} = (V_m^{(\mathcal{H})}, E_m^{(\mathcal{H})})$ be the symmetric graph, with no isolated vertices, defining $\mathcal{H}_m(S)$. Set $G_{\mathcal{E}} = G_1^{(\mathcal{H})}$, $V_{\mathcal{E}} = V_1^{(\mathcal{H})}$ and $E_{\mathcal{E}} = E_1^{(\mathcal{H})}$. We claim there exists a mapping $f_m : V_m^{(\mathcal{H})} \rightarrow (V_{\mathcal{E}})^m$ such that for every edge $\mathbf{e} = e_0 e_1 \dots e_{m-1} \in E_m^{(\mathcal{H})}$, with each $e_i \in E_{\mathcal{E}}$,

$$f_m(\sigma(\mathbf{e})) = (\sigma(e_0), \dots, \sigma(e_{m-1})) \text{ and } f_m(\tau(\mathbf{e})) = (\tau(e_0), \dots, \tau(e_{m-1})). \quad (10)$$

This mapping is defined as follows. For a vertex $v \in V_m^{(\mathcal{H})}$, pick an incoming edge $\mathbf{e} = e_0 e_1 \dots e_{m-1} \in E_m^{(\mathcal{H})}$ and define $f_m(v)$ as $(\tau(e_0), \dots, \tau(e_{m-1}))$. This mapping is uniquely-defined: indeed if $\mathbf{e}' = e'_0 \dots e'_{m-1} \in E_m^{(\mathcal{H})}$ is another incoming edge of v , and $\mathbf{g} = g_0 \dots g_{m-1} \in E_m^{(\mathcal{H})}$ is an outgoing edge of v , then clearly, for every $i \in [m]$, both $e_i g_i$ and $e'_i g_i$ are paths in $G_{\mathcal{E}}$; consequently $\tau(e_i) = \tau(e'_i)$. It is easy to check that f_m satisfies the conditions in (10). Now, replace the definition of $\mathbf{z}_{m_k}^\phi$ in (6) with

$$(\mathbf{z}_{m_k}^\phi)_{\mathbf{v}} = \prod_{i=0}^{k-1} \phi(\mathbf{u}_{i\alpha:i\alpha+\mu+\alpha-1}) \quad ; \quad \mathbf{u} = f_{m_k}(\mathbf{v}), \mathbf{v} \in V_{m_k}^{(\mathcal{H})}.$$

With this new definition and the aid of (10), it can be verified that Lemma 1 and consequently Theorem 1 still hold.

Remark 3. Clearly, it is sufficient, for the theorem to hold, that $\mathcal{H}_m(S)$ is symmetric for large enough m .

We now give a sufficient condition for the constraint $S = T \otimes \mathcal{E}$ to have symmetric horizontal edge-constrained strips. For this to happen, we (generally) must have that $G_{\mathcal{E}}$ is symmetric. This means that there exists a “matching” between edges, where each edge is matched with an edge in the “reverse” direction. More precisely there is a bijection $R : E_{\mathcal{E}} \rightarrow E_{\mathcal{E}}$ such that for all $e \in E_{\mathcal{E}}$, $(\sigma(e), \tau(e)) = (\tau(R(e)), \sigma(R(e)))$ and $R(R(e)) = e$. We call such a bijection an *edge-reversing matching*, and we denote by $\mathcal{R}(G_{\mathcal{E}})$ the set of all edge-reversing matchings of $G_{\mathcal{E}}$. Clearly a graph G is symmetric iff it has an edge-reversing matching. Thus $T \otimes \mathcal{E}$ has symmetric horizontal edge-constrained strips iff for every m , $G_m^{(\mathcal{H})}$ has an edge-reversing matching. We present a sufficient condition for this to hold.

Proposition 3. *Let T, \mathcal{E} be 1-dimensional constraints over an alphabet Σ , with \mathcal{E} an edge constraint defined by a graph $G_{\mathcal{E}} = (V_{\mathcal{E}}, E_{\mathcal{E}})$ with $R \in \mathcal{R}(G_{\mathcal{E}})$ an edge-reversing matching. If for all positive integers m and words $e_1 \dots e_m \in E_{\mathcal{E}}^m$, if $e_1 \dots e_m$ satisfies T then $R(e_1) \dots R(e_m)$ satisfies T as well, then $T \otimes \mathcal{E}$ has symmetric horizontal edge-constrained strips.*

Proof. Let $G_m^{(\mathcal{H})} = (V_{\mathcal{E}}^m, E_m^{(\mathcal{H})})$ be the sub-graph of $G_{\mathcal{E}}^{\times m}$ that defines $\mathcal{H}_m(S)$. We show that $G_m^{(\mathcal{H})}$ is symmetric. Let $R^{\times m} : E_{\mathcal{E}}^m \rightarrow E_{\mathcal{E}}^m$ be defined by

$$R^{\times m}(e_1, \dots, e_m) = (R(e_1), \dots, R(e_m)).$$

Clearly, $R^{\times m}$ is an edge-reversing matching of $G_{\mathcal{E}}^{\times m}$. Recall that $E_m^{(\mathcal{H})}$ consists of all the edges in $E_{\mathcal{E}}^m$ that, when regarded as m -letter words over Σ , satisfy T . Therefore, by the assumption, it follows that for all $\mathbf{e} \in E_m^{(\mathcal{H})}$, $R^{\times m}(\mathbf{e}) \in E_m^{(\mathcal{H})}$ as well. Consequently, $R^{\times m}$ restricted to $E_m^{(\mathcal{H})}$, is an edge-reversing matching of $G_m^{(\mathcal{H})}$ and hence it is symmetric. ■

If \mathcal{G}_T is a presentation of T and $R \in \mathcal{R}(G_{\mathcal{E}})$, a sufficient condition for the hypothesis of Proposition 3 to hold, which may be easier to check, is the existence of a function $f : E_T \rightarrow E_T$ satisfying: 1) $\mathcal{L}_T(f(e)) = R(\mathcal{L}_T(e))$ for all $e \in E_T$ and 2) for any path $e_1 e_2$ of length 2 in \mathcal{G}_T , the sequence $f(e_1) f(e_2)$ is also a path in \mathcal{G}_T . Indeed, if such a function exists then any path $\epsilon_1 \epsilon_2 \dots \epsilon_m$ in G_T generating a word $e_1 e_2 \dots e_m$ has a corresponding path $f(\epsilon_1) f(\epsilon_2) \dots f(\epsilon_m)$ generating the word $R(e_1) R(e_2) \dots R(e_m)$, and thus the hypothesis of Proposition 3 is fulfilled. In fact, it can be shown that when \mathcal{G}_T is irreducible, deterministic and has the minimum number of states among all deterministic presentations of T , this condition is also necessary for the hypothesis of Proposition 3 to hold (see [13, Section 3.3]).

In Section 5 we use Proposition 3 to show that the method described in this section can be used to compute lower bounds on $\text{CHG}(b_1) \otimes \text{CHG}(b_2)$ for any positive integers b_1 and b_2 .

4 Constraints with symmetric vertex-constrained strips.

In this section we present an analog to Theorem 1 that gives lower bounds on the capacities of constraints for which every horizontal or every vertical strip is a symmetric vertex-constraint. We do this by transforming a 2-dimensional constraint with symmetric vertex-constrained strips to a 2-dimensional constraint with symmetric edge-constrained strips, having the same capacity.

Fix an alphabet Σ , and let S be a 2-dimensional constraint over Σ . We say that S has *horizontal vertex-constrained strips* if for every positive integer m , the constraint $\mathcal{H}_m(S)$ is a vertex-constraint. If, in addition, every horizontal strip is symmetric, we say that S has *symmetric horizontal vertex-constrained strips*. The notions of a 2-dimensional constraint with vertical vertex-constrained strips and symmetric vertical vertex-constrained strips are defined analogously.

It turns out that RWIM and NAK do not have horizontal or vertical edge-constrained strips, and so the method in section 3 does not apply directly. We illustrate this only for horizontal strips for $S = \text{RWIM}$. Recall from section 2 that an edge constraint is a constraint of memory 1 such that any two follower sets are either disjoint or identical. We claim that this condition does not hold for $\mathcal{H}_m(S)$. To see this, given any m , let w be the all-zeros word of length m and $u \neq w$ be any other word of length m . Now, the $m \times 2$ arrays

$$\begin{array}{cccccc} w_0 & w_0 & w_0 & u_0 & u_0 & w_0 \\ w_1 & w_1 & w_1 & u_1 & u_1 & w_1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ w_{m-1} & w_{m-1} & w_{m-1} & u_{m-1} & u_{m-1} & w_{m-1} \end{array},$$

belong to S , yet the $m \times 2$ array

$$\begin{array}{cc} u_0 & u_0 \\ u_1 & u_1 \\ \dots & \dots \\ u_{m-1} & u_{m-1} \end{array},$$

does not. Thus, w and u , viewed as $m \times 1$ columns have different but non-disjoint follower sets. Thus, RWIM does not have horizontal edge-constrained strips.

However, it is not hard to show that RWIM and NAK have both symmetric horizontal vertex-constrained strips and symmetric vertical vertex-constrained strips. For instance, for $S = \text{RWIM}$, $\mathcal{H}_m(S)$ is the vertex constraint defined by the graph $G = (V, E)$, where V consists of all binary vectors $u_0 \dots u_{m-1}$ of length m and E consists of a single edge from $u \in V$ to $v \in V$ iff for all i , whenever $u_i = 1$, then $v_{i+1} = v_i = v_{i-1} = 0$ (with the obvious modification when $i = 0$ or $m - 1$). And $\mathcal{V}_m(S)$ is the vertex constraint defined by the graph of $G' = (V', E')$, where V' consists of all binary vectors $u_0 \dots u_{m-1}$ of length m which do not contain two adjacent '1's and E' consists of a single edge from $u \in V$ to $v \in V$ iff for all i , whenever $u_i = 1$, then $v_{i+1} = v_{i-1} = 0$ (again, with the obvious modification when $i = 0$ or $m - 1$). Clearly, both G and G' are symmetric.

Now, let S be a 2-dimensional constraint over Σ . For a finite $m \times n$ array Γ with $m \geq 1$ and $n \geq 2$ over Σ its $[1 \times 2]$ -higher block recoding or $[1 \times 2]$ -recoding

is an $m \times (n - 1)$ array $\hat{\Gamma}$ over $\Sigma^{1 \times 2}$ with entries given by

$$\hat{\Gamma}_{i,j} = (\Gamma_{i,j} \Gamma_{i,j+1}) ; i = 0, \dots, m - 1, j = 0, \dots, n - 2.$$

We denote by $S^{[1 \times 2]}$ the set of all $[1 \times 2]$ -recodings of arrays in S and refer to it as the $[1 \times 2]$ -higher block recoding of S . The $[1 \times 2]$ -higher block recoding of a constraint is a constraint. This is stated in the following proposition.

Proposition 4. *Let S be a 2-dimensional constraint over Σ . Then $S^{[1 \times 2]}$ is a 2-dimensional constraint over $\Sigma^{1 \times 2}$.*

Remark . We may, of course, define, in a similar manner, the $[s \times t]$ -higher block recoding of S , for any positive integers s and t , and the $[s \times t]$ -higher block recoding of a 2-dimensional constraint S is a 2-dimensional constraint.

Proof. Set $S' = S^{[1 \times 2]}$ and let $(\mathcal{G}_V, \mathcal{G}_H)$ be a presentation of S with $\mathcal{G}_V = (V_V, E, \mathcal{L})$ and $\mathcal{G}_H = (V_H, E, \mathcal{L})$. We construct labeled graphs $\mathcal{G}'_V = (V_V \times V_V, E', \mathcal{L}')$ and $\mathcal{G}'_H = (E, E', \mathcal{L}')$ as follows. The set of edges E' is defined as

$$E' = \{(e_0 e_1) \in E^{1 \times 2} : e_0, e_1 \text{ is a path in } \mathcal{G}^{(H)}\},$$

and the labeling function $\mathcal{L}' : E' \rightarrow \Sigma^{1 \times 2}$ is given by

$$\mathcal{L}'(e_0 e_1) = (\mathcal{L}(e_0) \mathcal{L}(e_1)) ; (e_0 e_1) \in E'.$$

For every edge $(e_0 e_1) \in E'$ we define

$$\begin{aligned} \sigma_{\mathcal{G}'_V}(e_0 e_1) &= (\sigma_{\mathcal{G}_V}(e_0), \sigma_{\mathcal{G}_V}(e_1)) \\ \tau_{\mathcal{G}'_V}(e_0 e_1) &= (\tau_{\mathcal{G}_V}(e_0), \tau_{\mathcal{G}_V}(e_1)) \\ \sigma_{\mathcal{G}'_H}(e_0 e_1) &= e_0 \\ \tau_{\mathcal{G}'_H}(e_0 e_1) &= e_1 \end{aligned}.$$

It's easy to verify that $(\mathcal{G}'_V, \mathcal{G}'_H)$ is a presentation of S' . ■

Clearly, recoding is an injective mapping, thus $|S_{m \times n}| = |S_{m \times (n-1)}^{[1 \times 2]}|$ for all positive integers $m \geq 1, n \geq 2$. It follows that $\text{cap}(S) = \text{cap}(S^{[1 \times 2]})$. The next proposition shows that the $[1 \times 2]$ -higher block recoding of a constraint with symmetric horizontal vertex-constrained strips has symmetric horizontal edge-constrained strips.

Proposition 5. *Let S be a 2-dimensional constraint with horizontal vertex-constrained strips.*

1. $S^{[1 \times 2]}$ has horizontal edge-constrained strips. Moreover, $S^{[1 \times 2]}$ has symmetric horizontal edge-constrained strips iff S has symmetric horizontal vertex-constrained strips.
2. $S^{[1 \times 2]} = \mathcal{V}_1(S^{[1 \times 2]}) \otimes \mathcal{H}_1(S^{[1 \times 2]})$.

Proof. (1). Let m be a positive integer, $G_m^{(\mathcal{H})} = (V_m^{(\mathcal{H})}, E_m)$ be the graph defining the vertex-constraint $\mathcal{H}_m(S)$ and set $\Delta = \Sigma^{1 \times 2}$, where Σ is the alphabet of S . We define a labeling $\mathcal{L} : E_m \rightarrow \Delta^{m \times 1}$ of the edges of $G_m^{(\mathcal{H})}$. For $e \in E_m$,

we regard $\sigma(\mathbf{e})$ and $\tau(\mathbf{e})$ as $m \times 1$ arrays over Σ , and define $\mathcal{L}(\mathbf{e})$ to be the array in $\Delta^{m \times 1}$ with entries given by

$$\mathcal{L}(\mathbf{e})_{(i,0)} = (\sigma(\mathbf{e})_{i,0} \tau(\mathbf{e})_{i,0}) ; \quad i \in [m].$$

It's easily verified that the word generated by every path in the labeled graph $(G_m^{(\mathcal{H})}, \mathcal{L})$ is the $[1 \times 2]$ -higher block recoding of the array formed by concatenating the vertices along the path horizontally in sequence. It follows that $(G_m^{(\mathcal{H})}, \mathcal{L})$ is a presentation of $\mathcal{H}_m(S^{[1 \times 2]})$. Since the labels of the edges in $(G_m^{(\mathcal{H})}, \mathcal{L})$ are distinct, we may identify each edge with its label, and it follows that $\mathcal{H}_m(S^{[1 \times 2]})$ is an edge-constraint. Since the same graph defines both $\mathcal{H}_m(S)$ and $\mathcal{H}_m(S^{[1 \times 2]})$ (the former as a vertex-constraint and the latter as an edge-constraint), it follows that $\mathcal{H}_m(S)$ is symmetric iff $\mathcal{H}_m(S^{[1 \times 2]})$ is. This completes the proof.

(2). Clearly, $S^{[1 \times 2]} \subseteq \mathcal{V}_1(S^{[1 \times 2]}) \otimes \mathcal{H}_1(S^{[1 \times 2]})$. As for the reverse inclusion, let $\hat{\Gamma} \in \mathcal{V}_1(S^{[1 \times 2]}) \otimes \mathcal{H}_1(S^{[1 \times 2]})$ be an $m \times n$ array over $\Sigma^{1 \times 2}$. Since every row of $\hat{\Gamma}$ is in $\mathcal{H}_1(S^{[1 \times 2]})$, every row has a unique $1 \times (n+1)$ pre-image under the recoding map. Let Γ be the $m \times (n+1)$ array over Σ , whose i th row is the pre-image under the recoding map of the i th row of $\hat{\Gamma}$, for $i = 0, \dots, m-1$. Clearly, $\hat{\Gamma}$ is the $[1 \times 2]$ -higher block recoding of Γ . Thus, it suffices to show that $\Gamma \in S$. For $i = 0, 1, \dots, n-1$ clearly, the $m \times 2$ array $\Gamma_{*,i:i+1}$ over Σ recodes to the column $\hat{\Gamma}_{*,i}$. By our assumption this column is in $\mathcal{V}_1(S^{[1 \times 2]})$. Since recoding is injective, $\Gamma_{*,i:i+1}$ must be in S . Since this holds for all $i = 0, 1, \dots, n-1$ and since $\mathcal{H}_m(S)$ has memory 1, it follows that $\Gamma \in S$ and therefore $\hat{\Gamma} \in S^{[1 \times 2]}$. ■

We can now use the method described in Section 3 to get lower bounds on 2-dimensional constraints with symmetric horizontal vertex-constrained strips. This is stated in the following theorem.

Theorem 2. *Let S be a 2-dimensional constraint over an alphabet Σ with symmetric horizontal vertex-constrained strips. Let $\mu \geq 0$, and $\alpha, p, q > 0$ be integers, $G_{\mathcal{E}} = (V_{\mathcal{E}}, E_{\mathcal{E}})$ be the graph defining the vertex-constraint $\mathcal{H}_1(S)$ (hence $V_{\mathcal{E}} \subseteq \Sigma$), and $\phi : (V_{\mathcal{E}})^{\mu+\alpha} \rightarrow [0, \infty)$ be a nonnegative function. For an integer $n \geq 2$, let \mathcal{G}_n be a labeled graph obtained from a deterministic presentation for $\mathcal{V}_n(S)$ by replacing each edge-label with its $[1 \times 2]$ -higher block recoding. Set $\hat{A}_{n,\phi} = A(\mathcal{I}(\mu, \alpha, n-1, \mathcal{G}_n, G_{\mathcal{E}}), \mathcal{W}_{\phi})$, where \mathcal{I} , \mathcal{W}_{ϕ} , and $A(\mathcal{I}, \mathcal{W}_{\phi})$ are as defined in Section 3. Then*

$$\text{cap}(S) \geq \frac{\log \lambda(\hat{A}_{p+2q+1,\phi}) - \log \lambda(\hat{A}_{2q+1,\phi})}{p\alpha}.$$

Proof. Let $S' = S^{[1 \times 2]}$. By Proposition 5, $S' = \mathcal{V}_1(S') \otimes \mathcal{H}_1(S')$, and S' has horizontal symmetric edge-constrained strips. Since $G_{\mathcal{E}}$ has no parallel edges, we may identify each edge $e \in E_{\mathcal{E}}$ with the pair $(\sigma(e), \tau(e))$; then, with this identification, $\mathcal{H}_1(S') = X(G_{\mathcal{E}})$. Also, note that \mathcal{G}_{2q+p+1} and \mathcal{G}_{2q+1} are deterministic presentations for $\mathcal{V}_{2q+p}(S')$ and $\mathcal{V}_{2q}(S')$, respectively. The result follows from Theorem 1 applied to S' . ■

5 Capacity bounds for axial products of constraints.

In this section we show how the method described in Section 3 can be applied to axial products of certain 1-dimensional constraints. Let S and T be two 1-dimensional constraints over an alphabet Σ . We wish to lower bound the capacity of the 2-dimensional constraint $T \otimes S$. To this end, we pick a lossless presentation $\mathcal{G}_S = (G_S, \mathcal{L}_S)$, with $G_S = (V_S, E_S)$, for S . We extend the function \mathcal{L}_S to multidimensional arrays over E_S in the manner described in Section 2, and for a set $A \subseteq \Sigma^*$, we denote by $\mathcal{L}_S^{-1}(A) \subseteq E_S^*$ the inverse image of A under this map, namely

$$\mathcal{L}_S^{-1}(A) = \{w \in E_S^* : \mathcal{L}_S(w) \in A\}.$$

The following proposition shows that we can reduce the problem of calculating the capacity of $T \otimes S$ to that of calculating the capacity of $\mathcal{L}_S^{-1}(T) \otimes X(G_S)$.

Proposition 6. *Let S, T be two 1-dimensional constraints and let $X(G_S)$ and $\mathcal{L}_S^{-1}(T)$ be as defined above. Then*

1. $\mathcal{L}_S^{-1}(T)$ is a 1-dimensional constraint.
2. $\text{cap}(T \otimes S) = \text{cap}(\mathcal{L}_S^{-1}(T) \otimes X(G_S))$.

Proof. 1. Let $\mathcal{G}_T = (V_T, E_T, \mathcal{L}_T)$ be a presentation of T . We shall construct a presentation $\mathcal{F} = (V_T, E_{\mathcal{F}}, \mathcal{L}_{\mathcal{F}})$ of $\mathcal{L}_S^{-1}(T)$. The set of edges is given by $E_{\mathcal{F}} = \{(e_T, e_S) \in E_T \times E_S : \mathcal{L}_T(e_T) = \mathcal{L}_S(e_S)\}$, and for an edge $(e_T, e_S) \in E_{\mathcal{F}}$, $\sigma_{\mathcal{F}}(e_T, e_S) = \sigma_{\mathcal{G}_T}(e_T)$, $\tau_{\mathcal{F}}(e_T, e_S) = \tau_{\mathcal{G}_T}(e_T)$ and $\mathcal{L}_{\mathcal{F}}(e_T, e_S) = e_S$. It is easily verified that $\mathcal{L}_S^{-1}(T)$ is presented by \mathcal{F} , and therefore it is a 1-dimensional constraint.

2. We set $R = T \otimes S$, $U = \mathcal{L}_S^{-1}(T) \otimes X(G_S)$. For an array $\Delta \in R_{m \times n}$, define $P_{\Delta} = \{\Gamma \in U_{m \times n} : \mathcal{L}_S(\Gamma) = \Delta\}$, we claim that

$$1 \leq |P_{\Delta}| \leq |V_S|^{2m}. \quad (11)$$

Indeed, it's easily verified that an array $\Gamma \in E_S^{m \times n}$ is in P_{Δ} iff for all $i \in [m]$ the row $(\Gamma_{i,j})_{j=0}^{n-1}$ is a path in \mathcal{G}_S that generates $(\Delta_{i,j})_{j=0}^{n-1}$. Since \mathcal{G}_S is a lossless presentation of S , for every $i \in [m]$, there is at least one path in \mathcal{G}_S generating $(\Delta_{i,j})_{j=0}^{n-1}$ and at most $|V_S|^2$ such paths; the claim follows. Now, clearly for any $\Gamma \in U_{m \times n}$ the array $\mathcal{L}_S(\Gamma)$ is in $R_{m \times n}$. It follows that the sets P_{Δ} , for $\Delta \in R_{m \times n}$ form a partition of $U_{m \times n}$, and we have

$$|U_{m \times n}| = \sum_{\Delta \in R_{m \times n}} |P_{\Delta}|.$$

Therefore, by (11), we get

$$|R_{m \times n}| \leq |U_{m \times n}| \leq |R_{m \times n}| |V_S|^{2m},$$

and it follows from (1) that $\text{cap}(R) = \text{cap}(U)$. ■

Therefore if $\mathcal{L}_S^{-1}(T) \otimes X(G_S)$ has symmetric horizontal edge-constrained strips, we can apply the method of Section 3 to obtain lower bounds on $\text{cap}(T \otimes S)$. In this case, it also follows from Remark 1 of Theorem 1, that the

method of [1] for obtaining upper bounds on the capacity of the hard-square constraint, can be used to obtain upper bounds on $\text{cap}(T \otimes S)$. Proposition 3 present a sufficient condition for $\mathcal{L}_S^{-1}(T) \otimes \mathsf{X}(G_S)$ to have symmetric horizontal edge-constrained strips. Here we give another stronger sufficient condition involving only the presentation \mathcal{G}_S . We say that a labeled graph (G, \mathcal{L}) , with $G = (V, E)$, is *symmetric as a labeled graph*, if there exists an edge-reversing matching $R \in \mathcal{R}(G)$ which preserves \mathcal{L} , that is $\mathcal{L}(R(e)) = \mathcal{L}(e)$ for all $e \in E$. We assume now that \mathcal{G}_S is symmetric as a labeled graph, and that $R \in \mathcal{R}(V_S, E_S)$ is an edge-reversing matching which preserves \mathcal{L}_S . Since for any positive integer m and $e_1 \dots e_m \in E_S^m$, the label $\mathcal{L}(e_1) \dots \mathcal{L}(e_m) = \mathcal{L}(R(e_1)) \dots \mathcal{L}(R(e_m))$, it follows that $e_1 \dots e_m \in \mathcal{L}_S^{-1}(T)$ iff $R(e_1) \dots R(e_m) \in \mathcal{L}_S^{-1}(T)$. Consequently, the hypothesis of Proposition 3 holds and we have the following corollary.

Corollary 1. *If \mathcal{G}_S is symmetric as a labeled graph then $\mathcal{L}_S^{-1}(T) \otimes \mathsf{X}(G_S)$ has symmetric horizontal edge-constrained strips.*

Since the presentation in Figure 1a is symmetric as a labeled graph, we can apply the method of Section 3 to get lower bounds on the capacity of all constraints $T \otimes \text{EVEN}$ for any 1-dimensional constraint T .

Let $S = \text{CHG}(b_1)$ and let $T = \text{CHG}(b_2)$ for integers $b_1, b_2 \geq 2$. Let $\mathcal{G}_S = (G_S, \mathcal{L}_S)$, with $G_S = (V_S, E_S)$ be the presentation given in Figure 1c for $b = b_1$. Evidently, \mathcal{G}_S is symmetric with exactly one edge-reversing matching, $R : E_S \rightarrow E_S$. Fix a positive integer m and let $\mathbf{e} = e_1 e_2 \dots e_m \in E_S^m$. Obviously, T is closed under negation of words (i.e., negating each symbol), and we have

$$\begin{aligned}
& e_1 e_2 \dots e_m \in \mathcal{L}_S^{-1}(T) \\
\iff & \mathcal{L}_S(e_1) \mathcal{L}_S(e_2) \dots \mathcal{L}_S(e_m) \in T \\
\iff & (-\mathcal{L}_S(e_1))(-\mathcal{L}_S(e_2)) \dots (-\mathcal{L}_S(e_m)) \in T \\
\iff & \mathcal{L}_S(R(e_1)) \mathcal{L}_S(R(e_2)) \dots \mathcal{L}_S(R(e_m)) \in T \\
\iff & R(e_1) R(e_2) \dots R(e_m) \in \mathcal{L}_S^{-1}(T).
\end{aligned}$$

Consequently, it follows by Proposition 3 that $\mathcal{L}_S^{-1}(T) \otimes \mathsf{X}(G_S)$ has symmetric horizontal edge-constrained strips and we can apply the method of Section 3 to obtain lower bounds on the capacity of $\text{CHG}(b_2) \otimes \text{CHG}(b_1)$.

The reader will note a similarity in the constructions in proofs of Propositions 1 and 6. Indeed, as an alternative approach, one may be able to use the construction in Proposition 1 to obtain bounds on $\text{cap}(S \otimes T)$: namely, if G_1 and G_2 are the underlying graphs of a capacity-preserving presentation $(\mathcal{G}_1, \mathcal{G}_2)$ of $S \otimes T$ and $\mathsf{X}(G_1) \otimes \mathsf{X}(G_2)$ has symmetric horizontal edge-constrained strips. However, the approach given by Proposition 6 seems to be more direct and simpler than the alternative approach.

6 Heuristics for choosing ϕ .

In this section, we use the notation defined in Section 3, and assume that $S = T \otimes \mathcal{E}$ is a 2-dimensional constraint with symmetric horizontal edge-constrained strips, where \mathcal{E} is an edge constraint. We describe heuristics for choosing the function ϕ to obtain “good” lower bounds on the capacity of S .

6.1 Using max-entropic probabilities.

Recall that a vertex of a directed graph is isolated if no edges in the graph are connected to it. Note, that since $G_m^{(\mathcal{H})}$ is symmetric, every vertex is either isolated or has both incoming and outgoing edges. We assume here that for every positive integer m , ignoring isolated vertices, $G_m^{(\mathcal{H})}$ is a primitive graph. In this case, the Perron eigenvector of H_m is unique up to multiplication by a scalar. Let \mathbf{r}_m be the right Perron eigenvector of H_m normalized to be a unit vector in the L_2 -norm. Observe, that substituting \mathbf{r}_m for \mathbf{x}_m satisfies (4) with an equality. This motivates us to choose ϕ so that the resulting vector \mathbf{z}_m^ϕ approximates \mathbf{r}_m . Since $G_m^{(\mathcal{H})}$ (without its isolated vertices) is irreducible, there is a unique stationary probability measure having maximum entropy on arrays of \mathcal{H}_m , namely the max-entropic probability measure on \mathcal{H}_m . We denote it here by $\text{Pr}^{*,m}$. It is given by

$$\text{Pr}^{*,m}(\Gamma) = \frac{(\mathbf{r}_m)_{\sigma(\Gamma)} (\mathbf{r}_m)_{\tau(\Gamma)}}{\lambda(H_m)^\ell}.$$

for $\Gamma \in S_{m \times \ell}$, for some positive integer ℓ , and where $\sigma(\Gamma), \tau(\Gamma) \in V_{\mathcal{E}}^m$ are given by

$$\begin{aligned} (\sigma(\Gamma))_i &= \sigma(\Gamma_{i,0}) \\ (\tau(\Gamma))_i &= \tau(\Gamma_{i,\ell-1}) \end{aligned} \quad ; \quad i = 0, 1, 2, \dots, m-1.$$

Let $\mathbf{V}^{(m)}$ be a random variable taking values in $V_{\mathcal{E}}^m$ with distribution given by

$$\text{Pr}(\mathbf{V}^{(m)} = \mathbf{v}) = \text{Pr}^{*,m} \{ \Gamma \in S_{m \times 1} : \sigma(\Gamma) = \mathbf{v} \} \quad ; \quad \mathbf{v} \in V_{\mathcal{E}}^m.$$

It's easily verified that

$$\text{Pr}(\mathbf{V}^{(m)} = \mathbf{v}) = ((\mathbf{r}_m)_{\mathbf{v}})^2. \tag{12}$$

Thus approximating $\text{Pr}(\mathbf{V}^{(m)} = \mathbf{v})$ and taking a square root will give us an approximation for $(\mathbf{r}_m)_{\mathbf{v}}$. Roughly speaking, $\text{Pr}(\mathbf{V}^{(m)} = \mathbf{v})$ is the probability of seeing the column of vertices \mathbf{v} in the “middle” of an $m \times \ell$ array chosen uniformly at random from $S_{m \times \ell}$, for large ℓ . Fix integers $\mu \geq 0, \alpha \geq 1$ as in Section 3, and assume now that $m = m_k = \mu + k\alpha$, for a positive integer k . For an integer $0 \leq s < m$ and vectors $\mathbf{u} \in V_{\mathcal{E}}^\ell, \mathbf{w} \in V_{\mathcal{E}}^r$, with lengths satisfying $\ell \leq m-s, r \leq s$, denote by $p_s^{(m)}(\mathbf{u})$ and $p_s^{(m)}(\mathbf{u}|\mathbf{w})$ the probabilities given by

$$\begin{aligned} p_s^{(m)}(\mathbf{u}) &= \text{Pr}(\mathbf{V}_{s:s+\ell-1}^{(m)} = \mathbf{u}) \\ p_s^{(m)}(\mathbf{u}|\mathbf{w}) &= \text{Pr}(\mathbf{V}_{s:s+\ell-1}^{(m)} = \mathbf{u} | \mathbf{V}_{s-r:s-1}^{(m)} = \mathbf{w}). \end{aligned}$$

Then by the chain rule for conditional probability we have, for any vector $\mathbf{v} \in V_{\mathcal{E}}^m$,

$$\text{Pr}(\mathbf{V}^{(m)} = \mathbf{v}) = p_0^{(m)}(\mathbf{v}_{0:\mu-1}) \prod_{i=0}^{k-1} p_{\mu+i\alpha}^{(m)}(\mathbf{v}_{\mu+i\alpha:\mu+(i+1)\alpha-1} | \mathbf{v}_{0:\mu+i\alpha-1}).$$

A plausible way to approximate $\text{Pr}(\mathbf{V} = \mathbf{v})$, is by treating \mathbf{V} as the outcome of a Markov process. Here we use a Markov process with memory μ , and

assume that $p_s(\mathbf{u}|\mathbf{w})$ can be “well” approximated by $p_s(\mathbf{u}|\mathbf{w}_{r-\mu:r-1})$, for vectors $\mathbf{u} \in V_{\mathcal{E}}^{\ell}$, $\mathbf{w} \in V_{\mathcal{E}}^r$, with r, ℓ as above, and $r \geq \mu$. Using this approximation we get

$$\Pr(\mathbf{V}^{(m)}=\mathbf{v}) \approx p_0^{(m)}(\mathbf{v}_{0:\mu-1}) \prod_{i=0}^{k-1} p_{\mu+i\alpha}^{(m)}(\mathbf{v}_{\mu+i\alpha:\mu+(i+1)\alpha-1} | \mathbf{v}_{i\alpha:\mu+i\alpha-1}).$$

We hypothesize that for fixed vectors $\mathbf{u} \in V_{\mathcal{E}}^{\alpha}$, $\mathbf{w} \in V_{\mathcal{E}}^{\mu}$, as m gets large, the conditional probabilities $p_s^{(m)}(\mathbf{u}|\mathbf{w})$, for $0 \ll s \ll m-1$, are “approximately equal” to the value when s is in the “middle” of the interval $[0, m-1]$. We hypothesize that this holds for “most” of the integers s in that interval and moreover that this middle value converges as m gets large. Accordingly, we try to approximate the conditional probability $p_s^{(m)}(\mathbf{u}|\mathbf{w})$ by the conditional probability found in the “middle” of a “tall” horizontal strip. More precisely, we fix an integer $\delta \geq 0$, set $\omega = 2\delta + \mu + \alpha$, and approximate $p_s^{(m)}(\mathbf{u}|\mathbf{w})$ by $p_{\delta+\mu}^{(\omega)}(\mathbf{u}|\mathbf{w})$. We also approximate $p_0^{(m)}(\mathbf{w})$ by $p_0^{(\omega)}(\mathbf{w})$. This gives us

$$\Pr(\mathbf{V}^{(m)}=\mathbf{v}) \approx p_0^{(\omega)}(\mathbf{v}_{0:\mu-1}) \prod_{i=0}^{k-1} p_{\delta+\mu}^{(\omega)}(\mathbf{v}_{\mu+i\alpha:\mu+(i+1)\alpha-1} | \mathbf{v}_{i\alpha:\mu+i\alpha-1}),$$

which, by (12), implies that

$$(\mathbf{r}_m)_{\mathbf{v}} \approx \sqrt{p_0^{(\omega)}(\mathbf{v}_{0:\mu-1})} \prod_{i=0}^{k-1} \sqrt{p_{\delta+\mu}^{(\omega)}(\mathbf{v}_{\mu+i\alpha:\mu+(i+1)\alpha-1} | \mathbf{v}_{i\alpha:\mu+i\alpha-1})} ; \quad \mathbf{v} \in V_{\mathcal{E}}^{m_k}. \quad (13)$$

Set $F_m = |V_{\mathcal{E}}|^{m_k}$, and denote by $\widetilde{\mathbf{r}}_{m_k} \in \mathbb{R}^{F_m}$ the nonnegative real vector with entries indexed by $V_{\mathcal{E}}^{m_k}$ and given by the RHS of equation (13). Let $\phi : (V_{\mathcal{E}})^{\mu+\alpha} \rightarrow [0, \infty)$ be given by

$$\phi(\mathbf{u}) = \sqrt{p_{\delta+\mu}^{(\omega)}(\mathbf{u}_{\mu:\mu+\alpha-1} | \mathbf{u}_{0:\mu-1})} ; \quad \mathbf{u} \in (V_{\mathcal{E}})^{\mu+\alpha}, \quad (14)$$

and let $\mathbf{z}_{m_k}^{\phi} \in \mathbb{R}^{F_m}$ be the vector with entries indexed by $V_{\mathcal{E}}^{m_k}$ and given by (6). Setting $\mathbf{x}_{m_k} = \mathbf{z}_{m_k}^{\phi}$, we obtain

$$(\mathbf{r}_{m_k})_{\mathbf{v}} \approx (\widetilde{\mathbf{r}}_{m_k})_{\mathbf{v}} = (\mathbf{x}_{m_k})_{\mathbf{v}} \sqrt{p_0^{(\omega)}(\mathbf{v}_{0:\mu-1})} ; \quad \mathbf{v} \in (V_{\mathcal{E}})^{m_k}.$$

Now for $m_k \geq \omega$, if $\mathbf{v} \in V_{\mathcal{E}}^{m_k}$ is not an isolated vertex in G_{m_k} , then clearly, $\mathbf{v}_{0:\omega-1}$ is not an isolated vertex in G_{ω} as well. Therefore $(\mathbf{r}_{\omega})_{\mathbf{v}_{0:\omega-1}} > 0$, which implies that $p_0^{(\omega)}(\mathbf{v}_{0:\omega-1}) > 0$ and thus $p_0^{(\omega)}(\mathbf{v}_{0:\mu-1}) > 0$. Let $p_{\min} = p_{\min}^{(\omega)} = \min\{p_0^{(\omega)}(\mathbf{w}) : \mathbf{w} \in V_{\mathcal{E}}^{\mu} \text{ and } p_0^{(\omega)}(\mathbf{w}) > 0\}$. It follows that for all vertices $\mathbf{v} \in (V_{\mathcal{E}})^{m_k}$ of G_{m_k} that are not isolated, we have

$$(\widetilde{\mathbf{r}}_{m_k})_{\mathbf{v}} \geq \sqrt{p_{\min}} (\mathbf{x}_{m_k})_{\mathbf{v}}.$$

Now, for any positive integer ℓ and $F_{m_k} \times 1$ -real vector, the product $\mathbf{y}^{\dagger} H_{m_k}^{\ell} \mathbf{y}$ depends only on the values of the entries of \mathbf{y} indexed by non-isolated vertices of G_{m_k} . Consequently, we may write

$$p_{\min} \mathbf{x}_{m_k}^{\dagger} H_{m_k}^{\ell} \mathbf{x}_{m_k} \leq \widetilde{\mathbf{r}}_{m_k}^{\dagger} H_{m_k}^{\ell} \widetilde{\mathbf{r}}_{m_k} \leq \mathbf{x}_{m_k}^{\dagger} H_{m_k}^{\ell} \mathbf{x}_{m_k},$$

for all positive integers ℓ . Taking the log, dividing by m_k , and taking the limit as k approaches infinity, we obtain

$$\lim_{k \rightarrow \infty} \frac{\log \widetilde{\mathbf{r}}_{m_k}^\top H_{m_k}^\ell \widetilde{\mathbf{r}}_{m_k}}{m_k} = \lim_{k \rightarrow \infty} \frac{\log \mathbf{x}_{m_k}^\top H_{m_k}^\ell \mathbf{x}_{m_k}}{m_k},$$

where by Lemma 1, the limit in the RHS exists. Thus, choosing ϕ as given by (14) and computing the lower bound by the method described in Section 3 is equivalent to computing the limit of the lower bound in (4), with $\widetilde{\mathbf{r}}_m$ substituted for \mathbf{x}_m , as m approaches infinity. If $\widetilde{\mathbf{r}}_m$ approximates \mathbf{r}_m well enough, we expect to get good bounds. Note that we may use the heuristic described here even for constraints for which the graphs $G_m^{(\mathcal{H})}$ are not always irreducible. In this case, the geometric multiplicity of the Perron eigenvalue may be larger than 1, and there may be more than one choice of the vector \mathbf{r}_ω in the computation of $p_{\delta+\mu}^{(\omega)}(\cdot|\cdot)$. Regardless of our choice, we will get a nonnegative function ϕ and a lower bound on the capacity. In Section 7 we show numerical results obtained using the heuristic described here for several constraints.

6.2 General optimization.

We may also use general optimization techniques to find functions ϕ which maximize the lower bound on the capacity. Fix integers $\mu \geq 0$ and $p, q, \alpha > 0$, and for a positive integer ℓ , set $\mathcal{D}_\ell = (V_\mathcal{E})^\ell$. In this subsection, we identify a function $\phi: \mathcal{D}_{\mu+\alpha} \rightarrow \mathbb{R}$ with a real vector $\phi \in \mathbb{R}^{|\mathcal{D}_{\mu+\alpha}|}$ indexed by $\mathcal{D}_{\mu+\alpha}$; for each $\mathbf{j} \in \mathcal{D}_{\mu+\alpha}$ we identify $\phi(\mathbf{j})$ with the entry $\phi_{\mathbf{j}}$. For a positive integer n , let \mathcal{G}_n be a deterministic presentation for $\mathcal{V}_n(S)$, $\mathcal{I}_n = \mathcal{I}(\mu, \alpha, n, \mathcal{G}_n, G_\mathcal{E})$, and for a function $\phi: \mathcal{D}_{\mu+\alpha} \rightarrow [0, \infty)$, set $A_{n,\phi} = A(\mathcal{I}_n, \mathcal{W}_\phi)$. Observe that for a scalar $c \in [0, \infty)$, $A_{n,c\phi} = c^2 A_{n,\phi}$. It follows that using $c\phi$ in place of ϕ in equation (9) of Theorem 1 does not change the lower bound. Consequently, (as ϕ cannot be the constant 0 function), it's enough to consider functions ϕ whose images (of all vectors in $(V_\mathcal{E})^{\mu+\alpha}$) sum to 1. We thus have the following optimization problem.

$$\begin{aligned} & \text{maximize} && (\log \lambda(A_{2q+p,\phi}) - \log \lambda(A_{2q,\phi})) / (p\alpha) \\ & \text{subject to} && \phi \geq \mathbf{0}, \\ & && \phi \cdot \mathbf{1} = 1, \end{aligned} \tag{15}$$

where $\mathbf{0}$ and $\mathbf{1}$ denote the real vectors of size $|\mathcal{D}_{\mu+\alpha}|$ with every entry equal to 0 and 1 respectively, and for two real vectors of the same size, \mathbf{t}, \mathbf{r} we write $\mathbf{t} \geq \mathbf{r}$ or $\mathbf{t} > \mathbf{r}$ if the corresponding inequality holds entrywise.

Finding a global solution for a general optimization problem can be hard. We proceed to show that if we replace the constraint $\phi \geq \mathbf{0}$ with $\phi > \mathbf{0}$ in (15), thereby changing the feasible set and possibly decreasing the optimal solution, it can be formulated as an instance of a particular class of optimization problems known as ‘‘DC optimization’’ which may be easier to solve. Let d be a positive integer. A real-valued function $f: \mathbb{R}^d \rightarrow \mathbb{R}$ is called a *DC* (difference of convex) function, if it can be written as the difference of two real-valued convex functions on \mathbb{R}^d . An optimization problem of the form

$$\begin{aligned} & \text{maximize} && f(x) \\ & \text{subject to} && x \in X, \\ & && h_i(x) \leq 0; \quad i = 0, 1, \dots, \ell, \end{aligned}$$

where $X \subseteq \mathbb{R}^d$ is a convex closed subset of \mathbb{R}^d and the functions f, h_0, \dots, h_ℓ are DC functions, is called a *DC optimization* or *DC programming* problem. See [9] and the references within for an overview of the theory of DC optimization.

A nonnegative function $f : \mathbb{R}^d \rightarrow [0, \infty)$ is called *log-convex* or *superconvex*, if either $f(\mathbf{t}) > 0$ for all $\mathbf{t} \in \mathbb{R}^d$ and $\log f$ is convex in \mathbb{R}^d , or $f \equiv 0$. A log-convex function is convex, and in [12], it is shown that the class of log-convex functions is closed under addition, multiplication, raising to positive real powers taking limits, and additionally that for a square matrix $A(\mathbf{t}) = (a_{i,j}(\mathbf{t}))$ whose entries are log-convex functions $a_{i,j} : \mathbb{R}^d \rightarrow [0, \infty)$, the function $\mathbf{t} \rightarrow \lambda(A(\mathbf{t}))$ is log-convex as well.

Now, observe, that for a positive integer n , every entry of $A_{n,\phi}$ is a quadratic form in the entries $\phi(\mathbf{j})$, $\mathbf{j} \in \mathcal{D}_{\mu+\alpha}$, with nonnegative integer coefficients. Such a function is generally not log-convex. To fix this, we perform the change of variables $\phi = e^\psi$, where ψ is a real-valued function $\psi : \mathcal{D}_{\mu+\alpha} \rightarrow \mathbb{R}$. Note that by doing so, we added the constraint $\phi > \mathbf{0}$. Since every entry of ϕ is now positive, we may replace the constraint $\phi \cdot \mathbf{1} = 1$ by the constraint $\phi(\mathbf{v}_0) = 1$ or equivalently $\psi(\mathbf{v}_0) = 0$, for some fixed $\mathbf{v}_0 \in \mathcal{D}_{\mu+\alpha}$. Problem (15) with the additional constraint $\phi > \mathbf{0}$, can now be rewritten as

$$\begin{aligned} & \text{maximize} && (\log \lambda(A_{2q+p,e^\psi}) - \log \lambda(A_{2q,e^\psi})) / (p\alpha) \\ & \text{subject to} && \psi(\mathbf{v}_0) = 0. \end{aligned} \tag{16}$$

Obviously, we may substitute the maximization problem constraint, $\psi(\mathbf{v}_0) = 0$, above into the objective function, thereby reducing the number of variables by 1; however, this is not relevant for the discussion, so, for simplicity, we do not do so here. Now, for a positive integer n , the entries of the matrix A_{n,e^ψ} are of the form

$$\sum_{k=1}^{q_{i,j}} e^{\psi(\mathbf{w}_{k,i,j}) + \psi(\mathbf{u}_{k,i,j})},$$

where $q_{i,j}$ are nonnegative integers, and $\mathbf{w}_{k,i,j}$ and $\mathbf{u}_{k,i,j}$ are vectors in $\mathcal{D}_{\mu+\alpha}$, for all $i, j \in V_{\mathcal{I}_n}$ and integers $1 \leq k \leq q_{i,j}$. It can be verified that a function of such a form is log-convex in ψ . It follows that the function $\psi \rightarrow \lambda(A_{n,e^\psi})$ for $\psi \in \mathbb{R}^{|\mathcal{D}_{\mu+\alpha}|}$ is log-convex as well. Therefore either $\lambda(A_{n,e^\psi}) \equiv 0$, or $\lambda(A_{n,e^\psi}) > 0$ for all $\psi \in \mathbb{R}^{|\mathcal{D}_{\mu+\alpha}|}$. In particular, for $\psi \equiv 0$ the matrix $A_{n,\mathbf{1}}$ is the adjacency matrix of the graph \mathcal{I}_n . Since \mathcal{I}_n is deterministic, and for every nonnegative integer ℓ , the set of labels of its paths of length ℓ is $S_{\ell\alpha \times n}$, it follows that $(1/\alpha) \log \lambda(A_{n,\mathbf{1}}) = \text{cap}(\mathcal{V}_n(S)) \geq \text{cap}(S)$ (the latter inequality follows from (2)). Hence, if $\text{cap}(S) > -\infty$ (or equivalently $\text{cap}(S) \geq 0$), then $\lambda(A_{n,e^\psi}) > 0$ for all $\psi \in \mathbb{R}^{|\mathcal{D}_{\mu+\alpha}|}$ and $\log \lambda(A_{n,e^\psi})$ is a convex function of ψ . Clearly $\text{cap}(S) > -\infty$ iff. $S_{m \times n} \neq \emptyset$, for all positive integers m, n . We thus obtain the following theorem.

Theorem 3. *Let S be a constraint such that for all positive integers m, n , $S_{m \times n} \neq \emptyset$ then Problem (16) is a DC optimization problem.*

7 Numerical results for selected constraints.

In this section we give numerical lower bounds on the capacity of some 2-dimensional constraints obtained using the method presented in the sections above. The constraints considered are NAK, RWIM, EVEN $^{\otimes 2}$, and CHG(3) $^{\otimes 2}$.

Table 1 summarizes the best lower bounds obtained using our method. For comparison, we provide the best lower bounds that we could obtain using other methods. We also give upper bounds on the capacity of these constraints obtained using the method of [1]. Table 2 shows the lower bounds obtained using our max-entropic probability heuristic for choosing ϕ , described in Section 6.1. Table 3 shows the lower bounds obtained with our method by trying to solve the optimization problem described in Section 6.2. In this, we did not make use of the DC property of the optimization problem; instead, we used a generic sub-optimal optimization algorithm whose results are not guaranteed to be global solutions. Utilizing special algorithms for solving DC optimization problems may give better lower bounds. The rightmost column of each of these tables shows the lower bound calculated for the same values of p and q using the method of [1]. The largest lower-bound obtained for each constraint is marked with a ‘*’. In the next subsections we give remarks specific to some of these constraints.

The numerical results were computed using the eigenvalue routines in Matlab and rounded (down for lower bounds and up for upper-bounds) to 10 decimal places. Given accuracy problems with possibly defective matrices, we verified the results using the technique described in [16, Section IV].

7.1 The constraint RWIM

Observe that this constraint has both symmetric horizontal and symmetric vertical vertex-constrained strips. Thus, we can apply our method in the vertical as well as the horizontal direction to get lower bounds. Clearly, $\text{cap}(\text{RWIM}^\dagger) = \text{cap}(\text{RWIM})$, so we can obtain additional lower bounds on $\text{cap}(\text{RWIM})$ by using our method to get lower bounds on $\text{cap}(\text{RWIM}^\dagger)$. Some of these bounds are given in Tables 2 and 3.

7.2 The constraint $\text{EVEN}^{\otimes 2}$

We used the reduction described in Section 5 with $\mathcal{G}_{\text{EVEN}}$ being the presentation of EVEN given in Figure 1a, to get lower bounds on the capacity of $\text{EVEN}^{\otimes 2}$. Table 3 gives the results obtained with our method using the optimization described in Section 6.2. We also used the method with the max-entropic probability heuristic of Section 6.1 and the results are given in Table 2.

7.3 The constraint $\text{CHG}(b)^{\otimes 2}$

For this constraint, the case $b=1$ is degenerate. Indeed, there are exactly two $m \times n$ arrays in $\text{CHG}(1)^{\otimes 2}$ for all positive integers m and n , and consequently, $\text{cap}(\text{CHG}(1)^{\otimes 2})=0$. For $b=2$, we show in Theorem 4 in Section 8 that the capacity is exactly $1/4$. For $b=3$, we used the reduction of Section 5 with $\mathcal{G}_{\text{CHG}(3)}$ being the presentation of $\text{CHG}(3)$ given in Figure 1c, to get lower bounds on the capacity of $\text{CHG}(3)^{\otimes 2}$. Table 3 gives the results obtained with our method using the optimization described in Section 6.2. We also used the method with the max-entropic probability heuristic of Section 6.1 and the results are given in Table 2.

Table 1: Best known lower bounds on capacities of certain constraints.

Constraint	Prev. best lower bound	New lower bound	Upper bound
NAK	0.4250636891*	0.4250767745	0.4250767997*
RWIM	0.5350150†	0.5350151497	0.5350428519*
EVEN ^{⊗2}	0.4385027973*	0.4402086447	0.4452873312*
CHG(3) ^{⊗2}	0.4210209862*	0.4222689819	0.5328488954*

*Calculated using the method of [1].

†Appears in [17].

8 Exact Computation

While it seems difficult to compute the capacity exactly for constraints such as EVEN^{⊗ \mathfrak{D}} and CHG(3)^{⊗ \mathfrak{D}} , we can compute the capacities of constraints in related families:

Theorem 4. For all positive integers \mathfrak{D} ,

$$\text{cap}(\text{CHG}(2)^{\otimes \mathfrak{D}}) = \frac{1}{2^{\mathfrak{D}}}.$$

Theorem 5. For all positive integers \mathfrak{D} ,

$$\text{cap}(\text{ODD}^{\otimes \mathfrak{D}}) = \frac{1}{2}.$$

Proof of Theorem 4. Let $S = \text{CHG}(2)^{\otimes \mathfrak{D}}$. We first show that $\text{cap}(S) \geq 1/2^{\mathfrak{D}}$. Let $\Gamma^{(0)}, \Gamma^{(1)}$ be the \mathfrak{D} -dimensional arrays of size $2 \times 2 \times \dots \times 2$ with entries indexed by $\{0, 1\}^{\mathfrak{D}}$ and given by

$$\left(\Gamma^{(i)}\right)_{\mathbf{j}} = (-1)^{i+\mathbf{j} \cdot \mathbf{1}} \quad ; \quad \mathbf{j} \in \{0, 1\}^{\mathfrak{D}},$$

where as usual $\mathbf{1}$ denotes the \mathfrak{D} -dimensional vector with every entry equal to 1. Observe that the sum of every row of both of these arrays is zero. Now, let n be a positive integer. For any \mathfrak{D} -dimensional array of size $n \times n \times \dots \times n$ with entries in $\{0, 1\}$, it can be easily verified that replacing every entry equal to 0 with $\Gamma^{(0)}$ and every entry equal to 1 with $\Gamma^{(1)}$ results in a \mathfrak{D} -dimensional array of size $2n \times 2n \times \dots \times 2n$ that satisfies S . It follows that $|S_{2n \times 2n \times \dots \times 2n}| \geq 2^{n^{\mathfrak{D}}}$ for all positive integers n , which implies $\text{cap}(S) \geq 1/2^{\mathfrak{D}}$.

We now show that $\text{cap}(S) \leq 1/2^{\mathfrak{D}}$. For a positive integer $n \geq 2$, denote by $\mathcal{N}_n^{(0)}$ the set of all even integers in $\{0, 1, \dots, n-2\}$ and by $\mathcal{N}_n^{(1)}$ the set of all odd integers in $\{0, 1, \dots, n-2\}$. We shall make use of the following lemma.

Lemma 2. Fix a positive integer $n \geq 2$, and let $(a_i)_{i=0}^{n-1} \subseteq \{+1, -1\}$ be a sequence of length n . Then $a_0 \dots a_{n-1} \in \text{CHG}(2)$ if and only if at least one of the following statements hold.

1. For all $i \in \mathcal{N}_n^{(0)}$, $a_i = -a_{i+1}$.
2. For all $i \in \mathcal{N}_n^{(1)}$, $a_i = -a_{i+1}$.

Proof. We first show the “if” direction. Let $(a_i)_{i=0}^{n-1} \subseteq \{+1, -1\}$ be a sequence for which at least one of statements 1,2 of the lemma holds. Then clearly, for any integers $0 \leq i < j < n$, all the terms in the sum $\sum_{k=i}^j a_k$, with the possible exception of the first and last terms, cancel. Therefore $|\sum_{k=i}^j a_k| \leq |a_i| + |a_j| = 2$ and $a_0 \dots a_{n-1} \in \text{CHG}(2)$.

As for the “only if” direction, let $(a_i)_{i=0}^{n-1} \subseteq \{+1, -1\}$ such that $a_0 \dots a_{n-1} \in \text{CHG}(2)$, and consider the presentation of the CHG constraint given in Figure 1c for $b = 2$ (and vertices $\{0, 1, 2\}$). Let $(e_i)_{i=0}^{n-1}$ be a path in this presentation generating $a_0 \dots a_{n-1}$. It’s easily verified that if $\sigma(e_i) = 1$, for some $i \in [n-1]$, then $a_j = -a_{j+1}$ for all integers $i \leq j \leq n-2$ such that $j \equiv i \pmod{2}$. Evidently, either $\sigma(e_0) = 1$ and so statement 1 holds, or $\sigma(e_1) = 1$ implying statement 2. \blacksquare

We now return to the claim that $\text{cap}(S) \leq 1/2^{\mathfrak{D}}$. Fix a positive integer $n \geq 2$. For an integer $1 \leq i \leq \mathfrak{D}$, let $\mathbf{e}^{(i)} \in \{0, 1\}^{\mathfrak{D}}$, be the vector, indexed by $\{1, 2, \dots, \mathfrak{D}\}$, containing 1 in its i th entry and 0 everywhere else and let $\mathcal{J}_i \subseteq [n]^{\mathfrak{D}}$ denote the subset of all the vectors indexed by $\{1, 2, \dots, \mathfrak{D}\}$ with a 0 in the i th entry. For a vector $\mathbf{j} \in \mathcal{J}_i$, the sequence $(\mathbf{j} + k\mathbf{e}^{(i)})_{k=0}^{n-1}$, is a sequence of indices of entries of a row in direction i of a \mathfrak{D} -dimensional $n \times n \times \dots \times n$ array, and we shall say that it is a sequence in direction i . Let $\mathbf{R}(n, \mathfrak{D})$ be the set of all such sequences for all integers $1 \leq i \leq \mathfrak{D}$ and vectors $\mathbf{j} \in \mathcal{J}_i$, and let $\mathbf{r} \in \{0, 1\}^{|\mathbf{R}(n, \mathfrak{D})|}$ be a binary vector indexed by $\mathbf{R}(n, \mathfrak{D})$. For the purpose of this proof, let us refer to a sequence $(a_i)_{i=0}^{n-1} \subseteq \{+1, -1\}$ as a phase-0 sequence if statement 1 of Lemma 2 holds, and as a phase-1 sequence if statement 2 holds (note that a sequence may be both a phase 0 and a phase 1 sequence). Also, we denote by $\mathcal{A}(\mathbf{r}) \subseteq \{+1, -1\}^{*\mathfrak{D}}$, the set of all \mathfrak{D} -dimensional arrays Γ of size $n \times n \times \dots \times n$ for which the row $\Gamma_{\bar{\varrho}}$ is a phase- $\mathbf{r}_{\bar{\varrho}}$ sequence, for all $\bar{\varrho} = (\mathbf{j} + k\mathbf{e}^{(i)})_{k=0}^{n-1} \in \mathbf{R}(n, \mathfrak{D})$. Then by Lemma 2, we have

$$S_{n \times \dots \times n} = \bigcup_{\mathbf{r}} \mathcal{A}(\mathbf{r}). \quad (17)$$

We shall give an upper bound on the size of $\mathcal{A}(\mathbf{r})$. For a vector $\mathbf{v} \in [n]^{\mathfrak{D}}$, denote by $\rho(i, \mathbf{v})$ the unique sequence in direction i in $\mathbf{R}(n, \mathfrak{D})$ that has \mathbf{v} as one of its elements. Let $T_{\mathbf{r}, i} : [n]^{\mathfrak{D}} \rightarrow \mathbb{Z}^{\mathfrak{D}}$ be given by:

$$T_{\mathbf{r}, i}(\mathbf{v}) = \begin{cases} \mathbf{v} + \mathbf{e}^{(i)} & \text{if } v_i \equiv \mathbf{r}_{\rho(i, \mathbf{v})} \pmod{2} \\ \mathbf{v} - \mathbf{e}^{(i)} & \text{otherwise} \end{cases}, \quad \mathbf{v} \in [n]^{\mathfrak{D}}, \mathbf{v} = (v_1, \dots, v_{\mathfrak{D}}).$$

Next, we define the undirected graph $G_{\mathbf{r}} = (V, E_{\mathbf{r}})$ (without parallel edges), with vertices given by

$$V = [n]^{\mathfrak{D}},$$

and edges given by

$$E_{\mathbf{r}} = \{\mathbf{u} - \mathbf{v} : \mathbf{u}, \mathbf{v} \in V \text{ and } \mathbf{v} = T_{\mathbf{r}, i}(\mathbf{u}) \text{ for some integer } 1 \leq i \leq \mathfrak{D}\},$$

where $\mathbf{u} - \mathbf{v}$ denotes the undirected edge connecting vertices \mathbf{u}, \mathbf{v} . It’s easy to verify that an array $\Gamma \in \{+1, -1\}^{*\mathfrak{D}}$ of size $n \times n \times \dots \times n$ is in $\mathcal{A}(\mathbf{r})$ iff for every edge $\mathbf{u} - \mathbf{v} \in E_{\mathbf{r}}$, it holds that $\Gamma_{\mathbf{u}} = -\Gamma_{\mathbf{v}}$. Figure 3 shows an example of $G_{\mathbf{r}}$ for $\mathfrak{D} = 2$.

Let C_1, \dots, C_{ℓ} be the connected components of $G_{\mathbf{r}}$, and let $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(\ell)}$ be arbitrary vertices such that $\mathbf{v}^{(i)} \in C_i$, for $i = 1, 2, \dots, \ell$. It follows that for

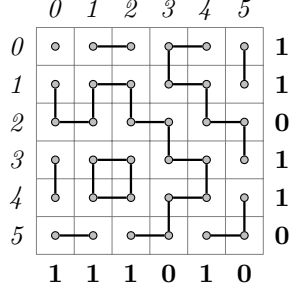


Figure 3: Example of the graph $G_{\mathbf{r}}$ for $\mathfrak{D} = 2$, $n = 6$. Each entry of \mathbf{r} corresponding to a row (column) is written to the right of it (below it). The index of each row (column) is written to its left (above it).

every vector $\mathbf{b} = (b_1, \dots, b_\ell) \in \{+1, -1\}^\ell$, there exists at most one array $\Gamma \in \mathcal{A}(\mathbf{r})$ satisfying $\Gamma_{\mathbf{v}^{(i)}} = b_i$ for all $i \in \{1, 2, \dots, \ell\}$, and consequently, $|\mathcal{A}(\mathbf{r})| \leq 2^\ell$ (in fact, while this is not needed for the proof, such an array $\Gamma \in \mathcal{A}(\mathbf{r})$ does exist, for any choice of \mathbf{b} , since each C_i is bipartite; thus $|\mathcal{A}(\mathbf{r})| = 2^\ell$).

Now, let $\mathbf{u} = (u_1, \dots, u_{\mathfrak{D}}) \in ([n] \setminus \{0, n-1\})^{\mathfrak{D}}$ be a vertex in the “interior” of $G_{\mathbf{r}}$. We show that the connected component of $G_{\mathbf{r}}$ containing \mathbf{u} has at least $2^{\mathfrak{D}}$ vertices. To this end, we match each word $\mathbf{w} = w_1 w_2 \dots w_{\mathfrak{D}} \in \{0, 1\}^{\mathfrak{D}}$, with a sequence of vertices $(\pi^{(\mathbf{w}, j)})_{j=0}^{\mathfrak{D}} \subseteq V$, defined recursively by

$$\pi^{(\mathbf{w}, j)} = \begin{cases} \mathbf{u} & \text{if } j=0 \\ \pi^{(\mathbf{w}, j-1)} & \text{if } j > 0 \text{ and } w_j = 0 \\ T_{\mathbf{r}, j}(\pi^{(\mathbf{w}, j-1)}) & \text{if } j > 0 \text{ and } w_j = 1 \end{cases} .$$

It’s easy to verify that since every $1 \leq u_i \leq n-2$, the sequence is well-defined and indeed $(\pi^{(\mathbf{w}, j)})_{j=0}^{\mathfrak{D}} \subseteq [n]^{\mathfrak{D}}$. Clearly, the sequence is contained entirely in the connected component containing \mathbf{u} , and so this component contains the vertex $\pi^{(\mathbf{w}, \mathfrak{D})}$. Write $\pi^{(\mathbf{w}, \mathfrak{D})} = (\pi_1^{(\mathbf{w}, \mathfrak{D})}, \dots, \pi_{\mathfrak{D}}^{(\mathbf{w}, \mathfrak{D})})$. Then for $i=1, 2, \dots, \mathfrak{D}$, it holds that $\pi_i^{(\mathbf{w}, \mathfrak{D})} = u_i$ if $w_i=0$, and $\pi_i^{(\mathbf{w}, \mathfrak{D})} = u_i \pm 1$ if $w_i=1$. Therefore, for two distinct words $\mathbf{w}, \mathbf{w}' \in \{0, 1\}^{\mathfrak{D}}$, the vertices $\pi^{(\mathbf{w}, \mathfrak{D})}$ and $\pi^{(\mathbf{w}', \mathfrak{D})}$ are distinct as well, and consequently there are $2^{\mathfrak{D}}$ such vertices. Thus, the connected component of $G_{\mathbf{r}}$ containing \mathbf{u} has at least $2^{\mathfrak{D}}$ vertices.

It follows that there are at most $n^{\mathfrak{D}}/2^{\mathfrak{D}}$ connected components of $G_{\mathbf{r}}$ containing a vertex in $\{1, 2, \dots, n-2\}^{\mathfrak{D}}$. There are at most $n^{\mathfrak{D}} - (n-2)^{\mathfrak{D}}$ connected components not containing a vertex in $\{1, 2, \dots, n-2\}^{\mathfrak{D}}$ and hence the total number of connected components, ℓ , in $G_{\mathbf{r}}$ satisfies $\ell \leq n^{\mathfrak{D}}/2^{\mathfrak{D}} + n^{\mathfrak{D}} - (n-2)^{\mathfrak{D}}$. Hence,

$$|\mathcal{A}(\mathbf{r})| \leq 2^{n^{\mathfrak{D}}/2^{\mathfrak{D}} + n^{\mathfrak{D}} - (n-2)^{\mathfrak{D}}} .$$

Since there are $2^{2n^{\mathfrak{D}}-1}$ binary vectors $\mathbf{r} \in \{0, 1\}^{|\mathbf{R}(n, \mathfrak{D})|}$, we obtain from (17)

$$\begin{aligned} |S_{n \times n \times \dots \times n}| &\leq \sum_{\mathbf{r}} |\mathcal{A}(\mathbf{r})| \\ &\leq 2^{n^{\mathfrak{D}}/2^{\mathfrak{D}} + n^{\mathfrak{D}} - (n-2)^{\mathfrak{D}} + \mathfrak{D}n^{\mathfrak{D}-1}} \\ &= 2^{n^{\mathfrak{D}}(1/2^{\mathfrak{D}} + 1 - (1-2/n)^{\mathfrak{D}}) + \mathfrak{D}n^{\mathfrak{D}-1}}, \end{aligned}$$

and the result follows from (1). ■

Proof of Theorem 5. Let S be the \mathfrak{D} -dimensional constraint $\text{ODD}^{\otimes \mathfrak{D}}$. We first show $\text{cap}(S) \geq 1/2$. For an integer n , let $\mathcal{B}_n \subseteq [2n]^{\mathfrak{D}}$ be the set of all vectors in $[2n]^{\mathfrak{D}}$ whose entries sum to an even number, and let \mathcal{A}_n be the set of all binary \mathfrak{D} -dimensional arrays Γ of size $2n \times 2n \times \dots \times 2n$, with entries satisfying $(\Gamma)_{\mathbf{j}} = 0$ for all $\mathbf{j} \in \mathcal{B}_n$. Then the number of zeros between consecutive ‘1’s, in any row of an array in \mathcal{A}_n is odd since it must be of the form $i-j-1$ for some integers i, j —either both odd, or both even. Thus, all such arrays satisfy the constraint S , and since $|\mathcal{A}_n| = 2^{(2n)^{\mathfrak{D}} - |\mathcal{B}_n|} = 2^{(2n)^{\mathfrak{D}}/2}$, we have $|S_{2n \times 2n \times \dots \times 2n}| \geq 2^{(2n)^{\mathfrak{D}}/2}$ for all positive integers n , which implies $\text{cap}(S) \geq 1/2$.

It remains to show that $\text{cap}(S) \leq 1/2$. For a positive integer d , let $T^{(d)} = \text{ODD}^{\otimes d}$. Since for any d positive integers m_1, \dots, m_d ,

$$\left| (T^{(d)})_{m_1 \times \dots \times m_d} \right| = \left| (T^{(d+1)})_{m_1 \times \dots \times m_d \times 1} \right|,$$

it follows from (2) that $\text{cap}(\text{ODD}^{\otimes d})$ is non-increasing in d . Thus, it’s enough to show $\text{cap}(S) \leq 1/2$ for $\mathfrak{D} = 1$. Let n be a positive integer. It can be easily verified that any 1-dimensional array $\Gamma \in \text{ODD}_n$ with entries indexed by $[n]$, satisfies either $\Gamma_j = 0$ for all even integers $j \in [n]$, or $\Gamma_j = 0$ for all odd integers $j \in [n]$. It follows that $|\text{ODD}_n| \leq 2^{\lceil n/2 \rceil} + 2^{\lfloor n/2 \rfloor}$ which implies the desired inequality ■

References

- [1] N. J. CALKIN AND H. S. WILF, “The number of independent sets in a grid graph,” *SIAM Journal of Discrete Math.*, Vol. 11, No. 1 (1998), 54–60.
- [2] M. COHN, “On the channel capacity of read/write isolated memory,” *Discrete Applied Mathematics*, Vol 56 (1995), 1–8
- [3] A. DESAI, “Subsystem entropy for \mathbb{Z}^d sofic shifts,” *Indagationes Mathematicae*, Vol 17 (2006), 353–359
- [4] M. J. GOLIN, X. YONG, Y. ZHANG AND L. SHENG, “New upper and lower bounds on the channel capacity of read/write isolated memory,” *Discrete Applied Mathematics*, Vol. 140 (2004), 35–48
- [5] S. FORSCHHAMMER AND J. JUSTESEN, “Bounds on the Capacity of Constrained Two-Dimensional Codes,” *IEEE Trans. Inform. Theory*, Vol. 46 (2000), 2659–2666
- [6] S. FRIEDLAND AND U. PELED “Theory of Computation of Multidimensional Entropy with an Application to the Monomer-Dimer Problem,” *Advances of Applied Math.*, Vol. 34 (2005), 486–522

- [7] S. HALEVI AND R. M. ROTH, “Parallel constrained coding with application to two-dimensional constraints,” *IEEE Trans. Inform. Theory*, Vol. 48 (2002), 1009–1020
- [8] A. HORN AND C. JOHNSON, “Matrix Analysis,” *Cambridge University Press*, 1985.
- [9] R. HORST AND N. V. THOAI, “DC programming: overview,” *Journal of Optimizaton Theory and Applications*, Vol. 103, No. 1 (1999), 1–43.
- [10] P.W. KASTELYN, The statistics of dimers on a lattice, *Physica A*, Vol. 27, (1961), 1209 – 1225.
- [11] A. KATO AND K. ZEGER, “On the capacity of two-dimensional run length constrained channels,” *IEEE Trans. Inform. Theory*, Vol. 45 (1999), 1527–1540.
- [12] J. F. C. KINGMAN, “A convexity property of positive matrices,” *The Quarterly Journal of Mathematics. Oxford. Second Series*, Vol. 12 (1963), 283–284.
- [13] D. LIND AND B. MARCUS, *An Introduction to Symbolic Dynamics and Coding*, *Cambridge University Press*, 1995 (reprinted 1999)
- [14] B. MARCUS, R. ROTH AND P. SIEGEL, *Constrained Systems and Coding for Recording Channels*, Chap. 20 in *Handbook of Coding Theory* (eds. V. S. Pless and W. C. Huffman), *Elsevier Science*, 1998.
- [15] N. MARKLEY AND M. PAUL, “Maximal measures and entropy of Z^{ν} subshifts of finite type,” *Classical mechanics and dynamical systems* (e.g., R. Devaney and Z. Nitecki), *Dekker Notes*, Vol. 70, 135 – 137.
- [16] Z. NAGY AND K. ZEGER, “Capacity Bounds for the Three-Dimensional (0,1) Run-Length Limited Channel,” *IEEE Trans. Inform. Theory*, Vol. 44 (2000), 1030–1033.
- [17] X. YONG AND M. J. GOLIN, “New techniques for bounding the channel capacity of read/write isolated memory,” *Proceedings of the Data Compression Conference* (2002).
- [18] W.WEEKS AND R. BLAHUT, “The capacity and coding gain of certain checkerboard codes,” *IEEE Trans. Inform. Theory*, Vol. 44 (1998), 1193–1203.

Table 2: Lower bounds using max-entropic probability heuristic (Section 6.1).

Constraint	δ	μ	α	p	q	Lower bound	Using [1]
NAK	3	1	1	1	5	0.4250766244	0.4248771038
	3	1	1	2	4	0.4250766446	0.4249055702
	6	1	1	1	5	0.4250767227	0.4248771038
	3	3	4	1	5	0.4250767590	0.4248771038
	7	1	1	1	5	0.4250767617	0.4248771038
	3	1	1	2	6	0.4250767647	0.4250636891
	3	1	4	1	5	0.4250767733	0.4248771038
	5	1	1	1	5	0.4250767744	0.4248771038
	3	1	4	2	6	0.4250767745	0.4250636891
	3	1	2	2	6	0.4250767745	0.4250636891
RWIM	3	1	3	1	6	0.5350147968	0.5235145644
	1	1	1	3	6	0.5350148753	0.5318753627
	3	2	2	1	5	0.5350148814	0.5160533001
	3	1	2	2	6	0.5350149069	0.5337927416
	2	1	2	2	6	0.5350149071	0.5337927416
	0	1	2	2	6	0.5350149136	0.5337927416
	2	2	2	1	5	0.5350149271	0.5160533001
	1	1	2	2	6	0.5350149462	0.5337927416
	1	1	3	1	6	0.5350149525	0.5235145644
	1	1	1	1	7	0.5350149707	0.5280406048
RWIM [†]	4	1	3	2	4	0.5350145937	0.5350144722
	1	1	1	1	5	0.5350146612	0.5350149478
	4	2	1	1	4	0.5350147212	0.5350142142
	3	1	1	1	5	0.5350147328	0.5350149478
	5	1	1	1	5	0.5350147619	0.5350149478
	2	2	1	1	4	0.5350147969	0.5350142142
	4	1	1	1	5	0.5350148255	0.5350149478
	2	1	1	1	5	0.5350148449	0.5350149478
	0	1	1	1	5	0.5350148814	0.5350149478
	0	2	1	1	4	0.5350148980	0.5350142142
EVEN ^{⊗2}	3	2	1	1	3	0.4383238232	0.4347423815
	3	1	1	1	4	0.4383243738	0.4367818624
	3	1	3	2	3	0.4383632350	0.4356897662
	3	1	2	4	3	0.4383838005	0.4364303826
	3	1	1	2	4	0.4384647082	0.4371709990
	3	1	3	3	3	0.4384906740	0.4360537982
	3	1	2	1	4	0.4385448358	0.4367818624
	3	1	2	2	4	0.4386655840	0.4371709990
	3	1	3	1	4	0.4387455520	0.4367818624
	CHG(3) ^{⊗2}	0	0	1	1	2	0.4188210386
0		0	1	1	4	0.4222689819*	0.4197053158

*Best lower bound.

Table 3: Lower bounds using optimization (Section 6.2).

Constraint	μ	α	p	q	Lower bound	Using [1]
NAK	2	1	2	4	0.4250767692	0.4249055702
	1	2	1	5	0.4250767736	0.4248771038
	1	1	3	4	0.4250767737	0.4248960814
	1	2	1	3	0.4250767739	0.4224650194
	1	1	4	4	0.4250767739	0.4249674993
	1	1	5	4	0.4250767740	0.4249783192
	1	1	6	4	0.4250767741	0.4249995626
	1	2	3	3	0.4250767743	0.4244240822
	1	2	6	3	0.4250767744	0.4247979797
	1	2	2	5	0.4250767745*	0.4250294285
RWIM	1	1	1	3	0.5350147515	0.4832292495
	1	1	2	3	0.5350148675	0.5300373650
	1	1	3	3	0.5350149371	0.5212673183
	1	1	1	4	0.5350150805	0.5037272248
	1	1	2	4	0.5350151001	0.5318663054
	1	1	3	4	0.5350151123	0.5265953036
	1	1	1	5	0.5350151372	0.5160533001
	1	1	2	5	0.5350151410	0.5330440001
	1	1	2	6	0.5350151491	0.5337927416
	1	2	1	5	0.5350151497*	0.5160533001
RWIM ^t	1	2	4	3	0.5350151364	0.5350130576
	1	2	3	4	0.5350151377	0.5350146307
	1	2	5	3	0.5350151392	0.5350134356
	2	1	1	4	0.5350151405	0.5350142142
	1	1	1	5	0.5350151442	0.5350149478
	1	2	1	4	0.5350151465	0.5350142142
	1	2	1	5	0.5350151481	0.5350149478
	1	2	2	4	0.5350151482	0.5350144722
	1	3	1	4	0.5350151483	0.5350142142
EVEN ^{⊗2}	1	1	1	3	0.4395381520	0.4347423815
	1	1	2	3	0.4397347451	0.4356897662
	1	1	1	4	0.4402086447*	0.4367818624
CHG(3) ^{⊗2}	0	1	1	2	0.4189237100	0.4101473707
	0	1	2	2	0.4197037681	0.4182017399
	0	1	3	2	0.4201450063	0.4176642274
	0	1	1	3	0.4210954837	0.4165892023
	0	1	2	3	0.4214748454	0.4210209862
				1	4	

*Best lower bound.