Capacity Lower Bounds and Approximate Enumerative Coding for 2-D Constraints

Erik Ordentlich
Hewlett–Packard Laboratories
Palo Alto, CA 94304
erik.ordentlich@hp.com

Ron M. Roth*
Computer Science Department
Technion, Haifa, Israel
ronny@cs.technion.ac.il

***Abstract*— We present a general method for obtaining lower bounds on the capacities of two-dimensional (2-D) constraints. We apply our method to the 2-D $(d{=}2, \infty)$ run-length limited (RLL) constraint and obtain the best known lower bound, .4423, on the capacity of this constraint. Our lower bounds are shown to be achievable by a fixed-rate, polynomial-complexity encoding–decoding algorithm based on enumerative coding with approximate counts.**

I. Introduction

Consider a finite alphabet Σ and two one-dimensional constraints with memory 1 on sequences over this alphabet determined by the $|\Sigma| \times |\Sigma|$ transition matrices H and V. The rows and columns of H and V are indexed by Σ and $H(x, y)$ (resp. $V(x, y)$) denotes the entry in H (resp. V) that is indexed by $x, y \in \Sigma$. In the sequel, the notation H and V is used to also mean the constraints themselves (not just the matrices). Let $\mathcal{H}_n$ and $\mathcal{V}_n$, respectively, denote the set of sequences of length n satisfying the constraints represented by H and V. Thus a sequence $x_1 x_2 \ldots x_n$ is in $\mathcal{H}_n$ if and only if $H(x_i, x_{i+1}) = 1$ for all $i = 1, 2, \ldots, n{-}1$. We are interested in the generic 2-D constraint on arrays over Σ in which rows and columns are constrained by H and V, respectively. Let $\mathcal{A}_{m,n}$ be the set of $m \times n$ arrays whose rows and columns, respectively, belong to $\mathcal{H}_n$ and $\mathcal{V}_m$. The capacity $C_{H,V}$ of this constraint is given by

$$C_{H,V} = \lim_{m,n\to\infty} \frac{1}{mn} \log_2 |\mathcal{A}_{m,n}|. \tag{1}$$

The limit in (1) is well known to exist and to be independent of how m and n are taken to infinity.

In Section II, we obtain a generic lower bound on $C_{H,V}$ in terms of $\nu_{H,V}$, an asymptotic lower bound on a sequence of conditional probabilities arising in a certain distribution on arrays over Σ. The approach is similar to that used in [8] to bound from above the redundancy of 2-D balanced binary arrays. While $\nu_{H,V}$ may be difficult to compute in general, we show in Section III how to compute it when H and V present the constraints on adjacent $r \times s$ sub-blocks of arrays that satisfy the 2-D (d, ∞)-RLL constraint, for certain values of d, r, and s. For $d = 2$, the resulting lower bound improves on the previously best published lower bounds [3], [4]. In Section IV, an efficient fixed-rate encoding scheme, based on approximate enumerative coding, is shown to achieve the rate bounds of the preceding sections. Enumerative coding is due to [2]. Approximate enumerative coding is relatively recent: an approximate enumerative encoding technique, very similar to that of Section IV, is applied in [8] to 2-D balanced arrays. A different approximate enumerative coding technique is used in [5] to reduce the complexity of exact enumerative coding of one-dimensional constraints. The technique of [5] serves as a building block in the 2-D constrained encoder of [10].

The following notation is used in the sequel. For any array x, let $x(i, j)$ denote the entry in the ith row and jth column and let $x(i_1{:}i_2, j_1{:}j_2)$ denote the sub-array $(x(i, j))_{i=i_1}^{i_2}{}_{j=j_1}^{j_2}$. An extension of this notation, $x(:, j_1{:}j_2)$ (resp. $x(i_1{:}i_2, :)$), refers to the sub-array encompassing all rows and columns j_1 through j_2 (resp. all columns and rows i_1 through i_2) of x. A sequence of ℓ symbols $x_1 x_2 \ldots x_\ell$ may also be denoted as x^ℓ. For a positive integer ℓ and a symbol $y \in \Sigma$, we define the subset $\mathcal{H}_\ell(y) \subseteq \mathcal{H}_\ell$ by

$$\mathcal{H}_\ell(y) = \{z^\ell \in \mathcal{H}_\ell : y z^\ell \in \mathcal{H}_{\ell+1}\}.$$

The subset $\mathcal{V}_\ell(y) \subseteq \mathcal{V}_\ell$ is defined in a similar manner. Vector transposition will be denoted by a superscript t. All logarithms are to the base 2.

II. Generic lower bound

For $x, y \in \Sigma$, let $p_{X|Y}(x|y)$ denote the limiting fraction of sequences in $\mathcal{H}_n(y)$ that start with the symbol x. Formally,

$$p_{X|Y}(x|y) = \lim_{n\to\infty} \frac{|\{z^n \in \mathcal{H}_n(y) : z_1 = x\}|}{|\mathcal{H}_n(y)|}. \tag{2}$$

It is well known from Perron-Frobenius theory that

$$p_{X|Y}(x|y) = \frac{H(y, x) v(x)}{\mu \cdot v(y)}, \tag{3}$$

where $v = (v(y))_{y\in\Sigma}$ is the right Perron eigenvector of H and μ is the corresponding Perron eigenvalue [6, Chapter 3].

Define $\nu_{H,V}$ as

$$\nu_{H,V} = \limsup_{m\to\infty} \frac{1}{m} \min_{y^m \in \mathcal{V}_m} \log \sum_{x^m \in \mathcal{V}_m} \prod_{j=1}^{m} p_{X|Y}(x_j|y_j). \tag{4}$$

Let C_H denote the capacity of the one-dimensional constraint defined by H. The following theorem gives a lower bound on $C_{H,V}$, in terms of C_H and $\nu_{H,V}$.

---
* Work done while visiting Hewlett–Packard Laboratories, Palo Alto, CA.

Theorem 2.1: The capacity $C_{H,V}$ of the 2-D constraint with rows and columns constrained, respectively, by H and V satisfies

$$C_{H,V} \geq C_H + \nu_{H,V}. \tag{5}$$

Proof: Let $X_{m,n}$ be uniformly distributed on the set of arrays whose rows belong to $\mathcal{H}_n$. This is equivalent to choosing the rows of $X_{m,n}$ independently and uniformly over $\mathcal{H}_n$. We then have

$$|\mathcal{A}_{m,n}| = |\mathcal{H}_n|^m P(X_{m,n} \in \mathcal{A}_{m,n})$$

and we proceed by bounding $P(X_{m,n} \in \mathcal{A}_{m,n})$. Clearly,

$$\begin{aligned} P(X_{m,n} \in \mathcal{A}_{m,n}) &= P(X_{m,n}(:,1) \in \mathcal{V}_m) \\ &\quad \cdot \prod_{i=2}^{n} P(X_{m,n}(:,i) \in \mathcal{V}_m | X_{m,n}(:,1{:}i{-}1) \in \mathcal{A}_{m,i-1}), \end{aligned}$$

where each conditional probability is given by

$$\begin{aligned} &P(X_{m,n}(:,i) \in \mathcal{V}_m | X_{m,n}(:,1{:}i{-}1) \in \mathcal{A}_{m,i-1}) \\ &\quad = \Big(\sum_{x_{m,i-1}} P(X_{m,n}(:,1{:}i{-}1) = x_{m,i-1}) \Big)^{-1} \\ &\qquad \cdot \Big(\sum_{x_{m,i-1}} P(X_{m,n}(:,i) \in \mathcal{V}_m | X_{m,n}(:,1{:}i{-}1) = x_{m,i-1}) \\ &\qquad\qquad \cdot P(X_{m,n}(:,1{:}i{-}1) = x_{m,i-1}) \Big), \end{aligned}$$

with $x_{m,i-1}$ ranging in each summation over all elements of $\mathcal{A}_{m,i-1}$. This implies

$$\begin{aligned} &P(X_{m,n}(:,i) \in \mathcal{V}_m | X_{m,n}(:,1{:}i{-}1) \in \mathcal{A}_{m,i-1}) \\ &\quad \geq \min_{x_{m,i-1}} P(X_{m,n}(:,i) \in \mathcal{V}_m | X_{m,n}(:,1{:}i{-}1) = x_{m,i-1}) \\ &\quad = \min_{y^m \in \mathcal{V}_m} \sum_{x^m \in \mathcal{V}_m} \prod_{j=1}^{m} \frac{|\{z^{n-i+1} \in \mathcal{H}_{n-i+1}(y_j) : z_1 = x_j\}|}{|\mathcal{H}_{n-i+1}(y_j)|}. \end{aligned}$$

Letting

$$p^{(\ell)}_{X|Y}(x|y) \triangleq \frac{|\{z^{\ell-1} \in \mathcal{H}_{\ell-1}(y) : z_1 = x\}|}{|\mathcal{H}_{\ell-1}(y)|} \tag{6}$$

and

$$p^{(\ell)}_X(x) \triangleq \frac{|\mathcal{H}_{\ell-1}(x)|}{|\mathcal{H}_\ell|},$$

we conclude that

$$\begin{aligned} &\frac{1}{n} \log P(X_{m,n} \in \mathcal{A}_{m,n}) \\ &\quad \geq \frac{1}{n} \log \sum_{x^m \in \mathcal{V}_m} \prod_{j=1}^{m} p^{(n)}_X(x_j) \\ &\qquad + \frac{1}{n} \sum_{\ell=2}^{n} \min_{y^m \in \mathcal{V}_m} \log \sum_{x^m \in \mathcal{V}_m} \prod_{j=1}^{m} p^{(\ell)}_{X|Y}(x_j|y_j) \end{aligned}$$

and, by continuity and (2), that

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{n} \log P(X_{m,n} \in \mathcal{A}_{m,n}) \\ &\quad \geq \min_{y^m \in \mathcal{V}_m} \log \sum_{x^m \in \mathcal{V}_m} \prod_{j=1}^{m} p_{X|Y}(x_j|y_j). \end{aligned}$$

Thus,

$$\begin{aligned} C_{H,V} &= \limsup_{m\to\infty} \frac{1}{m} \lim_{n\to\infty} \frac{1}{n} \log |\mathcal{A}_{m,n}| \\ &= \limsup_{m\to\infty} \frac{1}{m} \lim_{n\to\infty} \Big(\frac{m \log |\mathcal{H}_n|}{n} \\ &\qquad\qquad + \frac{1}{n} \log P(X_{m,n} \in \mathcal{A}_{m,n}) \Big) \\ &\geq C_H + \limsup_{m\to\infty} \frac{1}{m} \min_{y^m \in \mathcal{V}_m} \log \sum_{x^m \in \mathcal{V}_m} \prod_{j=1}^{m} p_{X|Y}(x_j|y_j) \\ &= C_H + \nu_{H,V}. \end{aligned}$$

□

III. 2-D (d,∞)-RLL CONSTRAINTS

The application of Theorem 2.1 to any given H and V requires the computation of $\nu_{H,V}$, which may be difficult in general. In this section, we consider 2-D (d,∞)-RLL constraints and apply Theorem 2.1 to the corresponding constraints induced on adjacent non-overlapping $r \times s$ sub-blocks of binary symbols. Though the method applies more generally, for concreteness, we focus on the case of $d \in \{1,2\}$ and show how to compute $\nu_{H,V}$ for moderate values of r and s. We remark that the capacity lower bounds reported at the end of this section are per binary symbol of the 2-D RLL constraint and are obtained by dividing the corresponding per sub-block lower bounds of the right-hand side of (5) (whose evaluation is detailed below) by $r \cdot s$.

Let $\mathcal{A}_{r,s}(d)$ denote the set of $r \times s$ binary arrays whose rows and columns all satisfy the one-dimensional (d,∞)-RLL constraint. We assume that $r, s \geq d$. The alphabet Σ of the previous section will be taken to be the set $\mathcal{A}_{r,s}(d)$. The (i,j)th entry of the matrix H (resp. V) indicates if the $r \times s$ arrays i and j can be adjacent horizontally (resp. vertically) according to the 2-D (d,∞)-RLL constraint. For each y in Σ, define also $D(y)$ to be the $|\Sigma| \times |\Sigma|$ diagonal matrix whose xth diagonal entry is $p_{X|Y}(x|y)$ as defined in (3). Also let $\mathbf{e}_1$ denote the column vector which is 1 in the component corresponding to the all-zero array element of Σ and zero in all other components. We can then express the sum-of-products in the definition of $\nu_{H,V}$ (4) as

$$\sum_{x^m \in \mathcal{V}^m} \prod_{j=1}^{m} p_{X|Y}(x_j|y_j) = \mathbf{e}_1^t \Big(\prod_{j=1}^{m} (VD(y_j)) \Big) \mathbf{1},$$

where $\mathbf{1}$ denotes the column vector of all 1's, so that

$$\nu_{H,V} = \limsup_{m\to\infty} \frac{1}{m} \min_{y^m \in \mathcal{V}_m} \log \mathbf{e}_1^t \Big(\prod_{j=1}^{m} (VD(y_j)) \Big) \mathbf{1}.$$

For the case of $(d{=}2,\infty)$ (as well as other values of d) and certain values of r and s, we are able to determine periodic sequences $(y_j^*)_{j=1}^{\infty}$ that provably achieve $\nu_{H,V}$, i.e., for which

$$\nu_{H,V} = \limsup_{m\to\infty} \frac{1}{m} \log \mathbf{e}_1^t \Big(\prod_{j=1}^{m} (VD(y_j^*)) \Big) \mathbf{1}.$$

Our approach is based on the following propositions.

Proposition 3.1: Let $\mathbf{q}$ be a probability (column) vector of dimension $|\Sigma|$. Then for any $r, s \geq d$ and $y^m \in \mathcal{V}_m$ the corresponding V and $D(y)$ satisfy

$$\mathbf{q}^t\Big(\prod_{j=1}^{m}(VD(y_j))\Big)\mathbf{1} \leq \mathbf{e}_1^t\Big(\prod_{j=1}^{m}(VD(y_j))\Big)\mathbf{1}$$

and

$$\begin{aligned}\nu_{H,V} &= \limsup_{m\to\infty}\frac{1}{m}\min_{y^m\in\mathcal{V}_m}\log \mathbf{e}_1^t\Big(\prod_{j=1}^{m}(VD(y_j))\Big)\mathbf{1}\\ &= \limsup_{m\to\infty}\frac{1}{m}\min_{y^m\in\mathcal{V}_m}\log \mathbf{q}^t\Big(\prod_{j=1}^{m}(VD(y_j))\Big)\mathbf{1}.\end{aligned}$$

Proposition 3.2: Given any $y' \in \Sigma$ and any positive integer N, let $\mathcal{W}_N = \mathcal{W}_N(y') \subseteq \mathcal{V}_N(y')$ be a set of sequences of length N such that for any $y^N \in \mathcal{V}_N(y')$ there exists $w^N \in \mathcal{W}_N$ satisfying

$$V(y_N,:) \leq V(w_N,:) \tag{7}$$

and

$$\mathbf{q}^t\Big(\prod_{j=1}^{N}(VD(w_j))\Big) \leq \mathbf{q}^t\Big(\prod_{j=1}^{N}(VD(y_j))\Big), \tag{8}$$

where the inequalities are componentwise. Then, for every $m > N$,

$$\begin{aligned}&\min_{y^m\in\mathcal{V}_m(y')}\mathbf{q}^t\Big(\prod_{j=1}^{m}(VD(y_j))\Big)\mathbf{1}\\ &\qquad = \min_{y^m\in\mathcal{V}_m : y^N\in\mathcal{W}_N}\mathbf{q}^t\Big(\prod_{j=1}^{m}(VD(y_j))\Big)\mathbf{1}.\end{aligned}$$

The condition (7), in words, requires that any sequence that can be preceded by y^N in the constraint V, can also be preceded by w^N. The proofs of these two propositions are straightforward and we omit them here.

Let $y_1^* y_2^* \ldots y_J^*$ be a sequence in $\mathcal{V}_J$ such that y_J^* can be followed by any symbol in Σ. Assume that $\prod_{j=1}^{J}(VD(y_j^*))$ is irreducible and aperiodic. This is readily seen to be the case for the (d,∞)-RLL constraints. Let $\mathbf{p}_1$ be the left Perron eigenvector of $\prod_{j=1}^{J}(VD(y_j^*))$ normalized to a probability vector, and for $i = 2, 3, \ldots, J$, let $\mathbf{p}_i^t = \mathcal{N}(\mathbf{p}_1^t \prod_{j=1}^{i-1}(VD(y_j^*)))$, where $\mathcal{N}(v) = v/(\mathbf{1}^t v)$ denotes normalization. Let λ be the Perron eigenvalue associated with $\mathbf{p}_1$. We then have the following lemma.

Lemma 3.3: Suppose for each $i = 1, 2, \ldots, J$ there exists a positive integer $N(i)$ and a subset $\mathcal{Y}(i) \subseteq \mathcal{V}_{N(i)}$ in which each sequence starts with y_i^*, such that the conditions of Proposition 3.2 are satisfied for $\mathbf{q} = \mathbf{p}_i$, $N = N(i)$, $y' = y_{i-1}^*$ (where we set $y_0^* = y_J^*$), and $\mathcal{W}_N = \mathcal{W}_N(y_{i-1}^*) = \mathcal{Y}(i)$. Then

$$\nu_{H,V} = \frac{1}{J}\log\lambda.$$

Proof: From the second part of Proposition 3.1 we know that

$$\nu_{H,V} = \limsup_{m\to\infty}\frac{1}{m}\min_{y^m\in\mathcal{V}_m}\log \mathbf{p}_1^t\Big(\prod_{j=1}^{m}(VD(y_j))\Big)\mathbf{1}. \tag{9}$$

Proposition 3.2 and the above assumption on y_J^* implies that for all sufficiently large m,

$$\begin{aligned}&\min_{y^m\in\mathcal{V}_m}\log \mathbf{p}_1^t\Big(\prod_{j=1}^{m}(VD(y_j))\Big)\mathbf{1}\\ &\quad= \min_{y^m\in\mathcal{V}_m(y_J^*)}\log \mathbf{p}_1^t\Big(\prod_{j=1}^{m}(VD(y_j))\Big)\mathbf{1}\\ &\quad= \min_{y^{m-1}\in\mathcal{V}_{m-1}(y_1^*)}\log \mathbf{p}_1^t VD(y_1^*)\Big(\prod_{j=1}^{m-1}(VD(y_j))\Big)\mathbf{1}.\end{aligned}$$

Since the last minimization is unaffected by normalizing $\mathbf{p}_1^t VD(y_1^*)$ and since $\mathcal{N}(\mathbf{p}_1^t VD(y_1^*)) = \mathbf{p}_2^t$, we can continue this for $i = 2$ and conclude that

$$\begin{aligned}&\min_{y^m\in\mathcal{V}_m}\log \mathbf{p}_1^t\Big(\prod_{j=1}^{m}(VD(y_j))\Big)\mathbf{1}\\ &\quad= \min_{y^{m-2}\in\mathcal{V}_{m-2}(y_2^*)}\log \mathbf{p}_1^t VD(y_1^*)VD(y_2^*)\Big(\prod_{j=1}^{m-2}(VD(y_j))\Big)\mathbf{1}.\end{aligned}$$

Indeed, this process can be repeated $m - \max_i N(i)$ times, cycling back to $i = 1$ after $i = \ell J$. This establishes that for all sufficiently large m, the minimum in (9) is achieved by repeating $y_1^* y_2^* \ldots y_J^*$, with the exception of a bounded suffix. The boundedness of this suffix and the aperiodicity and irreducibility of $\prod_{j=1}^{J}(VD(y_j^*))$ then implies the conclusion of the lemma. □

Thus, the basic method for determining $\nu_{H,V}$ for the 2-D (d,∞)-RLL constraint involves guessing a sequence $y_1^* y_2^* \ldots y_J^*$ satisfying the assumptions of Lemma 3.3, which then gives the value for $\nu_{H,V}$. The computations involved in establishing these assumptions, and then in evaluating $\nu_{H,V}$, for non-trivial block dimensions r and s, are greatly simplified through state merging. For example, when checking conditions (7)–(8), the sequence y^N (as well as the sequences in $\mathcal{W}_N(y')$) can be assumed to be over the alphabet $\mathcal{A}_{r,d}(d)$ (rather than $\mathcal{A}_{r,s}(d)$). Also, the matrix $VD(y_j)$ can be replaced by a smaller matrix whose rows and columns are indexed by the elements of $\mathcal{A}_{d,s}(d)$ (rather than $\mathcal{A}_{r,s}(d)$) and whose (x, x')th entry specifies the probability, under the conditional max-entropic distribution given y^N, of the set of $r \times s$ arrays that can occur below the $d \times s$ array x and whose bottom d rows coincide with the $d \times s$ array x'. Each such entry of this matrix can, in turn, be computed efficiently using a sequence of (smaller) matrix multiplications. Additional details of the use of state merging will be provided in the full paper.

Another key simplification in establishing the assumptions of Lemma 3.3 for a given guess $y_1^* y_2^* \ldots y_J^*$ involves a recursive procedure based on repeatedly applying Proposition 3.2. The idea is to find, for each i and N, the smallest dominating set $\mathcal{W}_N(y_{i-1}^*)$ satisfying the assumptions of Proposition 3.2 and noting that the prefixes of length N of the members of the smallest $\mathcal{W}_{N+1}(y_{i-1}^*)$ must belong to the smallest $\mathcal{W}_N(y_{i-1}^*)$, so that the search for the smallest $\mathcal{W}_{N+1}(y_{i-1}^*)$ can be restricted to subsets of the set obtained by appending Σ to each member of the smallest $\mathcal{W}_N(y_{i-1}^*)$. The process is stopped when all members of the smallest $\mathcal{W}_N(y_{i-1}^*)$ start with y_i^*.

We have successfully applied Lemma 3.3 to the 2-D (d, ∞)-RLL constraints with $d \in \{1, 2\}$ and various sub-block dimensions (r, s). Fortunately, $N(i)$ and $|\mathcal{Y}(i)|$ remain tractable for parameter values (r, s) that lead to reasonable bounds.

For $d = 1$, the best lower bound we obtain is for $(r, s) = (2, 15)$. In this case, $J = 1$ and $y_1^* = [00]^t$ (after state merging) and the lower bound on capacity is $.5865$. The best known lower and upper bounds on the capacity of this constraint are obtained using the method of Calkin and Wilf and are 0.5878911617 and 0.5878911619, respectively [1], [7], [11]. No efficient encoding is known, however, that achieves the Calkin–Wilf lower bound. In contrast, the present lower bound will be seen, in the next section, to be achievable by an efficient approximate enumerative coding scheme. For $r = 3$ and all values of s we could check we still have $J = 1$ but $y_1^* = [010]^t$ in this case. The resulting bounds are weaker, however, than for $r = 2$ and comparable s. As far as we can tell, no gains will be achieved for larger values of r as well.

For $d = 2$, the best lower bound we obtain is for $(r, s) = (4, 9)$. In this case, we find that $J = 3$ with

$$y_1^* = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad y_2^* = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{and} \quad y_3^* = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

and the lower bound on capacity is $.4423$. As in the case of $d = 1$, the bounds appear to deteriorate for larger values of r and comparable s. In the case of $d = 2$, our bound improves on the lower bound of $.4415$ obtained recently in [3] using a constraint-satisfying bit-stuffing process which, in turn, was an improvement over $.4267$ obtained as a lower bound on the entropy rate of a different constraint-satisfying bit-stuffing process in [4]. We note that simulations of the bit-stuffing process of [4] suggest that its entropy rate is considerably higher, on the order of $.4455$. Unfortunately, no rigorous proof of this is known.

We remark that, strictly speaking, the assumptions for Lemma 3.3 in general, and for the specific parameters mentioned above, involve irrational quantities, and hence cannot be verified directly with finite precision computations. Analytical verification seems hopelessly complex for the time being. Nevertheless, it is possible to justify the validity of the above lower bounds to the numerical precision given, even when based on finite precision verifications of the assumptions of Lemma 3.3. A detailed analysis of such numerical issues is deferred to the full paper.

IV. Fixed-rate enumerative encoder with approximate counts

In this section, we describe a fixed-rate encoding of messages into 2-D (d, ∞)-RLL constraints that asymptotically achieves the capacity lower bounds of the previous section. We begin with the description of an encoding framework for the general constraint of Section II and then specialize the framework to the 2-D RLL setting, leveraging the results of Section III.

For ease of exposition, we assume that there is at least one symbol in Σ which can be followed by any symbol, as constrained by H and V. For each ℓ, m, let $q_{\ell,m}$ satisfy

$$q_{\ell,m} \leq \min_{y^m \in \mathcal{V}_m} \sum_{x^m \in \mathcal{V}_m} \prod_{j=1}^{m} p_{X|Y}^{(\ell)}(x_j|y_j), \tag{10}$$

where $p_{X|Y}^{(\ell)}(x_j|y_j)$ is given by (6). For any $y^m \in \mathcal{V}_m$ define

$$\mathcal{A}_{m,\ell}(y^m) = \{x \in \mathcal{A}_{m,\ell} : y^m x \in \mathcal{A}_{m,\ell+1}\}.$$

An argument similar to that used in the proof of Theorem 2.1 shows that for any y^m,

$$|\mathcal{A}_{m,\ell}(y^m)| \geq \Big(\prod_{j=1}^{m} |\mathcal{H}_\ell(y_j)|\Big) \cdot \Big(\prod_{i=1}^{\ell} q_{i,m}\Big). \tag{11}$$

It follows from (10) that

$$q_{\ell,m} \leq \frac{\sum_{x^m \in \mathcal{A}_{m,1}(y^m)} \prod_{j=1}^{m} |\mathcal{H}_{\ell-1}(x_j)|}{\prod_{j=1}^{m} |\mathcal{H}_\ell(y_j)|}$$

which, in turn, implies, for any y^m,

$$\begin{aligned} &\Big(\prod_{j=1}^{m} |\mathcal{H}_\ell(y_j)|\Big) \cdot \Big(\prod_{i=1}^{\ell} q_{i,m}\Big) \\ &\quad \leq \sum_{x^m \in \mathcal{A}_{m,1}(y^m)} \Big(\prod_{j=1}^{m} |\mathcal{H}_{\ell-1}(x_j)|\Big) \cdot \Big(\prod_{i=1}^{\ell-1} q_{i,m}\Big). \end{aligned} \tag{12}$$

This consistency condition makes it possible to use the lower bound (11) on the number of constraint-satisfying extensions of any partial array to carry out approximate enumerative encoding.

The full encoding and decoding algorithms are detailed in Figures 1 and 2 (where it is assumed that $x(0,:)$ and $x(:,0)$ are set to the aforementioned symbol which can be adjacent anything in Σ). For a particular i and j, the function $L(w)$ in the decoder coincides with (13) in the encoder. The correctness of the encoding hinges on the consistency condition (12) and is established in a manner similar to that in [8].

The summation over an exponential number of products appearing in the expression (13) for $L(w)$ in the encoder can be computed efficiently using matrix multiplication. For each $y \in \Sigma$ define the $|\Sigma| \times |\Sigma|$ diagonal matrix $\Delta_\ell(y)$ whose xth

Input: Integer-valued message $M \in [0, |\mathcal{H}_n|^m \cdot \prod_{i=1}^{n} q_{i,m})$.
Output: Array $x \in \mathcal{A}_{m,n}$.

$M' \leftarrow M$
for $j \leftarrow 1$ to n
    for $i \leftarrow 1$ to m
        for each $w \in \Sigma$, let

$$L(w) \leftarrow \Bigg\lceil \Big(\prod_{\ell=1}^{n-j} q_{\ell,m}\Big)\Big(\prod_{v=1}^{i-1} |\mathcal{H}_{n-j}(x(v,j))|\Big) \cdot V(x(i-1,j),w)H(x(i,j-1),w) \cdot |\mathcal{H}_{n-j}(w)| \cdot \sum_{\substack{y^{m-i}\in \\ \mathcal{A}_{m,1}(x(i+1:m,j-1))}} V(w,y_1)\prod_{v=1}^{m-i}|\mathcal{H}_{n-j}(y_v)| \Bigg\rceil \quad (13)$$

        $x(i,j) \leftarrow \max\{y' : \sum_{w<y'} L(w) \le M' < \sum_{w\le y'} L(w)\}$
        $M' \leftarrow M' - \sum_{w<x(i,j)} L(w)$
        end
    end
end

Fig. 1. Encoder.

Input: Array $x \in \mathcal{A}_{m,n}$.
Output: Integer-valued message M.

$M \leftarrow 0$
for $j \leftarrow 1$ to n
    for $i \leftarrow 1$ to m
        $M \leftarrow M + \sum_{w<x_{i,j}} L(w)$
    end
end

Fig. 2. Decoder.

diagonal component is $H(y,x)\cdot|\mathcal{H}_\ell(x)|$. Then

$$\sum_{\substack{y^{m-i}\in \\ \mathcal{A}_{m,1}(x(i+1:m,j-1))}} V(w,y_1)\prod_{v=1}^{m-i}|\mathcal{H}_{n-j}(y_v)| = \mathbf{1}_w^t \prod_{v=i+1}^{m} (V\Delta_{n-j}(x(v,j-1)))\mathbf{1},$$

where $\mathbf{1}_w$ is the indicator column vector for w.

The encoding–decoding framework is specialized to the 2-D (d,∞)-RLL setting for $d \in \{1,2\}$ by again taking $\Sigma = \mathcal{A}_{r,s}(d)$ in the above generic description and leveraging the results of Section III. Let $y_1^* y_2^* \ldots y_J^*$ and $\mathbf{p}_1^t$ be as in Lemma 3.3, and let $D_\ell(y) = \alpha(\ell,y)D(y)$ where $\alpha(\ell,y)$ is set to $\min_x(p_{X|Y}^{(\ell)}(x|y)/p_{X|Y}(x|y))$, with $p_{X|Y}^{(\ell)}(x|y)$ defined in (6). We can then set $q_{\ell,m}$ to be

$$q_{\ell,m} = \min_{w^{N(\tilde\imath)}\in\mathcal{Y}(\tilde\imath)} \mathbf{p}_1^t\Big(\prod_{j=1}^{m-N(\tilde\imath)}(VD_\ell(y_j^*))\prod_{\ell=1}^{N(\tilde\imath)}(VD_\ell(w_\ell))\Big)\mathbf{1}, \quad (14)$$

with $y_j^* = y_{j-J}^*$ and $\tilde\imath \in \{1,2,\ldots,J\}$ uniquely satisfies $m = hJ+\tilde\imath-1+N(\tilde\imath)$ for some integer h (with $N(\cdot)$ and $\mathcal{Y}(\cdot)$ defined in Lemma 3.3). As noted above, $N(i)$ and $|\mathcal{Y}(i)|$ are not too large for the (d,r,s)-tuples leading to the bounds of Section III, implying that the minimization in (14) is computationally feasible. Rational parameters of appropriately growing precision to achieve polynomial complexity can be selected by "rounding down" in the above expressions so that the resulting $q_{\ell,m}$ continues to satisfy (10).

As long as the rounding precision grows suitably, Theorem 2.1 and Lemma 3.3 imply that the resulting asymptotic encoding rate for a given set of parameters d,r,s coincides with the corresponding capacity lower bound of Section III. In particular, for $(d,r,s)=(1,2,15)$, the asymptotic encoding rate of the above fixed-rate encoder is $.5865$, which improves on the asymptotic rate of $.581074$ of the efficient fixed-rate encoder of [9]. A variable-rate bit-stuffing procedure is also presented in [9] which achieves an asymptotic average rate of $.587277$. For $(d,r,s)=(2,4,9)$, the above encoding procedure achieves a rate of $.4423$, which is the best provable rate of any known encoding procedure, variable rate or fixed. The bit-stuffing processes of [3] and [4] can be converted to variable-rate bit-stuffing encoders achieving the corresponding process entropies. The only known fixed-rate encoders for this constraint are based on stripes separated by a sufficient amount of all-zero buffer space [11]. Such encoders, however, have significantly smaller rates for tractable stripe widths.

As in the determination of the minimizing periodic $(y_j^*)_j$, state merging can be used to gain additional computational savings in encoding and decoding. In particular, state merging can be used to determine the columns of each element $x(i,j)$ sequentially via an additional enumerative coding step. More details will be included in the full version of this work.

References

[1] N.J. Calkin, H.S. Wilf, *The number of independent sets in a grid graph, SIAM J. Discr. Math.,* 11 (1998), 54–60.
[2] T.M. Cover, *Enumerative source encoding, IEEE Trans. Inform. Theory,* 19 (1973), 73–77.
[3] S. Forchhammer, *Analysis of bit-stuffing codes and lower bounds on capacity for 2-D constrained arrays using quasistationary measures,* in *Proc. IEEE Int'l Symp. on Inform. Theory,* Chicago, IL (2004), p. 161.
[4] S. Halevy, J. Chen, R.M. Roth, P.H. Siegel, J.K. Wolf, *Improved bit-stuffing bounds on two-dimensional constraints, IEEE Trans. Inform. Theory,* 50 (2004), 824–838.
[5] K.A.S. Immink, *A practical method for approaching the channel capacity of constrained channels, IEEE Trans. Inform. Theory,* 43 (1997), 1389–1399.
[6] B.H. Marcus, R.M. Roth, P.H. Siegel, *Constrained systems and coding for recording channels,* in *Handbook of Coding Theory, Vol. II,* V.S. Pless, W.C. Huffman (Eds.), North-Holland, Amsterdam, 1998, pp. 1635–1764.
[7] Z. Nagy, K. Zeger, *Capacity bounds for the three-dimensional* $(0,1)$ *run length limited channel, IEEE Trans. Inform. Theory,* 46 (2000), 1030–1033.
[8] E. Ordentlich, R.M. Roth, *Two-dimensional weight-constrained codes through enumeration bounds, IEEE Trans. Inform. Theory,* 46 (2000), 1292–1301.
[9] R.M. Roth, P.H. Siegel, J.K. Wolf, *Efficient coding schemes for the hard-square model, IEEE Trans. Inform. Theory,* 47 (2001), 1166–1176.
[10] I. Tal, T. Etzion, R.M. Roth, *On row-by-row coding for 2-D constraints, Proc. IEEE Int'l Symp. on Inform. Theory,* Seattle, WA (2006), 1204–1208.
[11] W. Weeks IV, R.E. Blahut, *The capacity and coding gain of certain checkerboard codes, IEEE Trans. Inform. Theory,* 44 (1998), 1193–1203.