DTU Library

Bounds on the capacity of constrained two-dimensional codes

Forchhammer, Søren; Justesen, Jørn

Link back to DTU Orbit

[5] S. Kallel, "Analysis of a type II hybrid ARQ scheme with code combining," *IEEE Trans. Commun.*, vol. 38, pp. 1133–1137, Aug. 1990.

[6] Z. H. Lin, "Rate Compatible Convolutional Codes and Their Applications to Type II Hybrid ARQ Transmission," M.S. Thesis, Chalmers Univ. Technol., Göteborg, Sweden, 1998.

[7] P. Frenger, P. Orten, T. Ottosson, and A. Svensson, "Multi-Rate Convolutional Codes," Chalmers Univ. Technol., Göteborg, Sweden, Tech. Rep. 21, 1998.

[8] T. Ottosson, "Coding, Modulation and Multiuser Decoding for DS-CDMA Systems," Ph.D. dissertation, Chalmers Univ. Technol., Göteborg, Sweden, 1997.

[9] S. G. Wilson, *Digital Modulation and Coding*. Upper Saddle River, NJ: Prentice-Hall, 1996.

[10] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[11] P. Frenger, P. Orten, T. Ottosson, and A. Svensson, "Rate-compatible convolutional codes for multirate DS-CDMA systems," *IEEE Trans. Commun.*, vol. 47, pp. 828–836, June 1999.

[12] D. G. Daut, J. W. Modestino, and L. D. Wismer, "New short constraint length convolutional code constructions for selected rational rates," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 794–800, Sept. 1982.

[13] S. Falahati and A. Svensson, "Hybrid type-II ARQ schemes for Rayleigh fading channels," in *Proc. Int. Conf. Telecommunications*, vol. 1, 1998, pp. 39–44.

[14] S. Falahati, T. Ottosson, A. Svensson, and Z. H. Lin, "Hybrid type II ARQ schemes based on convolutional codes in wireless channels," in *Proc. IEEE Vehicular Technology Conf.*, 1999, pp. 2219–2224.

[15] W. C. Jakes, Ed., *Microwave Mobile Communications*, revised ed. New York: IEEE Press, 1993.

Bounds on the Capacity of Constrained Two-Dimensional Codes

Søren Forchhammer and Jørn Justesen, *Member, IEEE*

Abstract—Bounds on the capacity of constrained two-dimensional (2-D) codes are presented. The bounds of Calkin and Wilf apply to first-order symmetric constraints. The bounds are generalized in a weaker form to higher order and nonsymmetric constraints. Results are given for constraints specified by run-length limits or a minimum distance between pixels of a given value.

Index Terms—Capacity, constrained codes, entropy, two-dimensional (2-D) codes, 2-D fields.

I. INTRODUCTION

In this correspondence we consider the capacity of constrained two-dimensional (2-D) codes which is equivalent to the maximum entropy of 2-D fields defined by the given set of constraints. We shall derive bounds on the capacity.

We consider shift-invariant constraints of finite extent (N, M), in the sense that the constraints may be defined on an N row by M column rectangle. Each element is taken from an alphabet A of size $|A|$. The $|A|^{NM}$ possible configurations on the rectangle are divided into a set of admissible and a set of nonadmissible configurations.

Let $F(n, m)$ be the number of distinct configurations (or codewords) on an n row by m column rectangle not violating the constraints. We refer to these configurations of size n by m as admissible configurations. Let $E(n, m)$ be the set of admissible configurations of size n by m. The per-symbol capacity (or maximum entropy), H of the 2-D code defined by the constraints may be defined as follows.

Definition 1:

$$H = \lim_{n, m \to \infty} \frac{\log F(n, m)}{nm}. \tag{1}$$

A more formal treatment of the entropy definition and its existence is given in [1]–[3].

In one dimension (1-D), sequences satisfying a constraint on N consecutive symbols may be described by finite-state sources, where a state is characterized by $N - 1$ symbols, e.g., run-length limited sequences. The capacity in 1-D, defined as in Definition 1 but with $m = M = 1$ and $n \to \infty$, may be calculated following Shannon's approach [4]. The transfer (or adjacency) matrix $\boldsymbol{T}$ of the source indicates the possible transitions between two states. Each transition is derived from an admissible configuration of length N with the first $N - 1$ symbols as the starting state and the last $N - 1$ symbols as the ending state. Thus two states are combined with an overlap of $N - 2$ symbols in a transition. Each sequence of states formed by consecutive transitions corresponds to an admissible configuration. The largest eigenvalue Λ of $\boldsymbol{T}$ determines the growth rate of the number of configurations. Taking the logarithm gives the maximum entropy [4]:

$$H(1) = \log(\Lambda). \tag{2}$$

The one-dimensional approach is readily generalized to regions of finite width m. We refer to a rectangle of finite width m and arbitrary height n as a band. The admissible configurations of a band of width m may for all n be described by a finite-state source. For a constraint of extent (N, M), the states of the source are given by the symbols on the m by $N - 1$ segment which appear as the first or last $N - 1$ rows of an admissible configuration on an N by m rectangle, i.e., a configuration of $E(N, m)$. A transition from state i to state j is admissible if there is a configuration in $E(N, m)$, for which state i is identical to the top $N - 1$ rows and state j to the bottom $N - 1$ rows. States i and j have an overlap of $N - 2$ rows. The last row of j is generated by the transition from i to j and appended to the previous rows of the output. Any admissible configuration of $E(n, m)$ with fixed m and n ($> N - 1$) rows may be generated as an output by starting the source in the state specified by the first $N - 1$ rows and making $n - N + 1$ transitions appending one row to the output in each transition. The states and transitions above are derived from the set of configurations $S(T_m) = E(N, m)$ on the N by m segment of one transition of the finite-state source T_m. An output of the source is defined by the initial state and the output rows of a sequence of transitions. We are mainly interested in the set of possible outputs of the source. The transfer matrix $\boldsymbol{T}_m$ indicates transitions which satisfy the constraint by defining the elements $t_{ij} = 1$ if the transition from state i to j is admissible and $t_{ij} = 0$ if it is not admissible. The capacity of the source on a band of width m ($n \to \infty$) is given by

$$\frac{H(m)}{m} = \frac{\log(\Lambda_m)}{m} \tag{3}$$

where Λ_m is the largest eigenvalue of $\boldsymbol{T}_m$. Equation (3) is an upper bound on the capacity H defined by (1) [3]. For constraints where any two configurations on bands of width m may be admissibly concatenated by padding c columns between the two bands, the capacity is lower-bounded by $H(m)/(m + c)$.

Calkin and Wilf [5] presented a method giving tight bounds on capacity for a particular constraint, which we briefly describe in the following example.

Example 1: A well-known example is a binary field for which the maximum run-length of one of the symbols is one in both the n- and m-direction. This has been called, among other names, the hard square problem [6]. Formulated in terms of nonadmissible configurations, the constraint is that for any two 4-neighbors, i.e., horizontal and vertical neighbors, both of them cannot be "1." Therefore, by checking every $N = 2$ by $M = 2$ square we may check whether the constraint is satisfied.

The bounds for the constraint in Example 1 [5] are expressed by capacities which may be determined by the use of (3).

The method of [5] may be applied to other constraints, but it is restricted to constraints for which the transfer matrices are symmetric for all n or m. This is satisfied for first-order constraints which are symmetric in the n- or m-direction. We address the problem of bounding capacity for higher order and nonsymmetric constraints. We shall specify ways to describe supersets and subsets for which the transfer matrices are symmetric.

II. BOUNDS ON CAPACITY FOR FIRST-ORDER CONSTRAINTS

This section presents the upper and the lower bound of Calkin and Wilf [5], derived for the hard square constraint in Example 1. The approach will be initially to keep n fixed and find a bound on the number of configurations for $m \to \infty$. This bound is given by an expression for a finite m. For fixed m we can then define a finite-state source in the n-direction and let $n \to \infty$. For the upper bound we consider the transfer matrix $\boldsymbol{T}_n$ for a band extending in the horizontal direction, and apply the result that for any real symmetric matrix, the largest eigenvalue Λ satisfies [5]

$$\Lambda \le \mathrm{Trace}\left(\boldsymbol{T}_n^{2p}\right)^{1/2p} \tag{4}$$

for all positive integers p.

Thus the eigenvalue which determines the growth rate of the number of configurations for fixed n and large m is upper-bounded by the trace of $\boldsymbol{T}_n^{2p}$. Calkin and Wilf [5] made the observation that the trace counts the number of configurations starting and ending in the same state after $2p$ transitions. We may interpret these as solutions on a cylinder of finite circumference $2p$ and height n satisfying the constraint all the way around the cylinder. Keeping $2p$ fixed, we can now define a finite-state source with states of height $N - 1$ on the cylinder and use the one-dimensional result to derive a limit for $n \to \infty$

$$H \le \frac{H'(2p)}{2p} \tag{5}$$

where $H'(2p)$ is the capacity of the cylinder with circumference $2p$.

The prerequisite of (5) is that the transfer matrices $\boldsymbol{T}_n$ are symmetric. Therefore, the upper bound is valid for constraints which are symmetric and restricted to neighboring pixel interaction in the n- (or m-) direction. Constraints restricted to neighboring interaction vertically or horizontally are called first-order constraints, and they are characterized by $N = 2$ or $M = 2$.

The lower bound on capacity by Calkin and Wilf [5] for the hard square problem is derived by a similar idea using bands in the m-direction followed by a limit in the n-direction. Again we consider the transfer matrix for a band of height n and apply an inequality for the largest eigenvalue

$$\Lambda \ge \frac{(\boldsymbol{x}, \boldsymbol{T}_n \boldsymbol{x})}{(\boldsymbol{x}, \boldsymbol{x})} \tag{6}$$

where $(\boldsymbol{x}, \boldsymbol{y})$ denotes inner product.

Replacing $\boldsymbol{T}_n$ by $\boldsymbol{T}_n^p$ and letting $\boldsymbol{x} = \boldsymbol{T}_n^q \boldsymbol{1}$ we get

$$\Lambda^p \ge \frac{(\boldsymbol{T}_n^q \boldsymbol{1}, \boldsymbol{T}_n^p \boldsymbol{T}_n^q \boldsymbol{1})}{(\boldsymbol{T}_n^q \boldsymbol{1}, \boldsymbol{T}_n^q \boldsymbol{1})} = \frac{\left(\boldsymbol{1}, \boldsymbol{T}_n^{p+2q} \boldsymbol{1}\right)}{\left(\boldsymbol{1}, \boldsymbol{T}_n^{2q} \boldsymbol{1}\right)}. \tag{7}$$

The numerator and the denominator may be interpreted as the number of configurations after $p + 2q$ and $2q$ transitions. This gives $F(n, p + 2q + 1)$ and $F(n, 2q + 1)$. We can keep p and q fixed, take the logarithm, and let $n \to \infty$ to get

$$H > \frac{H(p + 2q + 1) - H(2q + 1)}{p} \tag{8}$$

where $H(m)$ is the capacity of a band source of width m, and $q \ge 0$ and $p \ge 1$.

This lower bound is valid for constraints which are of first order and symmetric in the n-direction, i.e., perpendicular to the direction of propagation of the bands of (8).

For higher order constraints, the states for a fixed n are composed of $M - 1 > 1$ columns. A single transition creates a single new column and two consecutive states overlap on $M - 2$ columns leading to a nonsymmetric transfer matrix. For $p \ge M - 1$ in (7) and (8) there is no overlap of columns between the initial and the final state after the p transitions. Still the direction of the transitions implies that, even for symmetric constraints the transfer matrix is usually nonsymmetric. If the constraint is symmetric, the reverse transition of every admissible transition is possible. In this case, the order of the columns in the states is also reversed, thereby changing the states.

III. UPPER BOUND FOR HIGHER ORDER CONSTRAINTS

We shall extend the upper bound presented in Section II to higher order and nonsymmetric constraints. Higher order constraints may be defined in terms of run-lengths or distances (>2) between symbols of a given value. For each symbol in the alphabet and each direction, we may specify that the number of consecutive elements (runs) is constrained to be within certain limits. The extent (N, M) must exceed the specified run-length. In order to achieve an upper bound we shall specify a source which has the required symmetric transfer matrices and can generate all configurations admissible by the original constraint as a subset.

Let $\boldsymbol{T}_n$ be the transfer matrix of a constraint with finite extent (N, M), having states characterized by the symbols of n rows and $(M - 1)$ columns. One new column is generated in each step. Let $\boldsymbol{A}_n = \boldsymbol{T}_n^p$ be the transfer matrix corresponding to p transitions. The matrices $\boldsymbol{A}_n$ and $\boldsymbol{T}_n$ have the same states. The element a_{ij} of $\boldsymbol{A}_n$ gives the number of admissible configurations on an n by m ($= p + M - 1$) rectangle starting in state i and ending in state j.

A new matrix $\boldsymbol{B}_n$ is introduced by adding $\boldsymbol{A}_n$ and its transpose $\boldsymbol{A}_n^*$. The matrix $\boldsymbol{B}_n = \boldsymbol{A}_n + \boldsymbol{A}_n^*$ is symmetric and it has elements $b_{ij} = a_{ij} + a_{ij}^*$, where $a_{ij}^* = a_{ji}$. These elements may be interpreted as the number of original configurations (a_{ij}) and the number of configurations (a_{ij}^*) having a reversed direction of transitions.

Using (4) we shall bound the growth rate of the number of configurations on a band of height n by the trace of $\boldsymbol{B}_n^2$. The element i of the main diagonal of $\boldsymbol{B}_n^2$ is found by summing

$$b_{ij} b_{ji} = (a_{ij} + a_{ij}^*)(a_{ji} + a_{ji}^*) = a_{ij} a_{ji} + a_{ij}^* a_{ji}^* + a_{ij} a_{ji}^* + a_{ij}^* a_{ji} \tag{9}$$

over all j. The trace of $\boldsymbol{B}_n^2$ is found by summing (9) over all i and j.

The sum $\sum_i \sum_j a_{ij} a_{ji}$ counts the configurations on the cylinder of circumference $2p$ as described by $\boldsymbol{A}_n$. The term $a_{ij} a_{ji}^*$ counts the number of configurations being output by a transition from i to j in the original direction concatenated with the output of a transition in the reverse direction from j to i. As $a_{ji}^* = a_{ij}$, the term $a_{ij} a_{ji}^* =$
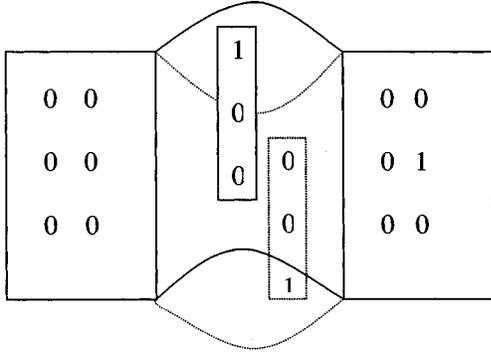
Fig. 1.   Two-seam cylinder ($n = 3$, $M = 3$).

$a_{ij}a_{ij}$ also counts the number of configurations being output by two transitions from i to j in the original direction. These configurations are referred to as two-seam cylinder configurations.

Definition 2: A configuration on a *two-seam cylinder* is given by a pair of configurations which are the output of two $\boldsymbol{A}_n$ transitions having identical starting states i and identical ending states j. The $\boldsymbol{A}_n$ configurations belong to the set of admissible configurations $E(n, p + M - 1)$.

The two $\boldsymbol{A}_n$ configurations each forms half of the two-seam cylinder having $p - (M - 1)$ columns between the i and the j state. Both transitions start in the same state and end in the same state. The starting state i and the ending state j may be seen as seams of the cylinder (Fig. 1).

As for the cylinders, we may introduce a finite-state source, in the n-direction, on the two-seam cylinder with states of height $N - 1$ and $2p$ distinct columns. For all n, the configurations comply with Definition 2. The two-seam cylinder specifies the neighborhood of the columns. The outer-most columns of the states only have one neighboring column as a band. The inner-most columns of the states, on the other hand, have three neighboring columns. The capacity of the two-seam cylinder ($n \to \infty$) of width $2p$ is denoted by $H''(2p)$.

Theorem 1: The capacity of a 2-D code specified by shift-invariant constraints of finite-extent (N, M) has the upper bound

$$H \leq \frac{H''(2p)}{2p} \qquad (10)$$

where H'' is the capacity of the two-seam cylinder of the constraint, p is a positive integer, and $2p \geq 2M - 2$. Further, $H''(2p) \geq H'(2p)$ and if $\boldsymbol{A}_n$ ($= \boldsymbol{T}_n^p$) is symmetric for all positive n, $H''(2p) = H'(2p)$.

Proof: The band with transfer matrix $\boldsymbol{B}_n$ generates as a subset all configurations of the band of $\boldsymbol{A}_n = \boldsymbol{T}_n^p$. The elements of $\boldsymbol{B}_n$

$$b_{ij} = a_{ij} + a_{ij}^* = a_{ij} + a_{ji} \qquad (11)$$

show that $b_{ij} = b_{ji}$ and, therefore, for all n $\boldsymbol{B}_n$ is symmetric and thus satisfies the prerequisite for (4). Therefore, the number of cylinder configurations, given by the trace of $\boldsymbol{B}_n^2$, gives an upper bound. Using (11) gives

$$\mathrm{Trace}\left(\boldsymbol{B}_n^2\right) = \sum_i \sum_j b_{ij} b_{ji}$$
$$= \sum_i \sum_j a_{ij} a_{ji} + a_{ij}^* a_{ji}^* + a_{ij} a_{ji}^* + a_{ij}^* a_{ji}. \qquad (12)$$

The first two terms of the right-hand side correspond to ordinary cylinders as described by $\boldsymbol{A}_n$ (or $\boldsymbol{A}_n^*$). The last two terms of (12) correspond to two-seam cylinders.

Let $F'(n, 2p)$ and $F''(n, 2p)$ denote the number of solutions on the ordinary and the two-seam cylinders of height n and circumference $2p$, respectively. Summing the right-hand terms of (12) gives

$$F'(n, 2p) = \sum_i \sum_j a_{ij} a_{ji} \qquad F''(n, 2p) = \sum_i \sum_j a_{ij}^2 \qquad (13)$$

as $a_{ji}^* = a_{ij}$.

Inserting (13) in (12) gives an expression of the number of $\boldsymbol{B}_n^2$ cylinder configurations in terms of cylinder and two-seam cylinder configurations for $\boldsymbol{A}_n^2$

$$\mathrm{Trace}\left(\boldsymbol{B}_n^2\right) = 2F'(n, 2p) + 2F''(n, 2p). \qquad (14)$$

The largest eigenvalue Λ_B of $\boldsymbol{B}_n$ is bounded by (4)

$$\Lambda_B^2 \leq \mathrm{Trace}\left(\boldsymbol{B}_n^2\right). \qquad (15)$$

For all pairs of states (i, j) in (13) we have

$$2a_{ij} a_{ji} \leq a_{ij}^2 + a_{ji}^2. \qquad (16)$$

Therefore, the number of two-seam cylinders $F''(n, 2p) \geq F'(n, 2p)$ will dominate when we go to the limit ($n \to \infty$).

Let Λ_{2p}'' denote the maximum eigenvalue of the two-seam cylinder. This eigenvalue determines the growth rate of $F''(n, 2p)$ for the two-seam cylinder for $n \to \infty$. By (15) this will bound the capacity of the band source given by $\boldsymbol{B}_n$. Combining (14) and (15), taking the nth root, the logarithm, and the limit ($n \to \infty$) gives

$$\lim_{n, m \to \infty} \frac{\log F_B(n, m)}{nm} = \lim_{n \to \infty} \frac{\log \Lambda_B^2}{2pn} \leq \frac{\log \Lambda_{2p}''}{2p} \qquad (17)$$

where $F_B(n, m)$ is the number of configurations of the band given by $\boldsymbol{B}_n$. As $F_B(n, m) \geq F(n, m)$ the bound in (17) also applies to the constraint in question.

Equation (16) leads to $F''(n, 2p) \geq F'(n, 2p)$. Taking the limit ($n \to \infty$) gives $H''(2p) \geq H'(2p)$. The equality in (16) applies if and only if $a_{ij} = a_{ji}$. Further, if and only if this equality applies for all (i, j) then $F''(n, 2p) = F'(n, 2p)$. Therefore, $\boldsymbol{A}_n$ symmetric for all n implies that $H''(2p) = H'(2p)$. $\qquad \square$

In Section VI, we consider constraints defined based on distances. Distances by r-norm between elements $f(n, m)$ and $f(n', m')$ may be expressed as

$$d = \left(|n - n'|^r + |m - m'|^r\right)^{1/r}.$$

We shall consider the 1-norm distance ($r = 1$) and the ∞-norm distance

$$d = \max(|n - n'|, |m - m'|).$$

Example 2A: Minimum Distance of 3 Between 1s (1-Norm): For a binary field (or 2-D code) $A = \{0, 1\}$, a minimum distance of 3 between 1s, measured by the 1-norm, is required. Therefore, for any two binary elements of the field $f(n, m)$ and $f(n', m')$ within a 1-norm distance of 2 of each other at most one of these can have a nonzero value

$$|n - n'| + |m - m'| \leq 2 \Rightarrow f(n, m)f(n', m') = 0. \qquad (18)$$

Around each "1," there is a (1-norm) "circle" of radius 2 of 0s (Fig. 2). This was referred to as the Diamond II constraint in [7].

$$
\begin{matrix}
& & 0 & & \\
& 0 & 0 & 0 & \\
0 & 0 & 1 & 0 & 0 \\
& 0 & 0 & 0 & \\
& & 0 & &
\end{matrix}
$$

Fig. 2.   Configuration ensuring a min. distance of 3 between 1s.

Consider a vertical band of width $m = 2$. There are five states of height $N - 1 = 2$

$$
s_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad s_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad s_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}
$$
$$
s_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad s_4 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.
$$

For transitions downwards from one state to the state below, the transition matrix (going from column to row) is

$$
\boldsymbol{T}_2 = \begin{bmatrix}
1 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0
\end{bmatrix}.
$$

The matrix $\boldsymbol{T}_2$ is nonsymmetric due to the overlap of one row for two consecutive states, e.g., the transition $s_3 \rightarrow s_2$ is admissible whereas $s_2 \rightarrow s_3$ is not.

For $p = 3$, $\boldsymbol{A}_2 = \boldsymbol{T}_2^3$ and $\boldsymbol{B}_2 = \boldsymbol{A}_2 + \boldsymbol{A}_2^*$ are given by

$$
\boldsymbol{A}_2 = \begin{bmatrix}
3 & 1 & 1 & 2 & 2 \\
2 & 1 & 1 & 2 & 1 \\
2 & 1 & 1 & 1 & 2 \\
1 & 0 & 1 & 1 & 1 \\
1 & 1 & 0 & 1 & 1
\end{bmatrix}
\qquad
\boldsymbol{B}_2 = \begin{bmatrix}
6 & 3 & 3 & 3 & 3 \\
3 & 2 & 2 & 2 & 2 \\
3 & 2 & 2 & 2 & 2 \\
3 & 2 & 2 & 2 & 2 \\
3 & 2 & 2 & 2 & 2
\end{bmatrix}.
$$

The number of cylinder configurations can be determined from these matrices. For the ordinary cylinder

$$
F'(6, 2) = \text{Trace}(\boldsymbol{A}_2^2) = \sum_i \sum_j a_{ij} a_{ji} = 37.
$$

The two-seam cylinder gives

$$
F''(6, 2) = \text{Trace}(\boldsymbol{A}_2 \boldsymbol{A}_2^*) = \sum_i \sum_j a_{ij}^2 = 49
$$
$$
F'_B(6, 2) = \text{Trace}(\boldsymbol{B}_2^2) = \sum_i \sum_j b_{ij} b_{ji}
$$
$$
= 2F'(6, 2) + 2F''(6, 2) = 172.
$$

The cylinder configurations for $m = 2$ are the states of the source on the horizontal cylinder $m \rightarrow \infty$. The two-seam cylinder thus has $F''(6, 2) = 49$ states for $p = 3$. Fig. 3 shows a configuration for one transition on the two-seam cylinder. This two-seam cylinder is identical to the one in Fig. 1 (rotated $90°$). In Fig. 3, the two-seam cylinder is unfolded into two bands, each having five rows, having identical top and bottom $N - 1 = 2$ rows. (Only $p = 3$ new rows are generated by each of the two configurations forming the two-seam cylinder so there are only six unique rows out of the ten rows.) This gives the upper bound (10), $H < H''(2p)/2p = H''(6)/6 = 0.3682$.

Fig. 3.   Two-seam cylinder configuration unfolded into two configurations on bands in the m-direction. The five top and bottom rows, each depict one configuration of the pair forming the two-seam cylinder. The two pairs of rows labeled i are identical for all m. This also applies to the rows labeled j.

IV. LOWER BOUND ON CAPACITY

The Calkin and Wilf lower bound requires that the transfer matrices of the bands in one direction are symmetric. For higher order constraints the matrices are not symmetric. We present a procedure for deriving a modified source which yields symmetric transfer matrices, generating a subset of the admissible configurations of the constraint. The procedure is not restricted to symmetric nor first-order constraints.

Given the transfer matrix $\boldsymbol{T}_m$, we introduce $\boldsymbol{A}_m = \boldsymbol{T}_m^p$ as the transfer matrix obtained after p transitions. The states are n by m elements, $n \geq N - 1$ and $m \geq M$. From $\boldsymbol{A}_m$ we derive a subset source on the band with transfer matrix $\boldsymbol{B}_m$. Let b_{ij} be the number of possible transitions from state i to j of $\boldsymbol{B}_m$. The symmetry required is given by $b_{ij} = b_{ji}$. A simple solution is only to maintain the elements of the main diagonal of $\boldsymbol{A}_m$, i.e., $b_{ii} = a_{ii}$ and $b_{ij} = 0$ if $i \neq j$. Obviously, $\boldsymbol{B}_m$ will be symmetric for all m and the configurations are all cylinders. Later we shall state the Calkin and Wilf lower bound (8) based on the cylinder capacity. We may hope to improve this lower bound by including elements outside the diagonal of $\boldsymbol{A}_m$ and enforce the symmetry of $\boldsymbol{B}_m$ by defining $b_{ij} = \min(a_{ij}, a_{ji})$. This may easily be done for any m, but it leads to a problem of determining the number of configurations on the $(p + n)$ by m rectangles as m increases.

Before proceeding, we generalize the way of defining a source on a band consistent with Section I. For brevity, we shall use the term band source. In the generalized version, each transition may output more than one row (or columns for horizontal bands). This also implies having multiple admissible outputs for a transition. As in Section I, a band source X is defined by a set of configurations $S(X)$ on a rectangle of given size along with the size and position of a starting and an ending state. The states are derived from the set of configurations $S(X)$ as in Section I. The elements x_{ij} of the transfer matrix $\boldsymbol{X}$ count the number of configurations of $S(X)$ starting in i and ending in j. Each of these configurations specifies a distinct transition from i to j, outputting one or more columns. Transitions admissible by the constraint are referred to as being constraint-admissible. The band configurations are constructed by sequentially combining the configurations of $S(X)$ such that the ending state of one configuration is identical to the starting state of the next configuration. Let $S(X^l)$ denote the set of configurations after l transitions on this band. For $l = 1$ we have $S(X^1) = S(X)$.

Using the generalized definition of a band source, we may specify a subset source. For a given m ($m \geq M - 1$), a reduced set of configurations $S(B_m)$ is defined by choosing b_{ij} out of a_{ij} configurations of $S(A_m) = E(n + p, m)$ for all i and j. This reduced set, $S(B_m) \subseteq E(n + p, m)$, is chosen to be the set of states of a new band source in the horizontal direction with transfer matrix $\boldsymbol{C}$. $\boldsymbol{C}$ specifes the transitions to a new state displaced one column in the horizontal direction. Thus a $\boldsymbol{C}$ transition covers a $(p + n)$ by $(m + 1)$ rectangle. For all positive l, the set of configurations $S(C^l)$ of size $(p + n)$ by $(m + l)$ is then used to define a vertical band source of width $m + l$ with transfer matrix $\boldsymbol{D}_l$ by defining $S(D_l) = S(C^l)$ for $l \geq 1$. The states of D_l are of size n by $(m + l)$. The problem is now to specify $S(C)$ such that the set of configurations $S(D_l)$ is a subset of the admissible configurations $E(p + n, m + l)$ and the matrices $\boldsymbol{D}_l$ are symmetric for $l \geq 0$. To ensure this $S(C)$ and thereby $\boldsymbol{C}$ is defined in two steps. First, admissibility of a transition is checked according to the constraint. Second, some transitions admissible by the constraint are set to 0 to ensure the symmetry of all $\boldsymbol{D}_l$. For each pair of starting (i) and ending (j) states of $S(B_m)$, we have b_{ij} configurations. Let the triple (i, x, j) denote such a configuration. A bijective mapping of the reduced set of b_{ij} configurations shall be introduced, *matching* each triple (i, x, j) with the unique triple (j, y, i) which is its image by the mapping. The starting and ending states of the configurations are reversed for each matched pair. All transitions admissible by the constraint from states s_1 to s_2 of $\boldsymbol{C}$ are thereafter evaluated as follows. If the horizontal transition between the states *matching* s_1 and s_2, respectively, is also constraint-admissible, the transition from s_1 to s_2 is set to 1 and otherwise to 0. The resulting matrix $\boldsymbol{C}$ is said to have matched states and transitions. Thus a transition from s_1 to s_2 is only set to 1 if the transition itself as well as the transition between the states matching s_1 and s_2 are constraint-admissible. This way the transitions as well as the states of the binary matrix $\boldsymbol{C}$ come in matched pairs.

To apply the approach of the lower bound of [5] in (7) and (8), we need to show that the matrices $\boldsymbol{D}_l$ are symmetric and that we can specify (horizontal) band sources consistent with $\boldsymbol{D}_l$. Lemmas 1 and 2, below, addresses the first and second issue, respectively.

Lemma 1: Given a matrix $\boldsymbol{C}$ having matched states and transitions, the transfer matrices $\boldsymbol{D}_l$ specified by $S(D_l) = S(C^l)$ are symmetric for all $l \geq 1$.

Proof: We may prove by induction that for any sequence of l transitions described by $\boldsymbol{C}^l$ there is a unique sequence of l matched transitions. Let s^l denote one configuration of $\boldsymbol{C}^l$ having $s_l = (i, x, j)$ as the last state, i.e., s^l is given by the sequence of states $s_1, s_2, \ldots, s_l$. We assume there is a matching sequence t^l with last state $t_l = (j, y, i)$. For any transition s_l to s_{l+1} there is by the construction of $\boldsymbol{C}$ a matching transition t_l to t_{l+1}, yielding t^{l+1} matched with s^{l+1}. For $l = 0$ we have the states of $\boldsymbol{C}$ which are matched by the definition of $\boldsymbol{C}$, thus completing the induction proof. The configurations $S(C^l)$ specify the configurations $S(D_l)$. Thus for each configuration in $S(D_l)$ there is a unique matching configuration in $S(C^l)$ with the $\boldsymbol{D}_l$ starting and ending states reversed. The source derived from reduced set $S(C)$, therefore, leads to the required symmetry properties for all matrices $\boldsymbol{D}_l$. $\square$

By Lemma 1, the matrices $\boldsymbol{D}_l$ are symmetric providing sufficient conditions for the lower bound. For any width l, the band source D_l, defining the sets of configurations $S(D_l^q)$ by q transitions, can generate a vertical band of infinite height ($q \rightarrow \infty$). To set up a consistent horizontal band source for $q > 1$, we return to the vertical band source D_1 of width $m + 1$ given by $S(D_1^1) = S(C_1^1)$. The configurations of this reduced set $S(D_1)$ may be combined sequentially, describing the set of configurations $S(D_1^q)$ of a vertical band source D_1 consistent with $\boldsymbol{D}_1^q$. Based on these sets, we may introduce band sources C_q in the

horizontal direction, defined by $S(C_q^1) = S(D_1^q)$. For any $q \geq 1$, the band source C_q can generate the configurations of a horizontal band of infinite length. Thus based on the reduced set of configurations $S(C)$, and the sources defined on this set C and D, a reduced source in the plane has been introduced. The configurations of $S(C_q^l)$ and $S(D_l^q)$ are of size $(n + pq)$ by $(m + l)$. The states of the horizontal band source C_q are of size $(n + pq)$ by m and the states of the vertical band source D_l are of size n by $(m + l)$. The consistency of combining C and D is shown in the following lemma.

Lemma 2: The set of configurations $S(D_l^q) = S(C_q^l)$ for $l, q \geq 1$.

Proof: A configuration in $S(D_l^2)$ is a combination of two configurations in $S(D_l^1)$ each being identical to a configuration in $S(C_1^l)$. These two configurations in $S(C^l)$ are uniquely combined to form a configuration in $S(C_2^l)$. Likewise the other way around. This generalizes readily to larger values of q yielding $S(D_l^q) = S(C_q^l)$. $\square$

The lower bound by Calkin and Wilf may be applied to the reduced source, in turn providing a lower bound for the specified constraint.

Theorem 2: The capacity of a 2-D code specified by shift-invariant constraints of finite-extent (N, M) is lower-bounded by

$$H \geq \max_{S(C)} \left\{ \frac{1}{rp} \log \frac{\Lambda(r + 2q)}{\Lambda(2q)} \right\} \qquad (19)$$

where $\Lambda(k)$ is the largest eigenvalue of $\boldsymbol{C}_k$ derived from $S(C), k \geq 0, p > 0, q \geq 0, r > 0$ and $\boldsymbol{C}_k$ has matched states and transitions. For the special case $k = 0$, $\Lambda(0)$ is the largest eigenvalue of $\boldsymbol{C}_0$ which is the transfer matrix of the band of height n describing the states of $\boldsymbol{D}_l$.

Proof: We may apply the lower bound of Calkin and Wilf. The real transfer matrix $\boldsymbol{D}_l^r$ is symmetric for $r > 0$ and $l \geq 1$ by Lemma 1. Therefore, the largest eigenvalue Λ_l of $\boldsymbol{D}_l$ is bounded by

$$\Lambda_l^r \geq \frac{(\boldsymbol{1}, \boldsymbol{D}_l^r \boldsymbol{1})}{(\boldsymbol{1}, \boldsymbol{1})} \qquad (20)$$

where $(\boldsymbol{x}, \boldsymbol{y})$ denotes inner product. Further following Calkin and Wilf

$$\Lambda_l^r \geq \frac{(\boldsymbol{D}_l^q \boldsymbol{1}, \boldsymbol{D}_l^r \boldsymbol{D}_l^q \boldsymbol{1})}{(\boldsymbol{D}_l^q \boldsymbol{1}, \boldsymbol{D}_l^q \boldsymbol{1})} = \frac{(\boldsymbol{1}, \boldsymbol{D}_l^q \boldsymbol{D}_l^r \boldsymbol{D}_l^q \boldsymbol{1})}{(\boldsymbol{1}, \boldsymbol{D}_l^q \boldsymbol{D}_l^q \boldsymbol{1})}$$
$$= \frac{(\boldsymbol{1}, \boldsymbol{D}_l^{r+2q} \boldsymbol{1})}{(\boldsymbol{1}, \boldsymbol{D}_l^{2q} \boldsymbol{1})} = \frac{(\boldsymbol{1}, \boldsymbol{C}_{r+2q}^l \boldsymbol{1})}{(\boldsymbol{1}, \boldsymbol{C}_{2q}^l \boldsymbol{1})} \qquad (21)$$

as the $\boldsymbol{D}$ transfer matrices are symmetric, $(\boldsymbol{1}, \boldsymbol{D}_l^q \boldsymbol{1}) = (\boldsymbol{1}, \boldsymbol{C}_q^l \boldsymbol{1})$ because the terms count the number of configurations on the same rectangle and the sets $S(D_l^q)$ and $S(C_q^l)$ on this rectangle are identical by Lemma 2 for all $l, q \geq 1$. Taking the logarithm and the limit

$$\lim_{l \rightarrow \infty} \frac{r \log(\Lambda_l)}{l} \geq \log \frac{\Lambda(r + 2q)}{\Lambda(2q)} \qquad (22)$$

where $\Lambda(k)$ is the largest eigenvalue of $\boldsymbol{C}_k$. This gives a lower bound on the subset of configurations in turn yielding the lower bound (19) for the specified constraint. The term $1/rp$ in (19) is because the band of $\boldsymbol{C}_{r+2q}$ has rp more rows than the band of $\boldsymbol{C}_{2q}$. For $q = 0$, the denominator $(\boldsymbol{1}, \boldsymbol{1})$ of (20) counts the number of states of $\boldsymbol{D}_l$, the growth rate ($l \rightarrow \infty$) of which may be determined by the largest eigenvalue $\Lambda(0)$ of $\boldsymbol{C}_0$ generating the states of $\boldsymbol{D}_l$.

For each i and j we delete $(\max(a_{ij}, a_{ji}) - \min(a_{ij}, a_{ji}))$ configurations and match the remaining configurations of $S(B_m)$ leading to $S(C)$. Under these restrictions we are free to choose the set $S(C)$ which yields the maximum lower bound. $\square$

Obviously, any choice of configurations $S(C)$ leading to the required symmetric matrices $\boldsymbol{D}_l$ in Lemma 1, gives a lower bound. It is an open question which set $S(C)$ leads to the best lower bound.
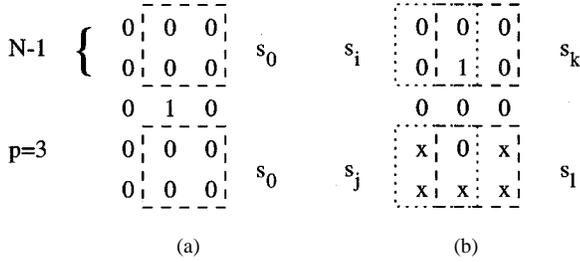
Fig. 4. Transitions horizontally. (a) $(s_0, x, s_0) \rightarrow (s_0, y, s_0)$. (b) $(s_i, 00, s_j) \rightarrow (s_k, 00, s_l)$. In (b) "x" marks symbols which will not conflict with the "1."

We may always include in $S(B_m)$ the configurations of the elements of the main diagonal of B_m. These cylinder configurations may be matched using the identity mapping. So Theorem 2 ($q = 0$) can always provide a lower bound at least as good as

$$H > \frac{(H'(p) - H(N - 1))}{p} \tag{23}$$

where the band entropy of the numerator in (19) is replaced by the cylinder entropy and the denominator by the band capacity of the states of D_l.

Example 2B: Minimum Distance of 3 Between $1s$ (1-Norm): For the constraint in Example 2A, B_2 is derived from A_2 ($= T_2^3$) by $b_{ij} = \min(a_{ij}, a_{ji})$

$$B_2 = \begin{bmatrix} 3 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

The states were defined in Example 2A. It may be noted that $b_{ij} \in \{0, 1\}$ except for $b_{00} = 3$. For the $s_0 \rightarrow s_0$ transitions we have the configurations (s_0, x, s_0) where $x \in \{00, 01, 10\}$. The configurations corresponding to the diagonal elements b_{ii} are matched with themselves. For the elements outside the diagonal, the subset of configurations has to be chosen and matched. For $b_{ij} = b_{ji} = 1$, the configurations (i, x, j) and (j, y, i) are matched choosing $x = y = 00$. This matching is always admissible as choosing x or $y = 00$ will never in itself introduce a violation of the constraint. y may be viewed as being equal to x or the mirrored configuration of x in this matching.

Based on the configurations of B_2^2 and the pairwise matching of these, the horizontal band with matrix C is defined. There are $\sum \sum b_{ij} = 23$ configurations and thereby 23 states of C. We can show that in this small example all constraint-admissible transitions between states of C also satisfy the requirement that the transition between their matching states is admissible.

The only nonzero values of x in (i, x, j) are for $i = j = s_0$, so $x = 01$ or 10 will only occur for $(s_0, x, s_0) \rightarrow (s_0, y, s_0)$ transitions (Fig. 4(a)). For all other transitions, x and y will be 00. This eliminates diagonal influence across the middle row. Therefore, for all constraint-admissible transitions, the matched transition does not violate the constraint when i and j are reversed to form the matching states (Fig. 4(b)). The resulting binary matrix C is of dimension 23 and all the 80 constraint-admissible transitions survive the check of the matching transition. Inserting $q = 0, r = 1, p = 3$ in (19) this gives a lower bound of $H \geq (H(5) - H(2))/3 = 0.297$.

V. UPPER BOUND ON CAPACITY—REVISITED

To supplement the upper bound of Section III, we may use the same approach as for the lower bound, but now specify a superset source. From A_m we derive a superset band source G_m with transfer matrix G_m given by the elements $g_{ij} = \max(a_{ij}, a_{ji})$. We shall set up the band sources J_q horizontally and K_l vertically corresponding to C_q and D_l, respectively, such that $S(J_q^l) = S(K_l^q)$. The rectangles of the configurations and states are defined as for C_q and D_l. The set of configurations $S(G_m)$ holds $S(A_m)$ as a subset and it defines the states of J. It is not necessary explicitly to define the additional states of $S(G_m)$ as we shall see. The states of J given by $S(G_m)$ and their transitions have to be matched to specify J as was the case for C. The set $S(J)$ shall form a superset of configurations on the $(p + n)$ by m rectangle. The matching is again specified by a bijective mapping. The transitions of J are specified by the the following matching of transitions: A transition between two states of J is set to 1 if the transition itself is constraint-admissible or the two matching state configurations and their transition are all constraint-admissible. Otherwise, the transition is set to 0. Based on J, the band sources J_q and K_l are defined. $\mathrm{Trace}(K_l^{2q})$ specifies the number of cylinder configurations which may be used to upper-bound the capacity of the superset source. Let J'_{2q} be the transfer matrix of the corresponding cylinder source, with circumference $2pq$ and states of size m by $2pq$, as $l \rightarrow \infty$. We may now apply the upper bound of Calkin and Wilf to the superset source which in turn gives an upper bound for the capacity of the code satisfying the constraints.

Theorem 3: The capacity of a 2-D code specified by shift-invariant constraints of finite extent (N, M) is upper-bounded by

$$H \leq \min_{S(J)} \frac{H'(2pq)}{2pq} \tag{24}$$

where $H'(2pq)$ is the capacity of the cylinder of circumference $2pq$ specified by J'_{2q}.

Proof: $S(J)$ is a superset containing all admissible configurations. The proof of Lemma 1 only requires the matching of states and transitions which J satisfies by construction. Therefore, the lemma applies and K_l is symmetric for all l. Given $S(J)$, the construction of J_q and K_l follows that of the lower bound, therefore, the consistency expressed by $S(J_q^l) = S(K_l^q)$ holds. Thereafter, the Calkin and Wilf upper bound may be applied. Any of the possible bijective mappings of the J states may be chosen, therefore, also the one leading to the $S(J)$ yielding the best upper bound. □

The two-seam cylinder approach of Section III may be described as a special case of the superset approach introduced above. The only slight modification is that G_m is defined by

$$g_{ij} = a_{ij} + a_{ij}^* = a_{ij} + a_{ji} \ (\geq \max(a_{ij}, a_{ji})).$$

Each of the a_{ij} configurations from i to j may be matched with itself in its reversed version as $a_{ij} = a_{ji}^*$. The configurations of $S(J^l)$ will be the ordinary and the reversed configurations. There is no crossover between the two sets for $l \rightarrow \infty$ due to the matching of transitions.

VI. NUMERICAL BOUNDS ON CAPACITIES

We may calculate the band and two-seam cylinder capacities H and H'' of the upper and lower bounds using (3).

A. Minimum-Distance Constraints

One class of constraints is given by generalizing the hard square problem by requiring a minimum distance between elements of the

TABLE I
CYLINDER CAPACITIES FOR MINIMUM DISTANCE 3 BETWEEN 1S (1-NORM)

$m = 2p$	$\frac{H''(2p)}{2p}$
4	0.3944
6	0.3682
8	0.3625
10	0.3595
12	0.3580
14	0.3569

TABLE II
LOWER BOUNDS ON CAPACITY FOR MINIMUM DISTANCE 3 BETWEEN 1S (1-NORM)

p	$H >$
8	0.3193
9	0.3228
10	0.3255
11	0.3277
12	0.3296
13	0.3312
14	0.3325

TABLE III
CYLINDER CAPACITIES FOR MINIMUM DISTANCE 3 BETWEEN 1S (∞-NORM)

$m = 2p$	$\frac{H''(2p)}{2p}$
4	0.3012
6	0.2554
8	0.2499
10	0.2494
12	0.2446
14	0.2432
16	0.2434

TABLE IV
CYLINDER CAPACITIES FOR MAXIMUM RUN-LENGTH 2

$m = 2p$	$\frac{H''(2p)}{2p}$
4	0.5700
6	0.4929
8	0.4790
10	0.4728

same value. Example 2 is completed below and another example belonging to this class is also presented.

Example 2C: Minimum Distance of 3 *Between* 1s *(1-Norm):* Table I gives the two-seam cylinder capacity results for the constraint of Example 2A. (The values given are rounded up as we seek an upper bound.) The best upper bound on capacity by (10) is $H < 0.3569 = H''(14)/14$.

This is an improvement of the previous best upper bound reported in [3] $H < H(14)/14 = 0.3597$. (The latter was achieved based on the entropy of a band with an additional constraint derived from requiring equal probabilities of a "1" in the border and inner columns of the band.)

Table II lists new lower bounds for $m = M = 3, q = 0$, and $r = 1$. The best of these is 0.3325, for $p = 14$. $\mathcal{C}$ was specified by choosing the subset of configurations forming the states and matching these as follows. 1) The configurations of the diagonal elements b_{ii} were matched by the identity mapping. 2) The configurations (i, x, j) and (j, x, i) were matched if both were constraint-admissible. 3) The remaining configurations with reversed states i and j were matched (somewhat arbitrarily) in an order given by increasing binary value. The new lower bound is an improvement of the lower bound by bands $H(15)/17 = 0.319$, but not close to the best lower bound reported in [3]: $H > 0.350306$. (This lower bound was obtained by bounding the entropy of a process using hidden states.)

It may be noted that the cylinder capacity $H'(6) = 0.3498$ is below this lower bound and thus cannot be an upper bound. The capacity may be estimated to be [3] $H \approx H(16) - H(15) = 0.35030719$. (For first-order processes the difference $H(16) - H(15)$ is a lower bound (8), but this does not readily generalize to higher orders.)

Example 3: Minimum Distance of 3 *Between* 1s *(∞-Norm):* For a binary field (or 2-D code), a minimum distance of 3 between 1s, measured by the ∞-norm, is required. Therefore, for any two elements $f(n, m)$ and $f(n', m')$ within ∞-norm distance 2 at most one of these can have a nonzero value

$$\max(|n - n'|, |m - m'|) \leq 2 \Rightarrow f(n, m)f(n', m') = 0. \quad (25)$$

This was called the Square II constraint in [7].

Table III gives two-seam cylinder capacity results for the constraint. The best upper bound on the capacity by (10) is $H < 0.2432 = H''(14)/14$.

In [7], upper and lower bounds are given, $0.25681 > H > 0.22257$. The new upper bound is a fair improvement. The capacity was estimated to be $H \approx 0.236$ (with an estimated precision of 10^{-3}) [7].

B. Run-Length Constraint

The hard square constraint may also be generalized by imposing run-length constraints in both the horizontal and vertical directions.

Example 4: Maximum Run-Length of 2*:* For a binary field (or 2-D code), a maximum run-length constraint of 2 is imposed for both values and in both directions.

Table IV gives two-seam cylinder capacity results for the constraint. The best upper bound by (10) is $H < 0.4728 = H''(10)/10$. This is a fair improvement of the upper bound achieved using bands, $H < H(10)/10 = 0.5079$. A lower bound of $H > 0.4650$ may be obtained

using the process with hidden states in [3]. The capacity may be estimated to be $H \approx H(10) - H(9) = 0.4682$.

VII. CONCLUSION

New bounds on the capacity of constrained 2-D codes were derived. The bounds are based on the transfer matrix of superset and subset sources, respectively. The bounds are expressed in terms of capacities of bands and cylinders, which may be determined using well-known one-dimensional results. Two upper and a lower bound applicable to any finite-order constraint were presented. One of the upper bounds was applied to three second-order constraints, improving previous upper bounds. The lower bound was applied to one of these constraints, improving previous results based on transfer matrices.

REFERENCES

[1] J. Justesen and Y. Shtarkov, "The combinatorial entropy of images," *Probl. Pered. Inform.*, vol. 33, no. 1, pp. 3–11, 1997. (In Russian).
[2] S. Friedland, "On the entropy of Z^d subshifts of finite type," *Linear Alg. Appl.*, vol. 252, pp. 199–220, 1997.
[3] S. Forchhammer and J. Justesen, "Entropy bounds for constrained two-dimensional random fields," *IEEE Trans. Inform. Theory*, vol. 45, pp. 118–127, Jan. 1999.
[4] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, July 1948.
[5] N. J. Calkin and H. S. Wilf, "The number of independent sets in a grid graph," *SIAM J. Discr. Math.*, vol. 11, no. 1, pp. 54–60, Feb. 1998.
[6] R. Burton and J. E. Steif, "Non-uniqueness of measures of maximal entropy for subshifts of finite type," *Ergod. Theory Dynam. Syst.*, vol. 14, pp. 213–235, 1994.
[7] W. Weeks IV and R. E. Blahut, "The capacity and coding gain of certain checkerboard codes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1193–1203, May 1998.

Partial Characterization of the Positive Capacity Region of Two-Dimensional Asymmetric Run Length Constrained Channels

Akiko Kato, *Member, IEEE,* and Kenneth Zeger, *Fellow, IEEE*

Abstract—A binary sequence satisfies a one-dimensional (d, k) run length constraint if every run of zeros has length at least d and at most k. A two-dimensional binary pattern is (d_1, k_1, d_2, k_2)-constrained if it satisfies the one-dimensional (d_1, k_1) run length constraint horizontally and the one-dimensional (d_2, k_2) run length constraint vertically. For given $d_1, k_1, d_2,$ and k_2, the asymmetric two-dimensional capacity is defined as

$$C_{d_1, k_1, d_2, k_2} = \lim_{m, n \to \infty} (1/(mn)) \log_2 N_{m,n}^{(d_1, k_1, d_2, k_2)}$$

where $N_{m,n}^{(d_1, k_1, d_2, k_2)}$ denotes the number of (d_1, k_1, d_2, k_2)-constrained $m \times n$ binary patterns. We determine whether the capacity is positive or is zero, for many choices of (d_1, k_1, d_2, k_2).

Index Terms—Capacity, constraints, magnetic, optical recording, run length coding.

I. INTRODUCTION

Run length constraints derive from digital storage applications [7]. For nonnegative integers d and k, a binary sequence is said to satisfy a one-dimensional (d, k)-*constraint* if every run of zeros has length at least d and at most k (if two ones are adjacent in the sequence we say that a run of zeros of length zero is between them). A two-dimensional binary pattern arranged in an $m \times n$ rectangle is said to be (d_1, k_1, d_2, k_2)-*constrained* if it satisfies a one-dimensional (d_1, k_1)-constraint horizontally and a one-dimensional (d_2, k_2)-constraint vertically. The two-dimensional (d_1, k_1, d_2, k_2)-*capacity* is defined as

$$C_{d_1, k_1, d_2, k_2} = \lim_{m, n \to \infty} \frac{\log_2 N_{m,n}^{(d_1, k_1, d_2, k_2)}}{mn}$$

where $N_{m,n}^{(d_1, k_1, d_2, k_2)}$ denotes the number of $m \times n$ rectangles that are (d_1, k_1, d_2, k_2)-constrained. If $d = d_1 = d_2$ and $k = k_1 = k_2$ (this is called the *symmetric constraint*) then the two-dimensional (d_1, k_1, d_2, k_2)-capacity is called the two-dimensional (d, k)-*capacity*, and is denoted by $C_{d,k}$. Two-dimensional run length constraints have recently become a focus of increased study [1], [2], [4]–[7], [9], [15], [16], [21]. A proof was given in [9] that shows the two-dimensional (d, k)-capacities exist, and essentially the same proof shows that the C_{d_1, k_1, d_2, k_2} exist.

The two-dimensional asymmetric *positive capacity region* is the set

$$\{(d_1, k_1, d_2, k_2) : C_{d_1, k_1, d_2, k_2} > 0\}.$$

It is of interest to determine the exact values of the capacities of the various two-dimensional constraints in the positive capacity region, or at least to find good approximations or bounds. A more basic question, however, is to determine which constraints actually lie in the positive capacity region and which do not. We provide here a partial answer to this question.

The exact value of the capacity C_{d_1, k_1, d_2, k_2} has been unknown for all but a few cases. In fact, in all cases when the capacity has been known exactly, its value has been zero and the constraints have been symmetric. The first exactly known two-dimensional capacity was shown in [1] to be $C_{1,2} = 0$ and a complete characterization of which (d, k) integer pairs yield positive capacities for symmetric constraints was given in [9] and is stated as the proposition below.

Proposition 1: $C_{d,k} > 0$ if and only if $k - d \geq 2$ or $(d, k) = (0, 1)$.

Fairly tight upper and lower bounds on the value of $C_{0,1}$ were given in [2], improved in [6], [12], and extended to three-dimensional run length constraints in [12]. In [15], an encoding procedure for the symmetric two-dimensional $(0, 1)$-constrained channel was given whose coding rating comes incredibly close to the capacity $C_{0,1}$. For other positive two-dimensional (d, k)-capacities various bounds were given in [9], [16], and approximations were given in [21]. Asymmetric two-dimensional (d_1, k_1, d_2, k_2)-constraints were studied in [4], which discussed mergings and the Hamming distances between (d_1, k_1, d_2, k_2)-constrained rectangles. Codes for certain other types of constraints in two dimensions were studied in [3], [6], [10], [11], [13], [14], [17]–[21].