New Bounds on the Capacity of Multidimensional Run-Length Constraints

Moshe Schwartz, *Senior Member, IEEE*, and Alexander Vardy, *Fellow, IEEE*

***Abstract*—We examine the well-known problem of determining the capacity of multidimensional run-length-limited constrained systems. By recasting the problem, which is essentially a combinatorial counting problem, into a probabilistic setting, we are able to derive new lower and upper bounds on the capacity of $(0,k)$-RLL systems. These bounds are better than all previously-known analytical bounds for $k \geqslant 2$, and are tight asymptotically. Thus, we settle the open question: what is the rate at which the capacity of $(0,k)$-RLL systems converges to 1 as $k \to \infty$? We also provide the first nontrivial upper bound on the capacity of general (d,k)-RLL systems.**

***Index Terms*—Constrained coding, multidimensional constraints, run-length limited coding, 2-D constrained coding.**

I. Introduction

The one-dimensional (d,k)-RLL constraint is the set of all finite binary sequences in which every two adjacent 1's are separated by at least d zeroes, and no more than k 0's appear consecutively. This constraint was first narrowly defined by Kautz [13], and later generalized to its current (d,k)-RLL form by Tang and Bahl [29].

The study of constrained systems was initiated by Shannon [24] who defined their capacity as

$$\mathsf{cap}(S) = \limsup_{n\to\infty} \frac{\log_2 |S(n)|}{n}$$

where $S(n)$ denotes the number length-n sequences with the constraint, and the constraint S is the union $S = \cup_{n\geqslant 1} S(n)$. These constraints have a variety of applications, especially in coding for storage devices (see [8], [16], [17] and references therein).

The emergence of 2-D recording systems brought to light the need for 2-D and even multidimensional constrained systems. A 2-D (d,k)-RLL constrained system is the set of all (finite-sized) binary arrays in which every row and every column obeys the 1-D (d,k)-RLL constraint. The generalization to the D-dimensional case is obvious, and we denote such a system as $S^D_{d,k}$. Though we consider in this paper mostly symmetrical constraints, i.e., the same d and k along every dimension, the results generalize easily to asymmetrical RLL constraints as well.

In the 1-D case, it is well known that $\mathsf{cap}(S^1_{d,k})$, for $0 \leqslant d \leqslant k$, is the base-2 logarithm of the largest positive root of the polynomial

$$x^{k+1} - x^{k-d} - x^{k-d-1} - \cdots - x - 1.$$

However, unlike the 1-D case, very little is known about the capacity of the 2-D case, defined for a general constraint S as

$$\mathsf{cap}(S) = \limsup_{n_1,n_2\to\infty} \frac{\log_2 |S(n_1,n_2)|}{n_1 n_2}$$

where $S(n_1,n_2)$ is the set of all $n_1 \times n_2$ arrays with the constraint. The definition easily generalizes to the D-dimensional case

$$\mathsf{cap}(S) = \limsup_{n_1,n_2\ldots,n_D\to\infty} \frac{\log_2 |S(n_1,n_2,\ldots,n_D)|}{n_1 n_2 \ldots n_D}.$$

Exact expressions for the capacity of nontrivial 2-D constraints (with nonzero capacity) are rare. For the $(1,\infty)$-RLL constraint on the hexagonal lattice, Baxter [2] gave an exact but not rigorous analytical solution for the capacity using the corner transfer matrix method. Schwartz and Bruck [23] described a rigorous method for finding the capacity of general 2-D constraints. The expressive power of this method is, however, yet unknown. An example of an exact rigorously-proved expression for the capacity of a 2-D generalization of the 1-D (0, 1)-RLL is shown in [23]. More recently, Louidor and Marcus [15] calculated the exact capacity of two specific multidimensional constraints: the 2-CC (Charge Constrained) system, and the ODD constraint.

In [3], Calkin and Wilf gave a numerical estimation method for the capacity of the 2-D (0, 1)-RLL constraint which gives

$$0.5878911617 \leqslant \mathsf{cap}(S^2_{0,1}) \leqslant 0.5878911618.$$

Their method ingeniously uses the fact that a certain matrix induced by the constraint is symmetric. Unfortunately, this happens only for the case of (0, 1)-RLL (and by inverting all the bits, the equivalent $(1,\infty)$-RLL case). Using the same method in the 3-D case, it was shown in [18] that

$$0.522501741838 \leqslant \mathsf{cap}\left(S^3_{0,1}\right) \leqslant 0.526880847825.$$

Several constructive methods were suggested which, by devising an appropriate encoder and analyzing its rate, yield lower bounds on the capacity of the constrained system. These works

Manuscript received September 10, 2010; accepted December 02, 2010. Date of current version June 22, 2011. The material in this paper was presented in part at the 16th Applied Algebra, Algebraic Algorithms, and Error Correcting Codes Symposium (AAECC 16), Las Vegas, NV, February 2006.

M. Schwartz is with the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel (e-mail: schwartz@ee.bgu.ac.il).

A. Vardy is with the Department of Electrical and Computer Engineering, the Department of Computer Science and Engineering, and the Department of Mathematics, University of California San Diego, La Jolla, CA 92093-0407 USA (e-mail: avardy@ucsd.edu).

Communicated by G. Cohen, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2011.2119464

TABLE I
COMPARISON OF LOWER BOUNDS (LB) AND UPPER BOUNDS (UB) ON $\mathsf{cap}(S^2_{0,k})$, FOR $2 \leqslant k \leqslant 10$. LOWER AND UPPER BOUNDS ARE ROUNDED DOWN AND UP, RESPECTIVELY, TO SIX DECIMAL DIGITS

k	LB by [12]	LB by [28]	LB by Theorem 6	UB by Theorem 16	UB by [12]
2	0.587891		0.758292	0.904373	0.935785
3	0.793945		0.893554	0.947949	0.976723
4	0.793945		0.950450	0.970467	0.990840
5	0.862630		0.976217	0.983338	0.996214
6	0.862630		0.988383	0.990816	0.998384
7	0.896972		0.994268	0.995068	0.999295
8	0.896972	0.943398	0.997155	0.997410	0.999687
9	0.917578	0.943398	0.998583	0.998663	0.999860
10	0.917578	0.981164	0.999293	0.999318	0.999936

include [22], [25], [26] as well as the bit-stuffing method described in [7], [27]. A few recursive constructions were also given by Etzion in [5]. A generalization of 2-D (d,∞)-RLL called checkerboard constraints, defined by a mandatory region of 0's around each of the 1's, was explored in [19], [30].

General *analytical* bounds on the capacity were given in [12]. Amazingly, we still do not know the exact capacity of the multidimensional RLL-constraint except when it is exactly zero [4], [9]. We also know the limit of the capacity of multidimensional (d,∞)-RLL constraints as the number of dimensions goes to infinity [20].

The analytical bounds we improve upon in this work are those of 2-D $(0,k)$-RLL, $k \geqslant 2$. The bounds are the following:

Theorem 1 (Theorem 3, [12]): For every positive integer k

$$\mathsf{cap}\left(S^2_{0,k}\right) \geqslant 1 - \frac{1-\mathsf{cap}\left(S^2_{0,1}\right)}{\lceil k/2 \rceil}.$$

Theorem 2 (in [28]): For all integers $k \geqslant 8$

$$\mathsf{cap}\left(S^2_{0,k}\right) \geqslant 1 + \frac{\log_2\left(1-(\lfloor k/2\rfloor+1)2^{-(\lfloor k/2\rfloor-1)}\right)}{(\lfloor k/2\rfloor+1)^2}.$$

Theorem 3 (Theorem 7, [12]): For every positive integer k

$$\mathsf{cap}\left(S^2_{0,k}\right) \leqslant 1 - \frac{1}{k+1}\log_2\left(\frac{1}{1-2^{-(k+1)}}\right).$$

Our new bounds are given in Theorem 6 and Theorem 16. A numerical comparison with the previously-best analytical bounds for $2 \leqslant k \leqslant 10$ is given in Table I. Our results also surpass other published nonanalytical bounds for all $k \geqslant 2$ except the lower bound for (0, 2)-RLL given in [25]. Furthermore, our lower and upper bounds agree asymptotically. This settles the open question of the rate of convergence to 1 of $\mathsf{cap}(S^D_{0,k})$ as $k \to \infty$ by showing that for any fixed D

$$1 - \mathsf{cap}\left(S^D_{0,k}\right) = \frac{D\log_2 e}{4\cdot 2^k}(1+o(1))$$

where, throughout the paper, $o(1)$ denotes a function tending to 0 when $k \to \infty$.

Our approach to the problem of bounding the capacity is to recast the problem from a combinatorial counting problem to a probability bounding problem. Suppose we randomly select a sequence of length n with uniform distribution. Let $A_n(S)$ denote the event that this sequence is in the constrained system S. Then the total number of length-n sequences in S can be easily written as

$$|S(n)| = \Pr[A_n(S)]\cdot 2^n.$$

It follows that

$$\begin{aligned}\mathsf{cap}(S) &= \limsup_{n\to\infty}\frac{\log_2|S(n)|}{n}\\ &= \limsup_{n\to\infty}\frac{\log_2(\Pr[A_n(S)]2^n)}{n}\\ &= \limsup_{n\to\infty}\frac{\log_2\Pr[A_n(S)]}{n}+1.\end{aligned}$$

This translates in a straightforward manner to higher dimensions as well. By calculating or bounding $\Pr[A_n(S)]$, we may get the exact capacity or bounds on it, which is the basis for what is to follow.

The work is organized as follows. In Section II we use monotone families to construct lower bounds on $\mathsf{cap}(S^D_{0,k})$ and an upper bound on $\mathsf{cap}(S^D_{d,k})$. While this method may also be used to lower bound $\mathsf{cap}(S^D_{d,\infty})$, the resulting bound is extremely weak. We continue in Section III by deriving an upper bound on $\mathsf{cap}(S^D_{0,k})$ using a large-deviation bound for sums of nearly-independent random variables. We generalize our results to the asymmetric case over a general alphabet in Section IV. We conclude in Section V by discussing the asymptotics of our new bounds and comparing them with the case of (d,∞)-RLL.

II. Bounds From Monotone Families

We use monotone increasing and decreasing families to find new lower bounds on the capacity of $(0,k)$-RLL, and a new upper bound on the capacity of (d,k)-RLL, $d \geqslant 1$. We start with the definition of these families.

Definition 4: Let Ω be some finite set. A family $\mathcal{F} \subseteq 2^\Omega$ is said to be *monotone increasing* if when $A \in \mathcal{F}$ and $A \subseteq A' \subseteq \Omega$, then $A' \in \mathcal{F}$. It is said to be *monotone decreasing* if when $A \in \mathcal{F}$ and $A' \subseteq A \subseteq \Omega$, then $A' \in \mathcal{F}$.

The following theorem is due to Kleitman [14]:

Theorem 5: Let Ω be some finite set. Also, let $\mathcal{A},\mathcal{B}$ be monotone increasing families, and $\mathcal{C},\mathcal{D}$ be monotone decreasing families, all over Ω. Let X be a random variable describing a uniformly-distributed random choice of a subset of Ω out of the $2^{|\Omega|}$ possible subsets. Then

$$\Pr[X \in \mathcal{A} \cap \mathcal{B}] \geqslant \Pr[X \in \mathcal{A}] \cdot \Pr[X \in \mathcal{B}] \tag{1}$$
$$\Pr[X \in \mathcal{C} \cap \mathcal{D}] \geqslant \Pr[X \in \mathcal{C}] \cdot \Pr[X \in \mathcal{D}] \tag{2}$$
$$\Pr[X \in \mathcal{A} \cap \mathcal{C}] \leqslant \Pr[X \in \mathcal{A}] \cdot \Pr[X \in \mathcal{C}]. \tag{3}$$

We can now apply Kleitman's theorem to $(0,k)$-RLL constrained systems:

Theorem 6: For all integers $k \geqslant 0$

$$\mathsf{cap}(S^2_{0,k}) \geqslant 2\mathsf{cap}(S^1_{0,k}) - 1.$$

Proof: The constrained system we examine is $S^2_{0,k}$, and let us denote by $A_{n_1,n_2}(S^2_{0,k})$ the event that a randomly chosen $n_1 \times n_2$ array is $(0,k)$-RLL.

We now define two closely related constraints. Let R denote the set of finite 2-D arrays in which every **row** is $(0,k)$-RLL, and C denote the set of finite 2-D arrays in which every **column** is $(0,k)$-RLL. Similarly we define the events $A_{n_1,n_2}(R)$ and $A_{n_1,n_2}(C)$. By definition

$$A_{n_1,n_2}(S^2_{0,k}) = A_{n_1,n_2}(R) \cap A_{n_1,n_2}(C).$$

The crucial observation is that both R and C are monotone increasing families. This is seen by defining

$$\Omega = \{(i_1,i_2) \,|\, 0 \leqslant i_1 < n_1 \text{ and } 0 \leqslant i_2 < n_2\}$$

and considering an $n_1 \times n_2$ binary array as the subset of Ω corresponding to the positions of 1's in the array. It follows that an array in R (or in C) corresponds to a subset of Ω whose supersets are also in R (or in C) since, obviously, runs of 0's may only get shorter by adding 1's to the array.

Hence, by Theorem 5

$$\begin{aligned}\Pr\left[A_{n_1,n_2}\left(S^2_{0,k}\right)\right] &= \Pr[A_{n_1,n_2}(R) \cap A_{n_1,n_2}(C)] \\ &\geqslant \Pr\left[A_{n_1,n_2}(R)\right] \cdot \Pr\left[A_{n_1,n_2}(C)\right].\end{aligned}$$

It follows that

$$\begin{aligned}&\mathsf{cap}(S^2_{0,k}) \\ &\quad= \limsup_{n_1,n_2\to\infty} \frac{\log_2 \Pr[A_{n_1,n_2}](S^2_{0,k})}{n_1 n_2} + 1 \\ &\quad\geqslant \limsup_{n_1,n_2\to\infty} \frac{\log_2(\Pr[A_{n_1,n_2}(R)]\Pr[A_{n_1,n_2}(C)])}{n_1 n_2} + 1.\end{aligned}$$

Now, both $\Pr[A_{n_1,n_2}(R)]$ and $\Pr[A_{n_1,n_2}(C)]$ may be easily expressed in terms of 1-D constrained systems. An $n_1 \times n_2$ binary array chosen randomly with uniform distribution is equivalent to a set of $n_1 n_2$ i.i.d. random variables for each of the array's bits, each having a "1" with probability $1/2$. Thus

$$\begin{aligned}\Pr\left[A_{n_1,n_2}(R)\right] &= \left(\Pr\left[A_{n_2}\left(S^1_{0,k}\right)\right]\right)^{n_1} \\ \Pr\left[A_{n_1,n_2}(C)\right] &= \left(\Pr\left[A_{n_1}\left(S^1_{0,k}\right)\right]\right)^{n_2}.\end{aligned}$$

Plugging this back into the expression for the capacity, we get

$$\begin{aligned}&\mathsf{cap}\left(S^2_{0,k}\right) \\ &\quad\geqslant \limsup_{n_1,n_2\to\infty} \frac{\log_2(\Pr[A_{n_1,n_2}(R)]\Pr[A_{n_1,n_2}(C)])}{n_1 n_2} + 1 \\ &\quad= \left(\limsup_{n_1\to\infty} \frac{\log_2 \Pr\left[A_{n_1}\left(S^1_{0,k}\right)\right]}{n_1} + 1\right) \\ &\qquad+ \left(\limsup_{n_2\to\infty} \frac{\log_2 \Pr\left[A_{n_2}\left(S^1_{0,k}\right)\right]}{n_2} + 1\right) - 1 \\ &\quad= 2\mathsf{cap}\left(S^1_{0,k}\right) - 1.\end{aligned}$$

■

This is generalized to higher dimensions in the following theorem.

Theorem 7: Let $D_1, D_2 \geqslant 1$ be integers, then

$$\mathsf{cap}\left(S^{D_1+D_2}_{0,k}\right) \geqslant \mathsf{cap}\left(S^{D_1}_{0,k}\right) + \mathsf{cap}\left(S^{D_2}_{0,k}\right) - 1.$$

Proof: The proof is a simple generalization of the proof of Theorem 6. Let $D = D_1 + D_2$, and let $\boldsymbol{n}$ denote the list $n_1, n_2, \ldots, n_D$. We define $A_{\boldsymbol{n}}(S^D_{0,k})$ to be the event that a randomly chosen $n_1 \times \cdots \times n_D$ array is $(0,k)$-RLL.

We now define S' to be the set of all $n_1 \times \cdots \times n_D$ arrays for which all projections onto the last D_2 coordinates created by fixing the first D_1 coordinates form $n_{D_1+1} \times \cdots \times n_D$ arrays obeying the $S^{D_2}_{0,k}$ constraint. Similarly, S'' is defined as the set of all $n_1 \times \cdots \times n_D$ arrays for which all projections onto the first D_1 coordinates created by fixing the last D_2 coordinates form $n_1 \times \cdots \times n_{D_1}$ arrays obeying the $S^{D_1}_{0,k}$ constraint. The corresponding events $A_{\boldsymbol{n}}(S')$ and $A_{\boldsymbol{n}}(S'')$ are defined appropriately.

Again, both S' and S'' are monotone increasing families, and also

$$A_{\boldsymbol{n}}\left(S^D_{0,k}\right) = A_{\boldsymbol{n}}(S') \cap A_{\boldsymbol{n}}(S'').$$

Thus

$$\Pr[A_{\boldsymbol{n}}\left(S^D_{0,k}\right)] \geqslant \Pr[A_{\boldsymbol{n}}(S')]\Pr[A_{\boldsymbol{n}}(S'')].$$

Furthermore

$$\begin{aligned}\Pr[A_{\boldsymbol{n}}(S')] &= \left(\Pr[A_{n_{D_1+1},\ldots,n_D}(S^{D_2}_{0,k})]\right)^{n_1 n_2 \ldots n_{D_1}} \\ \Pr[A_{\boldsymbol{n}}(S'')] &= \left(\Pr[A_{n_1,\ldots,n_{D_1}}(S^{D_1}_{0,k})]\right)^{n_{D_1+1} n_{D_1+2} \ldots n_D}.\end{aligned}$$

Like in Theorem 6, it now follows that

$$\begin{aligned}\mathsf{cap}\left(S^D_{0,k}\right) &\geqslant \limsup_{\boldsymbol{n}\to\infty} \frac{\log_2(\Pr[A_{\boldsymbol{n}}(S')]\Pr[A_{\boldsymbol{n}}(S'')])}{n_1 \ldots n_D} + 1 \\ &= \mathsf{cap}\left(S^{D_1}_{0,k}\right) + \mathsf{cap}\left(S^{D_2}_{0,k}\right) - 1.\end{aligned}$$

■

Corollary 8: For any integer $D \geqslant 1$

$$\mathsf{cap}\left(S^D_{0,k}\right) \geqslant D\Big(\mathsf{cap}\left(S^1_{0,k}\right) - 1\Big) + 1.$$

Proof: Results by iterating Theorem 7 while choosing, for example, $D_1 = 1$ and $D_2 = D - 1$ in each step. ■

As mentioned in [12], for the 1-D capacity of $(0,k)$-RLL, we have

$$1 - \mathsf{cap}\left(S^1_{0,k}\right) = \frac{\log_2 e}{4 \cdot 2^k}(1 + o(1)). \tag{4}$$

Corollary 9: For any integer $D \geqslant 1$

$$1 - \mathsf{cap}\left(S^D_{0,k}\right) \leqslant \frac{D \log_2 e}{4 \cdot 2^k}(1 + o(1)).$$

Proof: Substitute (4) into Corollary 8. ■

We note that similar lower bounds to that of Theorem 7 may be given for the (d,∞)-RLL constraint, since such arrays form a monotone decreasing family. However, the resulting bounds are very weak. We can, however, mix monotone increasing and decreasing families to get the following result.

Theorem 10: Let $D \geqslant 1$ be some integer, and $k \geqslant d$ also integers, then

$$\mathsf{cap}\left(S^D_{d,k}\right) \leqslant \mathsf{cap}\left(S^D_{d,\infty}\right) + \mathsf{cap}\left(S^D_{0,k}\right) - 1.$$

Proof: Let $\boldsymbol{n}$ denote $n_1, n_2, \ldots, n_D$. It is easy to verify that $S^D_{d,\infty}$ is monotone decreasing, while $S^D_{0,k}$ is monotone increasing. Also, we note that $S^D_{d,k} = S^D_{d,\infty} \cap S^D_{0,k}$, so by Theorem 5

$$\begin{aligned}\Pr\left[A_{\boldsymbol{n}}(S^D_{d,k})\right] &= \Pr\left[A_{\boldsymbol{n}}\left(S^D_{d,\infty}\right) \cap A_{\boldsymbol{n}}\left(S^D_{0,k}\right)\right] \\ &\leqslant \Pr\left[A_{\boldsymbol{n}}(S^D_{d,\infty})\right] \Pr\left[A_{\boldsymbol{n}}\left(S^D_{0,k}\right)\right].\end{aligned}$$

Hence, we get

$$\begin{aligned}\mathsf{cap}\left(S^D_{d,k}\right) &= \limsup_{\boldsymbol{n}\to\infty} \frac{\log_2 \Pr\left[A_{\boldsymbol{n}}\left(S^D_{d,k}\right)\right]}{n_1 \ldots n_D} + 1 \\ &\leqslant \limsup_{\boldsymbol{n}\to\infty} \frac{\log_2 \Pr[A_{\boldsymbol{n}}(S^D_{d,\infty})]}{n_1 \ldots n_D} \\ &\quad + \limsup_{\boldsymbol{n}\to\infty} \frac{\log_2 \Pr[A_{\boldsymbol{n}}\left(S^D_{0,k}\right)]}{n_1 \ldots n_D} + 1 \\ &= \mathsf{cap}\left(S^D_{d,\infty}\right) + \mathsf{cap}\left(S^D_{0,k}\right) - 1.\end{aligned}$$

■

III. New Upper Bounds

In this section we present upper bounds on the capacity of $(0,k)$-RLL. Unlike the previous section, these bounds are explicit. The method we use is similar to the one employed by Godbole *et al.* in [6], which relies on a probability bound by Roos [21]. The latter is an improvement of the well-known bound by Janson [10]. Since the bound by Roos is overly-parametrized for our needs, we provide a simpler more specific bound that still retains the essence of the improvement of [21] over [10].

Let $\mathcal{Q}$ be a finite index set, and let $\{J_i\}_{i\in\mathcal{Q}}$ be a set of independent Boolean random variables. We are given a family $\mathcal{A} \subseteq 2^{\mathcal{Q}}$ of subsets of $\mathcal{Q}$. For each $A \in \mathcal{A}$, define an indicator random variable $I_A = \prod_{i\in A} J_i$, then set

$$X = \sum_{A\in\mathcal{A}} I_A.$$

Thus, X counts the number of sets $A \in \mathcal{A}$ contained in the random set $\{i \in \mathcal{Q} \,|\, J_i = 1\}$. In particular, if the sets $A \in \mathcal{A}$ correspond to "bad" events, then $\Pr(X = 0)$ is the probability that no such "bad" events occur.

Following [10], given an $A \in \mathcal{A}$ and a $B \in \mathcal{A}$, let us write $A \sim B$ if $A \cap B \neq \varnothing$ and $A \neq B$. As in [10], let us also define

$$\begin{aligned}p_A &= E(I_A) \\ \mu &= E(X) = \sum_{A\in\mathcal{A}} p_A \\ \delta &= \frac{1}{\mu}\sum_{A\in\mathcal{A}}\sum_{B\sim A} E(I_A I_B).\end{aligned}$$

The δ is a measure of dependence between the random variables $\{I_A\}_{A\in\mathcal{A}}$. Janson's inequality of [10], [11] uses δ to estimate the relative error in approximating the distribution of X by the Poisson distribution with mean μ. In particular, this inequality implies that

$$\Pr(X = 0) \leqslant \exp\left(-\frac{\mu}{1+\delta}\right). \tag{5}$$

We strengthen the bound (5) by replacing δ with a more refined measure of dependence. As in [10], let us define $X_A = I_A + \sum_{B\sim A} I_B$ for each $A \in \mathcal{A}$. Thus, X_A is the sum of all those indicators that are dependent on I_A. The following theorem is our strengthened version of Janson's inequality.

Theorem 11: For each $A \in \mathcal{A}$, define

$$\Delta_A = E\left(\frac{1}{X_A}\middle| I_A = 1\right) = \sum_{i\geqslant 1} \frac{\Pr\left(X_A = i \,|\, I_A = 1\right)}{i}.$$

Then

$$\Pr(X = 0) \leqslant \exp\left(-\sum_{A\in\mathcal{A}} p_A \Delta_A\right).$$

Proof: The beginning of the proof closely follows the original argument of Janson [10]. For completeness, we include this part of the argument below. For real $t \geqslant 0$, consider the moment generating function $\psi(t) = E(e^{-tX})$. It is clear that $\Pr(X = 0) \leqslant E(e^{-tX})$ for all t, and therefore

$$\ln \Pr(X = 0) \leqslant \ln \psi(t) = \int_0^t \frac{d\psi(s)}{ds}\frac{1}{\psi(s)}\, ds \tag{6}$$

where the (second) equality follows from the fact that $\ln \psi(0) = 0$. To produce an upper bound on the right-hand side of (6), write

$$-\frac{d\psi(t)}{dt} = E(Xe^{-tX}) = \sum_{A\in\mathcal{A}} E(I_A e^{-tX}). \tag{7}$$

Now express X as $X_A + (X - X_A)$, and then also observe that $X - X_A$ does not depend on I_A. The random variables e^{-tX_A} and $e^{-t(X-X_A)}$ are decreasing functions of $\{J_i\}_{i\in\mathcal{Q}}$, and therefore, positively correlated by the FKG inequality [1, p. 85]. Thus, we have

$$\begin{aligned} E\left(I_A e^{-tX}\right) &= p_A E\left(e^{-tX_A}e^{-t(X-X_A)}\,\middle|\, I_A=1\right) \\ &\geqslant p_A E\left(e^{-tX_A}|I_A=1\right)E\left(e^{-t(X-X_A)}|I_A=1\right) \\ &= p_A E\left(e^{-tX_A}|I_A=1\right)E\left(e^{-t(X-X_A)}\right) \\ &\geqslant p_A E\left(e^{-tX_A}|I_A=1\right)\psi(t) \end{aligned}$$

where the first transition is by the definition of conditional expectation, the second follows from the FKG inequality (the condition $I_A=1$ fixes J_i for $i\in A$, but e^{-tX_A} and $e^{-t(X-X_A)}$ are still decreasing in the remaining J_i); the third holds because $e^{-t(X-X_A)}$ is independent of I_A, and the last follows from the fact that $e^{-t(X-X_A)}\geqslant e^{-tX}$ for all $t\geqslant 0$. Combining this with (6) and (7) we get

$$\ln \Pr(X=0)\leqslant -\sum_{A\in\mathcal{A}} p_A\int_0^t E(e^{-sX_A}|I_A=1)ds. \quad (8)$$

At this point, the standard argument of [10] uses Jensen's inequality twice to arrive at (5). We depart from this argument, and instead simply evaluate (8) exactly, as follows:

$$\begin{aligned} &\ln\Pr\left(X=0\right) \\ &\leqslant -\sum_{A\in\mathcal{A}} p_A\int_0^t\left(\sum_{i\geqslant 1}\Pr\left(X_A=i\mid I_A=1\right)e^{-si}\right)ds \\ &= -\sum_{A\in\mathcal{A}} p_A\sum_{i\geqslant 1}\frac{\Pr\left(X_A=i\,|\,I_A=1\right)}{i}(1-e^{-ti}) \quad (9)\end{aligned}$$

where the inner summation starts from $i=1$ because $\Pr(X_A=0\,|\,I_A=1)=0$. The theorem now follows by taking the limit as $t\to\infty$ in (9). ■

Since we have avoided invoking Jensen's inequality after (8), Theorem 11 is *always* strictly stronger than (5) (note that the exponential function $x\mapsto e^{-sx}$ used with the Jensen inequality in [10] is strictly convex).

It appears that Theorem 11 is most useful when Δ_A is the same for all $A\in\mathcal{A}$. In this case, we will write Δ instead of Δ_A, and our bound reduces to

$$\Pr(X=0)\leqslant e^{-\mu\Delta}. \quad (10)$$

Our goal is to use Theorem 11 to show an upper bound on the capacity of 2-D $(0,k)$-RLL systems. If $S(n_1,n_2)$ denotes the set of 2-D $(0,k)$-RLL arrays of size $n_1\times n_2$ then by definition

$$\mathsf{cap}\left(S^2_{0,k}\right)=\limsup_{n_1,n_2\to\infty}\frac{\log_2|S(n_1,n_2)|}{n_1n_2}.$$

However, it would be more convenient to work in a more symmetric setting. Intuitively speaking, positions which are close enough to the edge of the array are "less constrained" than others lying fully within the array since they have smaller neighborhoods. We overcome this difficulty by considering cyclic $(0,k)$-RLL arrays.

We say that a binary $n_1\times n_2$ array $\mathcal{M}$ is *cyclic* $(0,k)$*-RLL* if there do not exist $0\leqslant i_1<n_1$ and $0\leqslant i_2<n_2$ such that $\mathcal{M}_{i_1,i_2}=\mathcal{M}_{i_1+1,i_2}=\cdots=\mathcal{M}_{i_1+k,i_2}=0$ or $\mathcal{M}_{i_1,i_2}=\mathcal{M}_{i_1,i_2+1}=\cdots=\mathcal{M}_{i_1,i_2+k}=0$, where the indices are taken modulo n_1 and n_2 respectively. We denote the set of all such $n_1\times n_2$ arrays as $S_c(n,m)$. The next lemma shows that by restricting ourselves to cyclic $(0,k)$-RLL arrays, we do not change the capacity.

Lemma 12: For all positive integers k

$$\mathsf{cap}(S^2_{0,k})=\limsup_{n_1,n_2\to\infty}\frac{\log_2|S_c(n_1,n_2)|}{n_1n_2}.$$

Proof: Trivially, $|S(n_1,n_2)|\geqslant|S_c(n_1,n_2)|$, so

$$\begin{aligned}\mathsf{cap}(S^2_{0,k}) &= \limsup_{n_1,n_2\to\infty}\frac{\log_2|S(n_1,n_2)|}{n_1n_2} \\ &\geqslant \limsup_{n_1,n_2\to\infty}\frac{\log_2|S_c(n_1,n_2)|}{n_1n_2}.\end{aligned}$$

For the other direction, we note that any array from $S(n_1-1,n_2-1)$, with an additional row and column of all 1's, becomes an $n_1\times n_2$ cyclic $(0,k)$-RLL array. So $|S_c(n_1,n_2)|\geqslant|S(n_1-1,n_2-1)|$ and

$$\begin{aligned}&\limsup_{n_1,n_2\to\infty}\frac{\log_2|S_c(n_1,n_2)|}{n_1n_2} \\ &\quad\geqslant \limsup_{n_1,n_2\to\infty}\frac{\log_2|S(n_1-1,n_2-1)|}{(n_1-1)(n_2-1)}\cdot\frac{(n_1-1)(n_2-1)}{n_1n_2} \\ &\quad= \mathsf{cap}\left(S^2_{0,k}\right).\end{aligned}$$

■

We are now ready to describe the upper bound on the capacity. We start by considering a random $n_1\times n_2$ binary array, chosen with uniform distribution, which is equivalent to saying that we have an array of n_1n_2 i.i.d. 0-1 random variables J_{i_1,i_2}, $0\leqslant i_j<n_j$, with $J_{i_1,i_2}\sim Be(1/2)$.

For the remainder of this section, we invert the bits of the array, or equivalently, we say that an array is $(0,k)$-RLL if it does not contain $k+1$ consecutive 1's along any row or column. Furthermore, by Lemma 12, we consider only cyclic $(0,k)$-RLL arrays. Suppose we define the following subsets of coordinates of the arrays:

$$\begin{aligned}\mathcal{A}_{\mathrm{V}} &= \{\{(i_1,i_2),(i_1+1,i_2),\ldots,(i_1+k,i_2)\}\,|\,0\leqslant i_j<n_j\} \\ \mathcal{A}_{\mathrm{H}} &= \{\{(i_1,i_2),(i_1,i_2+1),\ldots,(i_1,i_2+k)\}\,|\,0\leqslant i_j<n_j\} \\ \mathcal{A} &= \mathcal{A}_{\mathrm{V}}\cup\mathcal{A}_{\mathrm{H}}\end{aligned}$$

where all the coordinates are taken with the appropriate modulo.

We further define the following indicator random variables:

$$I_A=\prod_{(i_1,i_2)\in A}J_{i_1,i_2}\quad\text{for all}\quad A\in\mathcal{A}.$$

If $I_A=1$ for some $A\in\mathcal{A}$, we have a *forbidden event* of $k+1$ consecutive 1's along a row or a column (see Fig. 1 for an example).

Finally, we count the number of forbidden events in the random array by defining $X=\sum_{A\in\mathcal{A}}I_A$. It is now clear that the probability that this random array is $(0,k)$-RLL is simply

$$\Pr\left[A_{n_1,n_2}\left(S^2_{0,k}\right)\right]=\Pr[X=0].$$

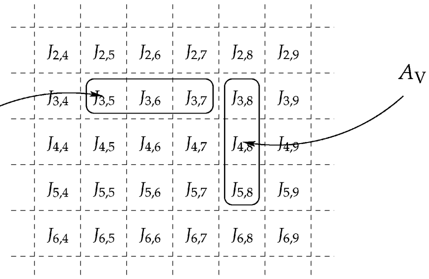

Fig. 1. Depiction of two subsets from $\mathcal{A}$ for $k=2$. A horizontal subset, $A_{\mathrm{H}} \in \mathcal{A}_{\mathrm{H}}$, induces $I_{A_{\mathrm{H}}} = J_{3,5}J_{3,6}J_{3,7}$. A vertical subset, $A_{\mathrm{V}} \in \mathcal{A}_{\mathrm{V}}$, induces $I_{A_{\mathrm{V}}} = J_{3,8}J_{4,8}J_{5,8}$.

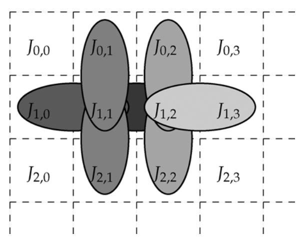

Fig. 2. Shown in different shades of gray are the following subsets for $k=1$: $A = \{(1,1),(1,2)\}$, $\mathcal{A}_{\mathrm{HL}} = \{\{(1,0),(1,1)\}\}$, $\mathcal{A}_{\mathrm{HR}} = \{\{(1,2),(1,3)\}\}$, $\mathcal{A}_{\mathrm{V},0} = \{\{(0,1),(1,1)\},\{(1,1),(2,1)\}\}$, and $\mathcal{A}_{\mathrm{V},1} = \{\{(0,2),(1,2)\},\{(1,2),(2,2)\}\}$.

It is easy to be convinced that this setting agrees with the requirements of Theorem 11, including the symmetry requirements of Δ which allow us to use (10). All we have to do now to upper bound $\Pr[X=0]$, is to calculate μ and Δ.

Let us start by calculating μ. We note that X is the sum of $2n_1n_2$ indicator random variables, so by linearity of expectation

$$\mu = E(X) = \frac{1}{2^{k+1}} \cdot 2n_1n_2 = \frac{n_1n_2}{2^k}$$

since each of the indicator random variables has probability exactly $1/2^{k+1}$ of being 1.

Calculating Δ is more tedious. Since Δ does not depend on the choice of A, we arbitrarily choose the horizontal set of coordinates

$$A = \{(k,k),(k,k+1),\ldots,(k,2k)\}.$$

We now have to calculate $\Pr[X_A = i \,|\, I_A = 1]$. We note that we can partition the set $\{B \in \mathcal{A} | B \sim A\}$ into the following disjoint subsets:

$$\{B \in \mathcal{A} \,|\, B \sim A\} = \left(\cup_{i=0}^{k} \mathcal{A}_{\mathrm{V},i}\right) \cup \mathcal{A}_{\mathrm{HL}} \cup \mathcal{A}_{\mathrm{HR}}$$

where

$$\begin{aligned}
\mathcal{A}_{\mathrm{HL}} &= \{B \in \mathcal{A}_{\mathrm{H}} - \{A\} | (k,k) \in B\} \\
\mathcal{A}_{\mathrm{HR}} &= \{B \in \mathcal{A}_{\mathrm{H}} - \{A\} | (k,2k) \in B\} \\
\mathcal{A}_{\mathrm{V},j} &= \{B \in \mathcal{A}_{\mathrm{V}} | (k,k+j) \in B\}, \quad \text{for all } 0 \leqslant j \leqslant k.
\end{aligned}$$

See Fig. 2 for an example.

We define $X_{\mathrm{HL}} = \sum_{B \in \mathcal{A}_{\mathrm{HL}}} I_B$, and in a similar fashion, X_{HR} and $X_{\mathrm{V},j}$ for all $0 \leqslant j \leqslant k$. Since the indicators for elements from different subsets are independent given $I_A = 1$ because their intersection contains only coordinates from A, it follows that X_{HL}, X_{HR} and $X_{\mathrm{V},j}$, $0 \leqslant j \leqslant k$, are independent given $I_A = 1$.

We now give two lemmas to help us determine the distribution function of X_A.

Lemma 13: Let $f_k^{||}(i)$ denote the number of binary strings of length $2k+1$ with their last $k+1$ positions 1's, and which contain exactly $1 \leqslant i+1 \leqslant k+1$ runs of $k+1$ 1's. Then

$$f_k^{||}(i) = \begin{cases} 2^{k-i-1}, & 0 \leqslant i \leqslant k-1 \\ 1, & i = k. \end{cases}$$

Proof: The only bits we can set to our liking are the first k bits. The case of $f_k^{||}(k)$ is trivial since it requires all the bits to be set to 1's. For the other cases it is easy to be convinced that setting the first $k-i-1$ bits to any value, followed by a single 0 and then i bits set to 1's, is the only way to create a string with exactly $i+1$ runs as required. ■

Since the distributions of X_{HL} and X_{HR} given $I_A = 1$ are the same, they may be now expressed as

$$\begin{aligned}
\Pr[X_{\mathrm{HL}} = i \,|\, I_A = 1] &= \Pr[X_{\mathrm{HR}} = i \,|\, I_A = 1] \\
&= \frac{f_k^{||}(i)}{2^k}. \qquad (11)
\end{aligned}$$

Lemma 14: Let $f_k^{\perp}(i)$ denote the number of binary strings of length $2k+1$ with their middle position a 1, and which contain exactly $0 \leqslant i \leqslant k+1$ runs of $k+1$ 1's. Then,

$$f_k^{\perp}(i) = \begin{cases} 2^{2k} - (k+2)2^{k-1}, & i = 0 \\ (k-i+4)2^{k-i-1}, & 1 \leqslant i \leqslant k \\ 1, & i = k+1. \end{cases}$$

Proof: We begin by noting that $f_k^{\perp}(k+1) = 1$ since only the all-ones string has $k+1$ runs of of 1's of length $k+1$ each. For $1 \leqslant i \leqslant k$ the basic observation is that a string of length $2k+1$ cannot have two nonoverlapping runs of 1's of length $k+1$ each. Hence, in order to get i runs of length $k+1$ 1's, we need exactly one run of 1's of length $k+i$ which has, either one 0 at each side, or is at the beginning or end of the string and has one 0 at its other side. Thus

$$f_k^{\perp}(i) = 2^{k-i+1} + (k-i)2^{k-i-1} = (k-i+4)2^{k-i-1}$$

for all $1 \leqslant i \leqslant k$. We also note that all such strings must have a 1 in their middle position.

Finally, $f_k^{\perp}(0)$ is given by subtracting from the total number of strings, all the previously counted strings, and the 2^{2k} strings having a 0 in their middle position. Thus

$$f_k^{\perp}(0) = 2^{2k+1} - 2^{2k} - \sum_{i=1}^{k+1} f_k^{\perp}(i) = 2^{2k} - (k+2)2^{k-1}.$$

■

Using this lemma, we can now say that

$$\Pr[X_{\mathrm{V},j} = i \,|\, I_A = 1] = \frac{f_k^{\perp}(i)}{2^{2k}}. \qquad (12)$$

Lemma 15: For the $(0,k)$-RLL setting described above

$$\Delta = \sum_{i\geqslant 1} \frac{1}{i} \sum_{\substack{i_L+i_R+i_0+\cdots+i_k=i-1\\ 0\leqslant i_L,i_R\leqslant k\\ 0\leqslant i_0,\ldots,i_k\leqslant k+1}} \frac{f_k^{||}(i_L)f_k^{||}(i_R)}{2^{2k}} \prod_{j=0}^{k} \frac{f_k^{\perp}(i_j)}{2^{2k}}.$$

Proof: By definition

$$X_A = X_{\mathrm{HL}} + X_{\mathrm{HR}} + \sum_{j=0}^{k} X_{\mathrm{V},j} + I_A.$$

As already mentioned before, given $I_A = 1$, we have that X_{HL}, X_{HR}, and $X_{\mathrm{V},j}$ are independent. Thus, the distribution function of X_A given $I_A = 1$ is

$$\begin{aligned}\Pr[X_A = i \,|\, I_A = 1] = &\sum_{\substack{i_L+i_R+i_0+\cdots+i_k=i-1\\ 0\leqslant i_L,i_R\leqslant k\\ 0\leqslant i_0,\ldots,i_k\leqslant k+1}} \Pr[X_{\mathrm{HL}} = i_L | I_A = 1]\\ &\cdot \Pr[X_{\mathrm{HR}} = i_R | I_A = 1]\\ &\cdot \prod_{j=0}^{k} \Pr[X_{\mathrm{V},j} = i_j | I_A = 1].\end{aligned}$$

Plugging (11) and (12) into the last expression and then using the definition of Δ_A we get the desired result. ■

We are now in a position to state the main upper bound.

Theorem 16: Let $k \geqslant 1$ be some integer, then

$$\mathsf{cap}(S_{0,k}^2) \leqslant 1 - \frac{\log_2 e}{2^k}\Delta$$

where Δ is given by Lemma 15.

Proof: We apply Theorem 11 for an array of size $n_1 \times n_2$ and get that

$$\Pr[X = 0] \leqslant e^{-\mu\Delta} = e^{-n_1 n_2 \Delta/2^k}$$

where Δ is given by Lemma 15. Then

$$\mathsf{cap}(S_{0,k}^2) = \limsup_{n_1,n_2\to\infty} \frac{\log_2(\Pr[X=0])}{n_1 n_2} + 1 \leqslant 1 - \frac{\log_2 e}{2^k}\Delta.$$

■

We can make the bound of Theorem 16 weaker for small values of k, but more analytically appealing for an asymptotic analysis. This is achieved by noting that $f_k^{\perp}(0)/2^{2k}$ is almost 1 for large values of k.

Theorem 17: For the 2-D $(0,k)$-RLL constraint

$$1 - \mathsf{cap}\left(S_{0,k}^2\right) \geqslant \frac{\log_2 e}{2\cdot 2^k}(1+o(1)).$$

Proof: We lower bound Δ from Lemma 15 by

$$\begin{aligned}\Delta &= \sum_{i\geqslant 1} \frac{1}{i} \sum_{\substack{i_L+i_R+i_0+\ldots+i_k=i-1\\ 0\leqslant i_L,i_R\leqslant k\\ 0\leqslant i_0,\ldots,i_k\leqslant k+1}} \frac{f_k^{||}(i_L)f_k^{||}(i_R)}{2^{2k}} \prod_{j=0}^{k} \frac{f_k^{\perp}(i_j)}{2^{2k}}\\ &\geqslant \sum_{i=1}^{k} \frac{1}{i} \sum_{\substack{i_L+i_R=i-1\\ 0\leqslant i_L,i_R\leqslant k-1}} \frac{f_k^{||}(i_L)f_k^{||}(i_R)}{2^{2k}} \prod_{j=0}^{k} \frac{f_k^{\perp}(0)}{2^{2k}}\\ &= \left(\frac{1}{2} - \frac{1}{2^{k+1}}\right)(1-(k+2)2^{-(k+1)})^{k+1}\\ &= \frac{1}{2}(1+o(1))\end{aligned}$$

and use Theorem 11. ■

We can generalize these results to higher dimensions.

Theorem 18: Let $D \geqslant 2$ and $k \geqslant 1$ be some fixed integers, then for the D-dimensional $(0,k)$-RLL constraint

$$\mathsf{cap}\left(S_{0,k}^D\right) \leqslant 1 - \frac{D\log_2 e}{2\cdot 2^k}\Delta$$
$$1 - \mathsf{cap}(S_{0,k}^D) \geqslant \frac{D\log_2 e}{4\cdot 2^k}(1+o(1))$$

where Δ is given by

$$\Delta = \sum_{i\geqslant 1} \frac{1}{i} \sum \frac{f_k^{||}(i_L)f_k^{||}(i_R)}{2^{2k}} \prod_{\substack{0\leqslant j\leqslant k\\ 1\leqslant \ell\leqslant D-1}} \frac{f_k^{\perp}(i_{j,\ell})}{2^{2k}}$$

where the inner summation is over all choices of i_L, i_R, and $i_{j,\ell}$ such that $0 \leqslant i_L, i_R \leqslant k$, $0 \leqslant i_{j,\ell} \leqslant k+1$, and also

$$i_L + i_R + \sum_{\substack{0\leqslant j\leqslant k\\ 1\leqslant \ell\leqslant D-1}} i_{j,\ell} = i - 1.$$

Proof: The proof is very similar to the proof of the 2-D case so we just give a brief sketch of it. We start by defining the sets of coordinates $\mathcal{A}_1, \ldots, \mathcal{A}_D$, where $\mathcal{A}_\ell$ contains $k+1$ adjacent positions along the ith dimension. We then define $\mathcal{A} = \cup_{\ell=1}^{D}\mathcal{A}_\ell$. For all $A \in \mathcal{A}$ we define the appropriate indicator random variable I_A and let $X = \sum_{A\in\mathcal{A}} I_A$.

For an $n_1 \times \cdots \times n_D$ array

$$\mu = E[X] = \frac{Dn_1 n_2 \ldots n_D}{2\cdot 2^k}.$$

Again, Δ_A is independent of the choice of $A \in \mathcal{A}$. For our convenience we choose

$$A = \{(k,\ldots,k,k),(k,\ldots,k,k+1),\ldots,(k,\ldots,k,2k)\}$$

which lies along the Dth dimension. We partition the events $B \sim A$ to the ones lying in parallel to A (along the Dth dimension), and those lying perpendicular to A (thus, intersecting A in a single position). The expressions for $f_k^{||}(i)$ and $f_k^{\perp}(i)$ remain the same and we get

$$\Delta = \sum_{i\geqslant 1} \frac{1}{i} \sum \frac{f_k^{||}(i_L)f_k^{||}(i_R)}{2^{2k}} \prod_{\substack{0\leqslant j\leqslant k\\ 1\leqslant \ell\leqslant D-1}} \frac{f_k^{\perp}(i_{j,\ell})}{2^{2k}}$$

where the inner summation is as claimed. Using μ and Δ like we did in Theorem 16 gives us

$$\mathsf{cap}\left(S_{0,k}^{D}\right) \leqslant 1-\frac{D\log_2 e}{2\cdot 2^k}\Delta.$$

Finally, a similar asymptotic analysis gives

$$\begin{aligned}\Delta &\geqslant \left(\frac{1}{2}-\frac{1}{2^{k+1}}\right)(1-(k+2)2^{-(k+1)})^{(D-1)(k+1)}\\ &=\frac{1}{2}(1+o(1)),\end{aligned}$$

which results in

$$1-\mathsf{cap}\left(S_{0,k}^{D}\right)\geqslant \frac{D\log_2 e}{4\cdot 2^k}(1+o(1)).$$

■

Corollary 19: For any integer $D\geqslant 1$

$$1-\mathsf{cap}(S_{0,k}^{D})=\frac{D\log_2 e}{4\cdot 2^k}(1+o(1)).$$

Proof: Simply combine Corollary 9 with Theorem 18. ■

The last corollary shows that we have finally closed the gap between the lower and upper asymptotic bounds.

IV. Generalizations

In this section we generalize the results of the previous section to the asymmetric case and to a general q-ary alphabet. To avoid unnecessary repetition, we state the results with brief sketches of the proofs. The results can be further generalized to the multidimensional asymmetric case, which we avoid here to keep the notation simple.

The q-ary 2-D $(0,k_1,0,k_2)$-RLL constrained system is the set of all (finite-sized) arrays over an alphabet of size q, such that no row contains more than k_1 consecutive zeroes, and no column contains more than k_2 consecutive zeroes. We denote this system as $S^2_{0,k_1,0,k_2}$.

We first note that in the q-ary case, by the definition of the capacity, say for a 1-D constraint S

$$\begin{aligned}\mathsf{cap}(S)&=\limsup_{n\to\infty}\frac{\log_2|S(n)|}{n}\\ &=\limsup_{n\to\infty}\frac{\log_2(\Pr[A_n(S)]q^n)}{n}\\ &=\limsup_{n\to\infty}\frac{\log_2\Pr[A_n(S)]}{n}+\log_2 q.\end{aligned}$$

Thus, in the following, we make sure we change the 1 in the expression for the capacity to $\log_2 q$. The lower bound of Theorem 6 is easily generalized.

Theorem 20: For all integers $k_1,k_2\geqslant 0$

$$\mathsf{cap}\left(S^2_{0,k_1,0,k_2}\right)\geqslant \mathsf{cap}\left(S^1_{0,k_1}\right)+\mathsf{cap}(S^1_{0,k_2})-\log_2 q.$$

Proof: In the proof of Theorem 6 we make sure that R is $(0,k_1)$-RLL in the rows, and C is $(0,k_2)$-RLL in the columns. The rest of the proof is the same. ■

For the upper bound, let us assume the setting as in Theorem 11 apart from the following: Let $\mathcal{A}$ be partitioned into disjoint subsets

$$\mathcal{A}=\mathcal{A}_1\cup\mathcal{A}_2\cup\cdots\cup\mathcal{A}_m$$

such that for any fixed $1\leqslant i\leqslant m$, all $A\in\mathcal{A}_i$ have the same Δ_A, which we will conveniently name $\Delta_{\mathcal{A}_i}$. We also define $\mu_{\mathcal{A}_i}=\sum_{A\in\mathcal{A}_i}E(I_A)$. In the setting of Theorem 11 we have

$$\Pr[X=0]\leqslant\exp\left(-\sum_{i=1}^{m}\mu_{\mathcal{A}_i}\Delta_{\mathcal{A}_i}\right).$$

We also need to generalize $f_k^{||}(i)$ and $f_k^{\perp}(i)$.

Lemma 21: Let $f_{k,q}^{||}(i)$ denote the number of q-ary strings of length $2k+1$ with their last $k+1$ positions 0's, and which contain exactly $1\leqslant i+1\leqslant k+1$ runs of $k+1$ 0's. Let $f_{k,q}^{\perp}(i)$ denote the number of q-ary strings of length $2k+1$ with their middle position a zero, and which contain exactly i runs of $k+1$ zeroes. Then

$$f_{k,q}^{||}(i)=\begin{cases}q^{k-i-1}, & 0\leqslant i\leqslant k-1\\ 1, & i=k\end{cases}$$
$$f_{k,q}^{\perp}(i)=\begin{cases}q^{2k}-\sum_{i=1}^{k+1}f_{k,q}^{\perp}(i), & i=0\\ q^{k-i-1}(q-1)((k-i)(q-1)+2q), & 1\leqslant i\leqslant k\\ 1, & i=k+1\end{cases}$$

Proof: The lemma uses the same counting arguments used in Lemma 13 and Lemma 14. ■

We partition the set of indicators into $\mathcal{A}=\mathcal{A}_{\mathrm{H}}\cup\mathcal{A}_{\mathrm{V}}$. It follows that for an $n_1\times n_2$ array

$$\begin{aligned}\mu_{\mathcal{A}_{\mathrm{H}}}&=\frac{n_1n_2}{q^{k_1+1}}\\ \mu_{\mathcal{A}_{\mathrm{V}}}&=\frac{n_1n_2}{q^{k_2+1}}\\ \Delta_{\mathcal{A}_{\mathrm{H}}}&=\sum_{i\geqslant 1}\frac{1}{i}\sum_{\substack{i_L+i_R+i_0+\cdots+i_{k_1}=i-1\\ 0\leqslant i_L,i_R\leqslant k_1\\ 0\leqslant i_0,\ldots,i_{k_1}\leqslant k_2+1}}\frac{f_{k_1,q}^{||}(i_L)f_{k_1,q}^{||}(i_R)}{q^{2k_1}}s\prod_{j=0}^{k_1}\frac{f_{k_2,q}^{\perp}(i_j)}{q^{2k_2}}\\ \Delta_{\mathcal{A}_{\mathrm{V}}}&=\sum_{i\geqslant 1}\frac{1}{i}\sum_{\substack{i_L+i_R+i_0+\cdots+i_{k_2}=i-1\\ 0\leqslant i_L,i_R\leqslant k_2\\ 0\leqslant i_0,\ldots,i_{k_2}\leqslant k_1+1}}\frac{f_{k_2,q}^{||}(i_L)f_{k_2,q}^{||}(i_R)}{q^{2k_2}}\prod_{j=0}^{k_2}\frac{f_{k_1,q}^{\perp}(i_j)}{q^{2k_1}}.\end{aligned}$$

which gives us:

Theorem 22: Let $k_1,k_2\geqslant 1$ be some integers, then

$$\mathsf{cap}\left(S^2_{0,k_1,0,k_2}\right)\leqslant\log_2 q-\log_2 e\left(\frac{\Delta_{\mathcal{A}_{\mathrm{H}}}}{q^{k_1+1}}+\frac{\Delta_{\mathcal{A}_{\mathrm{V}}}}{q^{k_2+1}}\right).$$

Numerical results for the lower bound of Theorem 20 and upper bound of Theorem 22 are given in Table II.

TABLE II
LOWER AND UPPER BOUNDS FOR THE q-ARY ASYMMETRIC $(0, k_1, 0, k_2)$-RLL CONSTRAINT
(IN EACH CELL THE UPPER BOUND APPEARS ABOVE THE LOWER BOUND)

	$k_1 \backslash k_2$	1	2	3	4	5
$q=2$	1	0.795511 0.388483	0.834995 0.573388	0.830916 0.641019	0.818966 0.669467	0.808408 0.682350
	2		0.904373 0.758292	0.919605 0.825923	0.919433 0.854371	0.915729 0.867255
	3			0.947949 0.893554	0.956258 0.922002	0.957665 0.934885
	4				0.970467 0.950450	0.975638 0.963333
	5					0.983338 0.976217
	$k_1 \backslash k_2$	1	2	3	4	5
$q=3$	1	1.534531 1.315006	1.547277 1.410812	1.546146 1.437644	1.544725 1.445965	1.543969 1.448657
	2		1.571080 1.506618	1.574573 1.533450	1.575099 1.541771	1.575125 1.544463
	3			1.580017 1.560282	1.581398 1.568603	1.581773 1.571295
	4				1.583175 1.576924	1.583718 1.579616
	5					1.584350 1.582308
	$k_1 \backslash k_2$	1	2	3	4	5
$q=4$	1	1.982025 1.845375	1.987358 1.905042	1.987194 1.918404	1.986981 1.921627	1.986893 1.922423
	2		1.996427 1.964708	1.997701 1.978070	1.997877 1.981293	1.997908 1.982089
	3			1.999131 1.991433	1.999431 1.994655	1.999500 1.995452
	4				1.999774 1.997878	1.999856 1.998674
	5					1.999943 1.999471

V. DISCUSSION

In this work we showed new lower and upper bounds on the multidimensional capacity of $(0, k)$-RLL systems, as well as a new upper bound on the capacity of (d, k)-RLL systems. We examined the rate of convergence to 1 of $\mathsf{cap}(S^D_{0,k})$ as $k \to \infty$. The best asymptotic bounds for the 2-D case were given in [12] (Corollary 3) as follows:

$$\frac{\log_2 e}{2(k+1)2^k} < 1 - \mathsf{cap}(S^2_{0,k}) \leqslant \frac{4\sqrt{2}\log_2 e}{(k+1)2^{k/2}} + \frac{8}{2^k}$$

for sufficiently large k. There are no previously-known bounds for the multidimensional case. In contrast, our results show asymptotically-matching lower and upper bounds giving

$$1 - \mathsf{cap}\left(S^D_{0,k}\right) = \frac{D \log_2 e}{4 \cdot 2^k}(1 + o(1)).$$

We conclude with an interesting comparison of the asymptotes of our new bounds with those of the best previously-known bounds for (d, ∞)-RLL. While $\mathsf{cap}(S^2_{d,\infty})$ converges to 0 as $\frac{\log_2 d}{d}$ (Corollary 4, [12]), just as it does in one dimension, for D-dimensional $(0, k)$-RLL the capacity converges to 1 slower than the 1-D case by a factor of D.

REFERENCES

[1] N. Alon and J. Spencer, *The Probabilistic Method*, 2nd ed. Hoboken, NJ: Wiley, 2000.

[2] R. J. Baxter, "Hard hexagons: Exact solution," *J. Phys. A: Math. Gen.*, vol. 13, pp. L61–L70, 1980.

[3] N. Calkin and H. Wilf, "The number of independent sets in the grid graph," *SIAM J. Discrete Math.*, vol. 11, pp. 54–60, 1998.

[4] K. Censor and T. Etzion, "The positive capacity region of two-dimensional run length constrained channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5128–5140, Nov. 2006.

[5] T. Etzion, "Cascading methods for runlength-limited arrays," *IEEE Trans. Inf. Theory*, vol. 43, no. 1, pp. 319–324, Jan. 1997.

[6] A. P. Godbole, L. K. Potter, and J. K. Sklar, "Improved upper bounds for the reliability of d-dimensional consecutive-k-out-of-n: F systems," *Naval Res. Log.*, vol. 45, no. 2, pp. 219–230, 1998.

[7] S. Halevy, J. Chen, R. M. Roth, P. H. Siegel, and J. K. Wolf, "Improved bit-stuffing bounds on two-dimensional constraints," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 824–838, May 2004.

[8] K. A. S. Immink, P. H. Siegel, and J. K. Wolf, "Codes for digital recorders," *IEEE Trans. Inf. Theory*, vol. 44, no. 10, pp. 2260–2299, Oct. 1998.

[9] H. Ito, A. Kato, Z. Nagy, and K. Zeger, "Zero capacity region of multidimensional run length constraints," *Elect. J. Comb.*, vol. 6, 1999.

[10] S. Janson, "Poisson approximation for large deviations," *Random Struct. Alg.*, vol. 1, pp. 221–230, 1990.

[11] S. Janson, T. Łuczak, and A. Ruciński, "An exponential bound for the probability of nonexistence of a specified subgraph in a random graph," in *Proc. 1987 Random Graphs*, 1990, pp. 73–87.

[12] A. Kato and K. Zeger, "On the capacity of two-dimensional run-length constrained channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1527–1540, Jul. 1999.

[13] W. H. Kautz, "Fibonacci codes for synchronization control," *IEEE Trans. Inf. Theory*, vol. IT-11, no. 2, pp. 284–292, Apr. 1965.

[14] D. J. Kleitman, "Families of non-disjoint subsets," *J. Combin. Theory*, vol. 1, pp. 153–155, 1966.

[15] E. Louidor and B. H. Marcus, "Improved lower bounds on capacities of symmetric 2D constraints using rayleigh quotients," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1624–1639, Apr. 2010.
[16] B. H. Marcus, P. H. Siegel, and J. K. Wolf, "Finite-state modulation codes for data storage," *IEEE J. Sel. Areas Commun.*, vol. 10, no. 1, pp. 5–37, Jan. 1992.
[17] B. H. Marcus, R. M. Roth, and P. H. Siegel, *Constrained Systems and Coding for Recording Channels*, V. S. Pless and W. C. Huffman, Eds. Amsterdam, The Netherlands: Elsevier, 1998.
[18] Z. Nagy and K. Zeger, "Capacity bounds for the 3-D (0, 1) run length limited channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 1030–1033, May 2000.
[19] Z. Nagy and K. Zeger, "Asymptotic capacity of two-dimensional channels with checkerboard constraints," *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2115–2125, Sep. 2003.
[20] E. Ordentlich and R. M. Roth, "Independent sets in regular hypergraphs and multi-dimensional runlength-limited constraints," *SIAM J. Discrete Math.*, vol. 17, no. 4, pp. 615–623, 2004.
[21] M. Roos, "An extension of Janson's inequality," *Random Struct. Alg.*, vol. 8, no. 3, pp. 213–227, 1996.
[22] R. M. Roth, P. H. Siegel, and J. K. Wolf, "Efficient coding schemes for the hard-square model," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1166–1176, Mar. 2001.
[23] M. Schwartz and J. Bruck, "Constrained codes as networks of relations," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 2179–2195, May 2008.
[24] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, Jul. 1948.
[25] A. Sharov and R. M. Roth, "Two-dimensional constrained coding based on tiling," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1800–1807, Apr. 2010.
[26] A. Sharov and R. M. Roth, "Fixed-rate tiling encoders for 2-D constraints," in *Proc. IEEE Int. Symp. Information Theory*, Austin, TX, Jun. 2010, pp. 1208–1212.
[27] I. Tal and R. M. Roth, "Bounds on the rate of 2-D bit-stuffing encoders," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2561–2567, Jun. 2010.
[28] R. Talyansky, "Coding for Two-Dimensional Constraints," M.Sc. thesis, Comput. Sci. Dep., Technion—Israel Inst. Technol., Haifa, Israel, 1997, (in Hebrew).
[29] D. T. Tang and L. R. Bahl, "Block codes for a class of constrained noiseless channels," *Inf. Control*, vol. 17, pp. 436–461, Dec. 1970.
[30] W. Weeks and R. E. Blahut, "The capacity and coding gain of certain checkerboard codes," *IEEE Trans. Inf. Theory*, vol. 44, no. 5, pp. 1193–1203, May 1998.

Moshe Schwartz (M'03–SM'10) was born in Israel in 1975. He received the B.A., M.Sc., and Ph.D. degrees from the Computer Science Department, Technion—Israel Institute of Technology, Haifa, in 1997, 1998, and 2004, respectively.

He was a Fulbright Postdoctoral Researcher in the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, and a Postdoctoral Researcher in the Department of Electrical Engineering, California Institute of Technology, Pasadena. He now holds a position with the Department of Electrical and Computer Engineering, Ben-Gurion University, Israel. His research interests include algebraic coding, combinatorial structures, and digital sequences.

Alexander Vardy (S'88–M'91–SM'94–F'98) was born in Moscow, U.S.S.R., in 1963. He earned the B.Sc. degree (*summa cum laude*) from the Technion, Israel, in 1985, and the Ph.D. degree from the Tel Aviv University in 1991.

During 1985–1990 he was with the Israeli Air Force, where he worked on electronic counter measures systems and algorithms. During 1992–1993 he was a Visiting Scientist at the IBM Almaden Research Center, San Jose, CA. From 1993 to 1998, he was with the University of Illinois at Urbana-Champaign, first as an Assistant Professor then as an Associate Professor. He is now a Professor in the Department of Electrical Engineering, the Department of Computer Science, and the Department of Mathematics, all at the University of California, San Diego (UCSD). While on sabbatical from UCSD, he has held long-term visiting appointments with CNRS, France, the EPFL, Switzerland, and the Technion, Israel. His research interests include error-correcting codes, algebraic and iterative decoding algorithms, lattices and sphere packings, coding for digital media, cryptography and computational complexity theory, and fun math problems.

Prof. Vardy received an IBM Invention Achievement Award in 1993, and NSF Research Initiation and CAREER awards in 1994 and 1995. In 1996, he was appointed Fellow in the Center for Advanced Study at the University of Illinois, and received the Xerox Award for faculty research. In the same year, he became a Fellow of the Packard Foundation. He received the IEEE Information Theory Society Paper Award (jointly with Ralf Koetter) for the year 2004. In 2005, he received the Fulbright Senior Scholar Fellowship, and the Best Paper Award at the IEEE Symposium on Foundations of Computer Science (FOCS). During 1995–1998, he was an Associate Editor for Coding Theory and during 1998–2001, he was the Editor-in-Chief of the IEEE TRANSACTIONS ON INFORMATION THEORY. He was also an Editor for the *SIAM Journal on Discrete Mathematics*. He has been a member of the Board of Governors of the IEEE Information Theory Society from 1998 to 2006, and again starting in 2011.