



## A Model for the Two-dimensional no Isolated Bits Constraint

**Forchhammer, Søren; Laursen, Torben Vaarby**

*Published in:*  
IEEE International Symposium on Information Theory, 2006

*Link to article, DOI:*  
[10.1109/ISIT.2006.261993](https://doi.org/10.1109/ISIT.2006.261993)

*Publication date:*  
2006

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Forchhammer, S., & Laursen, T. V. (2006). A Model for the Two-dimensional no Isolated Bits Constraint. In *IEEE International Symposium on Information Theory, 2006* IEEE. <https://doi.org/10.1109/ISIT.2006.261993>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# A Model for the Two-Dimensional No Isolated Bits Constraint

Søren Forchhammer and Torben V. Laursen  
 Department of Communications, Optics and Materials  
 Technical University of Denmark  
 DK-2800 Kongens Lyngby, Denmark  
 Email: {sf,tv1}@com.dtu.dk

**Abstract**—A stationary model is presented for the two-dimensional (2-D) no isolated bits (n.i.b.) constraint over an extended alphabet defined by the elements within 1 by 2 blocks. This block-wise model is based on a set of sufficient conditions for a Pickard random field (PRF) over an  $m$ -ary alphabet. Iterative techniques are applied as part of determining the model parameters. Given two Markov chains describing a boundary, an algorithm is presented which determines whether a certain PRF consistent with the boundary exists. Iterative scaling is used as part of the algorithm, which also determines the conditional probabilities yielding the maximum entropy for the given boundary description if a solution exists. Optimizing over the parameters for a class of boundaries with certain symmetry properties, an entropy of 0.9156 is achieved for the n.i.b. constraint, providing a lower bound. An algorithm for iterative search for a PRF solution starting from a set of conditional probabilities is also presented.

## I. INTRODUCTION

Constrained coding has found widespread use in optical and magnetic data storage devices [1]. Ideas for new storage systems such as holographic storage have caused interest for two-dimensional (2-D) constrained fields as models of data storage on a *surface*.

It appears to be difficult to find stationary probability distributions for 2-D constrained fields. The classical construction is the Pickard random field [2] based on properties of the probability distribution on  $2 \times 2$  elements. A variation yielding increased entropy was presented for analysis and optimization of bit-stuff coding for the hard-square constraint [3]. The hard-square constraint may be characterized using a neighborhood of  $2 \times 2$  elements. Recently models for constraints defined on larger neighborhoods have been considered [4]. We present a stationary and Markovian model for the no isolated bits (n.i.b.) constraint. The model is based on addressing the n.i.b. constraint by extending the alphabet by considering blocks of  $(1 \times 2)$  elements.

### A. Constrained fields

Two-dimensional fields specified by shift invariant constraints of finite extent  $(N, M)$  over some alphabet are considered. A constraint is defined by a list,  $\mathcal{F}$ , of forbidden configurations each of which is contained within a rectangle of maximum size  $N \times M$  made of symbols from the alphabet. A configuration on an  $n$  by  $m$  rectangle having no forbidden

configurations within the rectangle is called an admissible configuration.

Let  $E(n, m)$  be the set of admissible configurations on an  $n$  by  $m$  rectangle for a given field  $F$ . The capacity (or combinatorial entropy) of  $F$  is defined as

$$C(F) = \lim_{n, m \rightarrow \infty} \frac{\log_2 |E(n, m)|}{nm}. \quad (1)$$

The limit is indeed well-defined ([5], [6]).

We consider the “no isolated bits” (n.i.b.) 2-D constraint characterized by the two forbidden words:

$$\mathcal{F} = \left\{ \begin{array}{cccccc} & 0 & & 1 & & \\ 0 & 1 & 0, & 1 & 0 & 1 \\ & 0 & & 1 & & \end{array} \right\}.$$

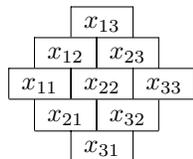
The capacity of the n.i.b. constraint is not known, but in [7] it was estimated to be 0.9238 and lower bounds on the entropy obtained by bit-stuffing were given. For a rectangle to represent the forbidden words, the n.i.b. constraint requires  $3 \times 3$  elements. We characterize this as a higher order constraint as it exceeds the  $2 \times 2$  elements of say the hard square constraint.

### B. Fields on lattices

In order to capture the forbidden words of a higher order constraint such as n.i.b. using a  $2 \times 2$  element model, an alphabet extension is introduced. The extended alphabet  $\mathcal{A}$ , having a finite number  $|\mathcal{A}|$  of elements, represents the distinct configurations of blocks of binary symbols. For the n.i.b. constraint, we shall use blocks of  $1 \times 2$  binary elements.

In addition we will consider configurations on rectangles of lattices rather than just horizontal-vertical rectangles. As an example consider the configuration, on a  $3 \times 3$  rectangle, over  $\mathcal{A}$  in Fig. 1. The entries of the configuration is labelled as (row, column), but here the “rows” and “columns” run along the diagonals.

In the sequel when we refer to an  $n \times m$  rectangle it will be oriented along the diagonal basis vectors of a rotated and scaled coordinate system as in Fig. 1, i.e. the basis vectors are  $e_1 = (1, -1)^T$  and  $e_2 = (1, 1)^T$ . The definition of entropy by (1) still stands in this setting, but now  $n$  and  $m$  refer to coordinates along diagonals as given by the lattice.

Fig. 1. A  $3 \times 3$  rectangle over  $\mathcal{A}$ .

### C. Measures on fields

Let  $F$  be the elements of an  $n \times m$  rectangle. Let  $\mu_F$  be a measure on  $\mathcal{A}^F$  that agrees with the constraint. That is, of all the  $|\mathcal{A}|^{n \times m}$  possible configurations those configurations, that contain forbidden words have probability zero according to  $\mu_F$ . The (measure theoretic) entropy of  $\mu_F$  is defined as

$$H(\mu_F) = -\frac{1}{nm} \sum_{x \in \mathcal{A}^{n \times m}} \mu_F(x) \log_2 \mu_F(x). \quad (2)$$

## II. A STATIONARY MODEL

We consider a Pickard random field [2]. Let  $A, B, C, D$  be random variables over  $\mathcal{A}$  in a  $2 \times 2$  rectangle.

$$\begin{array}{cc} A & B \\ C & D \end{array}$$

A set of conditions on the probability distribution  $(ABCD)$  shall be presented ensuring Markovian and stationary properties of extensions to measures on rectangles of arbitrary size.

Let  $X, Y, Z$  be random variables and let  $X \perp Y \mid Z$  denote that  $X$  and  $Y$  are independent given  $Z$ . The independence conditions  $B \perp C \mid A$  and  $B \perp C \mid D$  shall be assumed for the models considered.

The model is completely specified by the probability distribution on  $(A)$  as well as the three conditional probability distributions  $(B|A)$ ,  $(C|A)$  and  $(D|ABC)$ .

The probabilities of  $(ABCD)$  are expressed by

$$\Pr(ABCD) = \Pr(D|ABC) \Pr(ABC) \quad (3)$$

and due to the condition  $B \perp C \mid A$ ,

$$\Pr(ABC) = \Pr(B|A) \Pr(C|A) \Pr(A). \quad (4)$$

### A. Extending the measure

Given the distribution on the  $2 \times 2$  lattice,  $(ABCD)$ , one can extend this to a measure  $\mu_{n \times m}$  on an  $n \times m$  lattice  $x = (x_{ij})$  in the following manner. First the block  $x_{11}$  is drawn according to  $(A)$ . Then the first row  $x_{12} \dots x_{1m}$  is drawn according to the conditional distribution  $(B|A)$  one block at a time. Then the first column  $x_{21} \dots x_{n1}$  is drawn according to  $(C|A)$  one block at a time.  $x_{22}$  can then be drawn using  $(D|ABC)$ . Proceeding in this manner one has (using shorthand notation for probabilities given by the argument):

$$\begin{aligned} \mu_{n \times m}(x) &= \Pr(x_{11}) \cdot \prod_{j=2}^m \Pr(x_{1j} | x_{1(j-1)}) \\ &\quad \cdot \prod_{i=2}^n \Pr(x_{i1} | x_{(i-1)1}) \\ &\quad \cdot \prod_{i=2}^n \prod_{j=2}^m \Pr(x_{ij} | x_{(i-1)(j-1)}, x_{(i-1)j} x_{i(j-1)}). \end{aligned} \quad (5)$$

The union of the elements of the first row and the first column is called the boundary. The extended measure is *stationary* if the joint distribution of  $(ABCD)$  does not depend on which  $2 \times 2$  rectangle within the  $n \times m$  rectangle we regard.

In order for the measure  $\mu_{2 \times 2}$  to be stationary on the  $2 \times 2$  lattice, it is sufficient (and necessary) that the distributions on  $(AB)$  and  $(CD)$  be identical and the distributions on columns  $(AC)$  and  $(BD)$  be identical. This corresponds to the following two sets of  $|\mathcal{A}|^2$  equations each

$$P_{AB}(A = x, B = y) = P_{CD}(C = x, D = y), \forall (x, y) \in \mathcal{A}^2, \quad (6)$$

$$P_{AC}(A = x, C = y) = P_{BD}(B = x, D = y), \forall (x, y) \in \mathcal{A}^2. \quad (7)$$

The stationarity obviously implies *Condition 1*: The stationary distributions for  $(B|A)$  and  $(C|A)$  must be identical.

Furthermore we have that the probability distributions  $(AB)$ ,  $(AC)$ ,  $(CD)$ ,  $(BD)$ ,  $(D|abc)$  must each sum to 1.

The following Theorem due to Pickard [2] gives a sufficient condition on  $(ABCD)$  for the extended measures (3-5) to be stationary.

**Theorem 1** *Let  $\mu_{2 \times 2}$  be a stationary measure induced by  $(ABCD)$  satisfying  $B \perp C \mid A$ . If  $B \perp C \mid D$  then the extended measure  $\mu_{n \times m}$  based on (3-5) is Markovian and stationary for any  $n, m \geq 2$ .*

Theorem 1 provides sufficient conditions for the measure  $\mu_{2 \times 2}$  to be extended to a stationary measure. Since

$$\begin{aligned} B \perp C \mid D &\Leftrightarrow \\ \Pr(BCD) &= \Pr(D) \Pr(B|D) \Pr(C|D). \end{aligned} \quad (8)$$

Assuming stationarity (6-7), the right hand terms are readily derived from  $(ABC)$ .

It is a necessary condition for  $(ABC)$  and  $(BCD)$  to be marginal distributions of  $(ABCD)$ , that their marginal distributions on  $(BC)$  be identical. This is expressed by *Condition 2*:

$$\sum_{a \in \mathcal{A}} \Pr(A = a, bc) = \sum_{d \in \mathcal{A}} \Pr(bc, D = d), \forall (b, c) \in \mathcal{A}^2. \quad (9)$$

Further it follows that

**Theorem 2** *The entropy per symbol of a stationary measure  $\mu_{n \times m}$  defined by Theorem 1 and given by (2) is bounded by*

$$H(\mu_F) \geq H(D|ABC). \quad (10)$$

*Proof:* The  $n \times m$  rectangle, of the measure  $\mu_{n \times m}$ , is divided into the initial boundary  $x_{1i}, 1 \leq i \leq n$  and  $x_{1j}, 1 \leq j \leq m$  and the remaining interior part. The entropy of each element of the latter is given by  $H(D|ABC)$  due to the stationarity and the chain rule. The entropy of the initial boundary is not less than this as the distribution on the initial boundary is given by  $(A)$  and the Markov chains  $(B|A)$  and  $(C|A)$ , which are identical to the marginal distribution on the rows (and columns) of the remaining interior, due to the stationarity. ■

### III. ITERATIVE TECHNIQUES FOR STATIONARY SOLUTIONS

The stationarity conditions (6-7) and the independence condition (8) provide sufficient conditions for a probability distribution  $(ABCD)$  described by (3-4) to satisfy Theorem 1. Different approaches to determine parameters of the model satisfying (6-7) and (8) are considered.

#### A. Algorithm based on iterative scaling

Given a boundary description in terms of the distribution  $(ABC)$ , iterative scaling may be used to find the optimal conditional probabilities  $\Pr(D|ABC)$ .

For each configuration  $B = b, C = c$ , consider  $\Pr(AD|bc)$ . The distribution  $(ABC)$  determines  $\Pr(A|bc)$ ,

$$\alpha_i \equiv \Pr(A = i|bc) = \Pr(A = i, bc) / \sum_{j \in \mathcal{A}} \Pr(A = j, bc), i \in \mathcal{A}. \quad (11)$$

The distribution  $(ABC)$ , the stationarity (6-7), and the independence  $B \perp C | D$  (8) determines  $(BCD)$ , which in turn determines  $\Pr(D|bc)$ ,

$$\beta_j \equiv \Pr(D = j|bc) = \Pr(bc, D = j) / \sum_{i \in \mathcal{A}} \Pr(bc, D = i), j \in \mathcal{A}. \quad (12)$$

The probabilities of  $P(AD|bc)$ , must satisfy the linear relations due to (11-12):

$$\sum_{j \in \mathcal{A}} \Pr(A = i, D = j|bc) = \alpha_i, i \in \mathcal{A}. \quad (13)$$

$$\sum_{i \in \mathcal{A}} \Pr(A = i, D = j|bc) = \beta_j, j \in \mathcal{A}. \quad (14)$$

Thus we seek a solution in the intersection,  $\mathcal{L}$ , of the two linear families defined by (13) and (14), respectively, each determined by a partition (which may be described as row- and column-sums of a matrix given by elements  $p_{bc}(i, j) = \Pr(A = i, D = j|bc)$ ). (We note that  $p_{bc}(i, j) = 0$  if for  $a = i$  and  $b = j$ ,  $abcd$  is not an admissible configuration.)

Iterative scaling [8] may take an initial distribution,  $Q(x)$  and a class of distributions as  $\mathcal{L}$  and find a distribution,  $P^*$  which minimizes the divergence,  $D(P^*||Q)$  for  $P^* \in \mathcal{L}$ , where the divergence is given by  $D(P||Q) \equiv \sum P(x) \log(P(x)/Q(x))$ .

This leads to the scalings defining new probabilities,  $p_{bc}^*(i, j)$ , within each of the families,

$$p_{bc}^*(i, j) = c p_{bc}(i, j), c = \alpha_i / \sum_{i \in \mathcal{A}} p_{bc}(i, j), j \in \mathcal{A}. \quad (15)$$

$$p_{bc}^*(i, j) = d p_{bc}(i, j), d = \beta_j / \sum_{j \in \mathcal{A}} p_{bc}(i, j), i \in \mathcal{A}. \quad (16)$$

*Maximum entropy iterative scaling* of  $P(AD|bc)$  is defined by (15) and (16) with the initial distribution  $Q(AD|bc)$  set to a uniform distribution over the admissible configurations  $abcd$  for each  $bc$ . For each  $bc$  a sequence of distributions is generated by iterating (15) and (16). The entropy of the interior,  $H(D|ABC)$ , may be expressed by

$$H(D|ABC) = H(ABCD) - H(ABC) = H(BC) + H(AD|BC) - H(ABC). \quad (17)$$

**Theorem 3** Consider a distribution  $(ABC)$  satisfying  $B \perp C | A$  (4). Conditions 1 and 2 are necessary conditions for  $(ABC)$  to form the boundary of a PRF of the form in Theorem 1. If (4) and Conditions 1 and 2 are satisfied, maximum entropy iterative scaling of  $P(AD|bc), \forall (b, c)$ , determines whether such a PRF exists and if so determines the conditional probabilities,  $P(D|ABC)$ , which maximizes the entropy given the boundary distribution,  $(ABC)$ .

*Proof:* Condition 1 is implied by  $B \perp C | A$  (4) and the stationarity (6-7). Condition 2 is implied by  $B \perp C | D$  and the stationarity (6-7). Iterative scaling converges to a solution to (13-14) if it exists and this solution minimizes the divergence [8]. Starting with uniform distribution  $Q$ , implies that minimizing the divergence,  $D(P||Q)$ , maximizes the entropy  $H(P), P \in \mathcal{L}$ . The iterative scaling converges to a solution, if it exists, which defines a distribution  $(ABCD)$ , which provides stationary solutions for  $(ABC)$  and  $(BCD)$  satisfying the conditions of Theorem 1. Thus if a solution exists, it is a PRF in accordance with Theorem 1. As  $(ABC)$  is given, maximizing  $H(AD|BC)$  by the iterative scaling, maximizes  $H(D|ABC)$  (17). ■

#### B. An iterative search for a stationary solution

In this subsection, two iterative procedures are described as a method to search for a solution. Here the starting point shall be the conditional probabilities,  $P(D|ABC)$  instead of the boundary.

The first algorithm is part of the second and this combination provides a search for stationary solutions. The first procedure changes the boundary distributions  $(B|A)$  and  $(C|A)$  until  $(ABCD)$  is stationary. The second procedure extends this by changing  $(D|ABC)$  until  $B \perp C | D$ .

Let  $Q_B = (q_{ij})$  denote the  $|\mathcal{A}| \times |\mathcal{A}|$  transfer probability matrix for  $(B|A)$ , that is  $q_{ij} = \Pr(B = j|A = i)$ . Let  $\pi_A$  denote the stationary distribution for  $Q_B$ . We will use  $\pi_A$  as the distribution of  $(A)$ . For symmetry reasons, for n.i.b., we set the boundary distributions  $(B|A)$  and  $(C|A)$  to be identical. This way  $(B|A)$  and  $(C|A)$  automatically have the same stationary distribution,  $\pi_A$ , which is used for the distribution  $(A)$ . Let  $Q_D = (q_{ij})$  denote the  $|\mathcal{A}|^3 \times |\mathcal{A}|$  transfer probability matrix for  $(D|ABC)$ . Let  $Q_{BD}$  and  $Q_{CD}$  denote the  $|\mathcal{A}|^2 \times |\mathcal{A}|^2$  transfer probability matrices for the distributions  $(BD|AC)$  and  $(CD|AB)$ , respectively.

Algorithm 4 iterates to satisfy (6-7). Algorithm 5 iterates till (8) is also satisfied calling Algorithm 4 in each iteration.

In general, let  $P_X$  denote the probability vector of the marginal distribution  $X$  of the joint distribution  $(ABCD)$ . Let a superscript index the iterations, e.g.  $Q_B$  is initialized by  $Q_B^{(0)}$  and after  $n$  iterations  $Q_B^{(n)}$  is output. We set  $Q_C^{(n)} = Q_B^{(n)}$ .

**Algorithm 4** Assume that  $Q_B^{(0)}$  and  $Q_D$  are given.

- 1) Calculate  $\pi_A^{(n)}$  from  $Q_B^{(n-1)}$ . Calculate  $P_{ABCD}^{(n)}$  (3-4).
- 2) Calculate  $P_{AB}, P_{CD}, P_{AC}, P_{BD}$  from  $P_{ABCD}^{(n)}$ .
- 3) If  $\|P_{AB} - P_{CD}\| + \|P_{AC} - P_{BD}\| < \epsilon$  then goto 9.
- 4) Calculate  $Q_{CD}^{(n)}$  and  $Q_{BD}^{(n)}$  from  $P_{ABCD}^{(n)}$ .

- 5) Find the stationary distribution  $\pi_{AB}^{(n)}$  for  $Q_{CD}^{(n)}$  and  $\pi_{AC}^{(n)}$  for  $Q_{BD}^{(n)}$ .
- 6) Let  $\pi^{(n)}$  be the average of  $\pi_{AB}^{(n)}$  and  $\pi_{AC}^{(n)}$ .
- 7) Calculate  $Q_B^{(n)}$  from  $P_{AB}^{(n)} = \pi^{(n)}$ .
- 8) Goto 1.
- 9) Output  $Q_B^{(n)}$  and  $\pi_A^{(n+1)}$ .

Step 3 performs the test for the conditions (6-7) of stationarity on  $(ABCD)$ ,  $\epsilon$  controls when to stop. Symmetry expressed by  $Q_{BD} = Q_{CD}$  is given by the following conditions:  $Q_B$  and  $Q_C$  are identical and  $Q_D$  obeys the following symmetry constraint

$$\forall a, b, c \in \mathcal{A} : \Pr(D|abc) = \Pr(D|acb). \quad (18)$$

This symmetry and  $Q_B (= Q_C)$  being irreducible are beneficial towards the convergence.

Now we extend the procedure to search for the conditional independence  $B \perp C | D$ . The main condition is (8). We will write this in a slightly different form:

$$\sum_{x \in \mathcal{A}} \Pr(A = x, BCD) = \Pr(D) \Pr(B|D) \Pr(C|D). \quad (19)$$

Inspired by (19) we introduce a scale parameter for each context  $(bcd)$ . Define

$$\lambda_{bcd} = \frac{\Pr(D = d) \Pr(B = b|D = d) \Pr(C = c|D = d)}{P_{BCD}(bcd)}. \quad (20)$$

For each  $bcd$  configuration the two sides of (8) are calculated. Temporary parameters for the conditional probabilities of  $Q_D$  involved are modified by the scale parameter (20) to achieve equality for the given  $bcd$  configuration. This leads to the following algorithm.

**Algorithm 5** Given  $Q_B$  and  $Q_D^{(0)}$ .

- 1) Run Algorithm 4 until convergence.
- 2) Calculate  $P_{BCD}$  and  $P_t = P_D Q_{B|D} Q_{C|D}$
- 3) If  $\|P_{BCD} - P_t\| < \epsilon$  Goto 8.
- 4)  $\forall bcd \in \mathcal{A}^3 : \lambda_{bcd} = P_t / P_{BCD}(bcd)$ .
- 5)  $\forall abcd \in \mathcal{A}^4 : Q_D^{(n)}(d|abc) = \lambda_{bcd} Q_D^{(n-1)}(d|abc)$ .
- 6) Normalize  $Q_D^{(n)}$ .
- 7) Goto 1.
- 8) Output  $Q_D^{(n)}$ .

In Step 2, the terms are calculated from  $P_{ABCD}$  (3-4). In Step 5, the update of  $Q_D$  is given by considering the terms contributing to the difference,  $\|P_{BCD} - P_t\|$  in Step 3 (20). After one pass over the configurations, these parameters are normalized (Step 6) to define a new set of conditional probabilities  $Q_D^{(n)}$  appropriately summing to 1.

The search given by Algorithm 5 calls Algorithm 4 in its iterations, proceeding until both have converged. A solution satisfies the constraints of Algorithms 4 and 5, and thus (6-7) and (8).

#### IV. MODELS FOR N.I.B.

Consider the higher order alphabet,  $\mathcal{A}$ , over the binary alphabet given by  $\mathcal{A} = \{00, 01, 10, 11\}$ . In Fig. 2, an example of a  $3 \times 3$  rectangle over  $\mathcal{A}$  is shown.

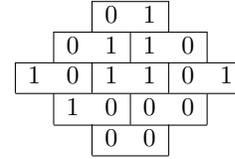


Fig. 2. An example of a n.i.b. configuration on a  $3 \times 3$  field.

Let  $A, B, C, D$  be random variables over  $\mathcal{A}$  in the  $2 \times 2$  rectangle along diagonals.

$$\begin{array}{ccc} & B & \\ A & & D \\ & C & \end{array}$$

##### A. A n.i.b. model based on iterative scaling

To define distributions on  $(ABC)$  which satisfy Conditions 1 and 2, three symmetries are introduced to reduce the parameter space: *S1*) The distributions on  $(AB)$  and  $(AC)$  are identical, thus only  $(AB)$  needs to be specified. *S2*) Let  $\bar{a}$  denote the configuration given by bitwise inverting the bits of configuration  $a$ . The probabilities of  $AB$  configurations are identical under bitwise inversion, i.e.  $\Pr(A = a, B = b) = \Pr(A = \bar{a}, B = \bar{b})$ . *S3*) The probabilities of  $AB$  configurations are identical under rotation, eg.  $P_{AB}(00, 01) = P_{AB}(10, 00)$ . Rewriting the latter equation gives

$$\Pr(01|00) \Pr(00) = \Pr(00|10) \Pr(10). \quad (21)$$

The symmetry under inversion of bits, *S2*, implies that  $\Pr(00) = \Pr(11)$  and  $\Pr(01) = \Pr(10) = (1 - \Pr(00))/2$ .

Four parameters are chosen:  $\Pr(00)$ ,  $\Pr(00|00)$ ,  $\Pr(00|01)$ , and  $\Pr(00|10)$ . Given these four parameters the rest of the parameters of  $\Pr(B|A)$  are determined by the stationarity of  $(B|A)$ ,  $\sum_b \Pr(B = b|a) = 1$ , the *S2* symmetry, (21), and one equation derived from (9) for  $bc = 00, 00$ :

$$\Pr(01|01) = c(\Pr(01|00) - \Pr(01|11)) + 1 - \Pr(00|01) - \Pr(11|01))/2, \quad (22)$$

$$c = [\Pr(00)(\Pr(00|00) - \Pr(00|11))]/[\Pr(01)(\Pr(00|10) - \Pr(00|01))].$$

By construction Condition 1 is satisfied. By inspection it is verified that Condition 2 is satisfied.

##### B. Using a bit-stuffer to induce $(D|ABC)$

Algorithm 5 was used to obtain a stationary solution for this model. First  $(B|A)$  was initialized to be an unbiased distribution,  $Q_B = (q_{ij})$  with  $q_{ij} = 1/16$  for all  $i, j \in \mathcal{A}$ . Then a bit-stuffer was used to determine  $Q_D$ , i.e.  $(D|ABC)$

as follows. The bit-stuffer works on the binary symbol level in order  $d_1, d_2$ .

$$\begin{array}{cc} & \begin{array}{cc} b_1 & b_2 \end{array} \\ \begin{array}{cc} a_1 & a_2 \end{array} & \boxed{\begin{array}{cc} d_1 & d_2 \end{array}} \\ & \begin{array}{cc} c_1 & c_2 \end{array} \end{array}$$

Whether bits are stuffed or written as is to positions  $d_1$  and  $d_2$  is decided solely in order not to violate the n.i.b. constraint. Based on the bit-stuffer one can induce the probabilities  $(D|ABC)$  at the block level. As an alternative a 'symbol-stuffer' may be applied. We consider introducing a bias for  $d_1$  given  $abc$ , such that the distribution on the admissible values of the symbol  $d = d_1, d_2$  is uniform. It may be noted that  $\Pr(D|abc) = \Pr(D|acb)$ .

### C. Numerical results for n.i.b.

1) *Results using iterative scaling:* The algorithm based on maximum entropy iterative scaling with the additional symmetry properties, *S1-S3* was applied to the n.i.b. constraint. Searching over the four free parameters gave a maximum value of the entropy (10) expressed by  $H(D|ABC) = 0.9156$  bits per binary symbol for the parameter values specifying the distribution  $(AB)$ :  $\Pr(A = 00) = 0.2798$ , and for  $\Pr(B|A)$ :  $\Pr(00|00) = 0.2300$ ,  $\Pr(00|01) = 0.2410$ , and  $\Pr(00|10) = 0.3070$ . This entropy is also a lower bound for the n.i.b. constraint. It is slightly higher than the lower bound of 0.9127 presented as a lower bound for bit-stuffing in [7].

2) *Results using iterative search:* Applying Algorithm 5 starting with the model with conditional probabilities induced by unbiased bit-stuffing, we obtained a stationary solution in accordance with Theorem 1, having the entropy (10) expressed by  $H(D|ABC) = 0.9037$  bits. Thereafter the symbol-stuffer above was used to initialize the conditional probabilities  $Q_D$ . This gave a stationary solution with an entropy of  $H(D|ABC) = 0.9073$  bits.

3) *Searching for a higher entropy:* Now taking our point of departure in a model derived by iterative searching, the parameters of this model may be perturbed slightly in order to search for a model with higher entropy.

To keep things simple we insist that the boundary distributions should still be identical. Furthermore the conditional distributions  $(D|ABC)$  should have the same symmetry as the bit-stuffing induced probabilities, that is  $\Pr(D|abc) = \Pr(D|acb)$ .

Let  $\delta > 0$ . For each context  $(abc)$  there are three probability parameters (of which 0, 1 or 2 may be 0):  $p_{abc} = \Pr(D = 00|abc)$ ,  $q_{abc} = \Pr(D = 01|abc)$ , and  $r_{abc} = \Pr(D = 10|abc)$ . For each context  $(abc)$  and for each parameter  $p \in \{p_{abc}, q_{abc}, r_{abc}\}$ , we perturb the parameter  $p = p + \delta$ , ensure that conditional probabilities  $\Pr(D|abc)$  still sum to 1 and then iterate the model towards stationarity according to the method described in Section III-B.

It should be noted that only a limited search has been carried out, from the two initial distributions, to demonstrate that increasing the entropy by a simple search is possible.

Starting from the distribution induced by the unbiased bit-stuffer, the model was modified by the searching method

described. This gave a stationary model having an entropy of  $H(D|ABC) = 0.9041$  bits per binary symbol. Starting from the symbol-stuffer an entropy of  $H(D|ABC) = 0.9082$  bits was obtained. For illustration, the parameters of  $(B|A)$  of the latter are given in Table I.

TABLE I

A MODEL FOR A N.I.B. BOUNDARY DISTRIBUTION  $(B|A)$ .

A	B			
	00	01	10	11
00	0.2351	0.2367	0.2224	0.3058
01	0.2574	0.1962	0.2735	0.2729
10	0.2796	0.2718	0.1919	0.2566
11	0.3073	0.2194	0.2409	0.2325

## V. CONCLUSION AND DISCUSSION

A stationary model for the no isolated bits constraint was presented. It is based on a Pickard random field, defined on blocks and thereby an extended alphabet, and applied to the higher order n.i.b. constraint. Sufficient conditions for stationarity and a Markovian property were stated based on the work on Pickard random fields. An expression for calculating the entropy was given. Starting from the distribution on the boundary, an algorithm based on iterative scaling was given which could decide whether a PRF with the given boundary could be specified and if so the iterative scaling could determine the conditional probabilities which would maximize the entropy. Applied to the n.i.b. constraint and optimized over the free parameters, an entropy of 0.9156 was obtained. It was also demonstrated that valid solutions to the set of sufficient conditions could be obtained by an iterative procedures starting from a set of conditional probabilities.

### ACKNOWLEDGMENT

The authors would like to thank Jørn Justesen for introducing iterative scaling to the problem.

### REFERENCES

- [1] K. A. S. Immink, P. H. Siegel, and J. K. Wolf, "Codes for digital recorders," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2260–2299, 1998.
- [2] D. Pickard, "Unilateral markov fields," *Adv. Appl. Probability*, vol. 12, pp. 655–671, 1980.
- [3] R. Roth, P. Siegel, and J. Wolf, "Efficient coding schemes for the hard-square model," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1166–76, 2001.
- [4] J. Justesen, "Finite state models of constrained 2D data," in *Proc. ISIT'04 Chicago*, p. 160, IEEE, 2004.
- [5] S. Friedland, "On the entropy of  $Z^d$  subshifts of finite type," *Linear algebra and its applications*, pp. 199–220, 1997.
- [6] A. Kato and K. Zeger, "On the capacity of two-dimensional run-length constrained channels," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1527–1540, 1999.
- [7] S. Halevy, J. Chen, R. Roth, P. Siegel, and J. Wolf, "Improved bit-stuffing bounds on two-dimensional constraints," *IEEE Transactions on Information Theory*, vol. 50, no. 5, pp. 824–38, 2004.
- [8] I. Csiszár and P.C. Shields, "Information theory and statistics: A tutorial," in *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 4, pp. 422–527, 2004.