



Maximizing Entropy of Pickard Random Fields for 2x2 Binary Constraints

Søgaard, Jacob; Forchhammer, Søren

Published in:

Proceedings of 2014 IEEE International Symposium on Information Theory

Link to article, DOI:

[10.1109/ISIT.2014.6875188](https://doi.org/10.1109/ISIT.2014.6875188)

Publication date:

2014

Document Version

Early version, also known as pre-print

[Link back to DTU Orbit](#)

Citation (APA):

Søgaard, J., & Forchhammer, S. (2014). Maximizing Entropy of Pickard Random Fields for 2x2 Binary Constraints. In *Proceedings of 2014 IEEE International Symposium on Information Theory* (pp. 2022-2026). IEEE. <https://doi.org/10.1109/ISIT.2014.6875188>

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.

Maximizing Entropy of Pickard Random Fields for 2x2 Binary Constraints

Jacob Søgaard and Søren Forchhammer

DTU Photonics

Technical University of Denmark

Kgs. Lyngby, Denmark

Abstract—This paper considers the problem of maximizing the entropy of two-dimensional (2D) Pickard Random Fields (PRF) subject to constraints. We consider binary Pickard Random Fields, which provides a 2D causal finite context model and use it to define stationary probabilities for 2x2 squares, thus allowing us to calculate the entropy of the field. All possible binary 2x2 constraints are considered and all constraints are categorized into groups according to their properties. For constraints which can be modeled by a PRF approach and with positive entropy, we characterize and provide statistics of the maximum PRF entropy. As examples, we consider the well known hard square constraint along with a few other constraints.

I. INTRODUCTION

We consider the problem of determining and maximizing the entropy of image models with constraints. In applications such as 2D data storage, the encoding of information might be subject to a constraint related to the physical properties of the media. This can lead to complex data structures and data sources with memory. If such data sources are constrained, the possible patterns will be limited and the per symbol information capacity of the source will be reduced, but the density may be increased. In the 2x2 binary setting the problem can be regarded as filling a $N \times M$ 2D grid with binary symbols with maximum information. Under these circumstances, we seek the maximum entropy and the corresponding coding model to fill the grid without violating the given constraint. For block Pickard models of higher order see [1].

The capacity of the binary source in terms of how much information that can be stored per symbol in the grid is defined by the combinatorial entropy C :

$$C = \lim_{N \rightarrow \infty} \log_2 \frac{E(N)}{N^2} \quad (1)$$

where N is the width of a square grid with $E(N)$ admissible outcomes [2, p. 122]. The maximum entropy of the Pickard Random Field (PRF) [3] subject to a constraint provides a causal model which provides a lower bound on the capacity of the constraint. The PRF also provides a probability distribution of the symbols that can be used for coding purposes.

We formulate the problem as follows; given a constraint in form of a set of binary 2x2 non-admissible configurations, decide if a PRF can be defined that does not have a non-admissible configuration as any subset of any possible outcome. If PRF can be defined, determine the maximum entropy of the PRF field and if the entropy of the field is greater than 0,

determine the parameters of the PRF. In this setting, we define *invalid* constraints as constraints where there is a possibility of reaching a configuration that has no solution when filling the grid. A constraint which is not invalid is considered to be valid. We consider this problem for all 2^{16} cases for 2x2 binary constraints and provide statistics for the cases with field entropy greater than 0. We also provide detailed examples.

II. PICKARD RANDOM FIELDS

PRF is defined as a 2D causal probabilistic model where the probability assignments for rows and columns share the property that the symbols are outcomes of the same irreducible Markov chain over a finite alphabet [3]. Our aim is to find a probability distribution for the symbols on a grid. If X denotes a random variable describing the field, we want to maximize the entropy $H(X)$. The entropy on a $N \times M$ rectangular grid can be expressed by the chain rule:

$$H(X) = \sum_j H(x_j | x_1, x_2, \dots, x_{j-1}) \quad (2)$$

where each symbol in the grid is assigned a one-dimensional index j e.g. in row-by-row order. However, in a 2D structure which can be expressed by a Markov Random Field (MRF), non-causal interaction might exist. Therefore we turn to PRFs, which provide us with causal 2D models in this environment.

PRFs are special cases of MRFs. PRFs have some desirable properties such as causality, sequential simulation of a random field and the property that the symbols in rows and columns are outcomes of irreducible Markov chains over a finite alphabet. A PRF is stationary and its joint probability can be determined by a 2x2 measure which must satisfy a constraint. The probability of each column or row is described by a Markov chain and transition probabilities which can be derived from the 2x2 measure. We assume the Markov chains for the rows to go from left to right and for the columns to go from bottom up. The PRF starting point is in the upper left corner. The Markov chains for rows and columns can be different.

We denote the 2x2 square consisting of random variables $X_{i,j}$ from a finite alphabet on the 2D grid as follows:

$$\begin{bmatrix} X_{i,j} & X_{i,j+1} \\ X_{i+1,j} & X_{i+1,j+1} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \quad (3)$$

After initializing the first row and column in the grid by using the Markov chains, the 2x2 square (3) can be moved one

position at a time and the symbol D can be coded using the conditional probability $P(D|ABC)$. The probabilities of the joint distribution $(ABCD)$ are expressed by:

$$P(ABCD) = P(D|ABC)P(ABC) \quad (4)$$

Since the contribution to the entropy will be dominated by $H(D|ABC)$ asymptotically [1], we seek to maximize this measure of field entropy thus maximizing the amount of information that can be stored in the grid by this approach.

A. Independence conditions

The Markov chain for columns defines the conditional probability distribution of the pair of symbols A, C . Likewise, the Markov chain for rows defines the conditional probability distribution of the pair of symbols A, B . An important assumption is that $P(B|AC) = P(B|A)$ which leads to [3]:

$$P(ABC) = P(B|A)P(A|C)P(C) \quad (5)$$

With this assumption the top row of the grid can be modeled as a single Markov chain by the distribution of (AB) . We use the condition $P(C|BD) = P(C|D)$ to assure that the other rows are described by the same Markov chain [3]:

$$P(BCD) = P(B|D)P(D|C)P(C) \quad (6)$$

The starting point can also be the lower left corner [1], [3], in which case (5-6) would be defined along the other diagonal.

B. Consistency conditions

We denote the matrices for the Markov chains of the rows and columns R and S , respectively. For consistency with the joint measure $(ABCD)$ the two distributions (ABC) and (BCD) must have identical marginal distribution on (BC) . Since the transition from C to B can be achieved in two ways (i.e. either by going right and then up or vice versa) due to (5-6) the transition matrices must commute [1]:

$$RS = SR \quad (7)$$

The last requirement is that the joint conditional probability $P(AD|bc)$ should be consistent with the marginal conditional probabilities $P(A|bc)$ (5) and $P(D|bc)$ (6), which gives us [1]:

$$\sum_d P(A, D = d|bc) = P(A|bc) \quad (8)$$

$$\sum_a P(A = a, D|bc) = P(D|bc) \quad (9)$$

The entropy of the PRF will converge to $H(D|ABC)$ since the interior of the field will dominate over the boundaries. In the general case this can be solved by iterative scaling [1].

$$H(D|ABC) = - \sum_{abcd} P(abcd) \log_2 P(d|abc) \quad (10)$$

For values (bc) of (BC) where $(abcd)$ is admissible for all (ad) , the maximum entropy is simply obtained by

$$P(D|Abc) = P(D|bc) = P(bcD)/P(bc) \quad (11)$$

C. Probabilistic aspects of binary PRF

We solve the PRF in general using (7-9) over the binary alphabet. With a binary alphabet the transition matrices R, S for rows and columns, respectively, can be defined as:

$$R = \begin{bmatrix} p & 1-p \\ 1-s & s \end{bmatrix} \quad S = \begin{bmatrix} q & 1-q \\ 1-r & r \end{bmatrix} \quad (12)$$

Since the matrices R and S have to commute (7) we can set each element in RS equal to the corresponding element in SR . Solving any of these equations gives us the following relation between the variables:

$$r + p - rp = s + q - sq \quad (13)$$

This relation can be used to express one of the variables by the other variables. This ensures that the commutation requirement (7) is fulfilled and it also reduces the number of free variables by one. Under certain constraints with positive entropy, one or two of the variables p, s, q, r can be set to 0 (but not 1) reducing the complexity of the entropy maximization; e.g. for the straightforward formulation of the hard-square constraint [4] one would set $r = s = 0$ and would only need to determine the value for $p = q$ that maximizes the entropy.

The stationary probabilities $p_r^* = p_r^* R$ and $p_s^* = p_s^* S$ of the two Markov chains should be equal. In the binary case this leads to the same relation as (13). Thus, in a 2x2 binary PRF the equality condition for the stationary probabilities is equivalent to the commutation condition (7).

Some requirements follows from this formulation of the problem. The variables p, s, q, r , including the one expressed by the relation in (13), should be constrained to the range $[0, 1]$. Let \tilde{a} be the inverted value of the binary symbol a .

By definition it is required that $(A, B = b, C = c, D)$ configurations for pairs of (bc) where $A = a, D = d$ is invalid, but $(AbcD)$ still contributes to the entropy should satisfy:

$$P(bcd) \leq P(\tilde{a}bc) \quad (14)$$

This inequality can be used to bound our choice of starting points for the free variables before applying maximization.

D. Calculating and maximizing the entropy

We characterize all possible contributions to the maximum entropy in a binary 2x2 PRF. The contributions to the entropy (10) can be calculated as follows. For pairs of (bc) for which all combinations are admissible following (11) the contribution to the maximum PRF entropy can be expressed by:

$$H_{bc}(D|ABC) = P(bc)H\left(\frac{P(bcd)}{P(bc)}\right) \quad (15)$$

For pairs of (bc) where one of the combinations of AD , $A = a, D = d$, is non-admissible we may solve (8-9) directly and the entropy contribution is calculated as:

$$H_{bc}(D|ABC) = P(A = \tilde{a}, bc)H\left(\frac{P(bc, D = d)}{P(A = \tilde{a}, bc)}\right) \quad (16)$$

as $P(\tilde{a}bcd) = P(bcd)$ and for abc the entropy contribution is 0. For pairs of (bc) where $A = a$ is non-admissible i.e. for

both $D = 0$ and $D = 1$, we get $P(\tilde{a}bcd) = P(bcd)$, $P(\tilde{a}bc) = P(bc)$ so (15) can be used again.

In other cases for the pairs of (bc) there is no contribution to the entropy. Using (5-6) the contributions in (15-16) can be calculated directly from the Markov chains R and S .

Finally the maximum conditional entropy is calculated as the sum of the non-zero contributions given by (15) and (16):

$$H(D|ABC) = \sum_{bc} H_{bc}(D|ABC) \quad (17)$$

For given parameters p, s, q, r this provides the maximum PRF entropy of the constraint. Depending on the constraint there will be 1 to 3 free parameters. We can maximize the conditional entropy (17) over the space given by the free parameters and restricted by (13) and (14).

III. CONSTRAINT ANALYSIS

Our goal is to analyze all 2^{16} possible 2×2 binary constraints and identify those for which we can define a PRF with positive entropy, i.e. in (10) $H(D|ABC) > 0$. To achieve this we can start by identifying constraint sets for which it is not possible to define a PRF with positive probabilities of all admissible $(abcd)$. We will take a closer look at what can be considered to be degenerate cases and why they occur.

We start out by introducing the concept of a constraint *family*. Some constraints can be characterized by having one or more configurations which can not appear in the PRF, but removing these provides a subset F for which all admissible configurations appear. We call this set the father set and the sets with identical fathers form a family. Let i, j be indexes, then a constraint c_i belongs to the same family as the constraint c_j if F of the non-admissible configurations in c_i is equal to F of the non-admissible configurations in c_j :

$$F_i = F_j \quad (18)$$

In other words, the set of non-admissible configurations N_i for the constraint c_i is a subset of F_i :

$$N_i \subseteq F_i \quad (19)$$

Inverting the binary symbols of a configuration may change the values but we consider the structure to be the same and extend the family definition such that the members of a family have F or the inverted set F' as their father set.

A. Analysis of configurations

We consider whether a PRF with positive entropy can be achieved for different categories of constraints and from the categories we derive necessary conditions. The categories are presented in the order which they are used to classify the constraints. Classified constraints are not considered again, thus the categories are disjoint. The first type of classification can be done solely on the constraints without considering Markov chains. They are denoted B_k for bit-stuffing, where k is an index. We define bit-stuffing as a PRF without imposing (5-9) and where all admissible configurations are possible.

Let N_i be the set of non-admissible configurations and the compliment set A_i the set of admissible configurations defining the constraint c_i . Let N_i^* be the union of N_i and the subset of configurations $(\hat{a}\hat{b}\hat{c}\hat{d})$ in A_i where:

$$\begin{aligned} \hat{c}\hat{d} &= ab && \text{if } abcd \in N_i, \forall ab \\ \hat{a}\hat{b} &= cd && \text{if } abcd \in N_i, \forall cd \\ \hat{b}\hat{d} &= ac && \text{if } abcd \in N_i, \forall ac \\ \hat{a}\hat{c} &= bd && \text{if } abcd \in N_i, \forall bd \end{aligned} \quad (20)$$

Let A_i^* be the complementary set of N_i^* . Repeat this process with N_i^* and A_i^* and denote the resulting sets N_i^{**} and A_i^{**} . A constraint c_i is considered to contain a *trap* if:

$$\begin{aligned} (abc, D = 0) \wedge (abc, D = 1) &\in N_i^{**} \\ \wedge \exists a'b' : (a'b', C = a, D = b) &\in A_i^{**} \\ \wedge \exists a'c' : (a'c', B = a, D = c) &\in A_i^{**} \end{aligned} \quad (21)$$

We consider constraints to be invalid if they contain a trap since there is a non-zero probability of reaching the trapping configuration. This set of constraints is denoted B_1 .

B_2 is defined as constraints which forces one or more parameters p, r, s, q in (12) to 1, in which case the underlying Markov chain only has one state or is an *absorbing Markov chain*. In the latter case, the entropy of the field degenerates to 0, since the absorbing Markov state cannot be left. This can be determined by looking at the rows or columns where this occurs separately i.e. either 0 or 1 can never change in one direction. Disregarding the two uniform cases, the constraints in this category are non-stationarity.

B_3 is defined by the constraint forcing one of the following pairs of parameters to {0,0}:

$$\{p, s\}, \{q, r\}, \{p, r\}, \{s, q\} \quad (22)$$

or by all four parameters

$$\{p, s, q, r\} \quad (23)$$

forced to {0,0,0,0}. This results in a trivial field with a checkerboard pattern, interchanging column/row pattern, or a combination hereof. In all cases the entropy of the field is equal to 0 since the only decisions are made either in the upper left corner or in the top/left edge. This can be seen directly from the admissible configurations since they will be limited to a subset of either $abcd = \{[0110], [1001], [1100], [0011]\}$ or $abcd = \{[0110], [1001], [0101], [1010]\}$.

B_4 are *absorbing fields* which are defined as constraints where $p = q = 0$ and the non-admissible configurations additionally contains $abcd = [a11d]$, where $s = r = 0$ and the non-admissible configurations additionally contains $abcd = [a00d]$, where $r = 0$ or $q = 0$ and the non-admissible configurations additionally contains $abcd = [0b1d] \wedge abcd = [1b0d]$, or where $s = 0$ or $p = 0$ and the non-admissible configurations additionally contains $abcd = [01cd] \wedge abcd = [10cd]$. In these cases the probability of the occurrence of a symbol using bit-stuffing depends on the distances to the starting edges; a structure which cannot be reflected by PRF. This category

implies non-stationarity. An example is a modified hard-square constraint where $(A, B = 0, C = 0, D = 1)$ is added to the non-admissible configurations. In this example the PRF will degenerate to uniform zeros.

Finally, B5 is defined by constraints c_i where the set of non-admissible configurations F_i includes all possible combinations of (ABC) without belonging to a prior category:

$$(abcd) \in F_i \quad \vee \quad (ab\tilde{c}\tilde{d}) \in F_i \\ \text{for } abc = \{[000], [001], \dots, [111]\} \quad (24)$$

In this case the field is defined only by the top and left edge and the per symbol field entropy is equal to 0.

A necessary condition to construct a PRF with positive entropy for a given constraint is that the constraint does not belong to any of the categories listed in this subsection. This can be determined without introducing probabilities. Fields constructed under the constraints in these categories can not contain all admissible configurations (B1), are not stationary (B1,B2,B4) and/or have a field entropy equal to 0 (B2,B3,B5). If the constraint is not categorized by now, we need to apply PRF analysis as described in the next section to determine whether the entropy of the PRF is non-zero.

B. PRF analysis

After analyzing the constraints within 2×2 blocks, we proceed and analyze any uncategorized constraints as PRF. For this reason the categories listed next are denoted P_k where k is an index. The first category P1 we define as the constraints where the only PRF solution is:

$$p = r = s = q = 0.5 \quad (25)$$

due to the requirements given by (13) and (14). In this case the entropy contribution calculated in (16) will be 0 and the calculated maximum entropy will only depend on contributions from (15) (of which there might be none). The 354 constraints in this category can be separated into 186 families, in which 3 family fathers have maximum entropy equal to 0.5, 34 family fathers have maximum entropy equal to 0.25 and the rest have maximum entropy equal to 0.

Category P2 is defined as constraints where the only configurations that can possibly contribute to the entropy is of the form (16) and satisfy $\tilde{a} = d, b = c$. Due to (13) this will lead to a maximum PRF entropy equal to 0. Constraints belonging to P1 or P2 can be analyzed by choosing e.g. the lower left corner as starting point for the PRF (corresponding to a rotation of the configurations), which might lead to a better PRF solution.

P3 contains the constraints which can be modeled by the PRF approach with positive field entropy less than 1. For these constraints we search for the maximum entropy over 1 to 3 of the free parameters from p, q, r, s . We provide test statistics and examples for this in Section IV.

Finally, P4 is the unconstrained case in which all configurations are admissible. In this case the maximum field entropy is equal to 1 with i.i.d. row and column distributions.

IV. RESULTS

We classify each of the 2^{16} possible constraints as belonging to one of the categories listed in Table I and as detailed in the previous section. In these categories, B1 is considered to be invalid, B2-5 have field entropy equal to 0, P1-2 are special cases for the PRF model and P4 is the single unconstrained case. We apply the PRF to the fathers of the families in P3.

TABLE I: Test statistics

Index	Short Description	Number of cases
B1	Invalid due to traps	47865
B2	$H = 0$, Absorbing Markov chain	9661
B3	$H = 0$, Checkerboard and/or rows/cols	275
B4	$H = 0$, Absorbing field	484
B5	$H = 0$, all (ABC)	600
P1	$p = r = s = q = 0.5$	354
P2	$H = 0$ due to commutation	324
P3	$H > 0$, PRF applied	5972
P4	$H = 1$, unconstrained	1

P3 consists of a total of 5972 constraints which were separated into 2645 families. For each family in P3, the optimal distributions and maximum PRF entropy were estimated as shown in Fig. 1. Most of the families in P3 have 2 members due to constraints being equal only through binary inversion. There exist 2584 of such families in P3. Only 38 constraints are completely unique (family size equal to 1) e.g. the no uniform squares configuration (where [0000] and [1111] are the non-admissible configurations). There are 22 families where one of the variables p, s, q, r is forced to 0 and they all have size equal to 30. The biggest family size is equal to 106 (of which there is only 1). This is the family corresponding to the hard square problem. It should be noted that families can have members which belongs to category B1, but these are not part of the analysis since they are considered invalid.

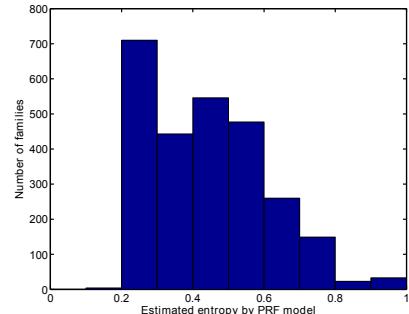


Fig. 1: Histogram of the estimated maximum PRF entropies.

V. EXAMPLES

In this section, some specific examples of the implemented PRF methodology is shown. To get an approximate value of the capacity we use bands, extending the binary alphabet to an alphabet with 2^M elements. We can now create the transfer (or adjacency) matrix T_M of the band and (when T_M is

symmetric) we calculate a lower bound for the combinatorial entropy of the field, first suggested by Calkin and Wilf [5]:

$$C \geq \log_2(\lambda_{2k+2}) - \log_2(\lambda_{2k+1}) \quad (26)$$

where λ_M is the greatest eigenvalue for T_M .

A. Hard-square Constraint

The first example is the hard-square constraint with the straightforward formulation as in [4]. This constraint forbids neighbours of 1s in the horizontal or vertical direction, giving nine non-admissible configurations. In this case the transfer matrix T is symmetric and the calculated lower bound (26) for the field entropy is 0.58789 for $k = 6$. The estimated maximum entropy with the PRF model is 0.58306 with parameters: $p = q = 0.71662$, $s = r = 0$. If we use the formulation of the hard-square constraint as suggested in [4] and also analyzed in [6] by shifting the bottom row of (3) to the left by one place, the non-admissible configurations are:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad (27)$$

With this formulation of the problem the transfer matrix T is not symmetric. The estimated maximum entropy with the PRF model improves to 0.58728 with the parameters: $p = 0.70901$, $s = 0$, $q = 0.81124$, $r = 0.35132$ (Fig. 2).



Fig. 2: The hard-square constraint in a 100x100 field.

B. No isolated 0-area

We consider the problem where no rectangles of all 0s are allowed and all 0s are connected to each other (when accepting diagonal connectivity). The non-admissible configurations are:

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \quad (28)$$

With this constraint the transfer matrix T is symmetric and the calculated lower bound (26) for the field entropy is 0.75949 for $k = 6$. The estimated maximum entropy with the PRF model is 0.75617 with parameters: $p = q = 0.65483$, $s = r = 0.19347$. This was used to generate the image in Fig. 3 (a).

C. No 1 columns

In this example, vertical neighboring 1s are not allowed. The non-admissible configurations are:

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad (29)$$

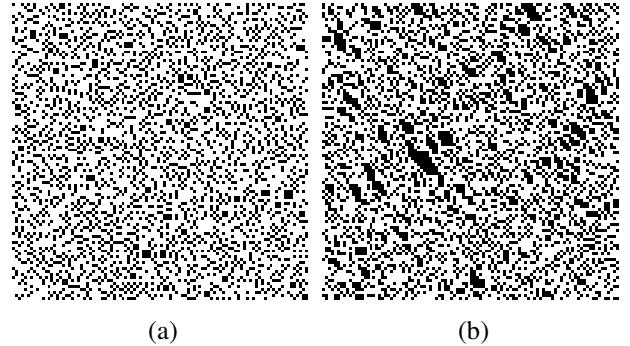


Fig. 3: An example of the "no isolated 0-area" constraint (a) and "no upper left 1-corner" constraint (b) in a 100x100 field.

With this constraint the transfer matrix T is symmetric and the calculated lower bound (26) for the field entropy is 0.69515 for $k = 6$. The estimated maximum entropy with the PRF model is 0.69424 with parameters: $p = 0.72361$, $s = 0.27639$, $q = 0.61803$, $r = 0$.

D. No upper left 1-corner

For this constraint the non-admissible configuration is $(abcd) = (1110)$. With this constraint the transfer matrix T is not symmetric. The maximum entropy with the PRF model is 0.82496. This provides a lower bound of the entropy for the constraint. The corresponding parameters were $p = q = 0.61504$, $s = r = 0.32161$ and were used to generate the image in Fig. 3 (b).

VI. CONCLUSION

Binary PRF have been analyzed and a closed form solution has been obtained for the maximum entropy of the binary PRF given any constraint defined on the 2×2 square. All binary constraints on the 2×2 square have been tested to determine whether a PRF got all admissible 2×2 configurations of the father constraint in its possible outcomes. Statistics have also been gathered on the corresponding maximum PRF entropies. Furthermore, invalid constraints and trivial constraints have been characterized and tested.

Future work could include a refined separation of categories with more strict description of the defining structure e.g. the category of constraints classified as invalid could be further divided depending on whether it is possible to fill the grid with a subset of the admissible configurations or not.

REFERENCES

- [1] S. Forchhammer and J. Justesen, "Block pickard models for two-dimensional constraints," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4626–4634, 2009.
- [2] J. Justesen and S. Forchhammer, *Two-Dimensional Information Theory and Coding*. Cambridge University Press, 2010.
- [3] D. K. Pickard, "Unilateral markov fields," *Advances in Applied Probability*, vol. 12, pp. 655–671, 1980.
- [4] R. Roth, P. Siegel, and J. Wolf, "Efficient coding schemes for the hard-square model," *IEEE Trans. Inf. Theory*, vol. 47, pp. 1166–1176, 2001.
- [5] N. J. Calkin and H. S. Wilf, "The number of independent sets in a grid graph," *SIAM J. Discr. Math.*, vol. 11, pp. 54–60, 1998.
- [6] S. Forchammer, M. Danieli, N. Burini, M. Zamarín, and A. Ukhanova, "Maximizing entropy of image models for 2-d constrained coding," in *Workshop on Inf. Theoretic Methods in Science and Engineering 3*, 2010.