

Hard Core via PCA: Entropy Bounds

Kari Eloranta

► **To cite this version:**

Kari Eloranta. Hard Core via PCA: Entropy Bounds. Jarkko Kari. 21st Workshop on Cellular Automata and Discrete Complex Systems (AUTOMATA), Jun 2015, Turku, Finland. Springer, Lecture Notes in Computer Science, LNCS-9099, pp.85-98, 2015, Cellular Automata and Discrete Complex Systems. <10.1007/978-3-662-47221-7_7>. <hal-01442484>

HAL Id: hal-01442484

<https://hal.inria.fr/hal-01442484>

Submitted on 20 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Hard Core via PCA: entropy bounds

Kari Eloranta

Institute of Mathematics
Aalto University, Finland
kari.v.eloranta@gmail.com

Abstract. We establish bounds for the entropy of the Hard Core Model/Independent Sets on a few 2-d lattices. Our PCA-based sequential fill-in method yields an increasing sequence of lower bounds for the topological entropy. Additionally the procedure gives some insight on the support of the measure of maximal entropy. The method also applies to other lattices and models with appropriate sublattice splitting.

Keywords: hard core model, independent sets, topological entropy

1 Introduction

The **Hard Core Model/Golden Mean Subshift/Independent Sets** is a highly useful model in various disciplines as witnessed by its many appearances under distinct names in fields like Statistical Mechanics/Symbolic Dynamics/Theoretical Computer Science respectively. Simply described, one distributes 0's and 1's on each vertex of a given graph and requires an **exclusion rule** to hold everywhere: no two 1's can be nearest graph neighbors ([1]).

On a k -partite graph it is natural to associate with it a probabilistic cellular automaton (PCA) which updates each of the k subgraphs from the others with a local rule as follows: in an all-0 neighborhood update the center vertex to 1 with probability $u \in (0, 1)$, otherwise update to 0. Running a sequential update with the PCA (from any input) through the subgraphs yields eventually any Hard Core configuration with positive probability (and no others).

The PCA behavior is of interest for different u -values. If u is near 1, we are in the **high density regime** and the characterization of the allowed Hard Core configurations is essentially a packing problem (studied e.g. in an earlier paper [6]). In the **dilute case** i.e. u is near 0, the configurations are essentially Bernoulli(\tilde{u})-fields (sometimes called Hard Square Gas), $0 < \tilde{u} < u$.

Here we concentrate on the intermediate **entropic regime**. The outstanding open question near $u = 1/2$ is the exponential size of the configuration space. In almost all two or higher dimensional set-ups the exact answer is unknown. We try to alleviate the situation a little bit by establishing a procedure to estimate the entropy from below and to characterize the typical configurations. For simplicity we restrict here the graphs to be 2-d lattices. This facilitates comparison to entropy approximations elsewhere via different methods (e.g. [2], [4], [7], [8], [9]). Our ideas also generalize to more complicated and higher dimensional set-ups. For a general percolation approach to Hard Core see e.g. [3].

1.1 Set-up

Let \mathbf{L} be a 2-dimensional lattice. Subsequently we will consider mostly one of the regular lattices (square (\mathbf{Z}^2), honeycomb (\mathbf{H}) or triangular (\mathbf{T})). In a few cases we illustrate the principles developed on more exotic stages like square lattice with the Moore neighborhood ($\mathbf{Z}^2\mathbf{M}$, eight nearest Euclidean neighbors, a non-planar graph) or the Kagomé lattice (\mathbf{K}). Our method is not intrinsically 2-d but for understanding the clear geometry of 2-d serves best.

A configuration on \mathbf{L} satisfying the **Hard Core Rule** is an element in $X = \{0, 1\}^{\mathbf{L}}$ where no two 1's can be nearest neighbors. This rule can naturally be viewed as a zero range infinite repulsive potential i.e. a hard exclusion rule not unlike that in hard sphere packing. Call the collection of configurations $X_{\mathbf{L}}^{hc}$.

The exclusion rule naturally imposes a sublattice split on \mathbf{L} if it is a k -partite graph. For example on \mathbf{Z}^2 , a bipartite graph, one can man all sites on $2\mathbf{Z}^2$ (\mathbf{Z}^2 rescaled by $\sqrt{2}$ and rotated by 45°) with 1's and the rest of \mathbf{Z}^2 must then be all 0's. Call the former the **even sublattice**, \mathbf{L}_e and the latter the **odd sublattice**, \mathbf{L}_o (it is a $(1/2, 1/2)$ -shifted copy of the former). In rendering these we will present the even/odd sublattices as **circle/dot sublattices**. In a similar fashion \mathbf{H} splits into two identical sublattices and \mathbf{T} (a tripartite graph) into three “thinned” copies of \mathbf{T} . Both in the dense packing regime of [6] and in the the entropic regime of this paper, this splitting will be highly relevant.

Let $X_0 \subset X = \{0, 1\}^{\mathbf{L}}$. $|\cdot|$ means the cardinality and $x|_A$ means the restriction of x to the set A . The standard measure of richness of the configuration set is

Definition 1. *The topological entropy of the set X_0 is*

$$h_{X_0}^{top} = \lim_{n \rightarrow \infty} \frac{1}{n} \ln |\{x|_{A_n} \mid x \in X_0\}|$$

where $|A_n| = n$ and the sequence $\{A_n\}$ grows in a sufficiently regular fashion.

Remark. For the full shift on any lattice $h_X^{top} = \ln 2$ (indicating two independent choices per lattice site). If $L = \mathbf{Z}$ the Hard Core model is explicitly solvable and a standard transfer matrix argument implies that $h^{top} = \ln \left(\frac{1+\sqrt{5}}{2} \right)$ (e.g. [11]). For two and higher dimensional lattices the matrix argument breaks down and the exact value of the Hard Core topological entropy remains an unsolved problem except for \mathbf{T} (see [1]). In this paper we try to approach and in particular approximate it in a novel way.

From the general theory of lattice dynamical systems ([11]) it is known that shift invariant probability measures on a space of configurations, \mathcal{M} , satisfy the maximum principle, $h^{top} = \sup_{\mathcal{M}} h_{\mu}$, where h_{μ} is the measure-entropy. The special measures yielding the equality are **measures of maximal entropy**. For two and higher dimensional systems they are in general not unique. In all our cases they are believed to be so, but we do not actually need to know this.

For the idea of our approach recall that given the entropy function $H(\mathcal{P}) = -\sum \mu(P_i) \ln \mu(P_i)$ for a partition \mathcal{P} , it holds $H(\mathcal{P} \vee \mathcal{P}') = H(\mathcal{P}) + H(\mathcal{P}'|\mathcal{P})$. In

the case of \mathbf{Z} let $\mathcal{P}^{(e)} = \{\{x|x_0 = i\}\}$ and $\mathcal{P}^{(o)} = \{\{x|x_1 = i\}\}$ be the (generating) partitions for the even/odd sublattice for the left shift σ . $\mathcal{P} = \mathcal{P}^{(e)} \vee \mathcal{P}^{(o)}$ generates on \mathbf{Z} and the conditional entropy formula above implies

$$h_\mu(\sigma) = \frac{1}{2}h_\mu(\sigma^2|\mathcal{P}) = \frac{1}{2}h_\mu^{(e)} + \inf_n \frac{1}{2n} H_\mu \left(\bigvee_{k=0}^{n-1} \sigma^{-2k} \mathcal{P}^{(o)} \middle| \bigvee_{k=0}^{n-1} \sigma^{-2k} \mathcal{P}^{(e)} \right). \quad (1.1)$$

The formula readily generalizes to \mathbf{Z}^d -actions, to multiple sublattices etc. The last expression is trivial only in the independent case (equals $\frac{1}{2}h_\mu^{(o)}$). It codes the additional entropy to be gained from the odd sublattice, *given the even*. In the case of Hard Core it is manageable since 1. the rule is of range one and 2. the hard exclusion greatly simplifies the computation of the conditional probabilities involved. This will become much more transparent in the subsequent analysis.

2 Lower bounds

We now proceed to establish lower bounds for the topological entropy using the sublattice partition representation (1.1) and a sequential fill-in scheme to circumvent the dependencies. To keep the ideas clear we first present them in the bipartite the case and then comment on the k -partite cases.

Let N_e denote an all-0 nearest neighbor neighborhood of a site on the odd lattice in the even lattice. In the case of \mathbf{Z}^2 lattice the sites in N_e form the vertices of an **even unit diamond**, \diamond_e (\diamond_o). On the honeycomb and triangular lattices these sites form triangular arrangements, \triangle or ∇ or a hexagon.

It will become quite useful to think the fill-in in terms of forming a tiling. The pieces are **0/1-tiles** which in \mathbf{Z}^2 case are either 0/1-diamonds (as above) depending on whether the center site carries 0 or 1. On the hexagonal and triangular lattices the tiles are 0/1-**(unit) hexes**. Once a sublattice is chosen, one can tile the plane using any combination of 0/1-tiles centered on the sublattice.

Recall that the Bernoulli measure with parameter p , $B(p)$, assigns 1's independently with probability p to each (sub)lattice site and 0's otherwise. Its entropy, denoted by $h_B(p)$, is $-p \ln p - (1-p) \ln(1-p)$.

Proposition 1. *The topological entropy of the hard core model on a lattice with a two-way sublattice split is given by*

$$h^{top} = \frac{1}{2} \left\{ h^{(e)} + \mathbf{P}(N_e) \ln 2 \right\}, \quad (2.1)$$

where $h^{(e)}$ is the entropy of the measure of maximal entropy computed from the even sublattice alone.

Proof. We attain the maximum entropy by first assigning the marginal of the measure of maximal entropy to the even lattice and then filling in the non-blocked sites on the odd lattice. These are centered at the even unit diamonds, which have probability $\mathbf{P}(N_e)$. $B(1/2)$ is the maximal entropy measure among $B(p)$, hence this on the allowed odd sites and the factor $h_B(1/2) = \ln 2$. QED

The principle in the Proposition can be directly applied to square and honeycomb lattices. A further argument is required to cover all regular lattices. In the following result we present these arguments and further extend to Kagomé lattice, \mathbf{K} (a tripartite graph), as well as to the square lattice with Moore neighborhood, $\mathbf{Z}^2\mathbf{M}$ (eight nearest neighbors in Euclidean metric, a 4-partite graph).

Theorem 1. *The topological entropy of the hard core model is bounded from below on the square ($m = 4$) and honeycomb ($m = 3$) lattices by*

$$h_{\mathbf{Z}^2/\mathbf{H}}(p) = \frac{1}{2} \left\{ h_B(p) + (1-p)^m \ln 2 \right\}, \quad (2.2)$$

on triangular ($m' = 3$) and Kagomé ($m' = 2$) lattices by

$$h_{\mathbf{T}/\mathbf{K}}(p, q) = \frac{1}{3} \left\{ h_B(p) + (1-p)^{m'} \left[h_B(q) + [1 - (1-p)q]^2 \ln 2 \right] \right\} \quad (2.3)$$

and on $\mathbf{Z}^2\mathbf{M}$ lattice by

$$h_{\mathbf{Z}^2\mathbf{M}}(p, q, r) = \frac{1}{4} \left\{ h_B(p) + (1-p)^2 \left[h_B(q) + [1 - (1-p)q]^4 h_B(r) + (1-p)^2 (1-q)^2 [1 - (1 - (1-p)q)^2 r]^2 \ln 2 \right] \right\} \quad (2.4)$$

where p, q and $r \in (0, 1)$.

Proof. The lower bounds (2.2) follow simply from (2.1) by assigning $B(p)$ to the even sublattice since then $\mathbf{P}(N_e) = (1-p)^{|N_e|}$ where the exponent is the number of elements in N_e in \mathbf{Z}^2 and \mathbf{H} respectively.

On the triangular lattice the sublattice partition is three-way. We call the parts the dot, circle and triangle sublattices. They are filled in three stages in the order $\circ \rightarrow \bullet \rightarrow \triangleright$. See Figure 1a and b for the notation and arrangement of the sublattices in a neighborhood of a triangle site.

Suppose the three sublattices are initially all empty. First fill-in the circle lattice with $B(p)$, hence the entropy contribution $\frac{1}{3}h_{B(p)}$. Then fill-in all dot sites centered at ∇ with $B(q)$, this implies the entropy increase $\frac{1}{3}(1-p)^3 h_{B(p)}$ from the dot lattice.

To update the center site which is a triangle we need to know that its value is not forced. Hence

$$\begin{aligned} & \mathbf{P}(\text{center triangle not forced by nearest neighbor circle or dot}) \\ &= \mathbf{P}(\text{no } 1\text{'s in the hexagon of nearest neighbors of the triangle}) \\ &= \mathbf{P}(\triangle = \underline{0} \text{ and } \nabla = \underline{0}) = \mathbf{P}(c_2 = c_4 = c_5 = 0 \text{ and } d_1 = d_2 = d_3 = 0) \\ &= \mathbf{P}(d_1 = d_2 = d_3 = 0 \mid c_2 = c_4 = c_5 = 0) \mathbf{P}(c_2 = c_4 = c_5 = 0) \\ &= \mathbf{P}(d_1 = d_2 = d_3 = 0 \mid c_2 = c_4 = c_5 = 0) (1-p)^3 \\ &= [\mathbf{P}(d_1 = 0 \mid c_2 = c_4 = c_5 = 0)]^3 (1-p)^3 \\ &= [\mathbf{P}(c_1 = 1 \text{ or } \{c_1 = 0 \text{ and } d_1 = 0\} \mid c_2 = c_4 = c_5 = 0)]^3 (1-p)^3 \\ &= [p + (1-p)(1-q)]^3 (1-p)^3 \end{aligned} \quad (2.5)$$

which together with the choice $B(1/2)$ on the non-blocked dots gives (2.3).

The Kagomé lattice argument is similar to the triangular one. There are three sublattices involved, all identical copies of the Kagomé, only thinned and reoriented. For the nearest neighbors of a triangle-site see Figure 1c. Again we fill in the order $\circ \rightarrow \bullet \rightarrow \triangleright$. In the last stage the probability of the triangle site being unforced is now

$$\begin{aligned} \mathbf{P}(c_2 = c_4 = 0 \text{ and } d_1 = d_2 = 0) \\ &= [\mathbf{P}(c_1 = 1 \text{ or } \{d_1 = 0 \text{ and } c_1 = 0\} \mid c_2 = c_4 = 0)]^2 (1-p)^2 \\ &= [p + (1-p)(1-q)]^2 (1-p)^2 \end{aligned}$$

In the case of the square lattice with Moore neighborhood there is a four-way sublattice partitioning. We denote and fill them as follows: $\circ \rightarrow \bullet \rightarrow \triangleright \rightarrow \diamond$ (see Fig. 1d).

The two first terms of the formula (2.4) are straightforward since circles are laid independently and each dot has exactly two circle neighbors. Furthermore as above we can show that $\mathbf{P}(\triangleright \text{ unforced}) = (1-p)^2 [p + (1-p)(1-q)]^4$.

For the diamond site at the center of Fig. 1d to contribute to the entropy we need to know the probability that it is unforced i.e. all entries in the punctured square S rendered with dotted line in Fig 1d. are 0's:

$$\begin{aligned} \mathbf{P}(S = \underline{0}) &= \mathbf{P}(\text{all } \triangleright, \bullet \in S \text{ are } 0 \mid \text{all } \circ \in S \text{ are } 0) (1-p)^4 \\ &= \mathbf{P}(\triangleright \in S \text{ are } 0 \mid \circ, \bullet \in S \text{ are } 0) (1-p)^4 (1-q)^2 \\ &= \mathbf{P}(d_1 = 1 \text{ or } \{d_1 = d_2 = 0 \text{ and } t_1 = 0\} \mid \circ, \bullet \in S \text{ are } 0) (1-p)^4 (1-q)^2 \\ &= [\mathbf{P}(d_1 = 1 \mid \circ, \bullet \in S \text{ are } 0) \\ &\quad + \mathbf{P}(d_1 = d_2 = 0 \text{ and } t_1 = 0 \mid \circ, \bullet \in S \text{ are } 0)]^2 (1-p)^4 (1-q)^2 \end{aligned} \tag{2.6}$$

One can compute the two probabilities in the last expression to be

$$\begin{aligned} 2p(1-p)q + (1-p)^2(2-q)q \quad \text{and} \\ [p^2 + 2p(1-p)(1-q) + (1-p)^2(1-q)^2] (1-r) \end{aligned}$$

respectively. From these the formula in the square brackets in (2.6) can finally be simplified to the form $1 - [1 - (1-p)q]^2 r$. QED

The entropy bounds in (2.2) - (2.4) can be maximized with respect to the parameters using standard optimization routines in a desktop machine.

In the square lattice case the topological entropy has been computed to a great accuracy (e.g. in [2] to some 40 decimal places) using the corner transfer matrix methods. These numerical studies attack the problem in a very different way. Our aim is not to compete in decimal count but rather present an alternative method applicable in many lattice set-ups to estimate the entropy which simultaneously yields some explicit information on the generic configurations/the measure of maximal entropy.

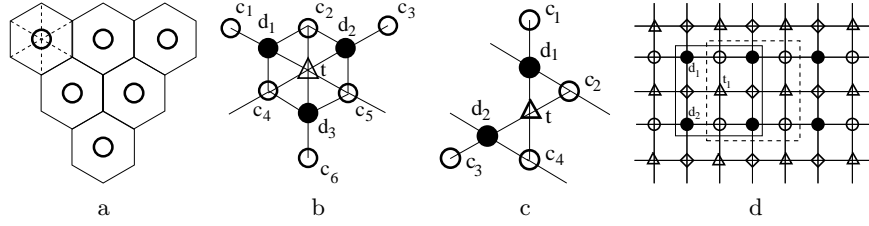


Fig. 1: Tiling of the triangular lattice with hexagonal 1-tiles and neighborhoods in triangular, Kagomé and $\mathbf{Z}^2\mathbf{M}$ cases.

Table 1: First lower bounds for Hard Core topological entropy and the corresponding sublattice densities for some 2-d lattices. To the right we have indicated the best numerical estimates for the entropy and corresponding density (in parenthesis) found in the literature.

\mathbf{L}	$\max \underline{h}_{\mathbf{L}}$	sublattice densities	best estimates
\mathbf{Z}^2	0.3924	(0.1702, 0.2370)	0.4075 (0.2266) _{[9],[2]}
\mathbf{H}	0.4279	(0.2202, 0.2371)	0.4360 (0.2424) _[2]
\mathbf{T}	0.3253	(0.1457, 0.1559, 0.1517)	0.3332 (0.1624) _[2]
\mathbf{K}	0.3826	(0.1944, 0.1948, 0.1866)	
$\mathbf{Z}^2\mathbf{M}$	0.2858	(0.119, 0.127, 0.130, 0.126)	

The measure of maximal entropy doesn't need to be unique for a 2-d lattice model but in the case of hard square gas it is. This follows from the Dobrushin criterion ([5], [10]). Using this knowledge and the results above we now establish bounds for the density of 1's in the generic configurations. The exact value of the upper bound in the following result is in the Proof but we prefer to give the statement in this more explicit form.

Proposition 2. *In the square lattice case the density of 1's at the equilibrium is in the interval (0.21367, 0.25806).*

Proof. Let ρ_e be the density of 1's on the even lattice and let c denote the expected number of 0's that a 1 forces on the odd lattice. Since exactly half of the non-forced sites will be 1's it must by the uniqueness of the measure of maximal entropy hold that $(2 + c)\rho_e = 1$. Hence under it

$$\mathbf{P}(x_i = 0) = \frac{1 + c}{2 + c}, \quad \mathbf{P}(x_i = 1) = \frac{1}{2 + c} \quad \text{and} \quad \mathbf{P}(\diamond_e) = \frac{2}{2 + c}$$

on both lattices. The last one is due to exactly half of \diamond_e giving rise to all the 1's on the other sublattice. (0-diamond as defined in the beginning of the section).

The entropy of any distribution on the even lattice with 1-density ρ_e is bounded from above by the entropy of the Bernoulli distribution with parameter

ρ_e . Hence the total entropy at that 1-density is bounded from above by

$$\frac{1}{2} \left(h_B \left(\frac{1}{2+c} \right) + \frac{2}{2+c} \ln 2 \right).$$

This expression bounded by h^{top} of Table 1 ([2] or [9]) yields an upper bound for c , 2.6801 which in turn gives the lower bound for ρ_e .

The upper bound for ρ_e follows from a lower bound for c which we establish using a monotonicity argument. The 1's on say the even lattice are $B(1/2)$ -distributed on the non-forced sites. Call this set F . Pick a site in it which has symbol 1. How many sites will this entry block? Let F' be a superset of F . Then clearly $\mathbf{E}(c|F) \geq \mathbf{E}(c|F')$ as in a bigger domain the 1 is more likely to share the blocking with a nearest neighbor 1 on the same sublattice. Hence a lower bound is obtained by calculating the blocking for a 1 with its eight nearest neighbors also in F . Enumerating the 2^8 possible neighborhood configurations and weighting them uniformly according to the $B(1/2)$ -distribution we get the lower bound for c : $15/8$. This in turn implies the upper bound for ρ_e , $8/31$. QED

Since our first estimate for the lower bound on \mathbf{Z}^2 is associated with densities incompatible with Proposition 2 we will try out a symmetric variant of the theme. The (near) equality of the densities on the sublattices should be a natural property of a measure corresponding to a good lower bound since the measure of maximal density is believed to be unique in all our cases. In the last three cases in Table 1, the non-equality of the densities isn't far off but for the first two we present an "equalization".

Proposition 3. *To achieve equal densities of 1's on each of the sublattices one needs to replace the $B(1/2)$ distribution in the last stage of the measure construction by $B(p')$ and thereby $\ln 2$ in (2.2) by $h_{B(p')}$, where $p' = p(1-p)^{-|N_e|}$.*

Proof. In the case of two sublattices after $B(p)$ distribution of 1's on the even lattice there are a density of $(1-p)^{|N_e|}$ unforcing neighborhoods on this sublattice. These have to produce the correct density of 1's on the odd lattice, hence we need the even lattice flip probability p' to satisfy $p'(1-p)^{|N_e|} = p$. QED

Using Proposition 3 one can optimize the square lattice topological entropy bound to (a slightly worse value) 0.3921 at common density level 0.2015. In view of Proposition 2 this indicates that the entropy generating 1's are not yet packed in densely enough. In the case of the honeycomb lattice the corresponding values are 0.427875 at 0.2284.

3 Higher order blocks

To improve the entropy bounds and more importantly to get some insight into the character of the measure of maximal entropy we now consider more complicated optimization schemes involving Bernoulli-distributed blocks on sublattices. We first illustrate the ideas on hexagonal and triangular lattices.

A **three-hex** is obtained by gluing together three unit hexes so that each has two joint sides. Figure 2a. illustrates three such three-hexes next to each other (for reference lattice edges are indicated as thin dotted lines in one of the unit hexes). Note that the unit tiles on each of them are all centered on the same sublattice, the circle lattice in this case (call the tile a **circle three-hex**). The dots of the other sublattice are all in the centers of the three-hexes or on their extremities (three of them are indicated). Three-hexes of the same orientation obviously tile the plane.

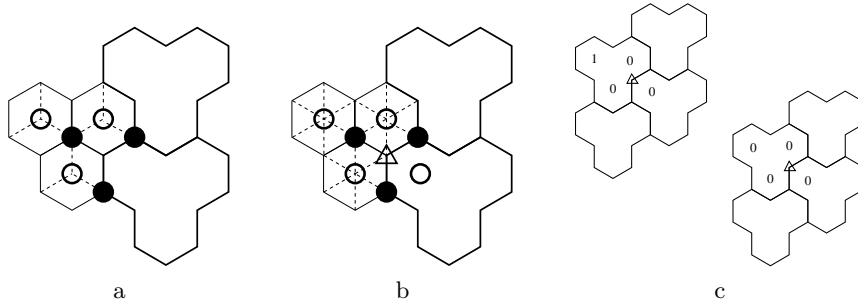


Fig. 2: 3-hex arrangements in hexagonal and triangular cases.

Let $B(\mathbf{p})$, $\mathbf{p} = (p_0, p_1, p_2, p_3)$ be the Bernoulli distribution on circle three-hexes with the probability that the three-hex has exactly k 1-tiles in it in a given orientation being p_k (so $p_0 + 3p_1 + 3p_2 + p_3 = 1$). Its entropy is then $h_B^{(3)}(\mathbf{p}) = -p_0 \ln p_0 - 3p_1 \ln p_1 - 3p_2 \ln p_2 - p_3 \ln p_3$.

Theorem 2. Let $a(\mathbf{p}) = p_0 + 2p_1 + p_2$. For the hexagonal lattice the Hard Core entropy is bounded from below by

$$\underline{h}_{\mathbf{H}}^{(3)}(\mathbf{p}) = \frac{1}{6} \left\{ h_B^{(3)}(\mathbf{p}) + [p_0 + 2a(\mathbf{p})^3] \ln 2 \right\} \quad (3.1)$$

and for the triangular lattice a corresponding bound is

$$\underline{h}_{\mathbf{T}}^{(3)}(\mathbf{p}, q) = \frac{1}{9} \left\{ h_B^{(3)}(\mathbf{p}) + [p_0 + 2a(\mathbf{p})^3] h_B(q) + 3[p_1 + p_0(1 - q)] a(\mathbf{p})^3 (2 - q)^2 \ln 2 \right\} \quad (3.2)$$

where $p_i, q \in (0, 1)$.

Proof. For the construction of the measure we will fill in the lattice in the order $\circ \rightarrow \bullet$. If the circle three-hexes are distributed Bernoulli with parameter \mathbf{p} the entropy contribution from the circle lattice will be $\frac{1}{2} h_B(\mathbf{p})$ where the factors

result from the sublattice density and the fact that we distribute triples. As in Proposition 1. in the next stage the maximal entropy choice for the unforced sites on the dot lattice is the $B(1/2)$ distribution. The total density of sites available is computed at two different types of dot sites (as in Fig. 2a, the three dots indicated) and is $\frac{1}{3} \left[p_0 + 2(p_0 + 2p_1 + p_2)^3 \right]$ where the coefficient 2 and the power 3 follow from the fact that at two of the three dot sites three adjacent three-hexes coincide. These formulas combined and simplified yield (3.1).

On the triangular lattice a third sublattice enters and the fill-in order is then $\circ \rightarrow \bullet \rightarrow \triangleright$. The entropy contribution from the Bernoulli circle three-hexes is now $\frac{1}{3} \frac{1}{3} h_B(\mathbf{p})$ since each sublattice is identical, hence of density $1/3$.

In the second stage the unforced dot sites are filled with $B(q)$ distribution. Their density is computed as above to be $\frac{1}{3} \left[p_0 + 2(p_0 + 2p_1 + p_2)^3 \right]$, hence the entropy contribution from dot lattice will be this expression multiplied by $\frac{1}{3} B(q)$.

In the final stage the unforced triangle sites are filled by $B(1/2)$. Their density in the full lattice is

$$\begin{aligned} & \frac{1}{3} \mathbf{P}(\text{nearest neighbor } \circ \text{ and } \bullet \text{ sites all } 0\text{'s}) \\ &= \frac{1}{3} \left\{ p_1(p_0 + 2p_1 + p_2) [(p_1 + 2p_2 + p_3) + (p_0 + 2p_1 + p_2)(1 - q)]^2 \right. \\ & \left. + p_0(p_0 + 2p_1 + p_2)(1 - q) [(p_1 + 2p_2 + p_3) + (p_0 + 2p_1 + p_2)(1 - q)]^2 \right\}, \end{aligned} \quad (3.3)$$

which results from considering the two different arrangements of four neighboring three-hexes as shown in Fig. 2c. (top and bottom cases for the top and bottom expressions in (3.3)). The formulas merged and simplified result in (3.2). QED

Table 2: Optimized lower bounds and densities for three-hex Bernoulli blocks.

L	$\max h_{\mathbf{L}}$	$(p_0, p_1, p_2, p_3), q$	sublattice densities
H	0.4304	(0.504, 0.110, 0.048, 0.021)	(0.2276, 0.2376)
T	0.3265	(0.64, 0.092, 0.025, 0.010), 0.25	(0.153, 0.155, 0.151)

Remarks. 1. The Kagomé lattice case is treated in an analogous fashion to **T**.
 2. Note that apart from improvements in the entropy bounds, almost all of the sublattice densities have increased (in comparison to values in Table 1) indicating a better packing of the 1's on the sublattices. Moreover they have significantly less variation which is to be expected since the densities are equal for the measure of maximal entropy.

Let us now return to our original motivation, the Hard Core on the square lattice. Compounding the principles above and some further ideas we will implement an

increasing sequence of lower bounds converging to the topological entropy. Along the way we'll get more explicit information on the configurations favored by the measure of maximal entropy.

1-tiles in the \mathbf{Z}^2 case are diamonds of side length $\sqrt{2}$ centered on either of the two sublattices. k -omino is formed by gluing together k such 1-tiles along edges. If $k = n^2$ and the 1-tiles are in a diamond formation we call them a $n \times n$ -blocks. There are 2^{n^2} of them. The optimization results in Section 2 were for the 1×1 -blocks.

Consider next 2×2 -blocks. There are 16 of them, but after assuming isotropy for them i.e. that blocks that are rotations of each other are distributed with equal probability (inevitable when measure of maximal entropy is unique), there are only five free parameters for Bernoulli distribution $B(\mathbf{p})$ on them ($\mathbf{p} = (p_0, p_1, p_{21}, p_{22}, p_3, p_4)$, $p_0 + 4p_1 + 4p_{21} + 2p_{22} + 4p_3 + p_4 = 1$. Here the first subindex of p refers to the number of 1's in the block and p_{22} and p_{21} denote the two different arrangement of two 1's in the block (side by side and across)).

The entropy contribution from the even lattice (on which we distribute first the 1's using $B(\mathbf{p})$ is now

$$-\frac{1}{4} \left\{ p_0 \ln p_0 + 4p_1 \ln p_1 + 4p_{21} \ln p_{21} + 2p_{22} \ln p_{22} + 4p_3 \ln p_3 + p_4 \ln p_4 \right\}. \quad (3.4)$$

The density of the unforced sites on the odd lattice can be computed from the three cases indicated in Figure 3a. and results in

$$\frac{1}{4} \left\{ p_0 + 2(p_0 + 2p_1 + p_{21})^2 + (p_0 + 3p_1 + 2p_{21} + p_{22} + p_3)^4 \right\}. \quad (3.5)$$

These combined yield a lower bound for h^{top} , which is optimized in Table 3 (second row).

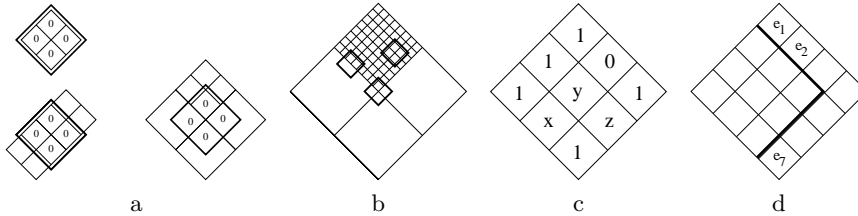


Fig. 3: $n \times n$ -blocks and update window in \mathbf{Z}^2 case. Reductions in a 3×3 -block and the extension.

In the block size 3×3 there are initially 511 free block probabilities to optimize. When rotational invariance is imposed the variable number is reduced and additionally we will expect blocks that are reflections of each other to have equal probabilities at the optimum. After these two types of symmetries are accounted the number of free variables will be 101.

Table 3: Optimized lower bounds and densities for a Bernoulli blocks on \mathbf{Z}^2 .

block size	max $h_{\mathbf{Z}^2}$	sublattice densities	final/initial variables
1×1	0.392421	(0.1702, 0.2370)	1/1
2×2	0.39877	(0.1993, 0.2254)	5/15
3×3	0.4014	(0.2073, 0.2254)	46/511

In this size and in larger blocks another feature appears which enables further variable weeding. Consider the block in Figure 3c. The symbol assignments in sites x , y and z are irrelevant in the sense that the existing 1's in the 3×3 -block already force all the odd sites (to carry 0's) that x , y and z might force if any of them were 1's. Hence there are 2^3 blocks of equal probability. This combined with the symmetry assumptions above yield the total of 64 blocks with identical probabilities at the optimum (this is actually the maximum reduction achievable in this block size). Combing through the set of all blocks for this feature will result in reduction by a factor about 11 to the final set of 46 variables. Their optimal values have been computed and the results are in Table 3.

Subsequently we call sites like x, y and z above **weak** with respect to the rest of the given block. Only the corner sites of a block cannot ever be weak.

The procedure of variable reduction is highly useful since the above rotational and reflection symmetry search as well as the weak site identification can be automated. Moreover the reduction improves significantly at every stage: for example in the next block size of 4×4 the initial variable number of 65.536 shrinks 66-fold to 991 final free variables.

Note also that the optima in block size $n \times n$ can be utilized as indicated in Figure 3d to initiate the search in the next larger block size. Once e.g. the 3×3 subblock optimum probability is known, the added half frame (e_1, \dots, e_7) should be assigned $B(p)$ entries with p computed from 3×3 blocks. With tailored optimization routines one should be able to deal with several thousands of variables in the larger block sizes. All the optimizations here were done with non-specialized code using *Mathematica*.

The optimal block probabilities satisfy a useful monotonicity property, that we establish next. For this let B_i , $i = 1, 2$ be $n \times n$ -blocks, whose subsets of 1's we refer to as $B_i^{(1)}$. There is a partial order on the blocks via $B_i^{(1)}$ using the ordinary set inclusion: $B_i \prec B_j$ if $B_i^{(1)} \subset B_j^{(1)}$. Let the optimal probabilities for the blocks be $\mathbf{p} = (p_0, p_1, p_2, p_3, \dots, p_l)$, $l = 2^{n^2}$ (no reductions done yet and no particular order in the coordinates).

Theorem 3. *Given two blocks B_1 and B_2 with optimal lower bound probabilities p_1 and p_2 , if $B_1^{(1)} \subset B_2^{(1)}$ then $p_1 \geq p_2$. If $B_2^{(1)} \setminus B_1^{(1)}$ contains only weak sites with respect to $B_1^{(1)}$ then $p_1 = p_2$, otherwise $p_1 > p_2$.*

Proof. The optimal lower bound is given by $\underline{h}(\mathbf{p}) = \frac{1}{n^2} \{-\sum_i p_i \ln p_i + \mathbf{P}(N_e) \ln 2\}$ where N_e is the even 2×2 -diamond of all 0's as in Section 2. Let B_i be such that $B_1^{(1)} \subset B_2^{(1)}$ and let $p_1 = p + \epsilon$, $p_2 = p - \epsilon$, $0 \leq |\epsilon| < p$. Denote by $h_\epsilon(\mathbf{p})$ the lower bound with the given p_1 and p_2 . To prove the result we will consider the entropy variation under the probability change of the two blocks: $\Delta h_\epsilon(\mathbf{p}) = h_\epsilon(\mathbf{p}) - h_0(\mathbf{p})$. More explicitly

$$\begin{aligned} \Delta h_\epsilon(\mathbf{p}) = \frac{1}{n^2} \left\{ \right. & \left[-(p + \epsilon) \ln(p + \epsilon) - (p - \epsilon) \ln(p - \epsilon) + 2p \ln p \right] \\ & + [\mathbf{P}_{1,\epsilon}(N_e) - \mathbf{P}_1(N_e) \\ & + \mathbf{P}_{2,\epsilon}(N_e) - \mathbf{P}_2(N_e) \\ & \left. + \mathbf{P}_{4,\epsilon}(N_e) - \mathbf{P}_4(N_e)] \ln 2 \right\}, \end{aligned} \quad (3.6)$$

where $\mathbf{P}_{k,\epsilon}(N_e)$ and $\mathbf{P}_k(N_e)$ are the N_e -diamond probabilities computed from the different arrangements involving $k = 1, 2$ or 4 $n \times n$ -blocks as in Figures 3a and b, for the block probability choices $p \pm \epsilon$ or p for both.

By $\ln(1+x) \approx x$ the first square bracket behaves for small ϵ like $c_1 \epsilon^2$, $c_1 < 0$.

If $B_2^{(1)} \setminus B_1^{(1)}$ contains only weak sites with respect to $B_1^{(1)}$ then the blocks B_i allow exactly the same sites to flip on the odd lattice hence each of the three last lines in (3.6) vanishes. The sole contribution to $\Delta h_\epsilon(\mathbf{p})$ then comes from the first square bracket and since this is negative for small but nonzero ϵ , it must be that $p_1 = p_2$ at the optimum.

If $B_2^{(1)} \setminus B_1^{(1)}$ contains non-weak sites with respect to $B_1^{(1)}$ let us first assume that they force k odd interior sites (recall that the odd sites are the vertices of the grids in Figure 3. There are $(n-1)^2$ such interior sites in a $n \times n$ -block). Let m be the number non-forced odd interior sites over block B_1 . Then

$$\begin{aligned} \mathbf{P}_{1,\epsilon}(N_e) - \mathbf{P}_1(N_e) &= \left(\dots + \frac{(p + \epsilon)m}{(n-1)^2} + \frac{(p - \epsilon)(m-k)}{(n-1)^2} + \dots \right) \\ &- \left(\dots + \frac{pm}{(n-1)^2} + \frac{p(m-k)}{(n-1)^2} + \dots \right) = \frac{k\epsilon}{(n-1)^2}, \end{aligned}$$

where the dots refer to the contributions from the other blocks. All these terms cancel out, since the other block probabilities are identical.

If non-weak sites only force odd interior sites then by geometry of the set-up the two last lines in (3.6) are immediately zero. If e extra odd edge, off-corner sites are forced, similar argument than above gives estimate $(c_2 + \frac{e\epsilon}{4(n-1)})^2 - c_2^2$, $c_2 > 0$ for $\mathbf{P}_{2,\epsilon}(N_e) - \mathbf{P}_2(N_e)$ so the next to last line in (3.6) has the first order behavior $c_3 \epsilon$, $c_3 > 0$. Some added bookkeeping yields $\mathbf{P}_{4,\epsilon}(N_e) - \mathbf{P}_4(N_e) = (c_4 + l\epsilon/4)^4 - c_4^4 \approx c_5 \epsilon$, $c_5 > 0$ (l is the number of odd corners forced).

The leading orders for the terms in the square brackets in (3.6) together yield $c_1 \epsilon^2 + d\epsilon$, $c_1 < 0$, $d \geq 0$. If there are non-weak sites in $B_2^{(1)} \setminus B_1^{(1)}$ with respect to $B_1^{(1)}$, then $d > 0$. Hence $p_1 > p_2$ must prevail at the optimum. QED

Remarks. 1. Intuitively the result says that if neither of two even blocks gives more subsequent choice on the odd lattice, for maximum entropy one should weight them equally. Otherwise one should favor the one giving more choice on the odd lattice.

2. One can readily see some chains imposed by the order in Figure 4: $0 \prec 12 \prec 23 \prec 31$ or $0 \prec 11 \prec 21/22 \prec 33$ etc. The monotonicity can be utilized in limiting the number of $n \times n$ -blocks optimized for larger values of n (dropping blocks with least probability as dictated by the Theorem and with least multiplicity (most symmetric)).

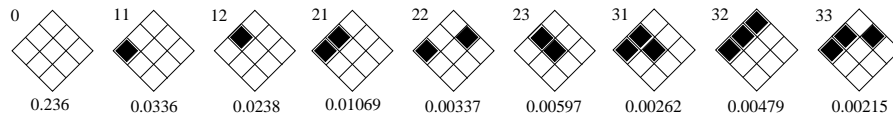


Fig. 4: Prevalent 3×3 -blocks with optimal probabilities without multiplicities.

The correlation structure inside the measure of maximal entropy gradually presents itself in the Bernoulli approximations when we consider higher order blocks. Correlations between the blocks are zero because of independence, but within the blocks it is worth making comparisons.

By adding the optimum probabilities of all 3×3 blocks at a given density level $k/9 = 0, 1/9, \dots, 1$ we obtain the “density profile” of this measure (here k is the number of 1’s in the block).

Suppose next that we generate the 3×3 blocks from 1×1 Bernoulli entries with the appropriate optimal p for 1’s (as found above). By adding these up we again obtain a density profile, this time for the 1×1 optimal Bernoulli measure at the resolution level of the block size 3×3 . The 3×3 blocks can of course be generated using the optimal 2×2 blocks as well and yet another density profile results. These three discrete plots are rendered as curves in Figure 5.

Perhaps the most notable feature here is the flattening of the distributions, as the block size increases i.e. the total block probabilities move towards the tails (while their means stay constant around 0.22). The curves cross between density levels $1/3 - 4/9$: below this cross over the shorter range Bernoulli measures favor light 3×3 blocks, above it they discount heavier blocks in comparison to the optimal 3×3 Bernoulli measure.

When examined closer one will see that the total probability of 3×3 blocks at a given density level essentially comes from at most three different kinds of local configurations (up to reductions above that is). These seem to be “grown”: when moving from density level d to level $d + 1/9$ the high probability blocks are generated by adding a (contiguous) 1 into an existing high probability block. This mechanism cannot prevail when the 3×3 blocks are generated independently from smaller blocks. Consequently the small block curves in Figure 5. have suppressed tails. We expect this phenomenon to prevail in the higher order

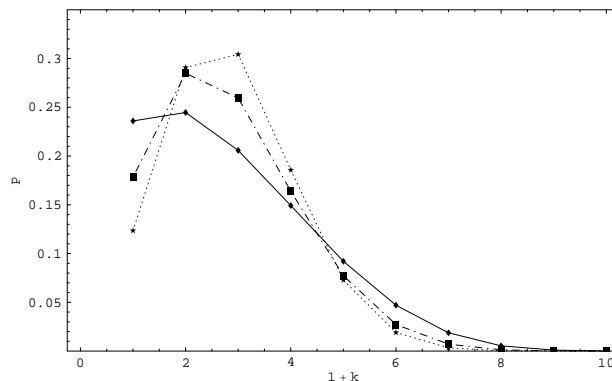


Fig. 5: 3×3 -block occupation probabilities from Bernoulli blocks of size 3×3 (diamond), 2×2 (square) and 1×1 (star). $k \in \{0, 1, \dots, 9\}$ is the number of 1's in the block.

Bernoulli blocks as well and thereby to be a significant feature in the long range correlations of the measure of maximal entropy.

Acknowledgments The author would like to thank the referees for valuable suggestions that led to the improvement of the exposition.

References

1. Baxter, R.J.: Exactly solvable models in statistical mechanics. Academic Press (1982)
2. Baxter, R.J.: Planar lattice gasses with nearest-neighbor exclusion. *Annals of Combinatorics* 3, no. 2-4, 191–203 (1999), MR1772345 (2001h:82010)
3. van den Berg, J., Steif, J.: Percolation and the hard-core lattice gas model. *Stoc. Proc. and Appl.* 49, 179-197 (1994)
4. Calkin, N., Wilf, H.: The number of independent sets in a grid graph. *SIAM J. Discr. Math.*, 11, 54-60 (1998)
5. Dobrushin, R., Shlosman, S. in *Statistical Physics and Dynamical System.*, J. Fritz, A. Jaffe and D. Szasz, eds., pp. 347-370, Birkhäuser (1985)
6. Eloranta, K.: Dense packing on uniform lattices. *Journal of Stat. Phys.* 130, no. 4, 741-55 (2008), [arXiv:math-ph/0907.4247](https://arxiv.org/abs/math-ph/0907.4247), MR2387564 (2009a:52016)
7. Engel, K.: On the Fibonacci number of an $M \times N$ lattice. *Fibonacci Quarterly*, 28, 1, (1990)
8. Forchhammer, S., Justesen, J.: Entropy bounds for constrained two-dimensional random fields. *IEEE Tr. Info. Theory*, 45, 1, 118-127 (1999)
9. Milosevic, S., Stosic, B., Stosic, T.: Towards finding exact residual entropies of the Ising ferromagnets. *Physica A*, 157, pp. 899-906 (1989)
10. Radulescu, D., Styer, D.: The Dobrushin-Shlosman Phase Uniqueness Criterion and Applications to Hard Squares. *Journal of Stat. Phys.* 49, 281-95 (1987), MR0923859 (89d:82010)
11. Walters, P.: *Ergodic theory*. Springer (1982)