

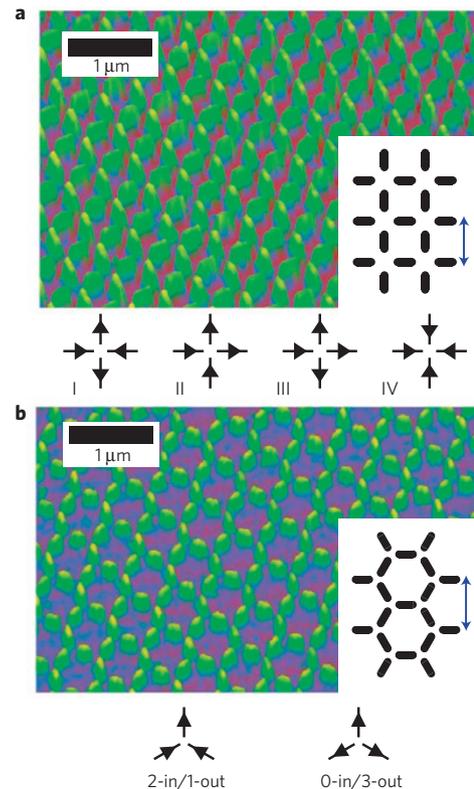
# Direct entropy determination and application to artificial spin ice

Paul E. Lammert<sup>1\*</sup>, Xianglin Ke<sup>1</sup>, Jie Li<sup>1</sup>, Cristiano Nisoli<sup>1,2</sup>, David M. Garand<sup>1</sup>, Vincent H. Crespi<sup>1</sup> and Peter Schiffer<sup>1</sup>

From thermodynamic origins, the concept of entropy has expanded to a range of statistical measures of uncertainty, which may still be thermodynamically significant<sup>1,2</sup>. However, laboratory measurements of entropy continue to rely on direct measurements of heat. New technologies that can map out myriads of microscopic degrees of freedom suggest direct determination of configurational entropy by counting in systems where it is thermodynamically inaccessible, such as granular<sup>3–8</sup> and colloidal<sup>9–13</sup> materials, proteins<sup>14</sup> and lithographically fabricated nanometre-scale arrays. Here, we demonstrate a conditional-probability technique to calculate entropy densities of translation-invariant states on lattices using limited configuration data on small clusters, and apply it to arrays of interacting nanometre-scale magnetic islands (artificial spin ice)<sup>15</sup>. Models for statistically disordered systems can be assessed by applying the method to relative entropy densities. For artificial spin ice, this analysis shows that nearest-neighbour correlations drive longer-range ones.

Our artificial spin-ice<sup>15</sup> systems are arrays of lithographically defined single-domain ferromagnetic islands (25 nm thick and 220 nm × 80 nm in area) on the links of square and honeycomb lattices (Fig. 1). Shape anisotropy forces island moments to point along the long axes, forming effective Ising spins. The coercive field is about 770 Oe (that is, a barrier of order 10<sup>5</sup> K), whereas the field from one island on a neighbour is only of order 10 Oe. The arrays are demagnetized by rotating in an in-plane external magnetic field  $H_{\text{ext}}$ , initially strong enough to produce complete polarization, subsequently reduced to zero in small increments  $\Delta H_{\text{ext}}$ , with reversal of the field at each step. For small step sizes, the result is a statistically reproducible macrostate, operationally defined by the demagnetization protocol<sup>15,16</sup>, which is probed by magnetic force microscopy to obtain the static moments of individual islands. We want the entropy of a single macrostate, but distinct runs might produce distinct macrostates (for example, a residual magnetization at larger step size). In most cases, data are collected in a single run, averting the problem, but the entropy of a macrostate mixture would be relatively unimportant anyway, as is shown in Supplementary Section S3. For large structurally regular systems such as ours, it is more appropriate to work not with total entropy, but with entropy density (see equation (1)) having units of bits per island, a value of 1 corresponding to complete disorder.

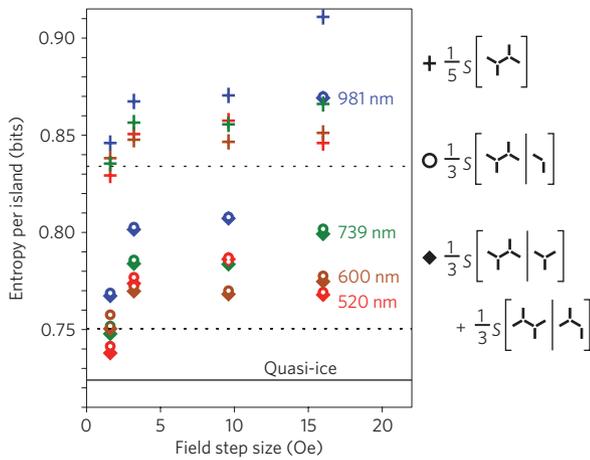
The strongest interactions, between islands meeting at a vertex, favour head-to-tail moment alignment. However, not all of these can be satisfied simultaneously, resulting in a kind of frustration. Still, for the square lattice, the ground state is only twofold degenerate<sup>15</sup>, because type-I vertices, as defined in Fig. 1, are lowest in energy. That the ordered ground state is never found experimentally<sup>16,17</sup> suggests that the evolution is kinetically



**Figure 1 | Artificial spin-ice arrays.** **a**, Atomic force micrograph of a square lattice of lattice spacing 400 nm, with inset showing lattice spacing, and schematics of vertex types. **b**, Similar to **a**, for a honeycomb lattice of 520 nm lattice spacing.

constrained<sup>18,19</sup>. For instance, one spin flip converts a type-II to a type-III; flips of two perpendicular islands are required to reach type-I. In contrast to the square lattice, the honeycomb lattice has a macroscopically degenerate ground state when only nearest- or next-nearest-neighbour interactions are effective (longer-range interactions break the degeneracy<sup>20</sup> at a much lower energy scale). The interactions prefer a 2-in/1-out or 1-in/2-out arrangement at every vertex (quasi-ice rule). This constraint alone produces a state, ideal quasi-ice, with an entropy density of 0.724 bits per island. Interactions between (mono)pole strengths  $Q$  at neighbouring vertices<sup>21,22</sup> reduce the ground-state degeneracy to 0.15 bits per island by favouring  $Q = -1$  (2-in/1-out) next to  $Q = +1$ . The contrast between the square- and honeycomb-lattice ground states—twofold degenerate versus macroscopically degenerate—provides an

<sup>1</sup>Department of Physics and Materials Research Institute, The Pennsylvania State University, 104 Davey Lab, University Park, Pennsylvania 16802-6300, USA, <sup>2</sup>CNLS and T-Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA. \*e-mail: pel1@psu.edu.



**Figure 2 | Entropy density upper bounds for honeycomb artificial spin ice at four lattice constants as a function of the demagnetization step size  $\Delta H_{\text{ext}}$ .** All bounds are derived using configuration statistics for the five-island cluster  $\Lambda_5$  shown on the right. Crosses are the direct estimate  $S(\Lambda_5)/5$ , whereas filled diamonds use inequality (5) and open circles use inequality (4). The dashed lines are the result of our technique applied to ideal quasi-ice (every vertex type-I with no other restrictions), the upper from the simple cluster-estimate, and the lower two (indistinguishable) from the two conditioning approximations. The simple cluster-estimate applied to a single vertex would give 0.86 bits per island. The solid line at 0.724 shows the actual entropy density of ideal quasi-ice. At the smallest lattice constants and smallest step sizes, subtle signs of longer-ranged monopole correlations appear.

opportunity to investigate the interplay between the strictures of kinetic constraint and the freedom of massive degeneracy.

We now develop a method to extract the entropy densities on our lattices from the measured configurations of the island magnetic moments. Consider a finite cluster  $\Lambda$  of islands, for example, the five-island cluster comprising two adjacent vertices (see Fig. 2 legend) and the collection of random variables  $\sigma_\Lambda$  that are the spins belonging to  $\Lambda$ . The Shannon–Gibbs–Boltzmann entropy<sup>23–25</sup> of  $P_\Lambda(\sigma_\Lambda)$ , the distribution of  $\sigma_\Lambda$ , is

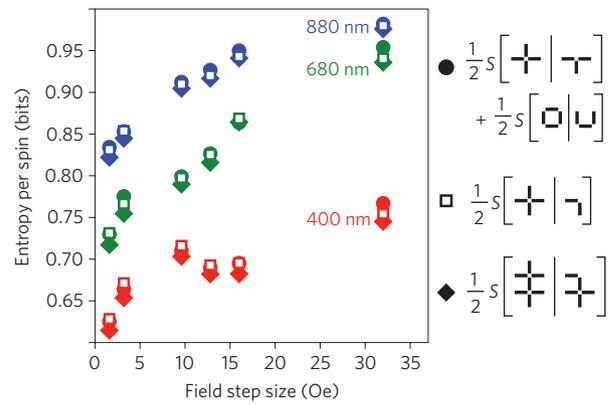
$$S(P_\Lambda) = - \sum_{\sigma_\Lambda} P(\sigma_\Lambda) \log_2 P(\sigma_\Lambda)$$

where the sum runs over all possible values of the random variable(s)  $\sigma_\Lambda$ . Note that  $S$  is rendered dimensionless by omitting Boltzmann’s constant, and the base of the logarithm is 2, so that the units are bits. If  $\Lambda$  is taken ever larger while the fraction of islands on the edge tends to zero (van Hove limit), we obtain the bulk entropy density  $s$ :

$$s = \lim_{|\Lambda| \rightarrow \infty} \frac{S(P_\Lambda)}{|\Lambda|} \quad (1)$$

If each island moment independently points either way with probability 1/2, then the entropy density is one bit per island, the largest possible. Lower entropy density indicates correlations in a generic sense. For example, the fully polarized initial state created by a large  $H_{\text{ext}}$  has zero entropy density.

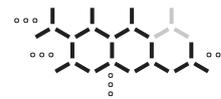
The obvious approximation to  $s$  suggested by equation (1) is simply  $S(P_\Lambda)/|\Lambda|$  for the biggest practicable cluster. However, this ‘simple cluster-estimate’ is not very good because the configuration space grows exponentially with cluster size  $|\Lambda|$ , and boundary-crossing correlations are completely neglected. To understand the last point, suppose the entire lattice covered without gaps or overlap by translates of  $\Lambda$ . The state constructed from the marginals of  $P$  on those translates, taking them to be independent, has entropy density exactly  $S(P_\Lambda)/|\Lambda|$ . However, short-range



**Figure 3 | Entropy density upper bounds for square-lattice artificial spin ice.** The three approximations agree closely. Even extrapolated to zero step size, a.c.-demagnetized square spin ice never approaches the ground state. The diamond estimate is always lower than the square because the former retains more conditioning, whereas the added islands are the same. Data-set sizes are reported in Supplementary Information.

boundary-crossing correlations are the same as corresponding intracluster correlations, so are reflected in small-cluster data and can be properly counted using conditional entropy. The method resembles one proposed some years ago<sup>26,27</sup> for Monte Carlo simulations of lattice-spin models in equilibrium.

One way to think of the total entropy of a given macrostate is as the average uncertainty about the particular microstate at hand. Imagine a microstate of the honeycomb lattice revealed three islands (one vertex) at a time, row-by-row. One instant in the process looks like this (the grey vertex is about to be revealed):



Neglecting the (far distant) lattice edge, each newly revealed vertex bears the same spatial relation to those already known, so the revelation, on average, reduces the uncertainty by exactly three times the entropy per spin. Cast this way, the entropy density appears as a conditional entropy<sup>28</sup>. The conditional entropy of  $\sigma_\Lambda$  given  $\sigma_\Gamma$  is

$$S(\sigma_\Lambda | \sigma_\Gamma) = - \sum_{\sigma_\Lambda, \sigma_\Gamma} P(\sigma_\Lambda, \sigma_\Gamma) \log_2 P(\sigma_\Lambda | \sigma_\Gamma) \quad (2)$$

where  $P(\sigma_\Lambda | \sigma_\Gamma) = P(\sigma_\Lambda, \sigma_\Gamma) / P(\sigma_\Gamma)$  is the conditional probability of  $\sigma_\Lambda$  given  $\sigma_\Gamma$ . The joint entropy of  $\sigma_\Lambda$  and  $\sigma_\Gamma$  then has the pleasant decomposition  $S(\sigma_\Lambda, \sigma_\Gamma) = S(\sigma_\Lambda | \sigma_\Gamma) + S(\sigma_\Gamma)$ . (Learning  $\sigma_\Gamma$  and  $\sigma_\Lambda$  at once is the same as learning  $\sigma_\Gamma$  and then  $\sigma_\Lambda$ .) Note that if  $\Lambda$  and  $\Gamma$  overlap, common spins contribute zero to  $S(\sigma_\Lambda | \sigma_\Gamma)$ .

As a simple illustration, suppose  $\Lambda$  and  $\Gamma$  are single islands, with the probabilities for  $P(\sigma_\Lambda, \sigma_\Gamma)$  being given by  $P(\downarrow, \uparrow) = 0$  and  $P(\uparrow, \uparrow) = P(\uparrow, \downarrow) = P(\downarrow, \downarrow) = 1/3$ . If we know that  $\sigma_\Gamma = \uparrow$ , then the remaining uncertainty about  $\sigma_\Lambda$  is zero, but if we know that  $\sigma_\Gamma = \downarrow$ , then the uncertainty is total: 1 bit. Weighting by the probabilities of  $\sigma_\Gamma$  to be  $\uparrow$  or  $\downarrow$  gives the entropy of  $\sigma_\Lambda$  conditioned on  $\sigma_\Gamma$ :  $P(\sigma_\Gamma = \uparrow) \cdot (0) + P(\sigma_\Gamma = \downarrow) \cdot (1) = 2/3$  bit.

The Methods section explains how conditional entropy and other basic notions of information theory can be used to obtain good approximations to the entropy density  $s$  from limited data. The result of applying two such approximations to the experimental data for honeycomb lattices is plotted in Fig. 2 as a function of field step size  $\Delta H_{\text{ext}}$  for each lattice constant, along with one simple cluster-estimate for comparison. Data-set sizes are reported

**Table 1 | Relative entropy: experiment versus maximum-entropy square ice.**

Lattice (nm)	$s(\text{expt}  \text{ME})$ ( $10^{-3}$ bits per island)				
	$\Delta H_{\text{ext}}$ (Oe)				
	1.6	3.2	9.6	12.8	16
400	5.1	4.2	13.5	2.9	9.2
680	4.5	5.6	9.1	5.4	5.0
880	4.4	16.1	7.9	1.4	2.6

in Supplementary Section S1. The simple cluster-estimate  $S(\Lambda)/|\Lambda|$  using the five-island di-vertex (Fig. 2 legend) provides a very poor bound compared with our conditioning technique. Reducing the lattice constant or step size should lower the entropy because the first leads to stronger interactions, and the second gives interactions a better chance to be the decisive factor for island flips. The expected lattice constant trend is seen but there is an unexpected plateau with respect to  $\Delta H_{\text{ext}}$ .

It makes sense to compare the experimental states with ideal honeycomb quasi-ice through entropy densities. That of honeycomb quasi-ice is 0.724 bits per island (Supplementary Section S2). However, proper comparison requires using the same estimates for both systems. The dashed lines in the plot show the estimates for the model system, the upper for the simple cluster-estimate (compare crosses) and the lower two (indistinguishable) for the conditioning estimates. Hence, the 520 nm array at  $\Delta H_{\text{ext}} = 1.6$  Oe has lower entropy than ideal quasi-ice. This can be explained by correlations between net magnetic charge  $Q = \pm 1$  of nearest-neighbour vertices. Ideal quasi-ice has a weak anticorrelation:  $\langle Q_i Q_j \rangle \approx -1/9$ . In some experimental samples, this correlation reaches  $-0.25$ , reflected in a small entropy decrease. Complete sublattice ordering,  $\langle Q_i Q_j \rangle = -1$ , would reduce the entropy to  $s \approx 0.15$  bits per island (Supplementary Section S2). This extra correlation may explain the slightly better performance of the bound with a complete vertex in the conditioning data.

The entropy of honeycomb artificial spin ice reveals a state close to ideal quasi-ice, with slight antiferromagnetic vertex charge ordering. The contrasting failure of a.c. demagnetization of the square lattice to approach the completely ordered ground state is precisely quantified by entropy. We use three approximations (upper bounds) for the square-lattice entropy density. They are found by a procedure parallel to that for the honeycomb lattice and are shown in the legend of Fig. 3. The three agree well, and this rough convergence test suggests that they are close to the true entropy densities. As for the honeycomb lattice, we expect smaller entropy for smaller  $\Delta H_{\text{ext}}$  or smaller lattice spacing. In general, this seems to be the case, but the ground state is never approached. Even extrapolations  $\Delta H_{\text{ext}} \rightarrow 0$  have large entropy densities.

Closer inspection suggests jamming at  $\Delta H_{\text{ext}}$ . A kinetically constrained approach to ground states defines the behaviour of complex systems across many fields<sup>19</sup>, such as protein folding<sup>14</sup>, self-assembly, glasses and granular systems<sup>3-8</sup>. Ergodicity is thwarted by both tall energy barriers and configuration-space constrictions, combined into free-energy barriers. An ergodic system explores all of configuration space, whereas folding proteins live within a 'folding funnel'. This dynamic constriction of allowed configurations introduces many deep conceptual challenges. a.c.-demagnetized artificial spin ice puts the conceptual challenge of kinetic constraint into sharp relief: as the rotating external field weakens, islands one-by-one fall out of 'field following' mode, driven by inter-island interactions that suppress the local depolarization field and lock in the orientation of the fallen-away island. Thus, each spin probably makes only a single decision on how to point on escaping coercion, with no prospect of later surmounting barriers.

The system's approach to the ground state is essentially one-way. Thus, it is not surprising that only a macroscopically degenerate ground-state target can be 'hit'. Notwithstanding this failure of ergodicity, square-lattice artificial spin ice can still be described by a statistical model based, like thermodynamics, on maximum entropy<sup>16,17</sup>. As an extreme case of kinetic constraint, rotationally annealed artificial spin ice can afford unique insights into statistical mechanics of complex systems. For example, array topology can control the ground-state degeneracy, as seen here.

Even without detailed knowledge of how the final square-lattice states develop, a concisely descriptive model may be sought. We conjecture that the lattice state is fully determined by correlations between nearest-neighbour pairs diagonally or straight across a vertex, and thus model it by a constrained maximum-entropy state, which is as random as possible, consistent with those correlations. Adapting the conditioning techniques (see the Methods section), we can efficiently estimate the entropy density of the experimental states relative to the maximum-entropy state,  $s(\text{expt}||\text{ME})$ . This global measure of dissimilarity does not depend on identifying the 'right' correlations, and allows an assessment of the model. The results are given in Table 1.

Note that<sup>29</sup> the probability that a typical experimental state in region  $\Lambda$  is likely to be mistaken for a maximum-entropy state decays asymptotically as  $\exp[-|\Lambda|s(\text{expt}||\text{ME})]$ . Apparently, the entropy reduction below independent islands in the square lattices is well accounted for by nearest-neighbour correlations, and those they entail.

The reduction in the entropy of an interacting system below that of uncoupled degrees of freedom is due mostly to short-range correlations, even near a critical point. Thus, the efficient conditional entropy technique described here can be applied to a wide variety of resolvable complex systems such as granular media and colloidal systems that can now be spatially resolved in the required detail<sup>9-13</sup>. Entropy density is a general measure of order that is not tied to pre-identified correlations. Hence, it is especially valuable for states such as square ice or other jammed, glassy states that are far from identifiable landmarks.

## Methods

According to the discussion around equation (2), the entropy density  $s$  of the infinite honeycomb lattice is given by (ignore the colour for now)

$$3s = S \left[ \begin{array}{c} \dots \\ \dots \\ \dots \end{array} \left| \begin{array}{c} \dots \\ \dots \\ \dots \end{array} \right. \right] \quad (3)$$

Now we find small-cluster approximations to this entropy density, using two principles<sup>28</sup>. (1) If  $\sigma_\Gamma$  is known there is no uncertainty about it, so  $S(\sigma_\Lambda, \sigma_\Gamma | \sigma_\Gamma) = S(\sigma_\Lambda | \sigma_\Gamma)$  for arbitrary  $\Lambda$  and  $\Gamma$ . Thus, in pictorial equations, visual perspicuity will dictate retention or omission of conditioning variables on the left of the bar. (2) Providing more conditioning information lessens uncertainty:  $S(\sigma_\Lambda | \sigma_\Gamma) \geq S(\sigma_\Lambda | \sigma_\Gamma, \sigma_\Gamma)$ . Unlike a simple application of equation (1), our conditional entropy method fully accounts for short-range correlations without boundary error. Like the simple cluster-estimate, it provides upper bounds on the true entropy density. Dropping all but the red islands in equation (3) yields

$$3s \leq S \left[ \begin{array}{c} \dots \\ \dots \\ \dots \end{array} \left| \begin{array}{c} \dots \\ \dots \\ \dots \end{array} \right. \right] \quad (4)$$

Alternatively, if we add a vertex in two stages rather than all at once as in equation (3), we immediately arrive at

$$3s \leq S \left[ \begin{array}{c} \dots \\ \dots \\ \dots \end{array} \left| \begin{array}{c} \dots \\ \dots \\ \dots \end{array} \right. \right] + S \left[ \begin{array}{c} \dots \\ \dots \\ \dots \end{array} \left| \begin{array}{c} \dots \\ \dots \\ \dots \end{array} \right. \right]$$

