Direct Entropy Determination and Application to Artificial Spin Ice

Paul E. Lammert[1], Xianglin Ke[1], Jie Li[1], Cristiano Nisoli[1,2],
David M. Garand[1], Vincent H. Crespi[1] and Peter Schiffer[1]

[1] *Department of Physics and Materials Research Institute,*
The Pennsylvania State University, 104 Davey Lab,
University Park, PA 16802-6300, USA
[2] *CNLS and T-Division, Los Alamos National Laboratory,*
Los Alamos, NM 87545, USA

From thermodynamic origins, the concept of entropy has expanded to a range of statistical measures of uncertainty, which may still be thermodynamically significant[1, 2]. But, 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[3–8] and colloidal[9–13] materials, proteins[14] and lithographically fabricated nanoscale 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 nanoscale magnetic islands ("artificial spin-ice"[15]). 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 neighbor correlations drive longer-range ones.

Our artificial spin ice[15] 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 (i.e., a barrier of order 10^5 K), while the field from one island on a neighbor only of order 10 Oe (10^4 K). The arrays are demagnetized by rotating in an in-plane external magnetic field H_{ext}, initially strong enough to produce complete polarization, subsequently reduced to zero in small increments[15, 16] ΔH_{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[15, 16], 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
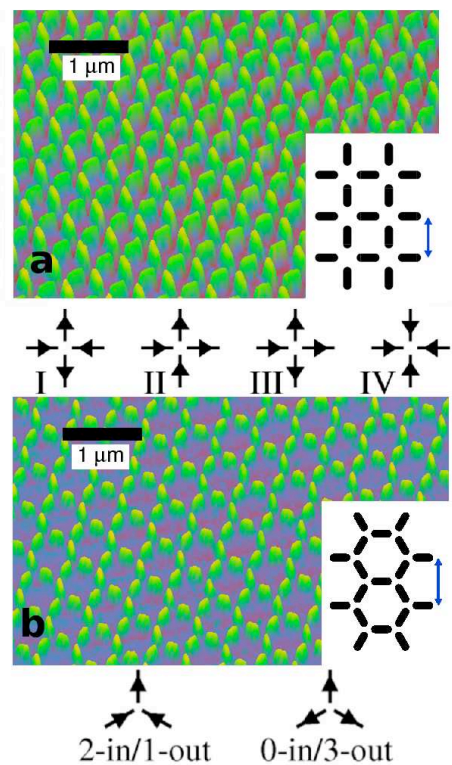

FIG. 1: **Artificial spin-ice arrays. a**. AFM image of 400 nm square lattice with inset showing lattice spacing, and schematics of vertex types. **b**. Similarly for 520 nm honeycomb lattice.

collected in a single run, averting the problem, but the entropy of macrostate mixture would be relatively unimportant anyway, as is shown in Supplementary Information §3. For large structurally regular systems such as ours, it is more appropriate to work not with total entropy, but with entropy density [See Eq. (2)] having units of bits/island, a value of 1 corresponding to complete disorder.

The strongest interactions, between islands meeting at a vertex, favor head-to-tail moment alignment. But not all these can be satisfied simultaneously, resulting in a kind of frustration. Still, for the square lattice,

the ground state is only two-fold degenerate[15], since Type-I vertices, as defined in Fig. 1, are lowest in energy. That the ordered ground state is never found experimentally[16, 17] suggests that the evolution is kinetically constrained[18, 19]. 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 neighbor interactions are effective (longer-range interactions break the degeneracy[20] 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/island. Interaction between (mono)pole-strengths Q at neighboring vertices[21, 22] reduce the ground state degeneracy to 0.15 bits/island by favoring $Q = -1$ (2-in/1-out) next to $Q = +1$. The contrast between the square and honeycomb lattice ground states – two-fold degenerate versus macroscopically degenerate – provides an 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 Λ of islands, for example, the 5-island cluster comprising two adjacent vertices (See Fig. 2 legend) and the collection of random variables σ_Λ which are the spins belonging to Λ. The Shannon(-Gibbs-Boltzmann) entropy[23–25] of $P_\Lambda(\sigma_\Lambda)$, the distribution of σ_Λ, is

$$S(P_\Lambda) = -\sum_{\sigma_\Lambda} P(\sigma_\Lambda) \log_2 P(\sigma_\Lambda), \qquad (1)$$

where the sum runs over all possible values of the random variable(s) σ_Λ. 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 Λ 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:
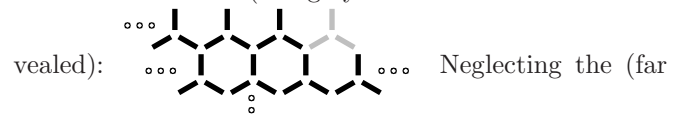
$$s = \lim_{\Lambda \nearrow \infty} \frac{S(P_\Lambda)}{|\Lambda|}. \qquad (2)$$

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_{ext} has zero entropy density.

The obvious approximation to s suggested by Eq. (2) is simply $S(P_\Lambda)/|\Lambda|$ for the biggest practicable cluster. But this "simple cluster-estimate" is not very good because the configuration space grows exponentially with cluster size $|\Lambda|$, while boundary-crossing correlations are

completely neglected. To understand the latter point, suppose the entire lattice covered without gaps or overlap by translates of Λ. The state constructed from the marginals of P on those translates, *taking them independent*, has entropy density exactly $S(P_\Lambda)/|\Lambda|$. However, short-range boundary-crossing correlations are the same as corresponding intra-cluster correlations, so are reflected in small-cluster data and can be properly counted using *conditional entropy*. The method resembles one proposed some years ago[26, 27] 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 3 times the entropy per spin. Cast this way, the entropy density appears as a *conditional* entropy [28]. The conditional entropy of σ_Λ given σ_Γ is

$$S(\sigma_\Lambda | \sigma_\Gamma) = -\sum_{\sigma_\Lambda, \sigma_\Gamma} P(\sigma_\Lambda, \sigma_\Gamma) \log_2 P(\sigma_\Lambda | \sigma_\Gamma), \qquad (3)$$

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

As a simple illustration, suppose Λ and Γ 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 σ_Λ is zero, but if we know that $\sigma_\Gamma = \downarrow$, then the uncertainty is total: 1 bit. Weighting by the probabilities of σ_Γ to be $\uparrow$ or $\downarrow$ gives the entropy of σ_Λ conditioned on σ_Γ: $P(\sigma_\Gamma = \uparrow) \cdot (0) + P(\sigma_\Gamma = \downarrow) \cdot (1) = 2/3$ 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 are plotted in Fig. 2 as a function of field step size ΔH_{ext} for each lattice constant, along with one simple cluster-estimate for comparison. Data-set sizes are reported in Supplementary Information §1. The simple cluster-estimate $S(\Lambda)/|\Lambda|$ using the five-island di-vertex (Fig. 2 legend) provides a very poor bound compared to our conditioning technique. Reducing lattice constant or
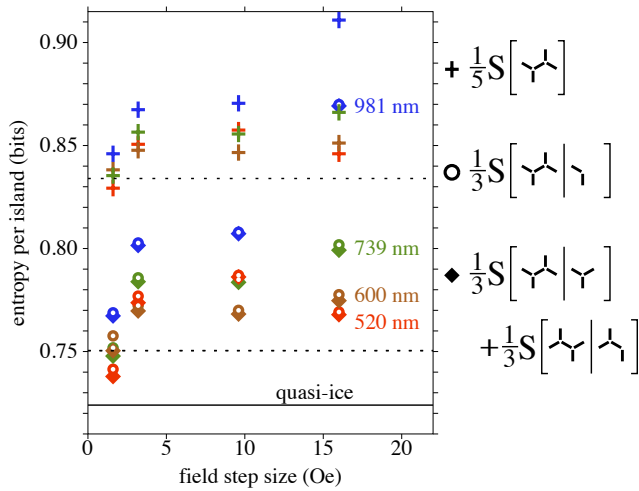
FIG. 2: **Entropy density upper bounds for honeycomb artificial spin ice at four lattice constants as a function of the demagnetization step size** ΔH_{ext}. All bounds are derived using configuration statistics for the 5-island cluster Λ_5 shown in the inset. Crosses are the direct estimate $S(\Lambda_5)/5$, while filled diamonds use inequality (7) and open circles use inequality (5). 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 bit/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.
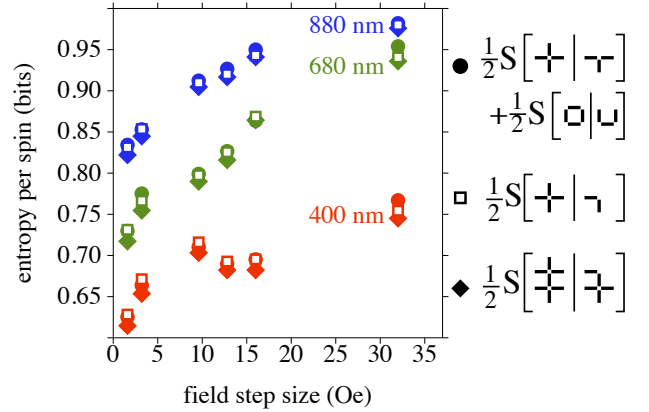
FIG. 3: **Entropy density upper bounds for square lattice artificial spin ice.** The three approximations agree closely. Even extrapolated to zero step size, AC demagnetized square spin ice never approaches the ground state. The diamond estimate is always lower than the square because the former retains more conditioning, whereas the added islands are the same. Dataset sizes are reported in Supplementary Information.

step size should lower the entropy since 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 ΔH_{ext}.

It makes sense to compare the experimental states to ideal honeycomb quasi-ice through entropy densities. That of honeycomb quasi-ice is 0.724 bit/island (Supplementary Information §2). But proper comparison requires using the same estimates for both systems. 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 *less* entropy than ideal quasi-ice. This can be explained by correlations between net magnetic charge $Q = \pm1$ of nearest-neighbor vertices. Ideal quasi-ice has a weak anticorrelation: $\langle Q_i Q_j \rangle \approx -1/9$. In some 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 $s \approx 0.15$ bit/island (Supplementary Information §2). 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 AC 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 ΔH_{ext} or smaller lattice spacing. In general, this seems to be the case, but the ground state is never approached. Even extrapolations $\Delta H_{\text{ext}} \to 0$ have large entropy densities.

Closer inspection suggests jamming at ΔH_{ext}. Kinetically constrained approach to ground states defines behavior of complex systems across many fields[19], such as protein folding[14], self-assembly, glasses and granular systems[3–8]. 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. AC demagnetized artificial spin ice puts the conceptual challenge of kinetic constraint into sharp re-

lief: 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 likely makes only a single decision on how to point upon 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[16, 17]. 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-neighbor 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 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. Results are given in the following table.

$s(\text{expt}||\text{ME})$ (10^{-3} bit/island)

lattice (nm)	ΔH_{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

Note that[29] the probability that a typical experimental state in region Λ 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-neighbor 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 which can now be spatially resolved in the required detail [9–13]. Entropy density is a general measure of order which is not tied to pre-identified correlations. Hence it is especially valuable for states such as square ice or other jammed, glassy states which are far from identifiable landmarks.

METHODS

According to the discussion around Eq. (3), the entropy density s of the infinite honeycomb lattice is given by (ignore the color for now)

$$3s = S\left[\ \vcenter{\hbox{}}\ \middle|\ \vcenter{\hbox{}}\ \right] \tag{4}$$

Now we find small-cluster approximations to this entropy density, using two principles[28]. (A) If σ_Γ is known there is no uncertainty about it, so $S(\sigma_\Lambda, \sigma_\Gamma | \sigma_\Gamma) = S(\sigma_\Lambda | \sigma_\Gamma)$ for arbitrary Λ and Γ. Thus, in pictorial equations, visual perspicuity will dictate retention or omission of conditioning variables on the left of the bar. (B) Providing more conditioning information lessens uncertainty: $S(\sigma_\Lambda | \sigma_\Gamma) \geq S(\sigma_\Lambda | \sigma_\Gamma, \sigma_{\Gamma'})$. Unlike a simple application of Eqn. (2), 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 Eqn. (4) yields

$$3s \leq S\left[\ \vcenter{\hbox{}}\ \middle|\ \vcenter{\hbox{}}\ \right] \tag{5}$$

Alternatively, if we add a vertex in two stages rather than all at once as in Eq. (4), we immediately arrive at

$$3s = S\left[\ \vcenter{\hbox{}}\ \middle|\ \vcenter{\hbox{}}\ \right] + S\left[\ \vcenter{\hbox{}}\ \middle|\ \vcenter{\hbox{}}\ \right] \tag{6}$$

The red islands again provide a visual cue. Following it, we get the bound

$$3s \leq S\left[\ \vcenter{\hbox{}}\ \middle|\ \vcenter{\hbox{}}\ \right] + S\left[\ \vcenter{\hbox{}}\ \middle|\ \vcenter{\hbox{}}\ \right] \tag{7}$$

Bounds for a translation-invariant state on the square lattice can be obtained in a similar way. We use the one-step bounds

$$2s \leq S\left[\ \vcenter{\hbox{}}\ \middle|\ \vcenter{\hbox{}}\ \right] \tag{8}$$

and

$$2s \leq S \left[\;\boxed{} \; \middle| \; \boxed{} \; \right] \tag{9}$$

and the two-step bound

$$2s \leq S \left[\;\boxed{} \; \middle| \; \boxed{} \; \right] + S \left[\;\boxed{} \; \middle| \; \boxed{} \; \right] \tag{10}$$

A constrained maximum entropy state on the square lattice with given correlations between diagonal and accross-the-vertex nearest neighbor correlations coincides with a Gibbs state for an Ising model with effective pair interactions for the two types of nearest neighbors of whatever strength is required to reproduce the required correlations. We have previously[16] studied specific pair correlations in such maximum entropy states using Monte Carlo simulation. Building on the conditioning techniques developed in this Letter, we compute $s(\text{expt}||\text{ME})$, the relative entropy density of an experimental state to the corresponding constrained maximum entropy state. In the case of two probability measures on the configuration space of a lattice system, the relative entropy of their restrictions to some finite region Λ is

$$
\begin{aligned}
S(Q_\Lambda || P_\Lambda) &= \sum_{\sigma_\Lambda} \log_2 \left(\frac{Q(\sigma_\Lambda)}{P(\sigma_\Lambda)} \right) Q(\sigma_\Lambda) \\
&= \left\langle \log_2 Q_\Lambda \right\rangle_Q - \left\langle \log_2 P_\Lambda \right\rangle_Q .
\end{aligned} \tag{11}
$$

The limiting relative entropy density which we want is

$$s(Q_\Lambda || P_\Lambda) = \lim_{\Lambda \nearrow \infty} |\Lambda|^{-1} S(Q_\Lambda || P_\Lambda). \tag{12}$$

The logarithm of the probability in Eq. (11) can be expanded in terms of conditionals just as was done for the entropy. For any collection $\{X_1, \ldots, X_N\}$ of random variables (the $m = 1$ term being read as an unconditioned probability),

$$\log_2 P(X_N, \ldots, X_1) = \sum_{m=1}^{N} \log_2 P(X_m | X_{m-1}, \ldots, X_1)$$

parallels exactly the formula

$$S(X_N, \ldots, X_1) = \sum_{m=1}^{N} S(X_m | X_{m-1}, \ldots, X_1).$$

The main difference is that $\log_2 P(\cdot)$ is a random variable, whereas $S(\cdot)$ is a number. Any of the class of approximations for the conditional entropy densities can now be applied to the conditional probabilities to obtain the relative entropy. However, we do not get bounds in this way, just ordinary estimates.

If P_{ME} is a maximum entropy state constrained to have expectations of specified observables match their expectations in P, then the relative entropy density $s(P||P_{ME})$ ought to equal the difference in the absolute densities. We believe that the method we have used is superior to this simple subtraction because it suppresses the unwanted effects of fluctuations in the counting of low-probability configurations. Due to limited experimental data, configurations which would be expected to have only a few occurences may have none at all, which has an anomalously large effect in the subtraction.

---

[1] Landauer, R. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development* **5**, 183–191 (1961). Reprinted in Leff and Rex.

[2] Leff, H. S. and Rex, A. F., editors. *Maxwell's Demon 2: entropy, classical and quantum information, computing.* Institute of Physics, Bristol, 2nd edition, (2003).

[3] Liu, A. J. and Nagel, S. R. Jamming is not just cool any more. *Nature* **396**, 21–22 (1998).

[4] Coniglio, A. and Nicodemi, M. The jamming transition of granular media. *J. Phys.: Condens. Matter* **12**, 6601–6610 (2000).

[5] O'Hern, C. S., Sibert, L. E., Liu, A. J., and Nagel, S. R. Jamming at zero temperature and zero applied stress: the epitome of disorder. *Phys. Rev. E* **68**, 011306 (2003).

[6] Corwin, E. I., Jaeger, H. M., and Nagel, S. R. Structural signatures of jamming in granular media. *Nature* **435**, 1075–1078 (2005).

[7] Majumdar, T. S., Sperl, M., Luding, S., and Behringer, R. P. Jamming transition in granular systems. *Phys. Rev. Lett.* **98**, 058001 (2007).

[8] Behringer, R. P., Daniels, K. E., Majumdar, T. S., and Sperl, M. Fluctuations, correlations and transitions in granular materials: statistical mechanics for a nonconventional system. *Philosophical Transactions of the Royal Society A* **366**, 493–504 (2008).

[9] Kegel, W. K. and van Blaaderen, A. Direct observation of dynamical heterogeneities in colloidal hard-sphere suspensions. *Science* **287**, 290–293 (2000).

[10] Weeks, E. R., Crocker, J. C., Levitt, A. C., Schofield, A., and Weitz, D. A. Three-dimensional direct imaging of structural relaxation near the colloidal glass transition. *Science* **287**, 627–631 (2000).

[11] Cui, B., Lin, B., and Rice, S. A. Dynamical heterogeneity in a dense quasi-two-dimensional colloidal liquid. *J. Chem. Phys.* **114**, 9142–9155 (2001).

[12] Gasser, U., Weeks, E. R., Schofield, A., Pusey, P. N., and Weitz, D. A. Real-space imaging of nucleation and growth in colloidal crystallization. *Science* **292**, 258–262 (2001).

[13] Alsayed, A. M., Islam, M. F., Zhang, J., Colllings, P. J., and Yodh, A. G. Premelting at defects within bulk colloidal crystals. *Science* **309**, 1207–1210 (2005).

[14] Pande, V. S., Grosberg, A. Y., and Tanaka, T. Heteropolymer freezing and design: Towards physical models of protein folding. *Reviews of Modern Physics* **72**, 259–314 (2000).

[15] Wang, R., Nisoli, C., Freitas, R. S., McConville, J. W., Cooley, B. J., Lund, M. S., Samarath, N., Leighton, C., Crespi, V. H., and Schiffer, P. Artificial 'spin ice' in a ge-

ometrically frustrated lattice of nanoscale ferromagnetic islands. *Nature* **439**, 303–306 (2006).

[16] Ke, X., Li, J., Nisoli, C., Lammert, P. E., McConville, W., Wang, R. F., Crespi, V. H., and Schiffer, P. Energy minimization and ac demagnetization in a nanomagnet array. *Phys. Rev. Lett.* **101**, 037205 (2008).

[17] Nisoli, C., Wang, R., Li, J., McConville, W. F., Lammert, P. E., Schiffer, P., and Crespi, V. H. Ground state lost but degeneracy found: the effective thermodynamics of artificial spin ice. *Phys. Rev. Lett.* **98**, 217203 (2007).

[18] Ritort, F. and Sollich, P. Glassy dynamics of kinetically constrained models. *Advances in Physics* **52**, 219–342 (2003).

[19] Janke, W., editor. *Rugged Free Energy Landscapes*, volume 736 of *Lecture Notes in Physics*. Springer-Verlag, Berlin, Heidelberg, (2008).

[20] Möller, G. and Moessner, R. Magnetic multipole analysis of kagome and artificial spin-ice dipolar arrays. *Phys. Rev. B* **80**, 140409(R) (2009).

[21] Tanaka, M., Saitoh, E., and Miyajima, H. Magnetic interactions in a ferromagnetic honeycomb nanoscale network. *Phys. Rev. B* **73**, 052411 (2006).

[22] Qi, Y., Brintlinger, T. E., and Cumings, J. Direct observation of the ice rule in an artificial kagomé spin ice. *Phys. Rev. B* **77**, 094418 (2008).

[23] Wehrl, A. general properties of entropy. *Rev. Mod. Phys.* **50**, 221–260 (1978).

[24] Penrose, O. *Foundations of Statistical Mechanics*. Pergamon, Oxford, (1970). Reprinted by Dover, 2005.

[25] Jaynes, E. T. Information theory and statistical mechanics. *Phys. Rev.* **106**, 620–630 (1957).

[26] Schlijper, A. G. and Smit, B. Two-sided bounds on the free energy from local states in monte carlo simulations. *J. Stat. Phys.* **56**, 247–259 (1989).

[27] Schlijper, A. G. and van Bergen, A. R. D. Local-states method for the calculation of free energies in monte carlo simulations of lattice nodels. *Phys. Rev. A* **41**, 1175–1178 (1990).

[28] Cover, T. M. and Thomas, J. A. *Elements of Information Theory*. Wiley, New York, (1991).

[29] Dembo, A. and Zeitouni, O. *Large Deviations Techniques and Applications*. Springer-Verlag, New York, 2nd edition, (1998).

Author contributions. P.S. and V.H.C. conceived the initial idea of this project, X.K., J.L. and D.M.G. performed the experiments and collected data. C.N. made theoretical contributions. P.E.L. analyzed data and developed theory.