Information flow in a kinetic Ising model peaks in the disordered phase

Lionel Barnett,[1,*] Joseph T. Lizier,[2,†] Michael Harré,[3,‡] Anil K. Seth,[1,§] and Terry Bossomaier[4,¶]

[1] *Sackler Centre for Consciousness Science, School of Informatics, University of Sussex, Brighton BN1 9QJ, UK*
[2] *CSIRO Computational Informatics, PO Box 76, Epping, NSW 1710, Australia*
[3] *School of Civil Engineering, The University of Sydney, Sydney NSW 2006, Australia*
[4] *Centre for Research in Complex Systems, Charles Sturt University, Panorama Ave, Bathurst NSW 2795, Australia*

(Dated: September 13, 2013)

There is growing evidence that for a range of dynamical systems featuring complex interactions between large ensembles of interacting elements, mutual information peaks at order/disorder phase transitions. We conjecture that, by contrast, information *flow* in such systems will generally peak strictly on the disordered side of a phase transition. This conjecture is verified for a ferromagnetic 2d lattice Ising model with Glauber dynamics and a transfer entropy-based measure of system-wide information flow. Implications of the conjecture are considered; in particular, that for a complex dynamical system in the process of transitioning from disordered to ordered dynamics (a mechanism implicated, for example, in financial market crashes and the onset of some types of epileptic seizures), information dynamics may be able to predict an imminent transition.

PACS numbers: 89.70.Cf, 89.75.Fb, 64.60.Cn, 64.60.De, 75.10.Hk, 02.50.Ey
Keywords: kinetic Ising model, Glauber dynamics, phase transition, information flow, transfer entropy

In a system comprising a large number of interacting elements with an order/disorder phase transition, it is easy to argue that mutual information between elements must peak at an intermediate order: for a highly ordered system there is little indeterminacy about the state of individual elements and hence mutual information between elements will be small, while for a highly disordered system elements will behave near-independently and again mutual information between elements will be small. It also seems reasonable (particularly if long-range interactions are considered) to expect that the peak will occur *at* the phase transition—i.e. where susceptibility peaks—and indeed this has been evidenced for a variety of complex systems, including the 2d lattice Ising model [1, 2], Vicsek's particle swarm model [3, 4], random Boolean networks (RBNs) [5, 6] and financial markets [7]. For a complex *dynamical* system, it is again easy to argue that information *flow* between elements must peak in an intermediate order regime; here, however, it is less clear that the peak should coincide with the phase transition. We show that for a 2d lattice Ising model with Glauber dynamics [8], information flow, as quantified by a global *transfer entropy* measure [9], attains a maximum strictly in the disordered (paramagnetic) phase. We conjecture that this phenomenon is universal for a class of complex dynamical systems and discuss implications.

We consider an isotropic ferromagnetic 2d lattice Ising model of size $N = L \times L$ with no external field . If the system is in state $\boldsymbol{s} = s_1 \dots s_N$, $s_i \in \{+1, -1\}$, then the Hamiltonian is given by [22]

$$\mathcal{H}(\boldsymbol{s}) = -\sum_{<i,j>} s_i s_j \,, \tag{1}$$

where $<i,j>$ denotes a sum over the $2N$ unique pairs of lattice neighbours. We impose periodic boundary conditions, so that the system is homogeneous: for any sites i, j there is a symmetry τ—i.e. a permutation of sites leaving the Hamiltonian invariant—with $\tau(i) = j$. At thermodynamic equilibrium, the Boltzmann-Gibbs probability to find the system in state $\boldsymbol{s}$ is

$$\Pi(\boldsymbol{s}) \equiv \mathbf{P}(\boldsymbol{S} = \boldsymbol{s}) = \frac{1}{Z} e^{-\beta\mathcal{H}(\boldsymbol{s})} \,, \tag{2}$$

where $\boldsymbol{S}$ denotes a random equilibrium state, $\beta \equiv 1/T$ is the inverse temperature (units are considered normalized so that Boltzmann's constant $= 1$) and $Z = \sum_{\boldsymbol{s}} e^{-\beta\mathcal{H}(\boldsymbol{s})}$ is the partition function. The magnetization per site is $\mathcal{M} = \frac{1}{N}\sum_i \langle S_i \rangle$, the free energy per site $\mathcal{F} = -\frac{1}{\beta N} \log Z$ and the internal energy per site $\mathcal{U} = -\frac{1}{N}\frac{\partial}{\partial\beta}\log Z = \frac{1}{N}\langle \mathcal{H}(\boldsymbol{S})\rangle$, where $\langle\cdots\rangle$ denotes ensemble average. The model is largely solved in the thermodynamic limit $N \to \infty$ [10, 11]; known results pertinent to our study are displayed in TABLE I.

For the kinetic model we consider discrete-time Glauber spin-flip dynamics [8]: at each time step a site i is selected uniformly at random and its spin flipped with probability.

$$P_i(\boldsymbol{s}) = \left[1 + e^{\beta\Delta\mathcal{H}_i(\boldsymbol{s})}\right]^{-1} \,. \tag{3}$$

where $\Delta\mathcal{H}_i(\boldsymbol{s}) \equiv \mathcal{H}(\boldsymbol{s}^i) - \mathcal{H}(\boldsymbol{s}) = 2s_i \sum_{j\in\nu(i)} s_j$ is the energy difference between the spin-flipped and original state. Here a superscript i denotes flipping the ith spin and $\nu(i)$ denotes the indices of the lattice neighbours of site i. This (Markov Chain Monte Carlo) scheme satisfies *detailed balance* [12] and thus yields the Boltzmann equilibrium probabilities (2) at stationarity.

Information-theoretic quantities may be framed in terms of statistical dependency. The basic measure of statistical dependency considered in this article is the *mutual information* $\mathrm{I}(X : Y \mid Z)$ between random

critical inverse temperature	$\beta_c = \frac{1}{2}\log\left(1+\sqrt{2}\right)$
magnetization	$\mathcal{M} = \pm\left(1-\sinh^{-4}2\beta\right)^{\frac{1}{8}}$ for $T < T_c$, $\mathcal{M} = 0$ for $T \ge T_c$
free energy	$-2\beta\mathcal{F} = \log\left(2\cosh^2 2\beta\right) + \frac{2}{\pi}\int_0^{\pi/2}\log\left(1+\sqrt{1-\kappa^2\sin^2\theta}\right)d\theta$
internal energy	$-\mathcal{U} = \coth 2\beta\left[1+\frac{2}{\pi}\left(\kappa\sinh 2\beta - 1\right)\int_0^{\pi/2}\frac{d\theta}{\sqrt{1-\kappa^2\sin^2\theta}}\right]$

TABLE I: Thermodynamic limits for the 2d lattice Ising model: here $\kappa \equiv 2\dfrac{\sinh 2\beta}{\cosh^2 2\beta}$.

variables X, Y, optionally conditional on a third variable Z [13]; it vanishes iff X, Y are independent conditional on Z. Given stationary *stochastic processes* $X(t), Y(t),\ \ t\ldots,0,1,2,\ldots,$ we take as a measure of past-conditional dependence—sometimes (albeit controversially [14]) interpreted as *information flow* from process Y to process X—the *transfer entropy* [9, 15] $\mathcal{T}_{Y\to X} \equiv \mathrm{I}\big(X(t) : Y^{(\ell)}(t) \,\big|\, X^{(\ell)}(t)\big)$. Here $X^{(\ell)}(t) \equiv X(t-1),\ldots,X(t-\ell)$ denotes the ℓ-length *history* of the process X at time t for $\ell = 1,2,\ldots$. $\mathcal{T}_{Y\to X}$ is zero iff process X, conditional on its own past, is independent of the past of Y. Our interest here is the behaviour of "global" statistical dependencies for the entire system of interacting spins. Below we define averaged pairwise (bivariate) as well as gross global measures for both static (mutual information-based) and dynamic (transfer entropy-based) dependencies.

We define the pairwise mutual information measure

$$I_{pw} \equiv \frac{1}{2N}\sum_{<i,j>}\mathrm{I}(S_i : S_j)\ ; \tag{4}$$

i.e. the lattice average mutual information between pairs of neighbouring sites. This is essentially the quantity previously considered in [1, 2]; [16] also considers the mutual information between two halves of a cylindrical 2d lattice. In [17], SEC. 1 we calculate that in the thermodynamic limit

$$I_{pw} \to -2\sum_\sigma p_\sigma\log p_\sigma + \sum_{\sigma,\sigma'} p_{\sigma\sigma'}\log p_{\sigma\sigma'} \tag{5}$$

where the sums are over $\sigma, \sigma' = \pm 1$, with (*cf.* [1])

$$p_\sigma = \tfrac{1}{2}(1+\sigma\mathcal{M}), \quad p_{\sigma\sigma'} = \tfrac{1}{4}[1+(\sigma+\sigma')\mathcal{M} - \tfrac{1}{2}\sigma\sigma'\mathcal{U}] \tag{6}$$

Note that for $T < T_c$ the sign of $\mathcal{M}$ does not affect this (and subsequent) results; i.e. our information measures are invariant under symmetry breaking. Our global measure of mutual information is the *multi-information* [18]

$$I_{gl} = \sum_i \mathrm{H}(S_i) - \mathrm{H}(\boldsymbol{S}) \tag{7}$$

where $\mathrm{H}(\cdots)$ denotes entropy, which may be considered a measure of gross statistical dependency among the S_i; it vanishes iff the S_i are all independent. It is known [11] that in the thermodynamic limit $\frac{1}{N}\mathrm{H}(\boldsymbol{S}) \to \beta(\mathcal{U}-\mathcal{F})$, so that

$$\frac{1}{N}I_{gl} \to -\sum_\sigma p_\sigma\log p_\sigma - \beta(\mathcal{U}-\mathcal{F})\,. \tag{8}$$

Next we consider the ($\ell = 1$ history) pairwise transfer entropy measure defined by

$$\mathcal{T}_{pw} \equiv \frac{1}{2N}\sum_{<i,j>}\mathcal{T}_{S_j\to S_i}\ . \tag{9}$$

In [17], SEC. 2 we calculate that in the thermodynamic limit

$$N\,\mathcal{T}_{pw} \to -q\sum_\sigma\log\frac{q}{p_\sigma} + \sum_{\sigma'} q_{\sigma'}\sum_\sigma\log\frac{q_{\sigma'}}{p_{\sigma\sigma'}}\,. \tag{10}$$

where

$$q = \tfrac{1}{2}\left\langle P_i(\boldsymbol{S})\right\rangle, \quad q_{\sigma'} = \tfrac{1}{4}(\left\langle P_i(\boldsymbol{S})\right\rangle + \sigma'\left\langle S_j P_i(\boldsymbol{S})\right\rangle), \tag{11}$$

for i, j arbitrary lattice neighbours. A symmetry argument ([17], SEC. 2) shows that $\langle S_j P_i(\boldsymbol{S})\rangle \equiv 0$ for $T \ge T_c$; i.e. when symmetry is unbroken. Finally, we define the ($\ell = 1$ history) global transfer entropy measure (*cf. collective transfer entropy* [19])

$$\mathcal{T}_{gl} \equiv \frac{1}{N}\sum_i\mathcal{T}_{\boldsymbol{S}\to S_i}\,; \tag{12}$$

i.e. the average information flow from the entire spin system to individual sites. This may be considered a measure of gross past-conditional statistical dependence of the S_i, insofar as it vanishes iff each S_i, conditional on its own past, does not depend on the past of spins at any other site. It may alternately be thought of as a measure of "information flow density", closely related to *causal density* [20]. In [17], SEC. 3 we show that in the thermodynamic limit

$$N\,\mathcal{T}_{gl} \to -q \sum_{\sigma} \log \frac{q}{p_\sigma} + \langle P_i(\boldsymbol{S}) \log P_i(\boldsymbol{S}) \rangle \ . \qquad (13)$$

In contrast to the mutual information measures, we do not have analytic expressions for $\mathcal{T}_{pw}$ and $\mathcal{T}_{gl}$.

It may be argued that the entire (i.e. $\ell \to \infty$) history of the process ought to be taken into account for a valid measure of information flow. Trivially, however, (10) and (13) will hold for any fixed finite history length ℓ, since in the theromodynamic limit there is only one spin update in the neighborhood of the considered spin. In this sense the measures $\mathcal{T}_{pw}, \mathcal{T}_{gl}$ are essentially short-range in time and space. Corresponding long-range measures (where the order of limits $N \to \infty$, $\ell \to \infty$ is reversed) are likely to be far more difficult to calculate.

Experimentally, the measures I_{pw} and I_{gl} at the thermodynamic limit were computed in accordance with their analytic expressions (5) and (8) respectively, while $\mathcal{T}_{pw}$ and $\mathcal{T}_{gl}$ were estimated in simulation. We simulated a kinetic Ising model of size $N = L \times L$ for $L = 512$. Each update comprised N (potential) spin-flips according to the Glauber transition probabilities (3). Initial spin configurations were uniform random ± 1 for $T \geq T_c$ and all $+1$ for $T < T_c$, where $T_c = 1/\beta_c \approx 2.269$ is the critical temperature. Simulations were initially run for a relaxation time of 10^4 updates and statistics then collated over a further 10^5 updates. Taking advantage of ergodicity of the Glauber dynamics, ensemble averages were calculated as means, over the 10^5 sample equilibrium states, of lattice averages. Estimates for $\mathcal{T}_{pw}$ and $\mathcal{T}_{gl}$ were then obtained according to (10) and (13) respectively. This procedure was performed for 200 runs at each of 100 temperature points enclosing the phase transition. In addition, at each temperature aggregate statistics based on a total sample size of 2×10^7 stationary states were calculated.

FIG. 1 displays $I_{pw}, I_{gl}, \mathcal{T}_{pw}$ and $\mathcal{T}_{gl}$ plotted against temperature. For I_{pw} and I_{gl} solid lines plot analytic values at the thermodynamic limit, while for $\mathcal{T}_{pw}$ and $\mathcal{T}_{gl}$ they plot aggregate statistics. The gradients of all measures appear to approach $+\infty$ as $T \to T_c$ (vertical dashed line) from below, consistent with critical behaviour of correlation statistics at a 2nd order phase transition. In [17], SEC. 4 we show that the gradients of I_{pw} and I_{gl} approach $-\infty$ logarithmically as $T \to T_c$ from above. Although not clear from the figures, we conjecture that the limiting gradient of $\mathcal{T}_{pw}$ is *negative* and that of $\mathcal{T}_{gl}$ *positive* as $T \to T_c$ from above. Indeed, we see that all measures peak at T_c except for $\mathcal{T}_{gl}$, which has a maximum at $T = 2.354 \pm 0.003 > T_c$. FIG. 2 displays the detailed behaviour of $\mathcal{T}_{pw}$ and $\mathcal{T}_{gl}$ beyond T_c. The vertical arrow marks the $\mathcal{T}_{gl}$ maximum. Due to *critical slowing down* [11], statistical error (indicated by 99% confidence intervals) is greatest around T_c. To establish that the post-critical maximum is not merely a finite-system size artifact, we repeated our simulations for smaller lattices of sizes $L = 64, 128$ and 256. At the temperature resolution deployed, it was observed that the position of the maximum does not change as system size is increased; i.e. it does not "creep" towards the critical temperature.

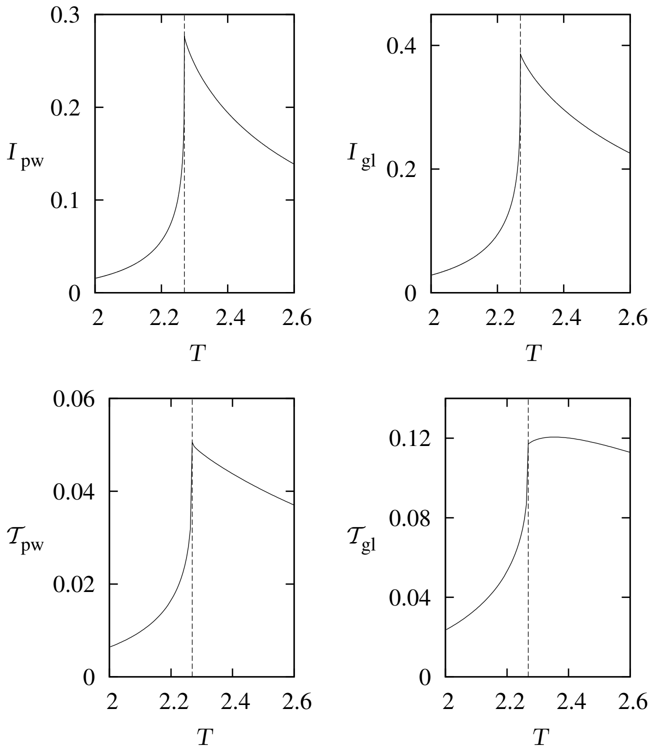

FIG. 1: Dependency measures (scaled by system size) plotted against temperature for $L = 512$. See main text for details.

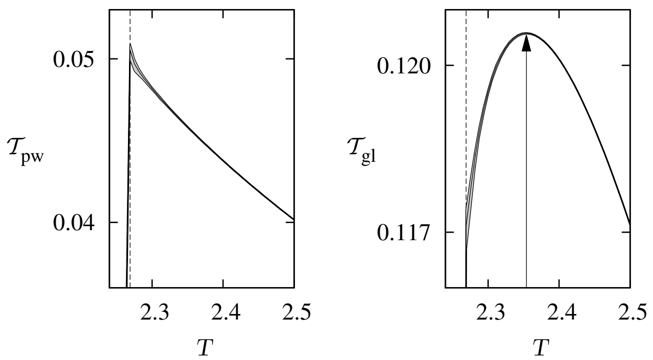

FIG. 2: $\mathcal{T}_{pw}$ and $\mathcal{T}_{gl}$ plotted against temperature for $L = 512$ (detail). The shaded regions indicate 99% confidence intervals based on the 200 trials, while the arrow marks the $\mathcal{T}_{gl}$ maximum at $T \approx 2.354$.

It has been established previously that pairwise mutual information peaks at the phase transition for the 2d lattice Ising model [1, 2]. Our results show that so too does pairwise transfer entropy. Note that both pairwise measures incorporate putative statistical dependencies intermediated by the joint distribution of the remaining system elements with the spin pair in question. Our global measures do not suffer from this effect; nonetheless, the static global measure I_{gl} also peaks at the phase transition, while the dynamic measure $\mathcal{T}_{gl}$ peaks strictly in the disordered regime. We conclude that a post-critical maximum is not simply a consequence of accounting for common influences, nor a consequence alone of incorporating past-conditional dependencies; both factors are required.

As to an intuitive explanation for the post-critical peak in $\mathcal{T}_{gl}$, preliminary research implicates a subtle interplay between differing contributions to $\mathcal{T}_{gl}$ from sites within and on the boundaries of same-spin domains, and the change in distribution of domain size as temperature increases and domains disintegrate [21]. A complementary perspective is offered in [19], where $\mathcal{T}_{gl}$ is regarded as a measure of *collective* information transfer, capturing both pairwise ($\mathcal{T}_{pw}$) and higher-order (multivariate) correlations to a site. Its peak is interpreted in terms of conflicting tendencies amongst these components as the level of disorder in the system increases away from the phase transition point (as empirically observed for RBNs [6]): that is, for pairwise correlations to decay, while higher-order multivariate effects become more prevalent.

If our conjecture holds that a post-critical peak in $\mathcal{T}_{gl}$ (and a critical peak in I_{gl}) is a universal phenomenon—at least, perhaps, for some class of 2nd order phase transitions—it raises the intriguing possibility of anticipation of an imminent phase transition in a system moving slowly towards criticality from the disordered regime. Specifically: *if $\mathcal{T}_{gl}$, estimated over time windows short compared to the time scale of change in a notional order parameter, is seen to peak whilst I_{gl} continues to increase, we might suspect that the system is approaching criticality*. We suggest possible application to complex systems such as financial markets and neural systems where disorder → order phase transitions are associated with pathological dynamics. Further research is required to test universality, both in complex dynamical models and on real-world data.

Finally, we remark that, of the measures presented, only I_{gl} is truly long-range, whereas critical phenomena typically manifest at all spatio-temporal scales. One might therefore question the relevance of our (short-range) measure $\mathcal{T}_{gl}$ for understanding information flow near criticality. However, since long-range interactions emerge from the cumulative effect of short-range interactions, our result establishes a crucial foothold for the understanding of the mechanism of this emergence. Secondly—in contrast to our measures—long-range information flow is unfeasible to estimate empirically from limited data and thus, while of theoretical interest, is of limited practical utility.

Acknowledgments: LB and AKS are grateful to the Dr. Mortimer and Dame Theresa Sackler Foundation for financial support; other support comes from the EPSRC (AKS; EP/G007543/1). The authors would also like to thank Intersect Australia Ltd. for use of their High Performance Computing facilities.

* Electronic address: l.c.barnett@sussex.ac.uk
† Electronic address: joseph.lizier@csiro.au
‡ Electronic address: michael.harre@sydney.edu.au
§ Electronic address: a.k.seth@sussex.ac.uk
¶ Electronic address: tbossomaier@csu.edu.au

[1] H. Matsuda, K. Kudo, R. Nakamura, O. Yamakawa, and T. Murata, Int. J. Theor. Phys. **35**, 839 (1996).
[2] S.-J. Gu, C.-P. Sun, and H.-Q. Lin, J. Phys. A **41**, 025002 (2008).
[3] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Phys. Rev. Lett. **75**, 1226 (1995).
[4] R. T. Wicks, S. C. Chapman, and R. O. Dendy, Phys. Rev. E **75**, 051125 (2007).
[5] A. S. Ribeiro, S. A. Kauffman, J. Lloyd-Price, B. Samuelsson, and J. E. S. Socolar, Phys. Rev. E **77**, 011901 (2008).
[6] J. T. Lizier, S. Pritam, and M. Prokopenko, in *Artificial Life (Special issue on Complex Networks)*, edited by M. A. Bedau (MIT Press, Cambridge, MA, 2011), vol. 17, pp. 293–314.
[7] M. Harré and T. Bossomaier, Europhys. Lett. **87**, 18009 (2009).
[8] R. J. Glauber, J. Math. Phys. (N.Y.) **4**, 294 (1963).
[9] T. Schreiber, Phys. Rev. Lett. **85**, 461 (2000).
[10] L.Onsager, Phys. Rev. **65**, 117 (1944).
[11] B. M. Mccoy and T. T. Wu, *The two-dimensional Ising model* (Harvard University Press, Cambridge, MA, 1973).
[12] N. V. Kampen, *Stochastic processes in physics and chemistry* (Elsevier Science, Amsterdam, 1992).
[13] T. M. Cover and J. A. Thomas, *Elements of information theory, 2nd edn.* (John Wiley & Sons, New Jersey, 1991).
[14] J. T. Lizier and M. Prokopenko, Eur. Phys. J. B **73**, 605 (2010).
[15] A. Kaiser and T. Schreiber, Physica D **166**, 43 (2002).
[16] H. W. Lau and P. Grassberger, Phys. Rev. E **87**, 022128 (2013).
[17] L. Barnett, T. Bossomaier, J. T. Lizier, M. Harré, and A. K. Seth, *See supplemental material at* `suppmat_url_here` *for mathematical details.* (2013).
[18] M. Studený and J. Vejnarová, in *Learning in Graphical Models*, edited by M. I. Jordan (MIT Press, Cambridge, MA, 1998), pp. 261–298.
[19] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, Chaos **20**, 037109 (2010).
[20] A. K. Seth, E. Izhikevich, G. N. Reeke, and G. M. Edelman, Proc. Natl. Acad. Sci. USA **103**, 10799 (2006).
[21] J. L. Cambier and M. Nauenberg, Phys. Rev. B **34**, 8071 (1986).
[22] Bold type $\boldsymbol{s}$ denotes a state vector of spins and normal/lower case Greek type s_i, σ denotes individual ± 1 spins. Capitals $\boldsymbol{S}, S_i$ denote random variables.