A LETTERS JOURNAL EXPLORING THE FRONTIERS OF PHYSICS

LETTER

Anomalous behaviour of mutual information in finite flocks

To cite this article: L. Barnett *et al* 2017 *EPL* **120** 38005

A LETTERS JOURNAL EXPLORING
THE FRONTIERS OF PHYSICS

November 2017

EPL, **120** (2017) 38005
doi: 10.1209/0295-5075/120/38005

www.epljournal.org

Anomalous behaviour of mutual information in finite flocks

L. BARNETT[1], J. BROWN[2] and T. BOSSOMAIER[3]

[1] *Sackler Centre for Consciousness Science, Department of Informatics, University of Sussex - Brighton, UK*
[2] *School of Computing & Mathematics, Charles Sturt University - Bathurst, NSW, Australia*
[3] *Centre for Research in Complex Systems, Charles Sturt University - Bathurst, NSW, Australia*

received 14 September 2017; accepted in final form 12 January 2018
published online 1 February 2018

PACS 89.70.Cf – Entropy and other measures of information
PACS 89.75.-k – Complex systems
PACS 64.60.De – Statistical mechanics of model systems (Ising model, Potts model, field-theory models, Monte Carlo techniques, etc.)

Abstract – The existing consensus is that flocks are poised at criticality, entailing long correlation lengths and a maximal value of Shannon mutual information in the large-system limit. We show, by contrast, that for *finite* flocks which do not truly break ergodicity in the long-observation-time limit, mutual information may not only fail to peak at criticality —as observed for other critical systems— but also diverge as noise tends to zero. This result carries implications for other finite-size, out-of-equilibrium systems, where observation times may vary widely compared to time scales of internal system dynamics; thus it may not be assumed that mutual information locates the phase transition.

Introduction. – From the 40000 strong murmurations of starlings to traffic jams, flocking occurs in many animal species, as well as many domains of human society. Recent developments in video pattern recognition and GPS technology have greatly increased our understanding of animal systems, such as fish [1,2], pigeons [3], starlings [4], midges [5] and sheep [6]. Flocks offer energy efficiency, reduced navigational effort and increased resilience to predation. For biological, finite-size flocks, however, it is under-appreciated that macroscopic statistics depend essentially on observation time scales.

Understanding of flocking dynamics owes much to abstract models, such as the Standard Vicsek model (SVM) [7] which, at large system size, exhibits phase-transition–like behaviour at a critical noise value. In other systems studied to date, such as the Ising spin model [8–10], cellular automata [11] and financial systems [12], mutual information (MI) [13] is the gold-standard marker of order-disorder (2nd order) phase transitions in equilibrium statistical mechanics: in the thermodynamic limit it typically tends to zero in the limits of low and high noise, peaking at criticality [14,15]. Less is known, however, of its behaviour in out-of-equilibrium and/or finite-size systems. The SVM exemplifies an out-of-equilibrium phase transition [16] between coordinated behaviour and random diffusion [7], thought to be in its own universality class [17]. The thermodynamic limit of large system size has been studied by Toner and Tu, both at the phase transition [18] and the low-noise, single-flock, limit [19]. At the limit, continuous rotational $O(\infty)$ symmetry is broken, leaving Goldstone modes, and thus large, long-range density fluctuations in two dimensions. In higher dimensions, the situation is more complicated.

Here, by exploiting an approximate isometry of the SVM, we obtain a novel closed-form dimensional reduction of the neighbour-pair MI between particle headings on the basis that, in a finite-size system at long observation times, rotational symmetry is never broken. This reveals a hitherto unnoticed behaviour of MI in such systems: *absence of a peak at the phase transition, and divergence at low noise*, contrary to the behaviour of the Ising model and other complex systems [20].

The standard Vicsek model. – The two-dimensional SVM comprises a set of N point particles (labelled $i = 1, \dots, N$) moving on a plane of linear extent L with periodic boundary conditions. Each particle moves with constant speed v, and interacts only with neighbouring particles within a fixed radius r, which we take to be 1. We denote the position of the i-th particle by $\boldsymbol{x}_i(t)$ and its velocity vector by $\boldsymbol{v}_i(t) = (v\cos\theta_i(t), v\sin\theta_i(t))$, where $\theta_i(t)$ is its *heading*[1]. Let $\nu_i(t) \equiv \{j : |\boldsymbol{x}_j(t) - \boldsymbol{x}_i(t)| < r\}$ be

[1]We consider headings as *circular* variables defined on $(-\pi, \pi]$ with arithmetic modulo 2π.

the index set of all particles neighbouring particle i at time t (including i itself, so that $\nu_i(t) \neq \emptyset$). The *neighbourhood-average velocity* of particle i is then given by

$$\bar{\boldsymbol{v}}_i(t) = \frac{1}{|\nu_i(t)|} \sum_{j \in \nu_i(t)} \boldsymbol{v}_j(t), \tag{1}$$

with heading $\bar{\theta}_i(t)$.

Particle positions and headings are updated synchronously[2] at discrete time intervals $\Delta t = 1$ according to

$$\boldsymbol{x}_i(t + \Delta t) = \boldsymbol{x}_i(t) + \boldsymbol{v}_i(t)\Delta t, \tag{2}$$

$$\theta_i(t + \Delta t) = \bar{\theta}_i(t) + \omega_i(t), \tag{3}$$

respectively, where $\omega_i(t)$ is a thermal fluctuation (white noise) uniform on the interval $[-\eta/2, \eta/2]$ with intensity $\eta \in (0, 2\pi]$.

Note that, since a particle travels a distance v in a single time increment $\Delta t = 1$, the SVM only approximates continuity in space and time in case $v \ll 1$ (the model is thus arguably unrealistic as a model for real-world flocking if particle velocities are large).

The SVM ensemble. We consider the SVM as a *statistical ensemble* of finite size N, parametrised by the velocity v, particle density $\rho = N/L^2$ and noise intensity η. For simplicity, particle density is fixed at $\rho = 0.25$ throughout, and noise intensity η is taken as a control parameter. We suppose that the ensemble is relaxed into a steady state, and use capitals $\boldsymbol{V}_i, \Theta_i$, etc., to indicate corresponding quantities sampled from the steady-state ensemble. In the limit $v \to 0$, the model is equivalent to an XY model, where particles do not move[3], while in the limit $v \to \infty$ particles become fully mixed between updates [7].

The full order parameter for the SVM ensemble is the 2D random vector

$$\boldsymbol{M} = \frac{1}{Nv} \sum_{i=1}^{N} \boldsymbol{V}_i, \tag{4}$$

with magnitude $M \equiv |\boldsymbol{M}|$ and heading Φ. We have $0 \leq M \leq 1$, with $M = 1$ if and only if all particles in the ensemble are aligned, and $M \to 0$ in the large-system limit $N \to \infty$. The ensemble variance

$$\chi = \langle M^2 \rangle - \langle M \rangle^2 \tag{5}$$

of the order parameter magnitude defines the susceptibility, where angle brackets $\langle \cdots \rangle$ denote ensemble averages. Although phase transitions only exist formally in the thermodynamic limit, for finite systems we consider a peak in susceptibility (with respect to a control parameter) as identifying the approximate location of a phase transition.

[2]We implement a "backward update" scheme, where both particle positions and velocities for time $t + \Delta t$ are updated on the basis of particle velocities at time t, as opposed to the "forward update" scheme which updates particle positions for time $t + \Delta t$ using the already updated velocity at $t + \Delta t$.

[3]Note that the limiting behaviour of the model as $v \to 0$ must be considered as distinct from models with $v = 0$, *e.g.*, the XY model [21].

Long-term vs. short-term statistics. In estimating ensemble statistics from simulated (steady-state) dynamics, it is commonplace to invoke *ergodicity* in some form: that is, the simulation is observed, and statistics collated, over a time window of length T, under the assumption that as $T \to \infty$ the statistic in question converges to its ensemble average value. This approach implicitly assumes that observation times are long in comparison to the internal dynamics of the system. In the case of the finite-size SVM, however, this assumption may well be violated, particularly at low-noise intensities. What we see, rather, is akin to what has been termed "continuous ergodicity-breaking" [22]: over short observation times, the system is confined to a comparatively small volume of phase space. As we observe the system over increasing lengths of time, progressively larger volumes of phase space are explored. Since a finite SVM is ergodic, the system eventually explores the entire phase space. At low noise, however, observation times necessary to obtain effectively ergodic behaviour become impractically large.

Our resolution to this issue is a pragmatic one: we consider ensemble statistics as essentially observation time-dependent. *Short-term* statistics are thus collated separately (with no ergodic assumptions) over ranges of observation times spanning several orders of magnitude. This affords insights into how the extent of phase-space exploration affects our statistics (and also neatly sidesteps the somewhat vexed issue as to whether the SVM features true ergodicity-breaking in the thermodynamic limit). In addition, to estimate the limiting ergodic behaviour of the system, below we exploit a rotational symmetry approximation to collate *long-term* statistics, under the assumption that in a finite-size SVM, symmetry —like ergodicity— is never truly broken.

Simulation details. Simulation models were written in C++ and run on the RAIJIN supercluster at the Australian National Computer Infrastructure Facility. Since the particle velocity (angle) is continuous, the *differential entropy* and mutual information were calculated using nearest-neighbour estimators which were developed for continuous variables [23,24]. The accuracy of the estimators was checked by: permutation testing —shuffling the source to remove any information sharing; and decimation —comparing the estimate with subsets of one-tenth of the number of events [25]. Theoretical work on the performance of these estimators is limited and is most relevant to smaller systems [26]. The entropy estimation by nearest neighbour is computationally demanding and was carried out *in situ* on RAIJIN. As an example of the data requirements, the interactions of particles in the large window simulation, with $N = 500, T = 5 \times 10^4$, at $\eta = 0.1$ produced approximately 2×10^9 points for the nearest-neighbours estimators, each of which required a search for the k nearest neighbours (k-nn) and fixed radius search.

Following Vicsek *et al.* [7], we employed a cooling regime to reduce computation times required for simulations to settle into a steady state, whereby simulations were started with the maximum noise ($\eta = 2\pi$) case, with particles uniformly distributed over the flat torus and headings uniformly distributed on $(0, 2\pi]$. Simulations were run for an initial number T_s of skip steps to allow the system to settle, followed by a data collection phase of T time steps, over which MI statistics were collated. On completion, η was decreased and another $T_s + T$ simulation steps run with the new η value. This technique enabled reduction of T_s by an order of magnitude, as compared to restarting simulations anew for each η. Appropriate settling time depends on η ($\eta = 2\pi$, for instance, requires zero settling time). We found that a satisfactory regime was to adjust T_s in tiers:

$$T_s = \begin{cases} 1000, & \eta \geq 3.0, \\ 50000, & \eta \leq 1.0, \\ 20000, & \text{otherwise.} \end{cases} \tag{6}$$

Neighbour-pair mutual information. – The neighbour-pair MI is defined as the ensemble statistic

$$\mathcal{I}_{pw} \equiv \mathcal{I}(\Theta_I : \Theta_J) = \mathbf{H}(\Theta_I) + \mathbf{H}(\Theta_J) - \mathbf{H}(\Theta_I, \Theta_J), \tag{7}$$

where $\mathbf{H}$ denotes differential entropy and (I, J) is uniform on the set of unique neighbour index-pairs[4]. While differential entropy may go negative, MI (in particular $\mathcal{I}_{pw}$) is strictly non-negative. Note that by particle indistinguishability, the marginal distributions of Θ_I and Θ_J are the same, so that $\mathbf{H}(\Theta_I) = \mathbf{H}(\Theta_J)$. In the short-term case, we estimate $\mathcal{I}_{pw}$ over multiple realisations of simulated SVMs. The SVMs are first relaxed/annealed to a steady state, and then headings $\theta_i(t)$ sampled over a further simulation period of T time steps, where T is the observation window.

Given that (as discussed above) ergodicity remains unbroken in the long-term–observation limit, near-isotropy of the SVM allows us to approximate eq. (7) in this case by a one-dimensional form, in which only particle heading differences $\theta_i - \theta_j$ appear. Specifically, we assume *rotational symmetry*: that for any fixed angle φ, the joint distribution of $(\Theta_1 + \varphi, \ldots, \Theta_N + \varphi)$ is the same as the joint distribution of $(\Theta_1, \ldots, \Theta_N)$. We note that the SVM on the 2D torus with periodic boundary conditions is not strictly isotropic, so that this is indeed an approximation. We tested the approximation by repeating our experiments with the frame of reference of the SVM rotated randomly between updates, thus enforcing isotropy [27]. We found that in a large, but finite, SVM the isotropy assumption introduces almost negligible error (the error only being discernible near the phase transition; see the inset in fig. 1).

Let $p(\theta_1, \theta_2)$ be the probability density function (pdf) of (Θ_I, Θ_J). Under assumption of rotational symmetry, we have

$$p(\theta_1, \theta_2) = \frac{1}{2\pi} q(\theta_1 - \theta_2), \tag{8}$$

[4] $\mathcal{I}_{pw}$ is essentially the same quantity as calculated in [10], although it was formulated somewhat differently there.

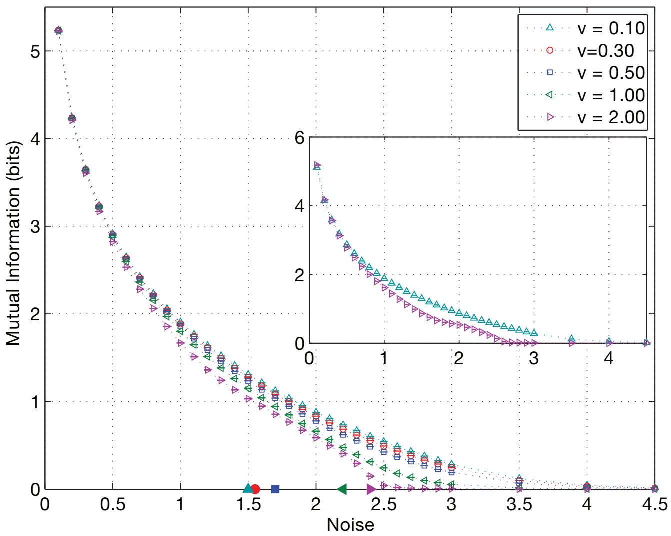

Fig. 1: (Colour online) Long-term MI $\mathcal{I}_{pw}^{LT}$ calculated according to eq. (10) for a range of particle velocities. System size $N = 1000$ particles, density $\rho = 0.25$ and velocities v as indicated. Simulation: 20 realisations at observation time $T = 500$ time steps. Error bars at 1 standard error (s.e.) (smaller than symbols) were constructed by 10 repetitions of the experiment. $\mathbf{H}(\Theta_I - \Theta_J)$ was calculated using a 512-bin histogram estimator. Filled symbols show estimated peaks in susceptibility χ. Inset: system using a rotated reference frame for $v = 0.10, 2.00$.

where $q(\theta)$ is the pdf of $\Theta_I - \Theta_J$. Noting that the marginal distributions of Θ_I and Θ_J are uniform on the unit circle, we obtain

$$\mathbf{H}(\Theta_I, \Theta_J) = \log 2\pi + \mathbf{H}(\Theta_I - \Theta_J), \tag{9}$$

leading to the expression

$$\mathcal{I}_{pw}^{LT} = \log 2\pi - \mathbf{H}(\Theta_I - \Theta_J) \tag{10}$$

for the approximate long-term neighbour-pair MI. Note that $\mathcal{I}_{pw}^{LT}$ vanishes precisely when $\Theta_I - \Theta_J$ is *uniform*. This is the case at maximum noise, when Θ_I, Θ_J are independent; that is, $\mathcal{I}_{pw}^{LT}$ vanishes at maximum noise, as we would expect. At very low noise, all particles nearly align so that the distribution of $\Theta_I - \Theta_J$ becomes sharply peaked. However, because $\mathbf{H}(\Theta_I - \Theta_J)$ is a *differential* entropy, it will generally diverge logarithmically to $-\infty$ as the variance decreases (*e.g.*, the differential entropy of a narrow uniform "notch" of width ε is $\log \varepsilon$). Thus, in the long-term–observation scenario, $\mathcal{I}_{pw}^{LT} \to +\infty$ as the noise intensity decreases to zero. That is to say, as particles align even more closely, Θ_I describes Θ_J with increasing precision —*i.e.*, the shared information. As the variables are continuous —and thus contain infinite precision— the MI increases, representing the extra bits required to encode the extra precision.

This diverging nature is demonstrated in the MI of two Gaussian variables, X, Y, with covariance r, which

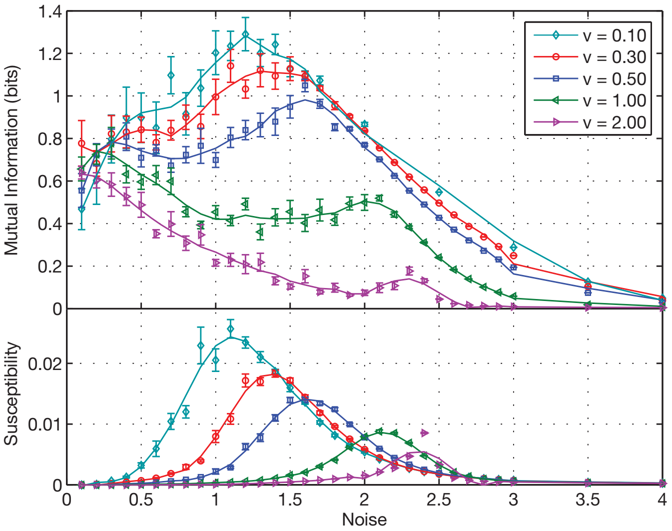

Fig. 2: (Colour online) $\mathcal{I}_{pw}$ (top plot) and susceptibility χ (bottom plot) estimated for a range of particle velocities (parameters as in fig. 1). $\mathcal{I}_{pw}$ was estimated according to eq. (7) over $T = 5000$ time steps after relaxation to steady state, using a nearest-neighbour estimator. χ was estimated over the same realisations. Error bars at 1 s.e. again constructed by 10 repetitions.

is exactly known [28]:

$$\mathcal{I}_{Gaussian}(X:Y) = -\frac{1}{2}\log(1-r^2)\,, \tag{11}$$

where $\mathcal{I}(X:Y)$ clearly diverges as X and Y become more and more correlated ($r \to 1$). However, we note that the case of $r = 1$ is different to $\eta = 0$ in the Vicsek model. At $\eta = 0$, ergodicity is truly broken —*i.e.*, $\Theta_I = \Theta_c$ for all T's after some settling time— and thus we have 0 MI, while the Gaussian variables are still "ergodic" (that is, any $x \in X$ can be drawn) therefore requiring infinite bits to encode.

Simulation results. – Figure 1 shows the long-term MI $\mathcal{I}_{pw}^{LT}$ estimated in sample according to eq. (10) for a range of particle velocities. Note that there is no evidence of a peak at the phase transition. For *short observation times,* by contrast, $\mathcal{I}_{pw}$ estimated according to eq. (7) (*i.e.*, with no assumption of rotational symmetry) does indeed peak at the phase transition, as reported by Wicks *et al.* [29]; see fig. 2. Some divergence at low noise is also in evidence. Figure 3 plots $\mathcal{I}_{pw}$ for a single fixed velocity at observation window size T varying over two orders of magnitude, along with the long-observation-time limit $\mathcal{I}_{pw}^{LT}$. As observation time increases, the MI peak flattens and divergence at low noise increases, approaching, as predicted, the long-observation-time limit.

Thus, long-term MI behaves in a distinctly different fashion to what has typically been observed in short-range MI studies of order-disorder transitions, where MI is seen to peak at the phase transition, and then tail off as noise tends to zero. Here, as the noise intensity is reduced

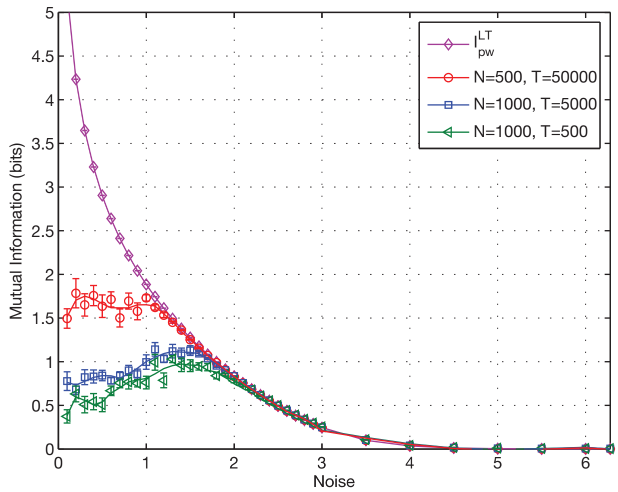

Fig. 3: (Colour online) $\mathcal{I}_{pw}$ estimates according to eq. (7) at fixed velocity $v = 0.30$ for a range of observation times T as indicated, along with the long-term $\mathcal{I}_{pw}^{LT}$ of eq. (10) as per fig. 1. System sizes N as indicated, other simulation details as for previous figures.

towards zero, particles align more and more strongly, so that the distribution of $\Theta_I - \Theta_J$ becomes more and more sharply peaked, resulting in divergence of $\mathcal{I}_{pw}^{LT}$. At the same time, non-ergodicity-breaking is evidenced by a random walk-like *precession* of the order parameter heading Φ around the unit circle (cf. fig. 4 below).

We remark that the continuous-state nature of the SVM is central to the divergence; in a discrete-state system MI cannot diverge. Nonetheless, a similar effect is seen for discrete systems, although divergence is capped by the number of distinct states. For a discretised Vicsek system, for example, where particle headings are constrained to m equispaced sectors, rotational symmetry remains unbroken (fluctuations due to noise still cause precession of Φ around the sectors) so that eq. (10) still holds, with the $\log 2\pi$ term replaced by $\log m$. Now $\mathbf{H}(\Theta_I - \Theta_J) \to 0$ as $\eta \to 0$, so that $\mathcal{I}_{pw}^{LT} \to \log m$.

Discussion. – Since the introduction of the SVM, in which the phase transition was originally claimed to be of second order, much controversy has surrounded its nature. Gregoire and Chaté [30] claimed on the basis of simulations that it was of first order, and much discussion ensued. Seemingly small details affect the nature of the transition: type of noise statistics [31]; forward *vs.* backward updating (especially at high particle velocities) [32]; boundary conditions associated with density bands or spin waves [33]; and the cone of influence on each particle [34,35]. In this study we utilise the original SVM model (backward updating, angular noise, periodic boundary conditions and low density) over a range of velocities.

But there is an additional aspect to the phase transition beyond the order controversy: the effect of finite size. In classical equilibrium systems, finite-size effects with $O(2)$ symmetry are known to exhibit a random walk behaviour

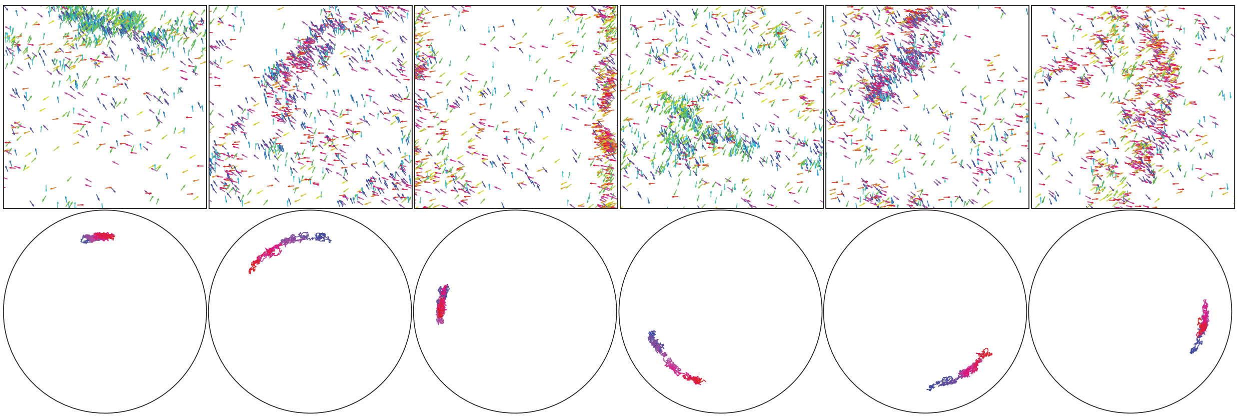

Fig. 4: (Colour online) Snapshots from a single simulation demonstrating precession of high density bands of a flock with $N = 1000$ particles at high velocity ($v = 2.0$) at $\eta = 1.5$ (below the peak in susceptibility). Snapshots taken at, from left to right, $t = 23 \times 10^3, 24 \times 10^3, 28 \times 10^3, 40 \times 10^3, 47 \times 10^3, 49 \times 10^3$. The top row shows the state of the flock, while the bottom row shows the two-dimensional order parameter $\boldsymbol{M}$ —that is, mean particle velocity— for the previous 1000 time steps going from blue ($t - 1000$) to red (t). Distance from the centre of the circle corresponds to the order parameter magnitude $M = |\boldsymbol{M}|$. Note that, as witnessed by the first two snapshots, precession can be rapid, with only 1000 time steps required for the band to precess $\pi/4$ radians.

along the Goldstone modes at low noise, but little is known about the active matter system considered here [36,37]. For $d = 2$, the Mermin-Wagner theorem (MW) would lead one to suspect that there is no order-disorder phase transition[5], but this strictly only holds for ergodic, equilibrium systems [40]. However, not only is the Vicsek system not an equilibrium model but it also has an effective dimension of $d = 4$ [41], thus the Mermin-Wagner theorem does not hold and an order-disorder transition is valid and present. Baglietto *et al.* [42] and Albano *et al.* [43], for example, discuss finite-size scaling, showing good agreement with theory for the susceptibility at the phase transition.

In the finite-size SVM, even at low (sub-critical) noise intensities, *neither ergodicity nor (approximate) rotational symmetry is broken* over large time scales. At short observation times, ergodicity is approximately broken, but as observation time increases the system becomes increasingly ergodic, exploring progressively larger volumes —and ultimately the entirety— of phase space. In the finite-size SVM this manifests as a stochastic precession of the order parameter heading Φ around the unit circle (fig. 4). We note too, that at (albeit physically implausible) high velocities, the SVM exhibits travelling "bands" of particles [32]; while it might be thought that this represents true symmetry breaking, detailed simulations (fig. 4) reveal that banding orientation, as well as Φ, precesses and through this, ergodicity is maintained.

The behaviour we see here has elements of classical thermodynamic equilibrium systems, although it is an active matter, far-from-equilibrium system. For active matter, the concept of equilibrium itself, is still not clearly defined [44]. The continuous symmetry implies that at even extremely low noise, the flock(s) can gradually change direction and cover the whole of phase space as observation time tends to infinity. This movement is analogous to the Goldstone modes of classical systems left behind when symmetry breaking occurs.

Goldstone modes have been discussed relative to flocking by Bialek *et al.* [45] and rejected as the source of information flow through the flock. On the other hand Melfo [46] claims that it is in fact the Goldstone mode which allows flock stability over a wide range of noise (system) parameters.

Although the unexpected behaviour of the MI has been demonstrated for the SVM —which is far from the only flocking model (see [47] for a recent alternative)— it seems likely that it will apply to many finite systems where symmetry only approximately breaks over short windows, but is restored over long observations: for such systems, MI may vary dramatically with observational time scale, diverging in the long-term limit as thermal noise approaches zero.

* * *

We thank Mike Harré, Joe Lizier and Guy Theroulaz for helpful discussions. Our attention was first drawn to the apparently anomalous behaviour of MI in the SVM by Dan Mackinlay. The National Computing Infrastructure (NCI) facility provided computing time for the simulations under project e004, with part funding under Australian Research Council Linkage Infrastructure grant LE140100002. JB would like to acknowledge the support of his PhD program and this work from the Australian Government Research Training Program Scholarship. We also thank an anonymous referee for comments on the Goldstone modes.

[5]See [38,39] for examples of phase transitions which are not forbidden at $d = 2$ by MW.

REFERENCES

[1] Calovi D. S., Lopez U., Schuhmacher P., Chaté H., Sire C. and Theraulaz G., *J. R. Soc. Interface*, **12** (2015) 20141362.

[2] Rosenthal S. B., Twomey C. R., Hartnett A. T., Wu H. S. and Couzin I. D., *Proc. Nat. Acad. Sci. U.S.A.*, **112** (2015) 4690.

[3] Nagy M., Vsrhelyi G., Pettit B., Roberts-Mariani I., Vicsek T. and Biro D., *Proc. Natl. Acad. Sci. U.S.A.*, **110** (2013) 13049.

[4] Cavagna A., Cimarelli A., Giardina I., Parisi G., Santagati R., Stefanini F. and Viale M., *Proc. Natl. Acad. Sci. U.S.A.*, **107** (2010) 11865.

[5] Attanasi A., Cavagna A., Del Castello L., Giardina I., Melillo S., Parisi L., Pohl O., Rossaro B., Shen E., Silvestri E. *et al.*, *PLoS Comput. Biol.*, **10** (2014) e1003697.

[6] Ginelli F., Peruani F., Pillot M.-H., Chat H., Theraulaz G. and Bon R., *Proc. Natl. Acad. Sci. U.S.A.*, **112** (2015) 12729.

[7] Vicsek T., Czirók A., Ben-Jacob E., Cohen I. and Shochet O., *Phys. Rev. Lett.*, **75** (1995) 1226.

[8] Matsuda H., Kudo K., Nakamura R., Yamakawa O. and Murata T., *Int. J. Theor. Phys.*, **35** (1996) 839.

[9] Lau H. W. and Grassberger P., *Phys. Rev. E*, **87** (2013) 022128.

[10] Barnett L., Lizier J. P., Harré M., Seth A. and Bossomaier T., *Phys. Rev. Lett.*, **111** (2013) 177203.

[11] Langton C. G., *Physica D*, **42** (1990) 12.

[12] Harré M. and Bossomaier T., *EPL*, **87** (2009) 18009.

[13] Shannon C., *Bell Syst. Tech. J.*, **27** (1948) 379.

[14] Vanni F., Luković M. and Grigolini P., *Phys. Rev. Lett.*, **107** (2011) 078103.

[15] Mora T. and Bialek W., *J. Stat. Phys.*, **144** (2011) 268.

[16] Ódor G., *Rev. Mod. Phys.*, **76** (2004) 663.

[17] Baglietto G. and Albano E. V., *Phys. Rev. E*, **78** (2008) 021125.

[18] Toner J. and Tu Y., *Phys. Rev. Lett.*, **75** (1995) 4326.

[19] Toner J. and Tu Y., *Phys. Rev. E*, **58** (1998) 4828.

[20] Erb I. and Ay N., *J. Stat. Phys.*, **115** (2004) 949.

[21] Baglietto G. and Albano E. V., *Comput. Phys. Commun.*, **180** (2009) 527.

[22] Mauro J., Gupta P. and Loucks R., *J. Chem. Phys.*, **126** (2007) 184511.

[23] Kraskov A., Stögbauer H. and Grassberger P., *Phys. Rev. E*, **69** (2004) 066138.

[24] Gómez-Herrero G., Wu W., Rutanen K., Soriano M. C., Pipa G. and Vicente R., *Entropy*, **17** (2015) 1958.

[25] Brown J. M., Bossomaier T. and Barnett L., *J. Comput. Sci.*, **23** (2017) 109.

[26] Gao W., Oh S. and Viswanath P., *Demystifying fixed k-nearest neighbor information estimators*, in *Proceedings of the 2017 IEEE International Symposium on Information Theory (ISIT)* (IEEE) 2017, pp. 1267–1271.

[27] Baglietto G. and Albano E. V., *Phys. Rev. E*, **80** (2009) 050103.

[28] Darbellay G. A. and Vajda I., *IEEE Trans. Inf. Theory*, **45** (1999) 1315.

[29] Wicks R. T., Chapman S. C. and Dendy R., *Phys. Rev. E*, **75** (2007) 051125.

[30] Grégoire G. and Chaté H., *Phys. Rev. Lett.*, **92** (2004) 025702.

[31] Chepizhko A., Kulinskii V., Holovatch Y., Berche B., Bogolyubov N. and Folk R., *AIP Conf. Proc.*, **1198** (2009) 25.

[32] Nagy M., Daruka I. and Vicsek T., *Physica A: Stat. Theor. Phys.*, **373** (2007) 445 .

[33] Aldana M., Dossetti V., Huepe C., Kenkre V. M. and Larralde H., *Phys. Rev. Lett.*, **98** (2007) 095702.

[34] Durve M. and Sayeed A., *Phys. Rev. E*, **93** (2016) 052115.

[35] Romensky M., Lobaskin V. and Ihle T., *Phys. Rev. E*, **90** (2014) 063315.

[36] Niel J. and Zinn-Justin J., *Nucl. Phys. B*, **280** (1987) 355.

[37] Goldschmidt Y. Y., *Nucl. Phys. B*, **280** (1987) 340.

[38] Kosterlitz J. M. and Thouless D. J., *J. Phys. C: Solid State Phys.*, **6** (1973) 1181.

[39] Grossmann R., Peruani F. and Bär M., *Phys. Rev. E*, **93** (2016) 040102.

[40] Mermin N. D. and Wagner H., *Phys. Rev. Lett.*, **17** (1966) 1133.

[41] Baglietto G., Albano E. V. and Candia J., *Interface Focus*, **2** (2012) 708.

[42] Baglietto G. and Albano E. V., *Phys. Rev. E*, **78** (2008) 021125.

[43] Albano E., Bab M. A., Baglietto G., Borzi R. A., Grigera T. S., Loscar E. S., Rodríguez D. E., Puzzo M. L. R. and Saracco G. P., *Rep. Prog. Phys*, **74** (2011) 026501.

[44] Fodor É., Nardini C., Cates M. E., Tailleur J., Visco P. and van Wijland F., *Phys. Rev. Lett.*, **117** (2016) 038103.

[45] Bialek W., Cavagna A., Giardina I., Mora T., Pohl O., Silvestri E., Viale M. and Walczak A. M., *Proc. Natl. Acad. Sci. U.S.A.*, **111** (2014) 7212.

[46] Melfo A., arXiv preprint, arXiv:1702.08067 (2017).

[47] Barberis L. and Peruani F., *Phys. Rev. Lett.*, **117** (2016) 248001.