Mutual information as a tool for identifying phase transitions in dynamical complex systems with limited data

R. T. Wicks, S. C. Chapman, and R. O. Dendy*
Centre for Fusion, Space and Astrophysics, University of Warwick, Coventry, CV4 7AL, United Kingdom
(Received 20 December 2006; published 30 May 2007)

We use a well-known model [T. Vicsek *et al.*, Phys. Rev. Lett. **15**, 1226 (1995)] for flocking, to test mutual information as a tool for detecting order-disorder transitions, in particular when observations of the system are limited. We show that mutual information is a sensitive indicator of the phase transition location in terms of the natural dimensionless parameters of the system which we have identified. When only a few particles are tracked and when only a subset of the positional and velocity components is available, mutual information provides a better measure of the phase transition location than the susceptibility of the data.

DOI: 10.1103/PhysRevE.75.051125 PACS number(s): 05.70.Fh, 89.75.−k, 05.45.Tp, 89.70.+c

I. INTRODUCTION

Order-disorder transitions are often found in complex systems. They have been identified in physical systems such as Bose-Einstein condensates and ferromagnets and in biological, chemical, and financial systems. Phase transitions are found, for example, in the behavior of bacteria [1], locusts [2], voting games, and utilization of resource in markets [3]. These systems have in common the property that there is competition between fluctuations driving the system towards disorder and inter-element interactions driving the system towards order. Insight into such systems can be gained using simple models. Although the dynamics of individual elements are difficult to predict, one can identify macroscopic parameters that characterize the behavior of the system. These can be approached through dimensional analysis—e.g., Buckingham's Π theorem [4].

A generic challenge in real-world measurements of physical, chemical, biological, or economic systems is that they yield data sets that are, in essence, sparse. Single elements such as tracer particles in turbulent flow, tagged birds or dolphins in a group, or a constituent of a financial index may, or may not, adequately sample the full underlying system behavior. In consequence, the behavior of a finite number of individual elements may, or may not, provide a proxy for the behavior of the entire system. If the system behavior is known to exhibit a phase transition, the question arises as to how this can best be captured from analysis of the dynamics of individual elements. Previously, for example, both mutual information (MI) [5–8] and susceptibility have been shown to be sensitive to the phase transition in the Ising spin model of ferromagnetism [9]. MI can also extract the correlation, or dependence, between causally linked but spatiotemporally separated observed parameters: for example, between *in situ* plasma measurements in the solar wind and the ionospheric response detected by ground-based measurements on Earth [10] or for example, within the brains of Alzheimer's disease patients [11].

Here we compare the use of MI and susceptibility to quantify the location of the phase transition in the dimensionless parameter space of the model of Vicsek *et al.* [12]. There are numerous statistical methods for analyzing systems with many degrees of freedom, dating back to work by Helmholtz and Boltzmann in the 19th century; see, for example, [13,14] and references therein. The microscopic behavior of the system of Vicsek *et al.*, described below, does not conserve energy or momentum, and this precludes a microscopic understanding of energy, momentum, or any related quantities in the system. The driving characteristic of the system is entropy: this is added by the random forcing of the particles and removed by their mutual interactions, which create correlation. With this in mind, we examine mutual information as a natural choice for capturing the entropy flow in the system and characterizing the system state with respect to its order parameters.

We find that when full knowledge of the system is available—that is, when all the particles are tracked—the susceptibility is an accurate method for estimating the position of the phase transition in the model of Vicsek *et al.* However, if the data are limited to a sample of just a few particles out of a large number, or a subset of the complete data, this method is less accurate.

We show that the mutual information of only a few particles, or of limited data from the whole system, can successfully locate the phase transition in dimensionless parameter space. For example, we find that the MI of a time series of components of particle position or velocity is sufficient. We thus show that MI can provide a practical method to detect order-disorder transitions when only a few particles, or elements, of the system are observed.

II. VICSEK MODEL

In 1995 Vicsek *et al.* [12] introduced the self-propelled particle model in which particles have a constant speed $|\underline{v}| = v_0$ and a varying direction of motion θ. In the discrete time interval $\delta t = t_{n+1} - t_n$ an isolated particle increments its vector position $\underline{x}_n \rightarrow \underline{x}_{n+1}$ by moving with constant speed v_0 in a direction θ_n which is in turn incremented at each time step. In the model, particles interact when they are within distance R of each other, such that the direction of their motion tends to become oriented with that of their neighbors. This interaction is implemented at each step, as shown in Fig. 1, by

*Also at UKAEA Culham Division, Culham Science Centre, Abingdon, Oxfordshire, OX14 3DB, UK.

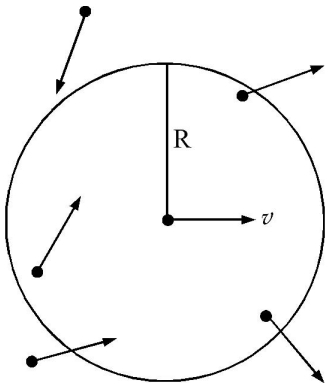

FIG. 1. Multiple particles interact if within a radius R of each other. Each of the N_R particles within R (here $N_R=4$) contributes its angle of propagation to the average $\langle \theta_n^{N_R} \rangle$, which is assigned to the particle at the center of R.

replacing the particle's direction of motion θ_n by the average of those particles, N_R, within distance R, so that $\theta_{n+1}=\langle \theta_n^{N_R} \rangle$ with a random angle $\delta\theta_n$ also added. The random fluctuation $\delta\theta_n$ is an independent identically distributed angle in the range $-\eta \leqslant \delta\theta_n \leqslant \eta$, where η characterizes the strength of the noise for the system. Thus for the ith particle in the system, after n time steps,

$$\underline{x}_{n+1}^i = \underline{x}_n^i + \underline{v}_n^i \delta t, \tag{1}$$

$$\theta_{n+1}^i = \langle \theta_n^{N_R} \rangle + \delta\theta_n^i, \tag{2}$$

$$\underline{v}_n^i = v_0(\cos\ \theta_n^i \underline{\hat{x}} + \sin\ \theta_n^i \underline{\hat{y}}). \tag{3}$$

Here, direction is defined by the angle from the x axis ($\underline{\hat{x}}$) and η is such that $\eta = \bar{\eta}\delta t$—that is, normalized to the time step δt.

There are two limiting cases for the system dynamics: disorder, where each particle executes a random walk, and order, where all particles move together with the same velocity. Figure 2 shows snapshots of the system dynamics for $\eta=0$, $\eta=2\pi/5 \simeq 1$, and $\eta=4\pi/5 > 1$. We see that $\eta \ll 1$ is highly ordered and $\eta \gg 1$ is highly disordered, and around $\eta \simeq 1$ there is a phase transition [15]. As with other critical systems it is possible to define an order parameter ϕ and a susceptibility χ [12,16–18]. For the Vicsek model, the magnitude of the average velocity of all the particles in the system provides a macroscopic order parameter and the variance of this speed is the susceptibility:

$$\phi = \frac{1}{Nv_0} \left| \sum_{i=1}^{N} \underline{v}^i \right|, \tag{4}$$

$$\chi = \sigma^2(\phi) = \frac{1}{N}(\langle \phi^2 \rangle - \langle \phi \rangle^2). \tag{5}$$

Here N denotes the total number of particles in an implementation of the model of Vicsek *et al.*

We plot ϕ and χ as a function of η in Fig. 3. In the thermodynamic limit $(N \to \infty, l \to \infty)$ where l is the system size, the susceptibility would tend to infinity at the critical noise η_c, where the phase transition occurs. In a finite-sized realization of the system, the susceptibility has a sharp but finite maximum at the critical noise at which the phase transition occurs. Finite-size effects make the peak location un-

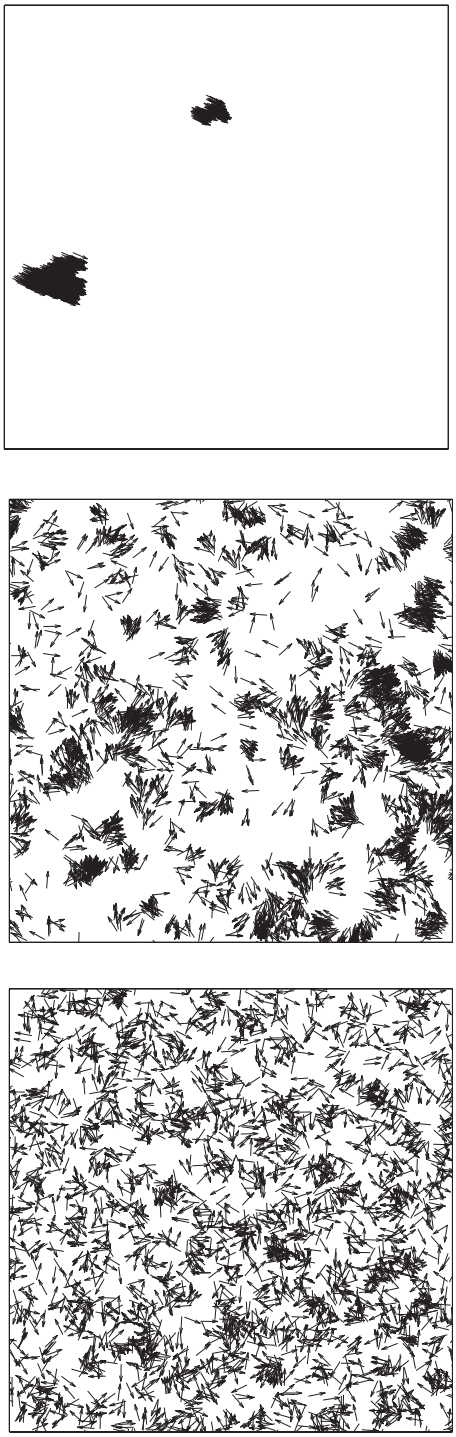

FIG. 2. The effect of increasing noise on a typical Vicsek system from ordered dynamics (top: $\eta=0$) to disordered dynamics (bottom: $\eta=4\pi/5$) and in the vicinity of the phase transition (middle: $\eta=2\pi/5$). Particle velocity vectors are plotted as arrows at the position of each particle in the x-y plane. The system has parameters $N=3000$, $|v|=0.15$, and $R=0.5$. This corresponds to $\Pi_2=0.3$, $\Pi_3=0.94$, and $\Pi_1=\eta$; see Eqs. (6)–(8).

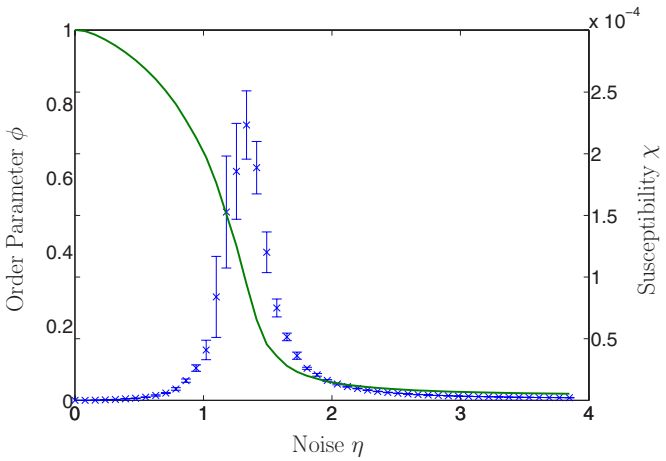

FIG. 3. (Color online) An example of a typical Vicsek system. The order parameter ϕ (line) is maximum for zero noise and falls to a constant small value at high noise. The susceptibility χ (crosses) peaks at the critical point $\eta_c \approx 1.33$ for the system. The system parameters are $\Pi_2=0.3$ and $\Pi_3=0.94$, with $N=3000$.

certain, but it is still possible to obtain an estimate of the critical noise η_c.

The system can be analyzed using Buckingham's Π theorem [4], and three independent dimensionless quantities can be found that characterize its behavior. The first of these (Π_1) is the amplitude of the normalized noise η, the second (Π_2) is the ratio of the distance traveled in one time step $v_0\delta t$ to the interaction radius R, and the third (Π_3) is the average number of particles within a circle of one interaction radius R:

$$\Pi_1 = \eta = \bar{\eta}\delta t, \tag{6}$$

$$\Pi_2 = v_0\delta t/R, \tag{7}$$

$$\Pi_3 = \pi R^2 \rho. \tag{8}$$

These three parameters determine the behavior of the system in the thermodynamic limit ($N\to\infty$, $l\to\infty$, R and ρ finite) where ρ denotes the number density of particles over the whole system.

The system size l affects the number of interactions that occur. If l is finite and the system is periodic as here, the finite system size increases the chance of two randomly chosen particles interacting, compared to the limit of infinite l. The system only approaches the thermodynamic limit when the finite interaction radius $R \ll l$. Conversely, for example, if the interaction radius is half the diagonal size of the system, then all the particles interact with each other at any given moment. This implies a fourth parameter reflecting the finite size of any computer-based realization of this model:

$$\Pi_4 = R/l. \tag{9}$$

In the thermodynamic limit we have $N\to\infty$, $l\to\infty$, while R and $\rho=N/l^2$ are finite, so that $\Pi_4\to 0$ and Π_{1-3} are finite and specify the system.

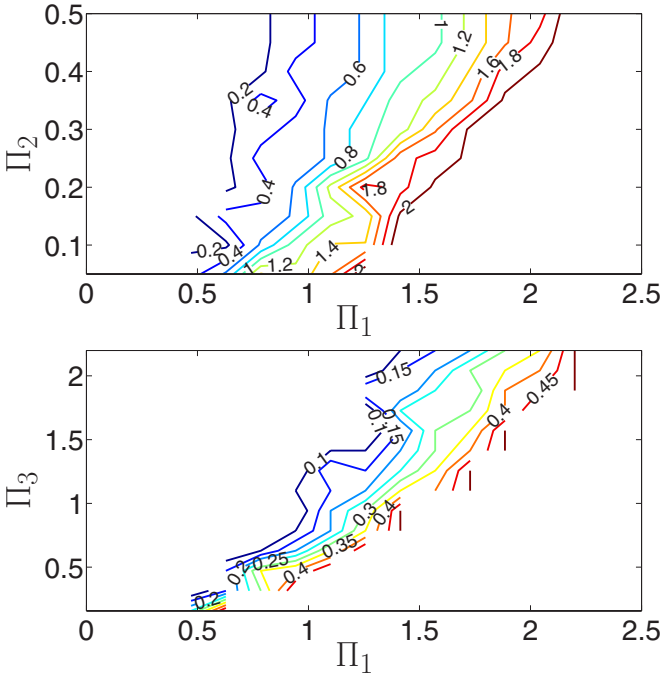

FIG. 4. (Color online) Phase transition diagram contours for the Vicsek model around $\Pi_3=1$. Top panel: the effect of changing Π_3, from $\Pi_3=0.2$ (dark blue contours, left-hand side) to $\Pi_3=2.0$ (dark red contours, right-hand side) in steps of 0.2, on the position of the phase transition in the Π_1, Π_2 plane. Bottom panel: the effect of changing Π_2, from $\Pi_2=0.05$ (dark blue, left-hand side) to $\Pi_2=0.5$ (dark red, right-hand side) in steps of 0.05, on the position of the phase transition in the Π_1, Π_3 plane.

III. SYSTEM PHASE SPACE

For given values of Π_2 and Π_3, we run simulations of the Vicsek system for a range of values of Π_1 to determine the value $\Pi_1=\Pi_1^c$ at which the susceptibility χ peaks and thus the phase transition occurs. By repeating this operation for a set of parameter values of Π_2 and Π_3, we obtain the full set of coordinates at which the phase transition is located for a region of the phase space around $\Pi_3=1$. We show this graphically in Fig. 4 where we plot contours of $\Pi_3^c(\Pi_1,\Pi_2)$ in the upper panel and $\Pi_2^c(\Pi_1,\Pi_3)$, in the lower panel. These plots confirm that there is a smooth, well-defined surface of Π_1^c, Π_2^c, and Π_3^c; they can be used to inform the choice of Π_1, Π_2, and Π_3 for the next section.

In relation to recent work by Nagy *et al.* [19], we see from Fig. 4 that the speed v_0 of the particles, which is a variable within Π_2, has a characteristic effect above $v_0\approx 0.3$. The bottom panel of Fig. 4 shows that, for constant Π_3, the phase transition becomes hard to detect and the contours start to break up as Π_2 is increased above approximately 0.3. This is a complementary demonstration of the statement made in [19] that the phase transition becomes first order in the high-velocity ($v_0 \geqslant 0.3$) regime.

IV. MUTUAL INFORMATION

Mutual information quantifies the information content shared by two signals A and B. For discrete signals we can write the MI as

$$I(A,B) = \sum_{i,j}^{m} P(a_i,b_j)\log_2\left(\frac{P(a_i,b_j)}{P(a_i)P(b_j)}\right). \quad (10)$$

Here the signal A has been partitioned into an alphabet (a library of possible values the signal can take) $A = \{a_1, \ldots, a_i, \ldots, a_m\}$ where a_1 and a_m are the extrema of A found in all data considered. The discretized signal takes value a_i with probability $P(a_i)$ and similarly for b_i we have $P(b_i)$, while $P(a_i,b_j)$ is the joint probability of a_i and b_j occurring together. The chosen base of the logarithm defines the units in which the mutual information is measured. Normally base 2 is used, so that the mutual information is measured in bits. If we define the entropy of a signal as:

$$H(A) = -\sum_{i}^{m} P(a_i)\log_2[P(a_i)], \quad (11)$$

then MI can be written as a combination of entropies [5]:

$$I(A,B) = H(A) + H(B) - H(A,B). \quad (12)$$

The calculation of the entropies needed to form the MI is not trivial, as there is some freedom in the method of discretization of the signals and in the method used to estimate the probabilities $P(a_i)$, $P(b_j)$, and $P(a_i,b_j)$. There are many different methods currently used, summarized and compared by Cellucci *et al.* [6] and Kraskov *et al.* [7].

MI has been used in the analysis of the two-dimensional Ising model by Matsuda *et al.* [9]. Importantly the critical temperature for the Ising model is identified precisely by the peak in the mutual information of the whole system. This peak survives the coarse graining of the system very well, which raises the possibility that mutual information can be used in the study of other complex systems.

V. IDENTIFYING THE PHASE TRANSITION

A. Full system mutual information

In the two-dimensional (2D) Vicsek system there are three variables for each of the N particles: their positions (x^i, y^i) and the orientation of their velocities θ^i, giving three signals X, Y, and Θ, each containing N measurements at every time step. The simplest discretization of these signals x^i, y^i, and θ^i is to cover the range of the signals with equally spaced bins, so for position coordinate X we have m bins X_i with width δX. Then, if n particles are in the range $(X_i - X_i + \delta X)$, we have probabilities

$$P(X_i) = \frac{n(X_i)}{N\delta X}, \quad (13)$$

$$\sum_i P(X_i) = 1. \quad (14)$$

The single and joint probabilities $P(Y_j)$, $P(\Theta_k)$, $P(X_i,\Theta_k)$, and $P(Y_j,\Theta_k)$ are calculated in a similar manner.

The key factor governing the accuracy with which MI is measured is to optimize the size of the bins used in the above procedure. If the bins are too large, then resolution is lost and the exact details of small scale structure and clustering cannot be identified. If the bins are too small, then at high noise the probability of finding a particle at a given point does not become smoothed over the whole system because individual particles can be resolved, giving $P(x_i,y_j) \neq P(x_i)P(y_j)$, even though the system is in a well-mixed random state.

There is no ideal bin structure determined for this method of MI calculation [6,7]. The Vicsek model has two natural length scales, R the interaction radius and l the box size, so that a good length scale to choose for discretization, when a snapshot of the whole system is being used, is the interaction radius R. Thus all our mutual information calculations made on the whole system use a bin size of $2R$, the diameter of the circle of interaction; the bins are therefore squares of size $4R$ in the (x,y) plane. When θ is discretized the same number of bins are used as for x or y because there is no natural size for bins in θ.

Given full knowledge of x_i, y_i, and θ_i for all N particles in the system over a large number of time steps, several different calculations of mutual information can be made. We find that the most accurate form of mutual information for the whole system is that calculated between the x or y position and θ. Thus we perform the following calculation at each time step n once the system has reached a stable state:

$$I(X,\Theta) = \sum_{i,j} P(X_i,\Theta_j)\log_2\left(\frac{P(X_i,\Theta_j)}{P(X_i)P(\Theta_j)}\right), \quad (15)$$

$$I(Y,\Theta) = \sum_{i,j} P(Y_i,\Theta_j)\log_2\left(\frac{P(Y_i,\Theta_j)}{P(Y_i)P(\Theta_j)}\right), \quad (16)$$

$$I = \frac{I(X,\Theta) + I(Y,\Theta)}{2}, \quad (17)$$

and average over all time steps for which MI is measured.

We compare the MI as calculated using the above method and the susceptibility as a function of normalized noise η in Fig. 5. At large η the MI falls to zero as X, Y, and Θ tend to uncorrelated noise (see also [9]). We would also expect the MI to fall to zero at sufficiently low η as the system becomes ordered and this behavior is also seen within the errors. The errors on our measurements of MI are calculated from the standard deviation of measurements of MI calculated over 50 simulations at each noise η. The error on the susceptibility is calculated in the same manner.

The error bars become larger at low η because the mutual information includes the signatures of spatial clustering as well as velocity clustering in the measurement. Thus at low η, when extended clusters form, the mutual information will give a higher value for the more spatially extended axis of the cluster and a lower value for the less extended axis of the cluster. This implies that the shape and orientation of the (usually single) large cluster formed at low noise influences the mutual information. Different measurements of MI thus arise for each implementation of the model, giving rise to the error seen at low η. This could be corrected by using other

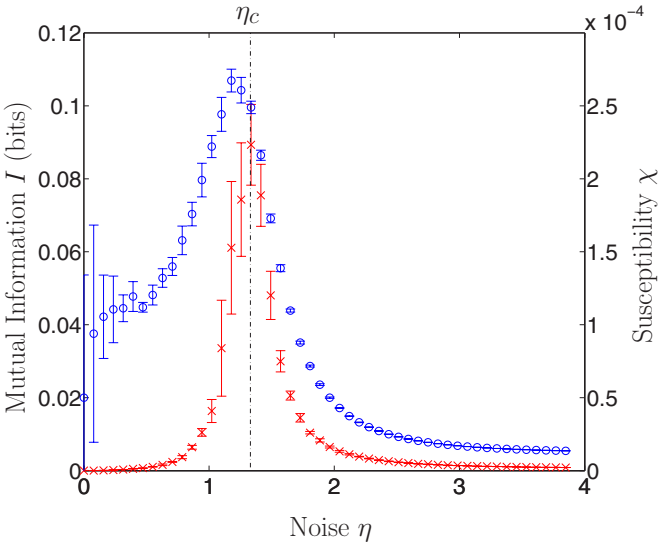

FIG. 5. (Color online) The mutual information I (circles) defined by Eq. (17) peaks at approximately the same point as the susceptibility χ (crosses) defined by Eq. (5). The critical noise $\eta_c \approx 1.33$ is marked. The system parameters are $\Pi_2=0.15$ and $\Pi_3=0.98$, with $N=3000$ particles. Error bars on the susceptibility are largest around η_c, unlike those on mutual information.

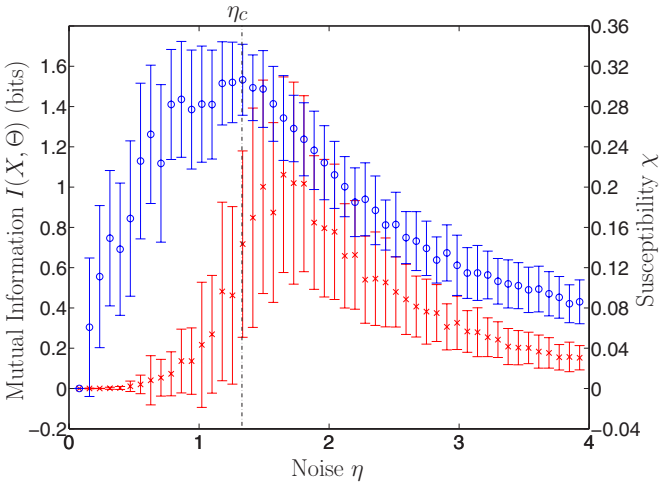

FIG. 6. (Color online) The mutual information I (circles) calculated using time series from only 10 particles for 5000 time steps, with $S=10$, compared to the average susceptibility $\bar{\chi}$ (crosses) for the same data using $S'=10$ subsections to calculate $\bar{\chi}$ and with the critical noise $\eta_c \approx 1.33$ marked. System parameters are $\Pi_2=0.3$ and $\Pi_3=0.94$, with $N=3000$ particles.

approaches to computing MI—for example, recurrence plots [8,10] or a different distribution of bins; these are more computationally intensive, however.

When estimated as the standard deviation over 50 repeated runs of the simulation, the error is found to be considerably larger, as a fraction of the overall measurement, for the susceptibility than for the MI. This arises since the susceptibility is simply an average fluctuation over all the velocity vectors of the system, whereas the MI also directly reflects the level of spatial "clumpiness" (that is, spatial correlation) of the particles. The detailed spatial distribution varies from one simulation to the next, but at fixed Π_{1-4} the degree of clumpiness does not. Mutual information is able to quantify clustering (correlation in space as well as velocity) in a simple dynamical complex system, in a manner that identifies the order-disorder phase transition.

B. Mutual information from limited data

Observations of many real-world systems typically provide only a subset of the full-system information, which here comprises the positions and velocities of all N interacting particles. We now consider results from the Vicsek model using only very limited amounts of data. The mutual information and susceptibility are now calculated on a $\tau=5000$ step time series of positional and velocity data for $n=10$ particles out of the $N=3000$ simulated. To optimize both methods, the data for each particle time series are cut into S sections, labeled $s=1,\ldots,S$, of length $N_s=\tau/S$ steps. This gives us nS pseudosystems, relying on the assumption that one particle over N_s steps is equivalent to N_s particles at one step. This is a reasonable assumption to make for the Vicsek model as it is ergodic while η remains constant.

To calculate the susceptibility, we need to estimate the variance of the average velocity of each of these nS pseudosystems. We therefore cut each section s into S' further subsections s', calculate the average velocity $\phi^i_{s'}$ of these subsections, and find their variance, giving χ^i_s the pseudosystem susceptibility. This is done for each pseudosystem individually to give χ^i_s and averaged over all nS pseudosystems to give $\bar{\chi}$, the average variance of the average velocity for all pseudosystems:

$$\phi^i_{s'} = \frac{SS'}{\tau v_0}\left|\sum_{k=1}^{\tau/SS'} \underline{v}^i_k\right|, \tag{18}$$

$$\chi^i_s = \frac{1}{S'}(\langle \phi^2_{s'}\rangle - \langle \phi_{s'}\rangle^2), \tag{19}$$

$$\bar{\chi} = \frac{1}{nS}\sum_{s=1}^{S}\sum_{i=1}^{n}\chi^i_s. \tag{20}$$

The result is shown in Fig. 6 where we also plot the mutual information $I(X,\Theta)$ from Eq. (15), but now as nS time series; the parameters used are $n=10$, $S=10$, and $S'=10$. The error bars are calculated as the standard deviation of the 100 measurements made using the different pseudosystems of length $\tau/S=500$ time steps. These values for n, S, and S' are chosen so as to limit the data in a realistic way. $n=10$ is a suitably small subset of the $N=3000$ particles. $S=10$ cuts the data into segments sufficiently long (500 time steps) to be treated independently. $S'=10$ is chosen so that each section s' is still long enough (50 time steps) to make as good an estimate of the average velocities $\phi^i_{s'}$ as possible, but allows enough of these measurements to be made to reduce the error in the measurement of χ^i_s.

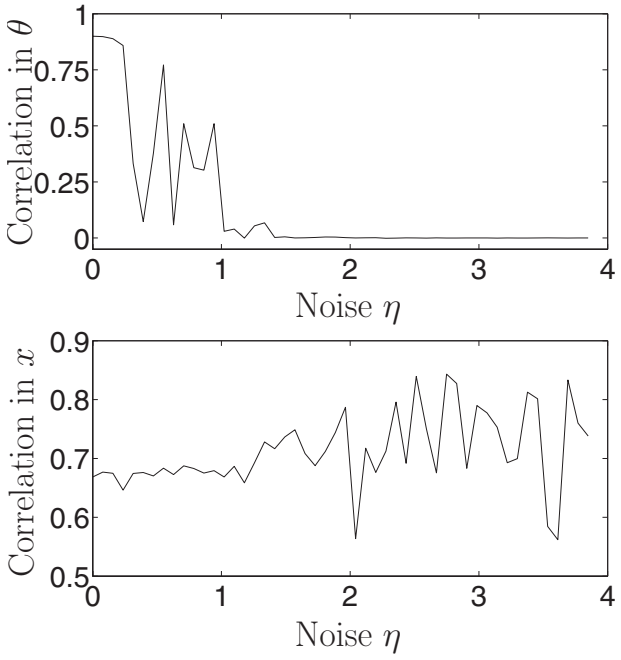

FIG. 7. The cross correlation between a randomly chosen particle and 9 others, calculated using a time series with 5000 steps. The top panel shows the average cross correlation between θ^1 and θ^k, $2 \leqslant k \leqslant 10$. The bottom panel shows the average cross correlation between x^1 and x^k, $2 \leqslant k \leqslant 10$. System parameters are $\Pi_2 = 0.3$ and $\Pi_3 = 0.94$, with $N = 3000$ particles.

The system is identical to that shown in Fig. 5, and the phase transition is at the same noise, $\eta_c \approx 1.33$. Near their respective peaks, the error in the mutual information remains smaller than that in the susceptibility and so MI better identifies the peak. The peak in the susceptibility no longer coincides with η_c and is shifted to the higher-noise side of the phase transition. This occurs because the susceptibility is now measured on too small a sample of data: only 50 angles θ_t^i are averaged to find each subsection velocity $\phi_{s'}^i$. Such a small ensemble average results in a large deviation in the average velocity from the expected value.

For comparison with a linear measure, we calculate the cross correlation for our ten trajectories. We choose one of the particles at random and compute its cross correlation with each of the remaining nine. The average of these is plotted in Fig. 7. The average cross correlation between angles θ^1 and θ^k, $2 \leqslant k \leqslant 10$, in the top panel shows strong correlation at low noise, as expected. This cross correlation declines as noise increases, but not smoothly, because the correlation depends on the exact dynamics of the particles considered. Angular cross correlation reaches zero around the phase transition, but does not provide an accurate location for the critical noise. In the bottom panel of Fig. 7 the cross correlation between x^1 and x^k, $2 \leqslant k \leqslant 10$, provides no reliable indication of the position of the phase transition. The cross correlation does become more variable on the higher-noise side of the graph but this effect cannot be used to accurately find the critical noise η_c.

The value of using mutual information can be seen when the available data are restricted still further. Let us consider

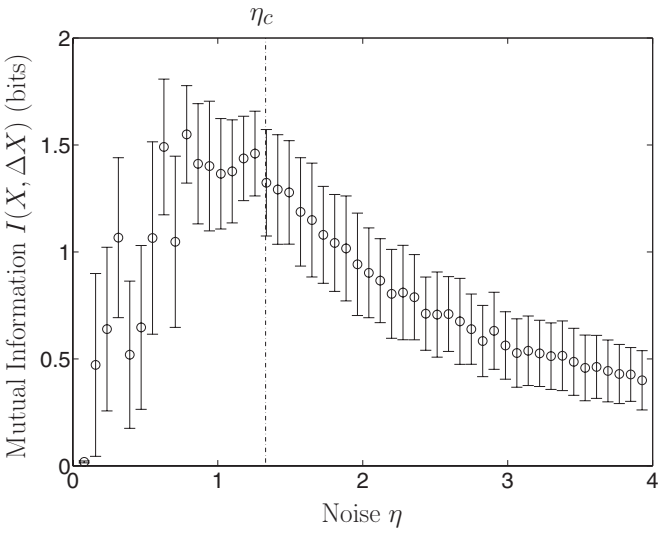

FIG. 8. The mutual information $I(X, \Delta X)$ (circles) calculated using a time series from only 10 particles for 5000 time steps, with $S = 10$ and the critical noise $\eta_c \approx 1.33$ marked. System parameters are $\Pi_2 = 0.3$ and $\Pi_3 = 0.94$, with $N = 3000$ particles.

signals derived from one component of the particle trajectory only, equivalent to a line-of-sight measurement. We then have one of the position coordinates x_k^i and the instantaneous x component of the velocity, $\Delta x_k^i = v_0 \cos(\theta_k^i)$. The susceptibility is calculated as in Eqs. (18)–(20), but using the average one-dimensional velocities Δx_k^i:

$$\phi_{s'}^i = \frac{SS'}{\pi v_0} \left| \sum_{k=1}^{\pi/SS'} \Delta x_k^i \right|, \tag{21}$$

$$\chi_s^i = \frac{1}{S'} (\langle \phi_{s'}^2 \rangle - \langle \phi_{s'} \rangle^2), \tag{22}$$

$$\bar{\chi} = \frac{1}{nS} \sum_{s=1}^{S} \sum_{i=1}^{n} \chi_s^i. \tag{23}$$

The mutual information is calculated for each section of the x only (and later y only) components of the time series for each particle using a suitable binning:

$$I(X, \Delta X) = \sum_{i,j} P(X_i, \Delta X_j) \log_2 \left(\frac{P(X_i, \Delta X_j)}{P(X_i) P(\Delta X_j)} \right), \tag{24}$$

$$I(Y, \Delta Y) = \sum_{i,j} P(Y_i, \Delta Y_j) \log_2 \left(\frac{P(Y_i, \Delta Y_j)}{P(Y_i) P(\Delta Y_j)} \right). \tag{25}$$

Figure 8 shows the mutual information calculated from the data in this manner with $S = 10$. The peak in the mutual information is at approximately the correct value of η ($\eta_c \approx 1.33$). Figure 9 shows for comparison the susceptibility calculated over the X data as in Eqs. (21)–(23). We see that although there is a peak, it no longer identifies $\eta \rightarrow \eta_c$ accurately. The peak is broader and has larger error bars than in Fig. 8, giving a large uncertainty in identifying η_c.

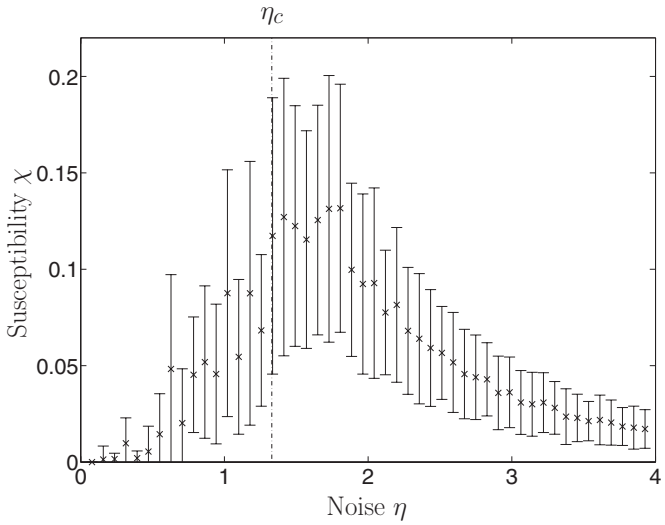

FIG. 9. The susceptibility χ (crosses) calculated using a time series of one-dimensional data $\{X,\Delta X\}$ from only 10 particles for 5000 time steps, with $S=10$ and the critical noise $\eta_c \approx 1.33$ marked. System parameters are $\Pi_2=0.3$ and $\Pi_3=0.94$, with $N=3000$ particles.

The peak in Fig. 8 is shifted to the low-noise side of the phase transition and shows some scatter. This can be understood by looking at the same data using a different value of the interval S. In Fig. 10 we show the same data analyzed using $S=1$; that is, we consider one time series of length 5000 time steps for each of 10 particles and obtain MI averaged over these 10. We plot $I(X,\Delta X)$ (circles) and $I(Y,\Delta Y)$ (squares). The measurements overlap within errors on the high-noise side of the phase transition but separate into two distinct branches, containing both $I(X,\Delta X)$ and $I(Y,\Delta Y)$, on the low-noise side.

One potential source of this behavior is that, as the system becomes ordered at $\eta < \eta_c$, the particles clump together. This

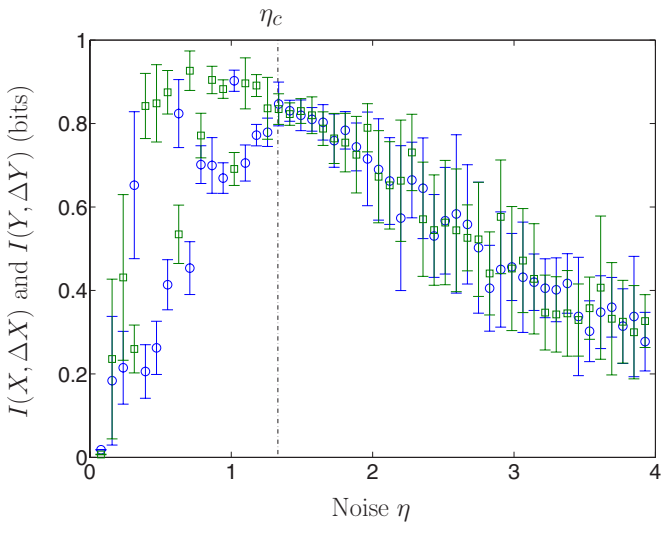

FIG. 10. (Color online) The mutual information $I(X,\Delta X)$ (circles) and $I(Y,\Delta Y)$ (squares) calculated using a time series from only 10 particles for 5000 steps, with $S=1$ and the critical noise $\eta_c \approx 1.33$ marked. System parameters are $\Pi_2=0.3$ and $\Pi_3=0.94$, with $N=3000$ particles.

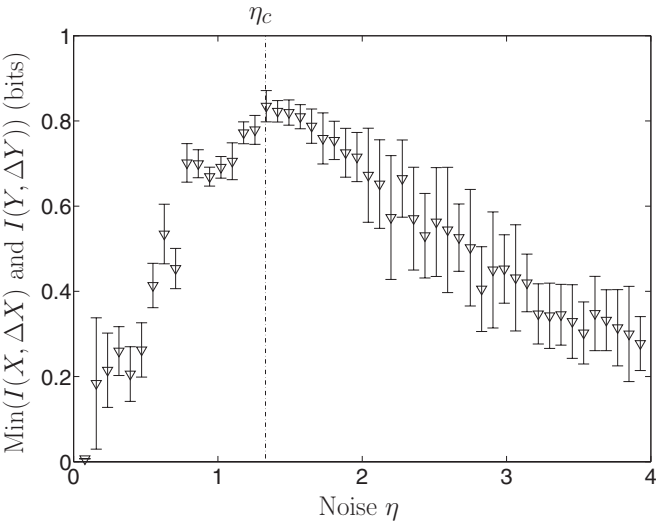

FIG. 11. The minimum results from mutual information measurements $I(X,\Delta X)$ and $I(Y,\Delta Y)$ calculated using a time series from only 10 particles for 5000 steps, with $S=1$ and the critical noise $\eta_c \approx 1.33$ marked. System parameters are $\Pi_2=0.3$ and $\Pi_3=0.94$, with $N=3000$ particles.

implies that the particles together take on a preferred direction of motion; in addition, a clump may be elongated in a particular spatial direction. The effectiveness of MI will then depend on whether our single-component (line-of-sight) data are aligned along, or perpendicular to, the characteristic directions of this clump. Mutual information measured in terms of coordinates aligned with the preferred direction of motion is increased by the dispersion of particle positions and velocities, whereas MI measured in term of perpendicular coordinates is decreased because the positional and velocity dispersions are smaller in this direction. Anomalously high MI measurements result from making measurements along the preferred direction of motion; large relative velocities lead to anomalously high peaks on the low-noise side of the phase transition, making it appear to be shifted towards $\eta=0$.

Finally in Fig. 11 we plot the minimum of $I(X,\Delta X)$ and $I(Y,\Delta Y)$ for each value of η from Fig. 10 and see that a clear peak emerges at $\eta=\eta_c$, where the error bars are the smallest. This outcome obviates the difficulty that arises if we only allow knowledge of $\{X,\Delta X\}$, for example, when it would be necessary to exclude high measurements of MI at low noise, as discussed above.

VI. CONCLUSIONS

The Vicsek model [12] is used here to test the potential of measurements of order and clustering that exploit mutual information in dynamical complex systems. We find that when complete knowledge of the system is available, the mutual information has a smaller error than the susceptibility (Fig. 5). Using Buckingham's Π theorem, the set of dimensionless parameters that capture the phase space of the Vicsek model have been presented as a complete set for the first time.

When data are limited to observations of only 10 particles out of 3000, the error in the mutual information remains comparatively small and the mutual information thus provides a better measurement than susceptibility of the position of the order-disorder phase transition (Fig. 6). When data are limited still further, such that only one line-of-sight component of the particle motion is available, the mutual information measurement remains sensitive enough to identify the critical noise of the phase transition, while the susceptibility does not (Figs. 8–11). In this case the mutual information also provides an indication of the preferred axial direction of clumped particle motion at low noise. Anomalously high mutual information estimates in this ordered phase indicate that the particles sampled are mostly moving along the dimension being measured; low estimates indicate that the particles are moving perpendicularly. This is remarkable given that the susceptibility does not contain this information and that the MI is a probabilistic measurement.

Real-world data are often in the form of the final data studied here, a limited sample from a much larger set, measured in fewer dimensions than those of the original system: for example, line-of-sight measurements of wind speed measured by an anemometer at a weather station, or satellite measurements of the solar wind. It has been shown here that mutual information can provide an effective measure of the onset of order and may provide a viable technique for real-world data with its inherent constraints.

ACKNOWLEDGMENTS

R.W. acknowledges support from EPSRC and CASE in association with UKAEA. The authors would like to thank K. Kiyani for valuable discussions and the Centre for Scientific Computing at the University of Warwick for providing the computing facilities.

[1] A. Czirók, E. Ben-Jacob, I. Cohen, and T. Vicsek, Phys. Rev. E **54**, 1791 (1996).
[2] J. Buhl, D. J. T. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller, and S. J. Simpson, Science **312**, 1402 (2006).
[3] R. Savit, R. Manuca, and R. Riolo, Phys. Rev. Lett. **82**, 2203 (1999).
[4] Malcolm Longair, *Theoretical Concepts in Physics: An alternative view of theoretical reasoning in physics*, 2nd ed. (Cambridge University Press, Cambridge, England, 2003).
[5] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).
[6] C. J. Cellucci, A. M. Albano, and P. E. Rapp, Phys. Rev. E **71**, 066208 (2005).
[7] A. Kraskov, H. Stögbauer, and P. Grassberger, Phys. Rev. E **69**, 066138 (2004).
[8] T. K. March, S. C. Chapman, and R. O. Dendy, Physica D **200**, 171 (2005).
[9] H. Matsuda, K. Kudo, R. Nakamura, O. Yamakawa, and T. Murata, Int. J. Theor. Phys. **35**, 839 (1996).
[10] T. K. March, S. C. Chapman, and R. O. Dendy, Geophys. Res. Lett. **32**, L04101 (2005).
[11] J. Jeong, J C. Gore, and B. S. Peterson, Clin. Neurophysiol. **112**, 827 (2001).
[12] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Phys. Rev. Lett. **75**, 1226 (1995).
[13] P. V. Coveney, Nature (London) **333**, 409 (1988).
[14] O. Penrose, Rep. Prog. Phys. **42**, 1939 (1979).
[15] G. Grégoire and H. Chaté, Phys. Rev. Lett. **92**, 025702 (2004).
[16] G. Grégoire, H. Chaté, and Y. Tu, Physica D **181**, 157 (2003).
[17] A. Czirók, H. E. Stanley, and T. Vicsek, J. Phys. A **30**, 1375 (1997).
[18] A. Czirók and T. Vicsek, Physica A **281**, 17 (2000).
[19] M. Nagy, I. Daruka, and T. Vicsek, Physica A **373**, 445 (2007).