

Local equilibrium in bird flocks

Thierry Mora^{1*}, Aleksandra M. Walczak², Lorenzo Del Castello^{3,4}, Francesco Ginelli⁵, Stefania Melillo^{3,4}, Leonardo Parisi^{4,6}, Massimiliano Viale^{3,4}, Andrea Cavagna⁴ and Irene Giardina^{3,4}

The correlated motion of flocks is an example of global order emerging from local interactions. An essential difference with respect to analogous ferromagnetic systems is that flocks are active: animals move relative to each other, dynamically rearranging their interaction network. This non-equilibrium characteristic has been studied theoretically, but its impact on actual animal groups remains to be fully explored experimentally. Here, we introduce a novel dynamical inference technique, based on the principle of maximum entropy, which accommodates network rearrangements and overcomes the problem of slow experimental sampling rates. We use this method to infer the strength and range of alignment forces from data of starling flocks. We find that local bird alignment occurs on a much faster timescale than neighbour rearrangement. Accordingly, equilibrium inference, which assumes a fixed interaction network, gives results consistent with dynamical inference. We conclude that bird orientations are in a state of local quasi-equilibrium over the interaction length scale, providing firm ground for the applicability of statistical physics in certain active systems.

Animal groups moving in concert, such as mammal herds, fish schools, and bird flocks show that in biology, just as in physics, local coordination can result in large-scale order^{1–3}. However, flocks differ from classical statistical physics in that their constituents are active: they constantly move by self-propulsion, pumping energy into the system and keeping it out of equilibrium^{4–7}. The key element is the rearrangement of the interaction network due to the active motion of individuals relative to each other, continuously changing their neighbours. Theoretical studies show that network rearrangement has major consequences, which include enhancing collective order, reducing the lower critical dimension from 3 to 2, and affecting the critical exponents^{4,8}.

However, the importance of activity must be assessed with respect to the relevant timescales of the system. The impact of network rearrangement depends on the interplay between its characteristic timescale, τ_{network} , defined as the average time it takes an individual to change its interaction neighbours, and the local relaxation timescale, τ_{relax} , defined as the time needed to relax locally the order parameter if the interaction network were fixed. If $\tau_{\text{network}} \leq \tau_{\text{relax}}$, the interaction network rearranges at least as fast as the order parameter relaxes, and the system remains far from equilibrium. If, on the other hand, $\tau_{\text{relax}} \ll \tau_{\text{network}}$, the relaxation of the order parameter is adiabatic, closely following the network as it slowly evolves. In this case, even though the system behaves in an out-of-equilibrium manner on the longest scales, it locally obeys a condition of equilibrium, and we expect some of the tools of equilibrium statistical physics to be applicable.

Here, we explicitly address the impact of network activity by developing a new inference method based on the exact integration of maximum-entropy dynamical equations, thus accounting for the reshuffling of the network. We apply the method to data of starling flocks of up to 600 individuals^{9–12} (see Methods and Supplementary Table 1 for data summary), inferring the relevant parameters of the interactions between individuals. We find that the alignment relaxation time, τ_{relax} , is more than one order of magnitude shorter

than the network rearrangement time, τ_{network} . Consistently, we show that the parameters learnt from the dynamics are consistent with those obtained by an equilibrium-like inference, which assumes a fixed network¹³. Our results suggest that natural flocks are in a state of local quasi-equilibrium over the interaction length scale, meaning that the relatively slow rearrangement of the local interaction network does not affect the ordering dynamics up to certain scales.

To compare the relevant timescales of the ordering process in flocks, we first need to learn the dynamical rules of their behaviour. Learning these rules usually relies on inferring the parameter of a chosen model directly from the data, as has been recently done in surf scoters¹⁴ and fish schools^{15–19}. Although in these studies the local rules of interaction were often learnt using small groups, in some cases they could also be used to predict large-group behaviour^{17,19}. Here, instead of assuming a model a priori, we apply the principle of maximum entropy to the trajectories of all birds in the group²⁰. We look for a distribution of the stochastic process that is as random as possible, while agreeing with the data on a key set of experimental observables.

In a flock of size N , we call $\vec{s}_i(t)$ the three-dimensional flight orientation of bird i at time t . The maximum-entropy distribution over possible flock trajectories that is consistent with the correlation functions $\langle \vec{s}_i(t) \cdot \vec{s}_j(t) \rangle$, as well as their derivatives $\langle d\vec{s}_i(t)/dt \cdot \vec{s}_j(t) \rangle$, can be exactly mapped, in the limit of strong polarization $P \equiv (1/N) \|\sum_i \vec{s}_i\| \approx 1$, onto the following stochastic differential equation (see Supplementary Information and ref. 20):

$$\frac{d\vec{s}_i}{dt} = \left(\sum_j J_{ij} \vec{s}_j + \vec{\xi}_i \right)_{\perp} \quad (1)$$

where $\vec{\xi}_i$ is a random white noise, and where the projection $\vec{x}_{\perp} \equiv \vec{x} - \vec{s}_i(\vec{x} \cdot \vec{s}_i)$ onto the plane perpendicular to \vec{s}_i ensures that \vec{s}_i remains of norm 1. Equation (1) can be viewed as a generalization of the Vicsek model²¹: each bird modifies its flight direction according

¹Laboratoire de physique statistique, CNRS, UPMC and École normale supérieure, 24, rue Lhomond, 75005 Paris, France. ²Laboratoire de physique théorique, CNRS, UPMC and École normale supérieure, 24, rue Lhomond, 75005 Paris, France. ³Dipartimento di Fisica, Università Sapienza, 00185 Rome, Italy. ⁴Istituto Sistemi Complessi, Consiglio Nazionale delle Ricerche, UOS Sapienza, 00185 Rome, Italy. ⁵SUPA, Institute for Complex Systems and Mathematical Biology, Kings College, University of Aberdeen, Aberdeen, AB24 3UE, UK. ⁶Dipartimento di Informatica, Università Sapienza, 00198 Rome, Italy. *e-mail: tmora@lps.ens.fr

to a weighted average of the directions of its neighbours. The interaction matrix J_{ij} encodes how much bird i is influenced by (that is, interacts with) bird j . Given the experimentally measured correlation functions, entropy maximization yields equations that fix the values of the noise amplitude and the interaction matrix J_{ij} . This matrix has too many parameters to be reliably determined from the data, but we can reduce its complexity by parametrizing it. It was shown in ref. 22 that the interaction decays exponentially with the topological distance k_{ij} between birds,

$$J_{ij} = J \exp(-k_{ij}/n_c) \quad (2)$$

where k_{ij} denotes the (time-dependent) rank of bird j among the neighbours of bird i ranked by distance. This interaction matrix has just two parameters: n_c is the topological interaction range, while J is the overall strength of the interaction. The noise is uncorrelated among birds and of uniform magnitude T , by analogy with physical temperature: $\langle \xi_i(t) \cdot \xi_j(t') \rangle = 2dT \delta_{ij} \delta(t - t')$, where d is the space dimension ($d = 3$ in the following).

In principle, to learn the parameters of equation (1), one needs actual continuous-time derivatives. In practice, we only have configurations separated by the finite experimental sampling time dt . A common solution is to use Euler's approximation:

$$\vec{s}_i(t + dt) \approx \vec{s}_i(t) + dt \sum_j J_{ij} \vec{s}_{j\perp} + \sqrt{2T dt} \vec{\eta}_{i\perp} \quad (3)$$

where $\vec{\eta}_i$ is a normally distributed vector of variance 1 in each direction. The conditional likelihood of the data given the model, $P[\{\vec{s}_i(t + dt)\} | \{\vec{s}_i(t)\}]$, can be written in Gaussian form after expanding equation (3) in the spin-wave approximation (see Methods). Maximizing this likelihood yields values for the alignment parameters n_c , J and T (see ref. 20 and Supplementary Information).

Euler's approximation is used by virtually all methods that try to fit a dynamical equation to a discrete time series^{15–17}. However, it is inappropriate when the experimental sampling time, dt , is larger than the intrinsic relaxation timescale, τ_{relax} . In this case information spreads between subsequent frames beyond the directly interacting neighbours and Euler's approximation overestimates the range of the interaction, as we shall see below. To overcome this issue, we rewrite equation (1) by formally subtracting $\sum_i J_{ij} \vec{s}_{i\perp} = 0$ from it:

$$\frac{d\vec{s}}{dt} = -J \mathbf{\Lambda} \vec{s}_{\perp} + \vec{\xi}_{\perp} \quad (4)$$

Bold symbols denote vectors and matrices over bird indices; the matrix $\Lambda_{ij} \equiv \delta_{ij} \sum_l n_{il} - n_{ij}$, where $n_{ij} = e^{-k_{ij}/n_c}$ is the connectivity matrix (2). $\mathbf{\Lambda}$ is analogous to a Laplacian defined on a lattice, and obeys the sum rule: $\sum_j \Lambda_{ij} = 0$. In the spin-wave approximation, where all orientations \vec{s}_i point in almost the same direction, this relation ensures that $\mathbf{\Lambda} \vec{s}$ has almost no contribution along the common direction of flight, implying $(\mathbf{\Lambda} \vec{s})_{\perp} \approx \mathbf{\Lambda} \vec{s}$ (see Methods and Supplementary Information). Equation (4) is now linear and it can be integrated exactly:

$$\vec{s}(t + dt) = e^{-J \mathbf{\Lambda} dt} \vec{s}(t) + \int_0^{dt} du e^{-J \mathbf{\Lambda} (dt-u)} \vec{\xi}_{\perp}(t + u) \quad (5)$$

This result assumes a constant J_{ij} in the interval dt , which is a good approximation if $dt \ll \tau_{\text{network}}$. Fortunately, this same condition is necessary for the very possibility to collect data: tracking requires one to follow each individual across time, which is possible only if individuals do not significantly change their neighbourhood between consecutive frames. The integrated noise in the right-hand side of (5) is Gaussian, of mean zero and covariance $4T \int_0^{dt} du e^{-J \mathbf{\Lambda} u} e^{-J \mathbf{\Lambda}^\top u}$. Using the exact solution (5) we can write an explicit expression for the (Gaussian) conditional likelihood $P[\{\vec{s}_i(t + dt)\} | \{\vec{s}_i(t)\}]$, which can then be maximized over the parameters of the model (see Methods).

We first tested our dynamical inference method on synthetic data simulated using the model of equation (1), with $\tau_{\text{relax}} \approx 0.7$, for various values of the interaction range n_c (see Methods). We infer the parameters of the model using either Euler's rule or the result of exact integration, for different values of the sampling time ranging from $dt = 0.2$ to $dt = 0.8$. The method based on exact integration predicts the interaction range n_c well, regardless of dt (Fig. 1a,b), while the method based on Euler's approximation largely overestimates n_c at large dt (Fig. 1b). We can now apply our dynamical inference to real flocks and learn the model parameters. First, we used data of natural flocks to check the effect of changing the sampling time dt , from the real sampling time of our set-up, $dt = 0.2$ s (see Methods), to 0.8 s. Although we cannot compare the inferred value of n_c to the ground truth as in simulations, we observe a similar trend as a function of dt (Fig. 1c), with the exact integration and Euler's approximation methods agreeing only at small dt . This suggests that the sampling time of 0.2 s is of the same order as the orientation relaxation time τ_{relax} , as we will confirm below. It also indicates that the inference method based on exact integration is extracting the parameters of alignment reliably.

Using the model parameters learnt from the data, we evaluate the two timescales of interest for activity, namely relaxation of the orientations and network rearrangement. We estimated the network rearrangement time τ_{network} experimentally for each flocking event as the characteristic decay time of its autocorrelation function $C_{\text{network}}(t) = \sum_{ij} n_{ij}(t_0) n_{ij}(t_0 + t)$, by fitting $C_{\text{network}}(t) \approx C_0 \exp(-t/\tau_{\text{network}})$ (Supplementary Fig. 1).

Working out the timescale of relaxation is more subtle. The relevant quantity is the product of the interaction strength J , which has units of inverse time, by the dimensionless connectivity matrix, $\mathbf{\Lambda}$, as can be seen from equation (4). Since there are n_c neighbours acting on each individual, the total alignment force is of order $J n_c$, suggesting that the characteristic timescale of relaxation of the orientations is $\tau_{\text{relax}} \sim (J n_c)^{-1}$. This result, however, seems at odds with the well-known fact that systems with spontaneously broken continuous symmetry—such as flocks—have correlation length and relaxation time that diverge with the system size L (Goldstone theorem²³). On the other hand, we do not expect the large-scale modes responsible for this divergence to affect the local relaxation dynamics and its interplay with network reshuffling. To clarify this issue we calculate the dynamical autocorrelation function of the fluctuations of the order parameter, $C_{\text{relax}}(t) = \langle \delta \vec{s}_i(t_0) \cdot \delta \vec{s}_i(t_0 + t) \rangle$, where $\delta \vec{s}_i = \vec{s}_i - \langle \vec{s}_i \rangle$. We consider a fixed lattice, because we need to gauge relaxation in the absence of network rearrangements, resulting in the autocorrelation function (see Supplementary Information):

$$C_{\text{relax}}(t) = \int_{1/L}^{1/a} d^d k \frac{e^{-J a^2 n_c k^2 t}}{J a^2 n_c k^2} \quad (6)$$

where a is the lattice spacing. The infrared divergence at small k , which correspond to large-scale modes, makes the integral divergent in the $L \rightarrow \infty$ limit for $d = 2$ (Mermin–Wagner theorem²⁴). For $d = 3$ the integral is finite, but the correlation function is a power law, so that the relaxation time diverges with L . The small k modes in (6) correspond to long-wavelength fluctuations spanning the entire flock, causing the local order parameter to relax slowly. However, these long-wavelength fluctuations do not contribute to the disordering of the local interaction network: if the wavelength of a fluctuation is much larger than the interaction range, all directions of motion in the interaction neighbourhood fluctuate in unison, causing no change in the mutual positions of the birds. We conclude that the autocorrelation function that impacts on local network rearrangements includes only contributions from wavelengths up to the local interaction range (let us call it r_c). This amounts to restricting the integral in (6) to the modes $r_c^{-1} \leq k \leq a^{-1}$, thus

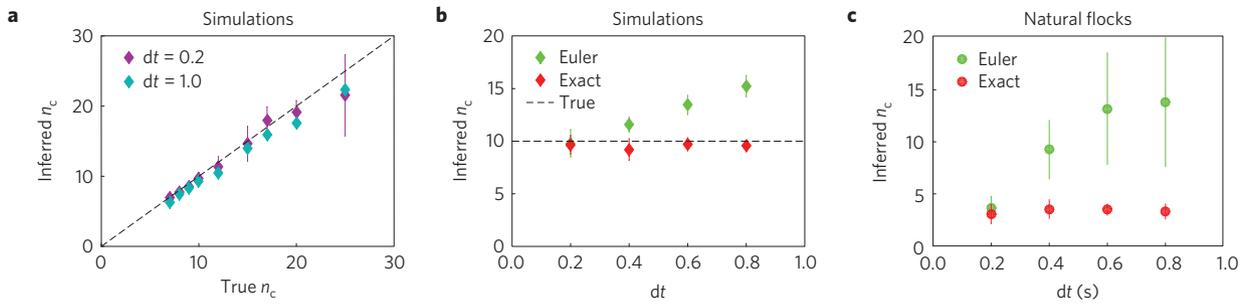


Figure 1 | Performance of the inference methods on the predicted interaction range n_c . **a**, Inferred versus real n_c obtained by applying our new inference method to simulated data generated with equation (1) at various interaction ranges. The method performs well for different values of the sampling rate dt . The dashed line is identity. **b**, Dependence of the inferred n_c on the sampling time dt . On simulated data with $n_c = 10$ (dashed line), the inference method based on exact integration (red points) performs well regardless of the sampling time dt . By contrast, the inference method based on Euler’s integration method (green points) overestimates the true interaction range at large dt . **c**, A similar trend is observed when we apply the two inference procedures to real flocking data, as illustrated here on one flocking event. Note that in this case the true value is not known. Error bars represent standard errors over time frames.

eliminating the infrared divergent modes $k \sim 1/L$. The resulting correlation function is exponentially decaying (see Supplementary Information for the calculation of the integral), with finite relaxation time equal to $\tau_{\text{relax}} = (Jn_c)^{-1}$, consistent with our initial guess. We note that, by considering wavelengths up to the interaction range, we are still dealing with a coarse-grained field theory, as in most biological systems the scale of interaction extends over tens of neighbours.

We can now proceed with the comparison of τ_{network} and τ_{relax} . Results are summarized in Fig. 2. The two timescales clearly separate, with local relaxation almost two orders of magnitude faster than network reshuffling. This separation of timescales suggests that flocks are in a state of local equilibrium. The network of interactions changes slowly enough for the dynamics of flight orientations to catch up before neighbours reshuffle. In other words, the orientation dynamics tracks network changes adiabatically. Note that this statement holds only locally, at the scale of the interaction range, as both τ_{network} and τ_{relax} are defined on that scale.

Since flocks behave as if they were in local equilibrium, an equilibrium inference procedure, which takes as input the local spatial correlation computed from a snapshot of the birds’ flight orientation¹³, should be consistent with the results of the dynamical inference. To check this prediction, we recall the equilibrium-like inference method of ref. 13. For symmetric J_{ij} , equation (1) is the Langevin equation derived from the Hamiltonian of the Heisenberg model

$$\mathcal{H} = -\frac{1}{2} \sum_{ij} J_{ij} \vec{s}_i \cdot \vec{s}_j \quad (7)$$

When J_{ij} varies slowly in time, the fluctuations of \vec{s}_i are in quasi-equilibrium and distributed according to Boltzmann’s law:

$$P(\vec{s}_1, \dots, \vec{s}_N) \sim \exp(-\mathcal{H}/T) \quad (8)$$

We recognize the maximum-entropy distribution consistent with the local correlation index $\sum_{ij} n_{ij} \langle \vec{s}_i \cdot \vec{s}_j \rangle$ fitted in ref. 13. In practice, the equilibrium inference consists in maximizing the likelihood of equation (8) over its parameters n_c and J/T (see Methods and Supplementary Information). If the variations of n_{ij} are slow compared to the dynamics of \vec{s}_i , $\tau_{\text{network}} \gg \tau_{\text{relax}}$, this inference procedure should give an accurate estimate of the alignment parameters. If, however, the two timescales are comparable, we expect the equilibrium inference to overestimate the true n_c , as the frequent exchange of neighbours results in an effective number of interaction partners that is larger than the instantaneous one. We verified both these expectations on simulated data, by showing

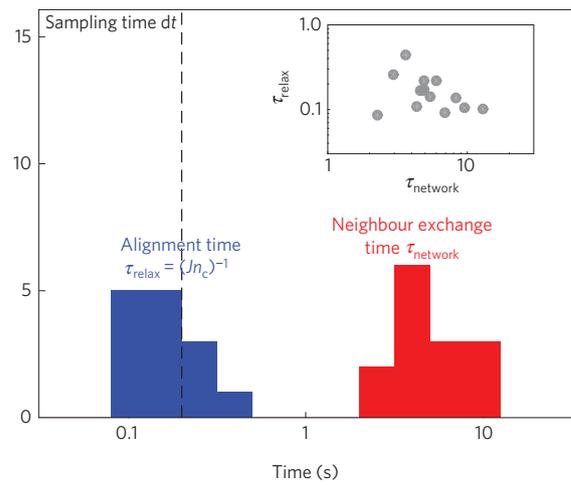


Figure 2 | Comparison between the two relevant timescales of active matter, as inferred in 14 natural flocks using our inference method based on exact integration. Histograms of the neighbour exchange time τ_{network} versus the local alignment time $\tau_{\text{relax}} = 1/Jn_c$, show that the relaxation of orientations is much faster than the turnover of neighbours. Note that the experimental sampling time $dt = 0.2$ s (dashed line) is of the same order as the alignment time, justifying the use of exact integration. Inset: the scatter plot of τ_{relax} versus τ_{network} shows no correlation between the two quantities.

that the equilibrium inference is accurate for $\tau_{\text{network}} \sim 100\tau_{\text{relax}}$, but overestimates n_c for $\tau_{\text{network}} \sim \tau_{\text{relax}}$ (see Supplementary Fig. 2). When applied to empirical data, the dynamical and equilibrium inferences give consistent results, and predict the same interaction range, n_c , and coupling-to-noise ratio, J/T (Fig. 3). Note that, while the dynamical inference provides the strength of the interaction, J , and the strength of the noise, T , separately, the equilibrium inference gives only the ratio J/T , which is the quantity to compare. To better appreciate this result, recall that the two inference procedures are based on independent pieces of information: the equilibrium inference uses instantaneous orientations, while the dynamical inference exploits how these orientations change in time. Their agreement confirms that the alignment dynamics of flocks are in an effective state of equilibrium over the range n_c .

Theoretical studies of active matter indicate that out-of-equilibrium effects induced by the rearrangement of the interaction network play a major role in the ordering of the system^{4,5}. In this light, any attempt to understand the properties of active biological systems based on equilibrium approaches may seem inappropriate.

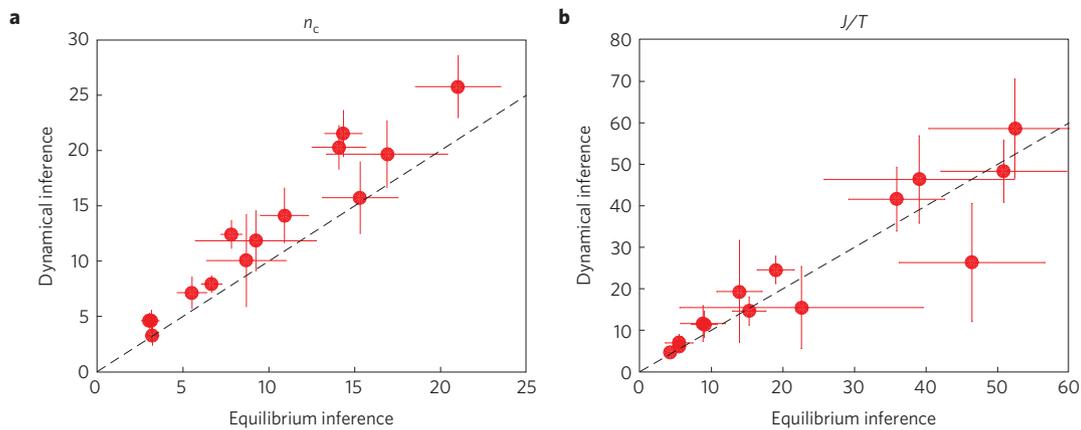


Figure 3 | Inference on natural flocks. a, b. For each of the 14 flocking events, the parameters of the model were inferred using either the dynamical inference method presented here, with $dt = 0.2$ s, or an equilibrium inference method as in ref. 13. Both methods agree well on the predicted value of the alignment range n_c (a). While the dynamical method infers the alignment strength J and the noise amplitude T separately, the equilibrium method infers only their ratio J/T , the value of which is consistent between the two methods (b). Error bars represent standard errors over time frames.

Does it mean that we should always relinquish the methods of equilibrium statistical mechanics when dealing with active systems? Our results address this question by showing that bird flocks are in a state of local equilibrium, due to the rapid relaxation of orientations compared to the slow rearrangement of the network, over the local scale of interaction. As a consequence, an equilibrium inference method, which assumes a fixed interaction network, gives equivalent results to a full dynamical treatment.

Equilibrium inference seems to be justified in this system, not only as a formal mathematical equivalence allowing for useful insights and predictions, but as a tool to extract *bona fide* biological parameters. The equilibrium approach is mathematically simpler and computationally less expensive than the dynamical one in the limit of strong polarization, making it easier to analyse larger groups. Although a dynamical approach such as the one presented here is still necessary for extracting the precise relaxation timescale of the ordering mechanism, there may be more straightforward ways to evaluate its order of magnitude and get a quick assessment of the local equilibrium hypothesis.

Our results do not mean that natural flocks are in global equilibrium and that network rearrangements play no role. The interaction network, far from being fixed as if individuals were linked by springs²⁵, completely reshuffles on long timescales²⁶. The directions of motion relax on a faster timescale than the network over the local scale of interaction, but the network does move on longer timescales, and over larger length scales, with important consequences. To appreciate this point we must stress again the difference between local, short-wavelength modes, which set the balance between relaxation and network rearrangement, and long-wavelength modes, which govern the long-time and long-distance correlations. Capturing these large-scale properties requires one to describe the active fluid using a hydrodynamic approach⁴. Equilibrium inference works despite the existence of these large-scale modes because it uses only information at the local scale of interaction, where relaxation is fast.

The local equilibrium we have uncovered in natural flocks is not merely the consequence of the high degree of polarization of this system. A high polarization certainly implies slow network rearrangements, but it does not constrain the relaxation time, which could be even slower, as illustrated in our simulations (Supplementary Fig. 2). Conversely, there may be unpolarized systems where local relaxation is faster than network rearrangement—a limit easily obtained theoretically by considering weakly interacting, slowly moving individuals. Midge swarms may be such an example: they are not polarized, poised below the

ordering transition²⁷, yet have been successfully analysed using standard equilibrium tools of critical phenomena²⁸. In general, one must carefully quantify these two timescales to determine to what degree the tools of equilibrium statistical mechanics may be applied to a given active system.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available in the [online version of this paper](#).

Received 12 February 2016; accepted 4 July 2016;
published online 15 August 2016

References

1. Camazine, S. *et al.* *Self-Organization in Biological Systems* (Princeton Univ. Press, 2001).
2. Krause, J. & Ruxton, G. D. *Living in Groups* (Oxford Univ. Press, 2002).
3. Sumpter, D. J. *Collective Animal Behavior* (Princeton Univ. Press, 2010).
4. Toner, J. & Tu, Y. Flocks, herds, and schools: a quantitative theory of flocking. *Phys. Rev. E* **58**, 4828–4858 (1998).
5. Ramaswamy, S. The mechanics and statistics of active matter. *Annu. Rev. Condens. Matter Phys.* **1**, 323–345 (2010).
6. Vicsek, T. & Zafeiris, A. Collective motion. *Phys. Rep.* **517**, 71–140 (2012).
7. Marchetti, M. *et al.* Hydrodynamics of soft active matter. *Rev. Mod. Phys.* **85**, 1143–1189 (2013).
8. Toner, J. & Tu, Y. Long-range order in a two-dimensional dynamical XY model: how birds fly together. *Phys. Rev. Lett.* **75**, 4326–4329 (1995).
9. Ballerini, M. *et al.* Empirical investigation of starling flocks: a benchmark study in collective animal behaviour. *Anim. Behav.* **76**, 201–215 (2008).
10. Cavagna, A. *et al.* The STARFLAG handbook on collective animal behaviour: 1. Empirical methods. *Anim. Behav.* **76**, 217–236 (2008).
11. Cavagna, A., Giardina, I., Orlandi, A., Parisi, G. & Procaccini, A. The STARFLAG handbook on collective animal behaviour: 2. Three-dimensional analysis. *Anim. Behav.* **76**, 237–248 (2008).
12. Attanasi, A. *et al.* Greta—a novel global and recursive tracking algorithm in three dimensions. *IEEE Trans. Pattern Anal. Mach. Intell.* **37**, 2451–2463 (2015).
13. Bialek, W. *et al.* Statistical mechanics for natural flocks of birds. *Proc. Natl Acad. Sci. USA* **109**, 4786–4791 (2012).
14. Lukeman, R., Li, Y.-X. & Edelstein-Keshet, L. Inferring individual rules from collective behavior. *Proc. Natl Acad. Sci. USA* **207**, 12576–12580 (2010).
15. Katz, Y., Tunstrom, K., Ioannou, C. C., Huepe, C. & Couzin, I. D. Inferring the structure and dynamics of interactions in schooling fish. *Proc. Natl Acad. Sci. USA* **108**, 18720–18725 (2011).
16. Herbert-Read, J. E. *et al.* Inferring the rules of interaction of shoaling fish. *Proc. Natl Acad. Sci. USA* **108**, 18726–18731 (2011).
17. Gautrais, J. *et al.* Deciphering interactions in moving animal groups. *PLoS Comput. Biol.* **8**, e1002678 (2012).

18. Strandburg-Peshkin, A. *et al.* Visual sensory networks and effective information transfer in animal groups. *Curr. Biol.* **23**, R709–R711 (2013).
19. Rosenthal, S. B., Twomey, C. R., Hartnett, A. T., Wu, H. S. & Couzin, I. D. Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion. *Proc. Natl Acad. Sci. USA* **112**, 4690–4695 (2015).
20. Cavagna, A. *et al.* Dynamical maximum entropy approach to flocking. *Phys. Rev. E* **89**, 042707 (2014).
21. Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I. & Shochet, O. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.* **75**, 1226–1229 (1995).
22. Cavagna, A. *et al.* Short-range interactions versus long-range correlations in bird flocks. *Phys. Rev. E* **92**, 012705 (2015).
23. Parisi, G. *Statistical Field Theory* (Frontiers in Physics, Addison-Wesley, 1988).
24. Mermin, N. D. & Wagner, H. Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic heisenberg models. *Phys. Rev. Lett.* **17**, 1133–1136 (1966).
25. Ferrante, E., Turgut, A. E., Dorigo, M. & Huepe, C. Elasticity-based mechanism for the collective motion of self-propelled particles with springlike interactions: a model system for natural and artificial swarms. *Phys. Rev. Lett.* **111**, 268302 (2013).
26. Cavagna, A., Duarte Queirós, S. M., Giardina, I., Stefanini, F. & Viale, M. Diffusion of individual birds in starling flocks. *Proc. Biol. Sci.* **280**, 20122484 (2013).
27. Attanasi, A. *et al.* Finite-size scaling as a way to probe near-criticality in natural swarms. *Phys. Rev. Lett.* **113**, 238102 (2014).
28. Attanasi, A. *et al.* Collective behaviour without collective order in wild swarms of midges. *PLoS Comput. Biol.* **10**, e1003697 (2014).

Acknowledgements

Work in Paris was supported European Research Council Starting Grant 306312. Work in Rome was supported by IIT-Seed Artswarm, European Research Council Starting Grant 257126, and US Air Force Office of Scientific Research Grant FA95501010250 (through the University of Maryland). F.G. acknowledges support from EU Marie Curie ITN grant n. 64256 (COSMOS) and Marie Curie CIG PCIG13-GA-2013-618399.

Author contributions

A.C., I.G., T.M. and A.M.W. designed the study. A.C., L.D.C., I.G., S.M., L.P. and M.V. acquired and processed the data. A.C., I.G., F.G., T.M. and A.M.W. developed the inference method. A.C., I.G., T.M. and A.M.W. wrote the paper.

Additional information

Supplementary information is available in the [online version of the paper](#). Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.M.

Competing financial interests

The authors declare no competing financial interests.

Methods

Flocking data. The three-dimensional (3D) trajectories of all birds were reconstructed using imaging techniques. Stereoscopic experiments on natural flocks of European starlings were performed in the field in Rome using three high-speed machine vision cameras shooting at 170 fps. The stereoscopic video acquisitions were then processed using a novel purpose-built 3D tracking algorithm based on a recursive global optimization method¹². This algorithm is extremely powerful, allowing for the reconstruction of full 3D trajectories of all individuals in groups of several hundreds of individuals. We collected 3D data from 12 flocking events with sizes ranging from 50 to 600 individuals, and lasting from 2 s to 6 s (for details on the experiments and the data set see Supplementary Table 1 and refs 10,29). To avoid interference from birds flapping, which occurs at frequency ≈ 10 Hz, we subsampled all the 3D sequences so that two snapshots are separated by $dt' = 0.1$ s. The instantaneous flight orientations were estimated by $\vec{s}_i(t) = [\vec{r}_i(t + dt') - \vec{r}_i(t)] / \|\vec{r}_i(t + dt') - \vec{r}_i(t)\|$. To avoid overlap between two subsequent evaluations of $\vec{s}_i(t)$, we used $dt = 2dt' = 0.2$ s. The lower sampling rates of Fig. 1c were obtained by taking $dt' = 0.2, 0.3$ and 0.4 s.

Simulated data. Data were simulated in three dimensions with the continuous Vicsek model of equation (1) with the interaction matrix of equation (2). The positions \vec{r}_i of individuals are updated according to $d\vec{r}_i/dt = v_0 \vec{s}_i$, with $v_0 = 1$. The simulations were set in a $8 \times 8 \times 8$ box with periodic boundary conditions, and $N = 512$ birds, so that density is exactly 1. We set $\sqrt{2T} = 0.15$ to obtain a polarization $P \approx 0.99$ similar to natural flocks. Equation (1) was integrated using Euler's method with a simulation step $dt_{\text{sim}} = 0.01$ that is much smaller than any other timescale in the system. The interaction range n_c varied from 7 to 25, and the interaction strength was picked so that $Jn_c = 1.5$, hence $\tau_{\text{relax}} = (Jn_c)^{-1} \sim 0.7$. The flocks were first brought to a steady state before taking snapshots for analysis.

Spin-wave approximation. The polarization P quantifies the level of order in the system. When $P \approx 1$, we can expand each \vec{s}_i around the common direction of flight $\vec{n} \equiv (1/NP) \sum_i \vec{s}_i$. This expansion gives $\vec{s}_i = \vec{\pi}_i + \sqrt{1 - \vec{\pi}_i^2} \vec{n} \approx \vec{\pi}_i + (1 - \vec{\pi}_i^2/2) \vec{n}$, with $\vec{n} \cdot \vec{\pi}_i = 0$. At leading order in $\vec{\pi}_i \ll 1$, equation (4) becomes

$$\frac{d\vec{\pi}_i}{dt} = -J \sum_j \Lambda_{ij} \vec{\pi}_j + \vec{\xi}_{i,\perp} \quad (9)$$

with $\langle \vec{\xi}_{i,\perp}(t) \cdot \vec{\xi}_{j,\perp}(t') \rangle = 4T \delta_{ij} \delta(t - t')$. Similarly, the equilibrium distribution (equation (8)) can be expanded into

$$P(\vec{\pi}) = \frac{1}{Z} e^{-\langle J/T \rangle \sum_{ij} \Lambda_{ij} \vec{\pi}_i \cdot \vec{\pi}_j} \quad (10)$$

Since this distribution is Gaussian, Z can be calculated analytically and reads: $Z = (2\pi T/J)^{(N-1)} \prod_{\lambda_k > 0} \lambda_k^{-1}$, where λ_k are the eigenvalues of the matrix Λ_{ij} .

Maximum likelihood inference. The equilibrium inference is performed by maximizing the likelihood of the data given by equation (10) over the parameters n_c and (J/T) (see Supplementary Information for detailed formulae).

The dynamical inference based on Euler's rule is implemented by maximizing the likelihood $P(\{\vec{\pi}_i(t + dt)\} | \{\vec{\pi}_i(t)\})$ calculated from Euler's formula (equation (3)). This likelihood reads

$$(4\pi T dt)^{-N} e^{-\frac{1}{4T dt} \sum_i [\vec{\pi}_i(t+dt) - \vec{\pi}_i(t) dt \sum_j \Lambda_{ij} \vec{\pi}_j]^2} \quad (11)$$

The dynamical inference based on exact integration uses equation (5), rewritten as $\vec{\pi}(t + dt) = e^{-J \Lambda dt} \vec{\pi}(t) + \vec{\epsilon}$, where $\vec{\epsilon}$ is a zero-mean Gaussian vector of covariance $\langle \vec{\epsilon} \cdot \vec{\epsilon}^\top \rangle = 4T \int_0^{dt} du e^{-J \Lambda u} e^{-J \Lambda^\top u} = 2\mathbf{X}^{-1}$. The conditional likelihood $P(\{\vec{\pi}_i(t + dt)\} | \{\vec{\pi}_i(t)\})$ now reads

$$\frac{\det(\mathbf{X})}{(2\pi)^N} e^{-\frac{1}{2} [\vec{\pi}(t+dt) - e^{-J \Lambda dt} \vec{\pi}(t)]^\top \mathbf{X} [\vec{\pi}(t+dt) - e^{-J \Lambda dt} \vec{\pi}(t)]} \quad (12)$$

Depending on whether one uses Euler's or exact integration rules, equation (11) or (12) is maximized over J , T and n_c (see Supplementary Information for detailed formulae).

In all three inference procedures, the parameters are learnt for each time t . Then the median and the associated standard error are calculated for each flocking event.

Data availability. The data that support the plots within this paper and other findings of this study are available from the corresponding author upon request.

References

29. Attanasi, A. *et al.* Information transfer and behavioural inertia in starling flocks. *Nature Phys.* **10**, 691–696 (2014).