

## Phase Coexistence in Insect Swarms

Michael Sinhuber and Nicholas T. Ouellette\*

*Department of Civil and Environmental Engineering, Stanford University, Stanford, California 94305, USA*  
(Received 4 April 2017; revised manuscript received 29 May 2017; published 24 October 2017)

Animal aggregations are visually striking, and as such are popular examples of collective behavior in the natural world. Quantitatively demonstrating the collective nature of such groups, however, remains surprisingly difficult. Inspired by thermodynamics, we applied topological data analysis to laboratory insect swarms and found evidence for emergent, material-like states. We show that the swarms consist of a core “condensed” phase surrounded by a dilute “vapor” phase. These two phases coexist in equilibrium, and maintain their distinct macroscopic properties even though individual insects pass freely between them. We further define a pressure and chemical potential to describe these phases, extending theories of active matter to aggregations of macroscopic animals and laying the groundwork for a thermodynamic description of collective animal groups.

DOI: [10.1103/PhysRevLett.119.178003](https://doi.org/10.1103/PhysRevLett.119.178003)

Collective behavior of social animals is ubiquitous in the natural world [1]. Birds [2,3], fish [4,5], insects [6,7], and many other species [8] routinely and spontaneously form aggregations that appear to possess an identity distinct from that of the underlying individuals, so that they are sometimes termed “superorganisms” [9]. However, descriptions of this kind are largely qualitative, and empirical characterizations of true emergent properties of animal groups are sparse [10–12]. Indeed, calling such groups “collective” at all makes tacit assumptions about their nature, and suggests that the group as a whole is somehow “more” than, or at least different from, the sum of its individuals. But, to paraphrase Williams [13], one must take care to distinguish between a “fleet herd of deer” and a “herd of fleet deer,” or, put another way, between “emergent” properties of the group that have no meaning at the individual level and “aggregate” properties that are simple statistical averages over the individuals [14]. In some special cases, it is straightforward to argue for collectivity; the social insects, for example, cannot survive and reproduce outside a colony [15]. But in general, demonstrating true collective, emergent behavior is surprisingly difficult.

It is often assumed that animal groups outperform individuals in accomplishing tasks due to the group dynamics [16]. Thus, one way to show collectivity would be to demonstrate this superior performance. Proving such enhancement, however, is subtle and difficult [17,18], and requires accurate knowledge of the task that is being optimized [19]. An alternative route would be to follow the prescriptions of condensed matter physics and directly define emergent properties that describe the group as a whole and are independent of the precise participants [20]. This is the approach taken, for example, in thermodynamics, where state variables such as pressure, temperature, and chemical potential, or response functions like viscosity or elastic moduli can be defined and related for bulk materials

without direct appeal to a molecular description. Although such ideas have recently begun to be applied to collective behavior in animals [10–12], there is not yet a unified “thermodynamic” theory.

In this Letter, we develop such a description for laboratory mating swarms of the nonbiting midge *Chironomus riparius*. From three-camera video measurements of swarms of various sizes, we extracted the three-dimensional time-resolved trajectories of each midge in the swarm [6,21], as well as their velocities and accelerations. Our methodology and the details of this data set have been described previously [6,10,22]. Using persistent homology, a topological data analysis method, we study the structure of the swarms, and find that they can be segmented into two distinct regions that we argue can be thought of as a condensed phase surrounded by a vapor phase. Although these phases have distinct macroscopic properties, midges move freely between them, suggesting that they are true collective, emergent states. Our results suggest powerful new ways of thinking about collectivity in animal groups, and help to bring macroscopic animal groups into the purview of theories of active matter.

Since we seek global features of the swarms rather than a detailed microscopic picture, we turned to topological data analysis, since topology naturally captures gross structure. Specifically, we used persistent homology [23–25]. This method treats a collection of points, such as the midge locations, as a discrete sampling of an underlying object. To quantify the topological structure of this object, we create a simplicial complex [24] from the discrete data points by associating each one with a sphere of radius  $\epsilon/2$ , where  $\epsilon$  is known as the proximity parameter.  $\epsilon$  is a free parameter, but one seeks topological features that persist over a range of  $\epsilon$ , as such features are likely to be meaningful. Simplicial complexes can be quantified in terms of their Betti numbers  $b_i$ , where  $b_0$  gives the number of connected components in

the complex,  $b_1$  the number of holes,  $b_2$  the number of topological circles, and so on. Here, we focus only on  $b_0$ , the number of connected components. We note that in practice there are many slightly different algorithmic ways to construct simplicial complexes. Here, we used the MATLAB implementation of JAVAPLEX [26] to construct Vietoris-Rips complexes from our swarm data and analyze their Betti numbers.

When  $\epsilon$  is very small, all the insects will appear to be isolated; thus, in that limit  $b_0 = N$ , where  $N$  is the number of individuals in the swarm. Likewise, when  $\epsilon$  is very large, all the insects will appear to be part of the same connected component, and so  $b_0 = 1$ . In between,  $b_0$  varies smoothly with  $\epsilon$  with a shape that is independent of  $N$  [Fig. 1(a)]. As  $b_0$  decreases, however, the sizes of the connected components do not uniformly increase. Instead, we typically observe a single large cluster surrounded by small unconnected components [Fig. 1(b)]. Unsurprisingly, this large cluster lies in the middle of the swarm [Fig. 1(c)], where the number density of midges tends to be somewhat higher. These results are insensitive to the particular value of  $\epsilon$  chosen. And while  $b_0$  fluctuates in time, we find that it is statistically stationary and so focus on time-averaged quantities here.

Persistent homology thus identifies a topological structure in the swarm, suggesting that the core of the swarm and its outer reaches are different. It cannot, however, tell us how to interpret this structure. For that, we turn to a statistical analysis of the physical variables. One might expect, for example, to see a kinematic difference between insects in the central cluster and the outer region. However, both velocity and acceleration statistics are indistinguishable in the two regions [Figs. 2(a), 2(b)]. But when we consider the number density and volume, the picture changes. The number density  $n$  is significantly larger in the central cluster than in the outer region [Fig. 2(c)], and, moreover, conditional statistics show that the number density in the central cluster is independent of the volume of the outer region, and vice versa [Fig. 2(c)]. In addition, the volume  $V$  occupied by the central cluster depends on the number of midges in it, while the volume of the outer region does not [Fig. 2(d)]. Similar results for other swarms are shown in the Supplemental Material [27]. Taken together, these results suggest that the central cluster and the outer region are distinct entities, since their properties are statistically independent, and that their properties are independent of the constituent insects, since the kinematic statistics are the same. Since, in addition, the volume of the central cluster depends on the number of constituent midges while the volume of the outer region does not but simply adjusts to fill the available space, we borrow terminology from thermodynamics and call the central cluster a “condensed phase” and the outer region a “vapor phase.” We note here, and discuss further below, that we

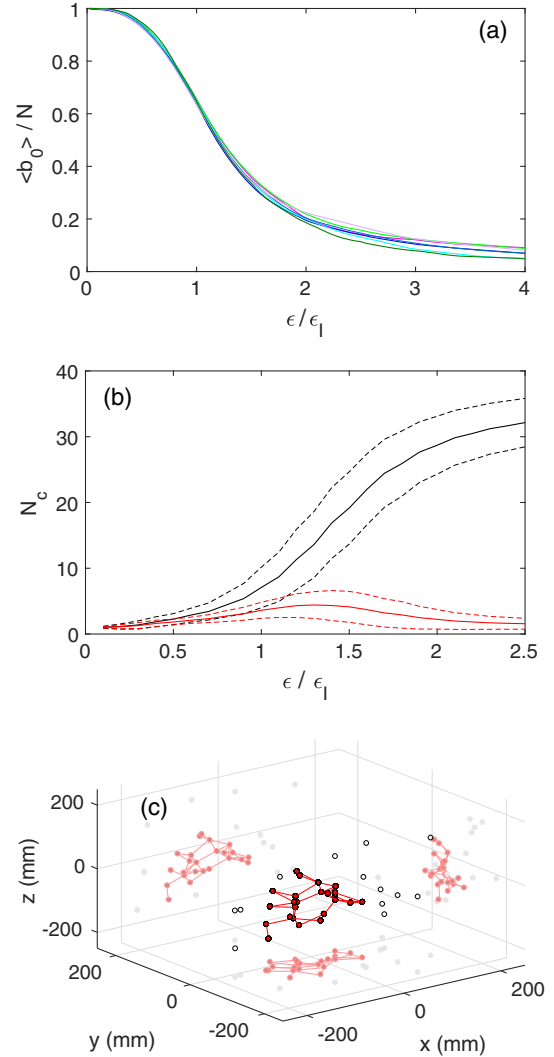


FIG. 1. Topological analysis of swarm structure. (a) Time-averaged zeroth Betti number  $\langle b_0 \rangle$  normalized by the number of insects in the swarm  $N$  for swarms with  $N = 18, 22, 26, 30, 34, 38,$  and  $42$ , as a function of the proximity parameter  $\epsilon$ . The standard error of the mean of the individual curves is less than 2%. Data from different swarms collapse when  $\epsilon$  is scaled by its value  $\epsilon_1$  at the curve’s inflection point, indicating that the topological structure of the swarms is preserved for different swarm sizes. (b) The number of individuals per connected component  $N_c$  for the largest (black) and second-largest (red) components as a function of  $\epsilon / \epsilon_1$ . As  $\epsilon / \epsilon_1$  increases, the largest component grows at the expense of the others. Solid lines indicate the mean (over time) of the component size, and dotted lines the standard deviation. (c) A snapshot from a single swarm with  $N = 38$  individuals showing the largest connected component (in red) computed for  $\epsilon = 1.5\epsilon_1$  (corresponding to about 3 to 4 times the typical nearest-neighbor distance), which we also use for all subsequent calculations. This connected component lies in the center of the swarm. For  $\epsilon = 1.5\epsilon_1$ , about half the midges typically lie in the largest cluster; our subsequent results, however, do not qualitatively depend on this choice.

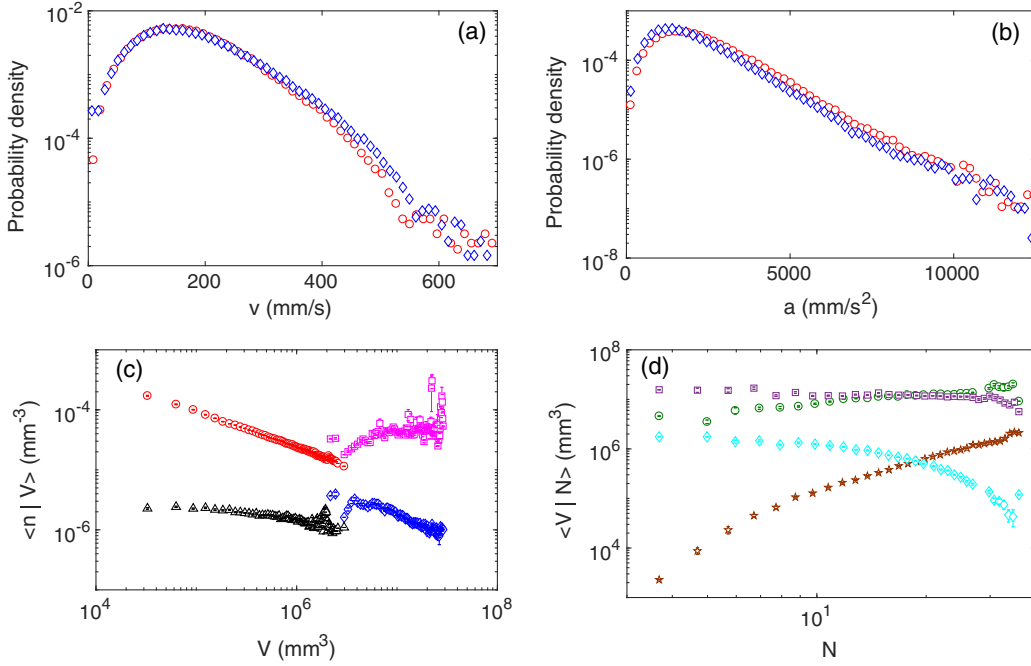


FIG. 2. Statistical characterization of the central cluster and outer region. (a) Probability density functions (PDFs) of midge speed  $v$  for the inner region (red circles) and outer region (blue diamonds) for the same swarm shown in Fig. 1(c). (b) PDFs of the midge accelerations for the inner and outer region. These kinematic statistics are nearly indistinguishable for the two regions. (c) Conditional averages of the number density  $n$  on the volume  $V$  as a function of  $V$  for four combinations of variables:  $\langle n_{\text{in}} | V_{\text{in}} \rangle$  (red circles),  $\langle n_{\text{out}} | V_{\text{out}} \rangle$  (blue diamonds),  $\langle n_{\text{in}} | V_{\text{out}} \rangle$  (magenta squares), and  $\langle n_{\text{out}} | V_{\text{in}} \rangle$  (black triangles). Note that using conditional medians rather than conditional averages leads to nearly identical results. We define the volume  $V_{\text{in}}$  as the volume of the convex hull of the midges in the central cluster; the volume  $V_{\text{out}}$  is then the difference between the volume of the convex hull of the entire swarm and  $V_{\text{in}}$ .  $n$  is much larger in the inner region than the outer; and while  $n$  for each region depends on the volume of that region, it is nearly independent of the volume of the other region. (d) Conditional averages of volume  $V$  on the number of midges  $N$  in each region for the corresponding combinations  $\langle V_{\text{in}} | N_{\text{in}} \rangle$  (brown stars),  $\langle V_{\text{out}} | N_{\text{out}} \rangle$  (green circles),  $\langle V_{\text{in}} | N_{\text{out}} \rangle$  (cyan diamonds), and  $\langle V_{\text{out}} | N_{\text{in}} \rangle$  (purple squares).  $V_{\text{in}}$  depends strongly on the number of insects in the swarm, while any dependence of  $V_{\text{out}}$  on the number is much weaker. Similar results for other swarms are shown in the Supplemental Material [27].

observe a continual exchange of individual insects between these two phases, just as one would expect at a liquid-vapor interface; the vapor phase, however, does not fill the entire midge enclosure, but instead occupies a smaller, and apparently self-regulated, region of space [6,22].

Thermodynamic phases are determined by more than just their density or volume; other state variables are needed to specify them completely. Defining such state variables cannot in general be done from first principles for active systems like insect swarms, due to the lack of conservation laws and knowledge of the microscopic interactions between individuals. We can, however, construct analogous quantities that play the same role. Starting from the virial relation, which we have shown holds for swarms [28], we define a pressure (per unit mass) as

$$\Pi = \frac{1}{NV} \sum_{i=1}^N \left( \frac{1}{3} (\mathbf{v}_i^2 + \mathbf{a}_i \cdot \mathbf{r}_i) \right), \quad (1)$$

where  $V$  is the volume of the phase,  $N$  is the number of constituent individuals,  $\mathbf{v}_i$  is the velocity of insect  $i$ ,  $\mathbf{a}_i$  is its acceleration vector, and  $\mathbf{r}_i$  is its distance from the swarm center. The acceleration term is motivated by our earlier

observations that insects in the swarm behave as if they are trapped in a harmonic potential well [6,28], and captures the work done (assuming the same mass for each insect) by a midge as it accelerates in this potential. We note that this formulation distinguishes our pressure from the classical definition as the stress exerted by a material on its confining walls [29], since our swarms, like most aggregations of macroscopic animals, are unconfined; rather, it is conceptually akin to the recently proposed idea of “swim pressure” [30,31]. Just as with number density and volume, the condensed and vapor phases are clearly distinguishable by their pressure statistics [Fig. 3(a)], and, as one would expect, the vapor phase exists at a much lower pressure. This pressure depends on the midge number density  $n$ , but in different ways for the two phases. In the condensed phase, the pressure is well fit by  $\Pi = An + Bn^2$  for constant  $A$  and  $B$  (Fig. 3(b) and Supplemental Material [27]), consistent with a second-order virial expansion along an isotherm. In the vapor phase, we observe behavior consistent with power-law scaling, and least-squares fits of the form  $\Pi = Cn^\xi$  to vapor-phase data for many swarms suggest that  $\xi \approx 1/2$  (Fig. 3(b) and Supplemental Material [27]).

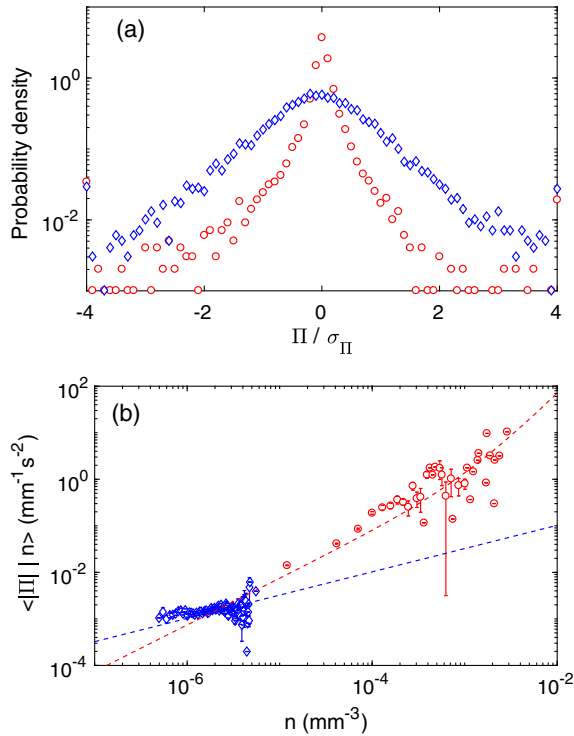


FIG. 3. Pressure statistics. (a) PDFs of the pressure scaled by its standard deviation  $\sigma_{\Pi}$  for the condensed (red circles) and vapor (blue diamonds) phases. The shape of the pressure PDFs is qualitatively different in the two phases. (b) Conditional averages of the absolute value of the pressure as a function of number density for each phase. The dashed lines are fits of  $\Pi = Cn^{1/2}$  (with constant  $C$ ) for the vapor phase and  $\Pi = An + Bn^2$  (with constant  $A$  and  $B$ ) for the condensed phase. Data for additional swarms are shown in the Supplementary Material [27].

Following further prescriptions of thermodynamics to define other state variables via conservation of energy is not possible, since energy is not conserved in active systems. But we can, at least quasistatically, appeal to conservation of mass, since the number of insects in the whole swarm changes very slowly compared with any dynamical time scale. We can thus define a chemical potential  $\mu$  via [31]

$$n \frac{\partial \mu}{\partial n} = (1 - \phi) \frac{\partial \Pi}{\partial n}, \quad (2)$$

where  $\phi$  is the volume fraction of midges. Our swarms are very dilute, with  $\phi \ll 1$ ; that condition combined with the measured dependence of  $\Pi$  on  $n$  allows us to integrate Eq. (2) for each phase and calculate  $\Delta\mu$ , the chemical potential difference between the two phases, though only up to a constant. In thermodynamics, two phases are considered to be in equilibrium when the chemical potential is uniform across the interface. In this case, since we only know the chemical potentials up to a constant, phase equilibrium would be indicated by a constant-in-time  $\Delta\mu$ . Consistently, we find that  $\Delta\mu$  fluctuates about a constant value with no temporal drift [Fig. 4(a)]. A more

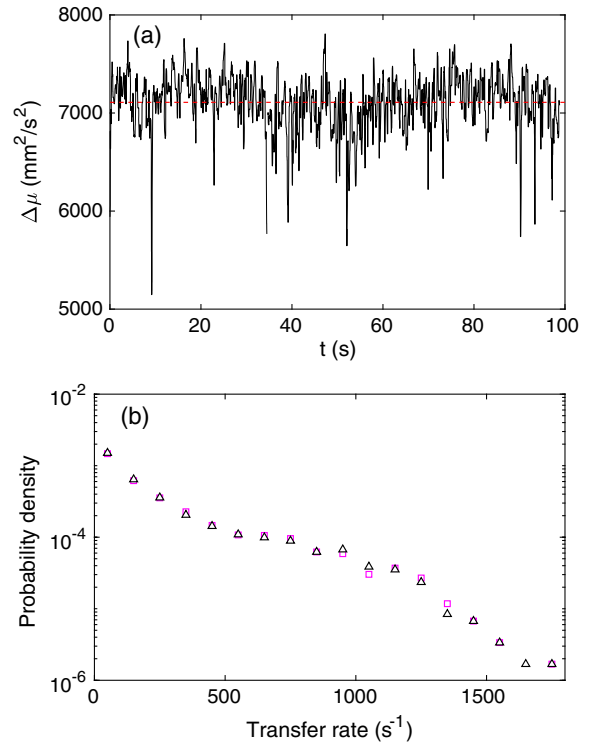


FIG. 4. Evidence for equilibrium phase coexistence. (a) Time series of the chemical potential difference  $\Delta\mu$  between the vapor and condensed phases. The red dashed line is the temporal mean. (b) PDFs of the instantaneous transfer rate (in number of midges per second) from vapor to condensed (black triangles) and condensed to vapor (magenta squares) phases computed over the full recording time for a single swarm. The two PDFs are indistinguishable, demonstrating detailed balance.

stringent condition for phase equilibrium at the microscopic level is detailed balance; in detailed balance, the likelihood of an individual midge moving from the condensed to the vapor phase would be the same as the likelihood of moving from the vapor to the condensed phase. When we measure the probability density functions of the transfer rates between phases, we find that detailed balance is indeed satisfied [Fig. 4(b)].

Our measurements provide strong evidence for two distinct “thermodynamic” phases in our swarms that coexist in equilibrium. These phases are true emergent phenomena, as they are not defined based on differences in the individual constituent midges and because individual midges pass freely across the phase boundary without changing the macroscopic properties of the phases. From a physical standpoint, these results provide a link for connecting recent theoretical work on active microparticles [29,31] to groups of macroscopic animals, and give us a framework for describing collective animal groups in nontrivial terms. Connecting collective behavior to thermodynamics also provides new evidence for the possibility of describing different kinds of group morphologies and dynamics as simply different phases of some underlying

unified framework, an idea that has been suggested previously [32] but has lacked a strong theoretical foundation. Key next steps toward this goal will be the proper definition of a temperaturelike variable for collective groups (which is often fraught in nonequilibrium systems [33]) and the construction of constitutive relations for different kinds of animals. Finally, our results may also have biological implications. The distinct and stable properties of the swarm core and outer region may provide a mechanism for the regulation of the swarm edge: an individual midge may be able to recognize that it has crossed the phase boundary, and that it is therefore time to turn around to remain in the swarm.

This work was supported by the U.S. Army Research Office under Grant No. W911NF-16-1-0185.

\*nto@stanford.edu

- [1] J. K. Parrish and L. Edelstein-Keshet, Complexity, pattern, and evolutionary trade-offs in animal aggregation, *Science* **284**, 99 (1999).
- [2] M. Nagy, Z. Ákos, D. Biro, and T. Vicsek, Hierarchical group dynamics in pigeon flocks, *Nature (London)* **464**, 890 (2010).
- [3] W. Bialek, A. Cavagna, I. Giardina, T. Mora, E. Silvestri, M. Viale, and A. M. Walczak, Statistical mechanics for natural flocks of birds, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4786 (2012).
- [4] Y. Katz, K. Tunstrøm, C. C. Ioannou, C. Huepe, and I. D. Couzin, Inferring the structure and dynamics of interactions in schooling fish, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18720 (2011).
- [5] J. E. Herbert-Read, A. Perna, R. P. Mann, T. M. Schaerf, D. J. T. Sumpter, and A. J. W. Ward, Inferring the rules of interaction of shoaling fish, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18726 (2011).
- [6] D. H. Kelley and N. T. Ouellette, Emergent dynamics of laboratory insect swarms, *Sci. Rep.* **3**, 1073 (2013).
- [7] A. Attanasi *et al.*, Collective behaviour without collective order in wild swarms of midges, *PLoS Comput. Biol.* **10**, e1003697 (2014).
- [8] I. D. Couzin and J. Krause, Self-organization and collective behavior in vertebrates, *Adv. Study Behav.* **32**, 1 (2003).
- [9] D. S. Wilson and E. Sober, Reviving the superorganism, *J. Theor. Biol.* **136**, 337 (1989).
- [10] R. Ni, J. G. Puckett, E. R. Dufresne, and N. T. Ouellette, Intrinsic Fluctuations and Driven Response of Insect Swarms, *Phys. Rev. Lett.* **115**, 118104 (2015).
- [11] M. Tennenbaum, Z. Liu, D. Hu, and A. Fernandez-Nieves, Mechanics of fire ant aggregations, *Nat. Mater.* **15**, 54 (2015).
- [12] R. Ni and N. T. Ouellette, On the tensile strength of insect swarms, *Phys. Biol.* **13**, 045002 (2016).
- [13] G. C. Williams, *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought* (Princeton University Press, Princeton, NJ, 1966).
- [14] S. Okasha, Emergence, hierarchy and top-down causation in evolutionary biology, *Interface Focus* **2**, 49 (2012).
- [15] S. N. Beshers and J. H. Fewell, Models of division of labor in social insects, *Annu. Rev. Entomol.* **46**, 413 (2001).
- [16] I. D. Couzin, Collective cognition in animal groups, *Trends Cognit. Sci.* **13**, 36 (2009).
- [17] A. Berdahl, C. J. Torney, C. C. Ioannou, J. J. Faria, and I. D. Couzin, Emergent sensing of complex environments by mobile animal groups, *Science* **339**, 574 (2013).
- [18] S. J. Portugal, T. Y. Hubel, J. Fritz, S. Heese, D. Trobe, B. Voelkl, S. Hailes, A. M. Wilson, and J. R. Usherwood, Upwash exploitation and downwash avoidance by flap phasing in ibis formation flight, *Nature (London)* **505**, 399 (2014).
- [19] J. R. Usherwood, M. Stavrou, J. C. Lowe, K. Roskilly, and A. M. Wilson, Flying in a flock comes at a cost in pigeons, *Nature (London)* **474**, 494 (2011).
- [20] J. P. Sethna, *Entropy, Order Parameters and Complexity* (Oxford University Press, Oxford, 2006).
- [21] N. T. Ouellette, H. Xu, and E. Bodenschatz, A quantitative study of three-dimensional Lagrangian particle tracking algorithms, *Exp. Fluids* **40**, 301 (2006).
- [22] J. G. Puckett and N. T. Ouellette, Determining asymptotically large population sizes in insect swarms, *J. R. Soc. Interface* **11**, 20140710 (2014).
- [23] R. Ghrist, Barcodes: The persistent topology of data, *Bull. Am. Math. Soc.* **45**, 61 (2008).
- [24] C. M. Topaz, L. Ziegelmeier, and T. Halverson, Topological data analysis of biological aggregation models, *PLoS One* **10**, e0126383 (2015).
- [25] B. T. Fasy and B. Wang, Exploring persistent local homology in topological data analysis, *Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (IEEE, Shanghai, 2016)*, pp. 6430–6434.
- [26] H. Adams, A. Tausz, and M. Vejdemo-Johansson, JAVAPLEX: A research software package for persistent (co)homology, in *Mathematical Software—ICMS2014*, edited by H. Hong and C. Yap (Springer, Berlin, 2014), pp. 129–136.
- [27] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.119.178003> for data for additional swarms.
- [28] D. Gorbonos, R. Ianculescu, J. G. Puckett, R. Ni, N. T. Ouellette, and N. S. Gov, Long-range acoustic interactions in insect swarms: an adaptive gravity model, *New J. Phys.* **18**, 073042 (2016).
- [29] A. P. Solon, Y. Fily, A. Baskaran, M. E. Cates, Y. Kafri, M. Kardar, and J. Tailleur, Pressure is not a state function for generic active fluids, *Nat. Phys.* **11**, 673 (2015).
- [30] S. C. Takatori, W. Yan, and J. F. Brady, Swim Pressure: Stress Generation in Active Matter, *Phys. Rev. Lett.* **113**, 028103 (2014).
- [31] S. C. Takatori and J. F. Brady, Towards a thermodynamics of active matter, *Phys. Rev. E* **91**, 032117 (2015).
- [32] I. D. Couzin, J. Krause, R. James, G. D. Ruxton, and N. R. Franks, Collective memory and spatial sorting in animal groups, *J. Theor. Biol.* **218**, 1 (2002).
- [33] L. F. Cugliandolo, The effective temperature, *J. Phys. A* **44**, 483001 (2011).