Research

Cite this article: Gorbonos D, Puckett JG, van der Vaart K, Sinhuber M, Ouellette NT, Gov NS. 2020 Pair formation in insect swarms driven by adaptive long-range interactions. *J. R. Soc. Interface* **17**: 20200367.
<http://dx.doi.org/10.1098/rsif.2020.0367>

Received: 14 May 2020

Accepted: 14 September 2020

Subject Category:

Life Sciences—Physics interface

Subject Areas:

computational biology, biophysics

Keywords:

swarming, long-range interactions,
insect swarm, adaptivity, pair formation

Author for correspondence:

Nir S. Gov

e-mail: nir.gov@weizmann.ac.il

[†]Present address: Max Planck Institute of Animal Behavior, Department of Collective Behavior, Universitätsstraße 10, 78464 Konstanz, Germany.

[‡]Present address: Department of Physics, Carl von Ossietzky Universität, 26111 Oldenburg, Germany.

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5143072>.

Pair formation in insect swarms driven by adaptive long-range interactions

Dan Gorbonos^{1,†}, James G. Puckett², Kasper van der Vaart³,
Michael Sinhuber^{3,‡}, Nicholas T. Ouellette³ and Nir S. Gov¹

¹Department of Chemical and Biological Physics, Weizmann Institute, Rehovot, Israel

²Department of Physics, Gettysburg College, Gettysburg, PA 17325, USA

³Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305, USA

JGP, 0000-0002-0886-268X; NSG, 0000-0001-7774-1139

In swarms of flying insects, the motions of individuals are largely uncoordinated with those of their neighbours, unlike the highly ordered motion of bird flocks. However, it has been observed that insects may transiently form pairs with synchronized relative motion while moving through the swarm. The origin of this phenomenon remains an open question. In particular, it is not known if pairing is a new behavioural process or whether it is a natural by-product of typical swarming behaviour. Here, using an ‘adaptive-gravity’ model that proposes that insects interact via long-range gravity-like acoustic attractions that are modulated by the total background sound (via ‘adaptivity’ or fold-change detection) and that reproduces measured features of real swarms, we show that pair formation can indeed occur without the introduction of additional behavioural rules. In the model, pairs form robustly whenever two insects happen to move together from the centre of the swarm (where the background sound is high) towards the swarm periphery (where the background sound is low). Due to adaptivity, the attraction between the pair increases as the background sound decreases, thereby forming a bound state since their relative kinetic energy is smaller than their pair-potential energy. When the pair moves into regions of high background sound, however, the process is reversed and the pair may break up. Our results suggest that pairing should appear generally in biological systems with long-range attraction and adaptive sensing, such as during chemotaxis-driven cellular swarming.

1. Introduction

Swarms are a form of collective animal behaviour that have caught the attention of physicists as self-organized non-equilibrium systems that remain cohesive yet exhibit no clear order parameter [1], as opposed to ‘flocking’ behaviour [2]. Such behaviour is observed in a variety of species, including fish [3,4], bats [5] and flying insects [6]. Theoretical models proposed to describe this collective phase often assume short-range (or near-neighbour) interactions [7,8] that contain a fine balance between attraction, repulsion, a tendency of the individuals to align their motion with that of their neighbours, and the effects of noise [9]. It has even been suggested that certain insect swarms may be finely tuned to be poised close to a critical point where global alignment of motion would commence [10].

In an alternative framework [11], we recently proposed that the interactions between flying insects (midges, in this case) are mediated by acoustics due to the sound they emit while flying, which gives rise to long-range power-law interactions. Furthermore, we suggested that the interactions are attractive, so that individuals tend on average to accelerate towards each other in proportion to the intensity of the sound received. For pure acoustics, the functional form of this acceleration is similar to gravity (i.e. proportional to r^{-2}), although similar behaviour can arise for other sensory modalities (such as vision [12]). In addition, swarms with similar characteristics arise even if the exponent of

the power-law interaction has a different value, as long as it remains long-range [13]. An additional, and crucial, component of this model is adaptivity, common to most sensory systems in biology [14], whereby the sensitivity of the midges to the received sound drops when there is a strong background sound. Exact adaptation means that the response of the organism to a change in the signal that it senses is normalized by the background, steady-state level of input, which is part of a fold-change detection mechanism [14]: when the background signal is strong, the sensitivity and response is diminished, while the response is enhanced when the background signal is low. It was previously shown that this ‘adaptive-gravity’ model reproduces many steady-state features of midge swarms [11], such as the observed reduction in the average accelerations of the midges towards the swarm centre in larger swarms [15]. More recently, the model was shown to account for the observed mass and velocity profiles within the swarms [16].

In addition to steady-state features, recent observations have found evidence for the dynamic formation of synchronized pairs of midges, which typically oscillate with respect to each other at a higher-than-normal frequency and maintain a small distance between them while they move together through the swarm (figure 1*a*) [17]. Pairs were identified in the laboratory swarms by analysing the relative distance between pairs of midges as they move through the swarm (figure 1*a,c*). The relative distance between any two unpaired midges exhibits oscillations with an amplitude on the order of the swarm radius and a period on the order of the typical time it takes a midge to fly across the swarm. Such behaviour, typical of unpaired midges, is seen in figure 1*c,e* outside of the time span indicated by the shading. However, occasionally the relative distance between the midges exhibited small-amplitude, high-frequency oscillations (when compared with the unpaired state). When these oscillations had a frequency that was higher than a threshold value and persisted longer than a threshold time (with both thresholds estimated using measured characteristics of the swarms), the two midges were identified as a pair [17]. Note that during pairing, the amplitude of the relative oscillations of the pair diminished (figure 1*c*), but remained much larger than the distance where midges might accelerate away from each other to avoid collision (a distance of about 12 mm, much smaller than the mean midge–midge distance of approximately 30–40 mm in our swarms) [18]. However, no mechanism for this phenomenon was proposed. In particular, it is not known if pairing is a result of additional, more complex behavioural rules or whether it can arise naturally as a passive by-product of swarming. Here, we report that the same model of adaptive long-range interactions (ALRI) [11] that captures many steady-state features of swarms indeed produces pairing without any modifications. Thus, pairing can be viewed as an emergent phenomenon and a natural outcome of the same interactions that lead to swarm formation.

Pairing is a rare event in systems with attractive long-range interactions, such as classical gravity, in the absence of adaptivity [19]. Such Hamiltonian systems conserve energy and momentum. Due to momentum conservation, the capture of two particles to form an orbiting pair must involve a third particle that will remove the excess momentum. Such situations are highly unlikely to occur, and indeed under classical gravity stellar pairs rarely form [20]. Adaptivity, however, means that the system does not obey energy or

momentum conservation [11], and the dynamics is not limited by these constraints. This has significant consequences for pairing, as we show below.

2. ALRI model

The basic equation of motion of the ALRI model gives the effective force on midge i due to the sum over all the other midges j as

$$\mathbf{F}_{\text{eff}}^i = C \sum_j \hat{\mathbf{r}}_{ij} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|^2 + \epsilon^2} \left(\frac{R_{\text{ad}}^{-2}}{R_{\text{ad}}^{-2} + \sum_k (|\mathbf{r}_i - \mathbf{r}_k|^2 + \epsilon^2)^{-1}} \right), \quad (2.1)$$

where $\mathbf{r}_i$ is the position vector for midge i , $\hat{\mathbf{r}}_{ij}$ is the unit vector pointing from midge i to midge j , C is a constant with dimensions of $\text{mass} \cdot \text{length}^3 / \text{time}^2$, ϵ is a constant with units of length, and R_{ad} is the length scale over which adaptivity occurs. For $r_{ij} \gg \sqrt{N} R_{\text{ad}}$, where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, and N is the number of midges in the swarm (that is, when the distance between a pair of midges far exceeds the range of adaptivity), the effective force reduces to a purely gravitational interaction. For comparison, we also considered an ϵ -gravity model, which is non-adaptive, classical gravity that is softened to prevent runaway accelerations that produce slingshots that break up the swarm too quickly [16]. In ϵ -gravity, the effective force on a midge is given by equation (2.1) when $R_{\text{ad}} \rightarrow 0$.

The adaptive term, in parentheses in equation (2.1), renormalizes the effective force that a midge feels towards the other midges: the directional (vector) sum of the sound intensities (the term outside the parentheses) is divided by the scalar sum of the sound intensities that appears in the denominator of the expression in the parentheses. This renormalization is effective as long as the distances between the midges are small compared to the adaptivity range R_{ad} . In this regime of ‘perfect adaptivity’, when $R_{\text{ad}}/r_{ij} \rightarrow \infty$, the forces are normalized by the background sound: they are diminished (enhanced) for strong (weak) background level. Outside of this regime, the forces approach regular (ϵ)-gravity.

Note that we consider here point particles (so that they do not collide), without any short-range repulsion. In our observations of real swarms, we found that midges in pairs do not approach each other close enough to feel the effective short-range repulsive interactions that prevent them from colliding [17]. Nevertheless, as shown in the electronic supplementary material, S9, we also tested the addition of a soft-core short-range (exponential) repulsive potential to the model [21], which acts to diminish the pair formation and pair stability (electronic supplementary material, figures S18 and S19).

We looked for evidence of pairing in both the ALRI and ϵ -gravity models by simulating the dynamics of particles of unit mass, obeying Newton’s equation of motion with the force given in equation (2.1). For details on the simulation technique and initial conditions [16], see electronic supplementary material, S1. We use $\epsilon^2 = 15$ and $C = 1$ throughout. Note that lengths in the observational data are given in millimetres, as there is no obvious length scale in the laboratory swarms that can be used for non-dimensionalization, while the simulations are necessarily conducted in dimensionless units. Given that our aim is only to compare the model and observations qualitatively to explore the underlying pairing mechanism rather

Figure 1. Pairs in laboratory observations of midge swarms (*a,c,e*) and simulations of the ALRI model (*b,d,f*). (*a*) Trajectories of two midges in a laboratory swarm that exhibited pairing (defined in the text [17]). The midges were identified as belonging to a pair between the blue and black points. Paired parts of the trajectories are coloured in red, while unpaired parts are in grey. Distances are in millimetres. (*b*) A pair as identified in a simulation of the ALRI. Symbols and colours are the same as in (*a*). Distances are in simulation unit lengths. Note that the simulated pair is defined by the energy ratio given in (*h*), but we plot here just the duration marked in (*h*) for clarity. (*c,d*) Distance between the members of the laboratory pair (*c*) and simulated pair (*d*) as a function of time. The blue and the black points correspond to the beginning and end of the red trajectories (*a,b*). The distance is normalized by the swarm size R_s , which is defined as the mean distance of a midge from the centre of mass of the swarm. Time is normalized by the typical orbit time around the centre of mass, defined as $T_{\text{com}} = 4R_s/\bar{v}$, where $\bar{v}$ is the mean midge speed (see electronic supplementary material). In the laboratory observations [17], R_s is typically on the order of 100 mm and $\bar{v}$ is roughly 100 mm s^{-1} . In the example here the laboratory swarm consists of $N = 21$ midges. The simulated swarm has $N = 30$, $R_{\text{ad}} = 15$, $R_s = 5.08$ and $\bar{v} = 0.028$. (*e*) Time-frequency analysis of the motion of the laboratory pair using a continuous wavelet transform [17]. Frequencies are normalized by the typical frequency $f_0 = 1/T_{\text{com}}$ and the amplitude by the time-resolved peak power $F_{\text{max}}(t)$. Higher powers are represented with brighter colours. During pairing, the peak power shifts to higher frequencies. (*f*) Time-frequency analysis of the motion of the simulated pair using a windowed Fourier transform with a window length of $0.14T_{\text{com}}$ (see electronic supplementary material, S2). The behaviour is similar to what is seen in the laboratory pair. (*g*) Probability density function (PDF) of the peak frequency of motion in the simulation during pairing (blue) and independent motion (red), where pairing is defined by the ratio of kinetic to potential energy being less than one. The increase in frequency during pairing seen in the example in (*f*) is statistically robust. See electronic supplementary material for more details. (*h*) The ratio of kinetic to potential energy for the example shown in (*b*). Note the decrease in the energy ratio during pairing, which is defined by having $|E_K/U| < 1$ when averaged over T_{com} .

than to look for quantitative agreement, this mismatch in units is not problematic. Thus, in the model, we are free to define the length, time and mass scales and so chose to take

the combination that appears in C to be 1 (equation (2.1)). This choice amounts to fixing the scale of the ‘mass’ in this gravity-like model. Particles were initialized in the simulations

at random locations within a cube of set size and with no initial velocity (zero kinetic energy). The particles then attract each other to form a stable swarm. We varied the initial conditions from simulation to simulation to maintain an approximately constant kinetic energy (within 10%, see details in electronic supplementary material, S1).

In classical gravity, as in any Hamiltonian system, a bound pair of objects is defined by having a kinetic energy with respect to each other that is smaller than the potential energy between them [22]. For the ALRI model in a swarm of $N > 2$ particles, we do not have a well-defined potential energy [11]. However, we may use as an approximation the expression for a pair of particles, assuming that the rest of the swarm contributes to leading order only a uniform background sound. Integrating the force acting on the two midges (equation (2.1)), we can calculate the effective two-body potential, which is related to the force by $F_{\text{eff}}^i(r_{1,2}) = -\partial U_{\text{pair}}(r_{1,2})/\partial r_{1,2}$ to be

$$U_{\text{pair}}(r_{1,2}) = \frac{C}{\gamma \sqrt{R_{\text{ad}}^2 + \epsilon^2 \gamma^2}} \left(\arctan \left(\frac{\gamma r_{1,2}}{\sqrt{R_{\text{ad}}^2 + \epsilon^2 \gamma^2}} \right) - \frac{\pi}{2} \right),$$

where $r_{1,2} \equiv |\mathbf{r}_1 - \mathbf{r}_2|$ and the background sound appears through the definition of γ (introduced to simplify the expressions) as

$$\gamma(\mathbf{r}_i) \equiv \sqrt{1 + R_{\text{ad}}^2 I_{\text{background}}(\mathbf{r}_i)} \quad i = 1, 2 \quad (2.2)$$

and

$$I_{\text{background}}(\mathbf{r}_i) = \sum_{j=3}^n \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|^2 + \epsilon^2}. \quad (2.3)$$

We approximate γ as a constant for a pair (see electronic supplementary material, S6). We thus take γ to be the average of $\gamma(\mathbf{r}_1)$ and $\gamma(\mathbf{r}_2)$, so that

$$\gamma \equiv \frac{1}{2} [\gamma(\mathbf{r}_1) + \gamma(\mathbf{r}_2)]. \quad (2.4)$$

Adaptivity weakens the interactions and therefore slows down the simulated particles (electronic supplementary material, figure S5), and so we normalize all times by the typical orbit time across the swarm $T_{\text{com}} = 4R_s/\bar{v}$, where $\bar{v}$ is the mean midge speed (see electronic supplementary material, S1) and R_s is the swarm size, defined as the mean distance of a midge from the centre of mass of the swarm.

3. Results

3.1. Pairs in the ALRI model

We use this approximate pair-potential energy (equation (2.2)) to define bound pairs in simulations of the ALRI model as pairs whose ratio of relative kinetic energy

$$E_k = \frac{1}{4} (\dot{\mathbf{r}}_1 - \dot{\mathbf{r}}_2)^2, \quad (3.1)$$

(for the reduced mass of two equal unit masses) and (approximate) potential energy is less than one: $|E_k/U_{\text{pair}}| < 1$ (figure 1h), averaged over a duration of T_{com} . We chose to define the pairs in the simulations differently from the those in the laboratory swarms, since we could take advantage of the additional information that we have in the simulations to

relate the pairing to the energetics of the particle pair, which is physically more meaningful. However, it was important to show (figure 1g) that this definition also agrees (qualitatively) with the observational definition that was necessarily based on threshold values for the oscillation frequency and amplitude [17]. Note that we chose Fourier analysis for the simulation data, as it is more accurate when the frequency range is smaller, which is the case in the simulations compared to the experiments. Both methods (wavelet and Fourier) give consistent results (see electronic supplementary material, S2 and figure S2).

While we cannot use the energy ratio criterion to analyse data from real midge swarms (since we do not know the quantitative strength of the interactions), we can compare other features of pairs between observational data and simulations. The common feature of diminished amplitude of the oscillations of the distance between the paired midges (or particles, in the simulations) (figure 1c,d), compared to the typical unpaired state, is found for most pairs in both observations [17] and simulations (with pairs defined using the energy ratio, figure 1h). Note that the relative-distance oscillations are much more regular for unpaired particles in the simulations when compared with unpaired midges in the laboratory observations. Similarly, we find that the frequency with which the midges (or particles, in the simulations) oscillate with respect to each other increases in the paired state compared to the typical unpaired state (figure 1e,f). In figure 1g, we show that although pairs in the simulations were defined according to the energy ratio (figure 1h), they are highly likely to exhibit the higher-than-normal frequency of mutual oscillations that were used to identify pairs in the observational data (for details of the frequency calculation see electronic supplementary material, S2). These similarities suggest that the mechanism driving pairing in the ALRI model may also be present in real swarms.

Now that we have shown that pairing exists in the ALRI, we can quantify some of its features. In figure 2a, we plot the fraction of time a midge spends in a pair (i.e. the ‘fraction of pairs’) as a function of the length scale of adaptivity R_{ad} for a simulated swarm of size $N=30$. The limit of $R_{\text{ad}} \rightarrow 0$ corresponds to the non-adaptive ϵ -gravity system. Although there are some pairs found in this limit, they are highly transient and their mean lifetime is small, on the order of a single passage across the swarm (figure 2b). As R_{ad} increases, we enter the ‘strong adaptivity’ regime where $R_{\text{ad}}/R_s \gg 1$ and the fraction of pairs saturates at a significantly higher value. In this regime, the pairs are clearly long-lived, on the order of several traversals of the swarm (figure 2b). Note that the values of R_s are emergent from the simulation rather than being set *a priori*, and change with R_{ad} [16].

In figure 2c, we plot the fraction of pairs as a function of the swarm size for a fixed $R_{\text{ad}} = 10$. In our previous analysis [11], we found that real swarms lie in the strong adaptivity regime (where $R_{\text{ad}} \gg R_s$). It is, therefore, highly satisfying that in this limit, where there is no free parameter in the model, we find that the fraction of pairs is similar to the observations [17] (figure 2c). Note that the fraction of pairs seems to become independent of swarm size for swarms larger than approximately 10 midges, both in laboratory observations and model simulations, similar to other properties such as mean speed and core density [23].

The distribution of pair lifetimes in the simulations (figure 2d), shows that long-lived pairs do not arise in ϵ -gravity.

Figure 2. Pairing statistics. (a) The mean percentage of time that midges spend in pairs (fraction of pairs), in a simulation of the ALRI model with $N = 30$ midges as a function of R_{ad}/R_s . The values of R_{ad} at the points (from left to right) are 0, 1, 5, 10, 20 and the corresponding mean R_s values are 6.93, 5.17, 6.03, 5.10, 5.55. (b) The mean lifetime of a pair (τ) in normalized time units as a function of R_{ad}/R_s for the same data as in (a). (c) The fraction of pairs as a function of the number of members of the swarm N for both simulations with $R_{ad} = 10$ (black) and laboratory swarm observations (green). In blue, we give the results from simulations of the non-adaptive ϵ -gravity system. The horizontal dashed lines denote the average percentage of pairs for swarms larger than 10. (d) PDF of the lifetime of pairs without adaptivity in blue ($R_{ad} = 0$, $R_s = 4.9$, $N = 36$), with adaptivity in black ($R_{ad} = 50$, $R_s = 5.1$, $N = 36$), and for laboratory swarm measurements in green ($N = 21$). All are in normalized time units. In the inset, we show that the mean lifetime of pairs is independent of the swarm size, from simulations with adaptivity in black and without adaptivity in blue.

However, ALRI is consistent with measurements from laboratory swarms, as we find there are pairs that survive for many orbits. Note that due to potential reconstruction difficulties in the observations leading to broken trajectories, the lifetimes of the pairs from the laboratory observations should be considered to be a lower bound.

Thus, the ALRI model naturally exhibits pairing, and these pairs bear a number of similarities to those observed in real swarms. It is thus natural to ask what features of the ALRI model produce this pairing, and whether these key features are likely to be present in real swarms. The critical component appears to be adaptivity. The difference in dynamics between the ALRI and ϵ -gravity models is strikingly apparent simply from watching movies of the two simulations (see electronic supplementary material, movies S1 and S2): in the ALRI case, pairs of particles are easily detectable by eye (figure 1*b,d*), while for ϵ -gravity no pairs are evident. In our simulations, the midges are not described as noisy self-propelled particles as is common in classical models of collective behaviour [24]; rather, they simply move inertially according to the effective forces (equation (2.1)) that they feel from the other midges. Effective stochasticity in the trajectories arises from the complexity of Hamiltonian many-body dynamics (e.g. [25]). In addition to this stochasticity, the trajectories of real midges seem to be affected by additional sources of high-frequency and

small-amplitude noise (compare, for example, figure 1*a,b*; electronic supplementary material, figures S3, S4 and S3), which does not seem to qualitatively change the large-scale dynamics of the midges. Since the ALRI is a minimal model, there are certainly additional effects in real swarms that it does not capture [16]. Nevertheless, it does not appear that these other effects, although important for determining the details in real swarms, are required to obtain pairing.

3.2. Pairing mechanism in the ALRI model

Due to adaptivity (equation (2.1)), it is clear that when a midge is close to another within a pair, the strong sound received from its partner acts to screen out the forces due to more distant midges in the swarm [11]. However, this observation does not explain how adaptivity induces the capture of two midges into a bound pair.

In figure 3*a*, we show schematically how this process happens. Suppose two midges are close to each other in the inner part of the swarm, where the background sound level is high and therefore their attraction towards each other is weak. If they happen to be moving together away from the swarm centre, they will experience decreasing background sound levels, resulting in stronger mutual attraction. These two midges initially moved toward each other in a regime of weak mutual attraction (high background sound), gaining

Figure 3. Pair formation. (a) Illustration of the proposed pair formation mechanism. When two interacting midges leave the dense region of the swarm (darker dashed lines), where the background sound γ is high, and move to a lower density region (such that $\gamma_1 > \gamma_2$), the mutual pull between them becomes stronger, their orbit gets tighter, and they become bound. (b) The background sound γ along the path of the simulated pair shown in figure 1b, showing the raw sound (green) and the signal-averaged over T_{com} (magenta). When the background sound γ is decreasing (for $t/T_{\text{com}} \in [3, 6]$), the amplitude of the pair oscillations also decreases (as in figure 1d), and finally the pair is formed (figure 1h, $t/T_{\text{com}} > 5$). (c) PDF of the gradient of the dimensionless background sound at the time when the pair is formed. Data are taken from 250 cases from simulated swarms with $R_{\text{ad}} = 50$, $R_s = 5.1$, and $N = 36$ where the mean energy ratio E_k/U was lower than 1 for a time segment of at least T_{com} . The background sound gradient (γ') tends to be negative during pair formation, in agreement with our theoretical predictions. The mean value of the dimensionless background sound gradient is $-0.65\sigma_{\gamma'}$. (d) The same statistics as in (c) for pair dissociations. Here the mean gradient is positive, with a mean value of $0.56\sigma_{\gamma'}$.

little kinetic energy in the process, but now find themselves in a regime of strong attraction (low background sound) that binds them together as a pair. A mathematical analysis of this process, whereby a decrease in the background sound leads to a tightening of the orbits of the pair, is given in electronic supplementary material, S6. And indeed, the radial distribution of pair-formation events is found in simulations to be concentrated in the high density region of the swarm (electronic supplementary material, figure S11), as there the particles are closest to each other and are likely to be moving from high to low background sound. The mechanism of pair formation in the ALRI model is therefore a many-body effect (since the two midges in this model are still described by Hamiltonian dynamics), but unlike capture in non-adaptive gravity, which hardly ever occurs, the production

of pairs happens robustly. This difference is further illustrated in electronic supplementary material, S7, figure S10.

Pairs form when particles move between the dense core to the less dense outer part of the swarm, thereby crossing gradients of the background sound. This is illustrated in electronic supplementary material, figure S14, showing that the rate of a particle forming a pair is maximal at the region of strong density gradients. This is also expected in an analytic calculation of the background gradients of an ideal Gaussian swarm (electronic supplementary material, figure S16). This property allows us to explain the dependence of the pairing fraction on the swarm size, shown in figure 2a,c. Let us compare the swarms at $R_{\text{ad}} = 10$ and $N = 30$ between the two panels, which differ by their initial conditions: in figure 2a, the initial conditions were of the particles in a cube of side 15, while in

figure 2c, the initial cube side was 35. As a result, the mean speed and size were $\bar{v} = 0.134 \pm 0.003$ and $R_s = 5.5 \pm 0.3$ for figure 2a, and $\bar{v} = 0.210 \pm 0.003$ and $R_s = 11.9 \pm 0.6$ for figure 2c ($N=30$), respectively. The percentage of pairs is found to increase from approximately 10% (figure 2a) to 20% (figure 2c) as the swarm gets larger and 'hotter'. Qualitatively we can explain this observed trend as follows: in small, dense swarms (figure 2a), with low average kinetic energy, the particles mostly remain in the dense core, and their trajectories do not extend to the lower-density outer parts (due to the low kinetic energy), and therefore do not cross the pairing-forming zone (electronic supplementary material, figure S14) very often. In a swarm of larger size, due to higher kinetic energy of the particles, the pairs cross this region more often, and the pairing fraction increases. At initial conditions with even larger spatial extent, and higher kinetic energy, the particles have trajectories that are mostly at distances outside the range of adaptivity, and the swarm loses its cohesiveness, as well as its pairing property.

In the ALRI model, the process of pair formation when moving from high to low background sound is reversed when pairs move from low to high background sound. Since the system obeys time-reversal symmetry, this reverse process acts to break up the pairs (illustrated in electronic supplementary material, figure S10). Note that in the regime of strong adaptivity ($R_{ad} \gg R_s$), the pairing behaviour should not be strongly dependent on the exponent of the power law of the long-range interaction [13]. In the simulations, we also find triplets that form, though at significantly lower proportions (electronic supplementary material, S5, figures S6 and S7).

We can test the validity of this proposed mechanism by computing the gradient of the background sound along the trajectory of the pair at the time of pair formation. In other words, we calculate $\gamma' = d\gamma/dr_{pair}$ where the background sound γ is defined in equations (2.2) and the gradient is calculated along the path of the pair's centre of mass r_{pair} . A specific example is shown in figure 3b for the simulated pair shown in figure 1b,d. We calculated the statistics of γ' at the time of pair formation (figure 3c). There is a clear asymmetry in the distribution of the background sound gradient, with a skewness towards decreasing values along the pair's trajectory at the time of pair formation. Similarly, when pairs break, the mechanism in the model is the increasing background sound (figure 3d).

4. Discussion

We have thus demonstrated that additional behavioural rules are not necessary to drive the formation of pairs in midge swarms, and that pairing (like swarming itself) can be viewed as an emergent phenomenon. Note, however, that our theory does not tell us about the biological function of pairing, since pairing and swarming are inexorably linked in this model. We find that the key ingredients that give rise to pairing are long-range interactions and adaptivity. These features were proposed in an adaptive-gravity model for the midge swarms designed only to describe large-scale features that affect all the midges and give rise to the swarm cohesion [11]. It is, therefore, highly satisfying that without any additional modifications the same model exhibits the pairing phenomenon (figure 1), with the right order of magnitude of the pairing fraction (figure 2c).

Long-range interactions and adaptivity appear in many different contexts and on different scales, including in cellular swarming [26] and insect swarming driven by chemical communication [27]. Our work, therefore, argues that pairing should be a general feature that emerges in biological collective systems that share these properties. Furthermore, recent work on non-biological active particles has also been exploring effective interactions that are long-range [28], and could allow testing for pairing under highly controlled conditions [29]. Pairing may also appear in engineered swarms of robots and drones, depending on the interactions between the autonomous agents [30].

Data accessibility. This article has no additional data.

Authors' contributions. D.G. and N.S.G. developed the theoretical model, and carried out its analysis. D.G. performed the simulations. D.G., N.T.O. and N.S.G. wrote the manuscript. J.G.P., K.v.d.V., M.S. and N.T.O. performed the experiments and analysis of the experimental data. All authors commented and edited the text.

Competing interests. We declare we have no competing interests.

Funding. The research at Stanford was sponsored by the Army Research Laboratory and accomplished under grant no. W911NF-16-1-0185. The views and conclusions in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the US government. K.v.V. acknowledges support from an Early Postdoc Mobility fellowship from the Swiss National Science Foundation, and M.S. acknowledges support from the Deutsche Forschungsgemeinschaft under grant no. 396632606. N.S.G. is the incumbent of the Lee and William Abramowitz Professorial Chair of Biophysics. This work is made possible through the historic generosity of the Perlman family.

References

1. Vicsek T, Zafeiris A. 2012 Collective motion. *Phys. Rep.* **517**, 71–140. (doi:10.1016/j.physrep.2012.03.004)
2. Ballerini M *et al.* 2008 Empirical investigation of starling flocks: a benchmark study in collective animal behaviour. *Anim. Behav.* **76**, 201–215. (doi:10.1016/j.anbehav.2008.02.004)
3. Lopez U, Gautrais J, Couzin ID, Theraulaz G. 2012 From behavioural analyses to models of collective motion in fish schools. *Interface Focus* **2**, 693–707. (doi:10.1098/rsfs.2012.0033)
4. Tunström K, Katz Y, Ioannou CC, Huepe C, Lutz MJ, Couzin ID. 2013 Collective states, multistability and transitional behavior in schooling fish. *PLoS Comput. Biol.* **9**, e1002915. (doi:10.1371/journal.pcbi.1002915)
5. Cvikel N, Berg KE, Levin E, Hurme E, Borissow I, Boonman A, Amichai E, Yovel Y. 2015 Bats aggregate to improve prey search but might be impaired when their density becomes too high. *Curr. Biol.* **25**, 206–211. (doi:10.1016/j.cub.2014.11.010)
6. Attanasi A *et al.* 2014 Collective behaviour without collective order in wild swarms of midges. *PLoS Comput. Biol.* **10**, e1003697. (doi:10.1371/journal.pcbi.1003697)
7. Couzin ID, Krause J, James R, Ruxton GD, Franks NR. 2002 Collective memory and spatial sorting in animal groups. *J. Theor. Biol.* **218**, 1–11. (doi:10.1006/jtbi.2002.3065)
8. Ballerini M *et al.* 2008 Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study. *Proc. Natl Acad. Sci. USA* **105**, 1232–1237. (doi:10.1073/pnas.0711437105)
9. Canizo J, Carrillo J, Rosado J. 2010 Collective behavior of animals: swarming and complex patterns. *Arbor* **186**, 1035–1049. (doi:10.3989/arbor.2010.746n1252)
10. Attanasi A *et al.* 2014 Finite-size scaling as a way to probe near-criticality in natural swarms. *Phys. Rev.*

- Lett.* **113**, 238102. (doi:10.1103/PhysRevLett.113.238102)
11. Gorbos D, Iancu R, Puckett JG, Ni R, Ouellette NT, Gov NS. 2016 Long-range acoustic interactions in insect swarms: an adaptive gravity model. *New J. Phys.* **18**, 073042. (doi:10.1088/1367-2630/18/7/073042)
12. Pearce DJ, Miller AM, Rowlands G, Turner MS. 2014 Role of projection in the control of bird flocks. *Proc. Natl Acad. Sci. USA* **111**, 10 422–10 426. (doi:10.1073/pnas.1402202111)
13. Gorbos D, Gov NS. 2017 Stable swarming using adaptive long-range interactions. *Phys. Rev. E* **95**, 042405. (doi:10.1103/PhysRevE.95.042405)
14. Shoval O, Goentoro L, Hart Y, Mayo A, Sontag E, Alon U. 2010 Fold-change detection and scalar symmetry of sensory input fields. *Proc. Natl Acad. Sci. USA* **107**, 15 995–16 000. (doi:10.1073/pnas.1002352107)
15. Kelley DH, Ouellette NT. 2013 Emergent dynamics of laboratory insect swarms. *Sci. Rep.* **3**, 1073. (doi:10.1038/srep01073)
16. Gorbos D, van der Vaart K, Sinhuber M, Puckett JG, Reynolds AM, Ouellette NT, Gov NS. 2020 Similarities between insect swarms and isothermal globular clusters. *Phys. Rev. Res.* **2**, 013271. (doi:10.1103/PhysRevResearch.2.013271)
17. Puckett JG, Ni R, Ouellette NT. 2015 Time-frequency analysis reveals pairwise interactions in insect swarms. *Phys. Rev. Lett.* **114**, 258103. (doi:10.1103/PhysRevLett.114.258103)
18. Puckett JG, Kelley DH, Ouellette NT. 2014 Searching for effective forces in laboratory insect swarms. *Sci. Rep.* **4**, 4766. (doi:10.1038/srep04766)
19. Tohline JE. 2002 The origin of binary stars. *Annu. Rev. Astron. Astrophys.* **40**, 349–385. (doi:10.1146/annurev.astro.40.060401.093810)
20. Bodenheimer P. 2011 *Principles of star formation*. Astronomy and Astrophysics Library. Berlin, Heidelberg: Springer.
21. D'Orsogna MR, Chuang Y -L, Bertozzi AL, Chayes LS. 2006 Self-propelled particles with soft-core interactions: patterns, stability, and collapse. *Phys. Rev. Lett.* **96**, 104302. (doi:10.1103/PhysRevLett.96.104302)
22. Goldstein H, Poole C, Safko J. 2002 *Classical mechanics*. Boston, MA: Addison-Wesley.
23. Puckett JG, Ouellette NT. 2014 Determining asymptotically large population sizes in insect swarms. *J. R. Soc. Interface* **11**, 20140710. (doi:10.1098/rsif.2014.0710)
24. Vicsek T, Czirók A, Ben-Jacob E, Cohen I, Shochet O. 1995 Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.* **75**, 1226–1229. (doi:10.1103/PhysRevLett.75.1226)
25. Meyer K, Hall G, Offin D. 2008 *Introduction to Hamiltonian dynamical systems and the N-body problem*, vol. 90. New York, NY: Springer Science & Business Media.
26. Daniels R, Vanderleyden J, Michiels J. 2004 Quorum sensing and swarming migration in bacteria. *FEMS Microbiol. Rev.* **28**, 261–289. (doi:10.1016/j.femsre.2003.09.004)
27. Slessor KN, Winston ML, Le Conte Y. 2005 Pheromone communication in the honeybee (*Apis mellifera* L.). *J. Chem. Ecol.* **31**, 2731–2745. (doi:10.1007/s10886-005-7623-9)
28. Lavergne FA, Wendehenne H, Bäuerle T, Bechinger C. 2019 Group formation and cohesion of active particles with visual perception-dependent motility. *Science* **364**, 70–74. (doi:10.1126/science.aau5347)
29. Saha S, Ramaswamy S, Golestanian R. 2019 Pairing, waltzing and scattering of chemotactic active colloids. *New J. Phys.* **21**, 063006. (doi:10.1088/1367-2630/ab20fd)
30. Vásárhelyi G, Virágh C, Somorjai G, Nepusz T, Eiben AE, Vicsek T. 2018 Optimized flocking of autonomous drones in confined environments. *Sci. Rob.* **3**, 1–13.