CANADIAN APPLIED
MATHEMATICS QUARTERLY
Volume 18, Number 3, Fall 2010

REFINING SELF-PROPELLED PARTICLE MODELS FOR COLLECTIVE BEHAVIOUR

CHRISTIAN A. YATES, RUTH E. BAKER, RADEK ERBAN
AND PHILIP K. MAINI

ABSTRACT. Swarming, schooling, flocking and herding are all names given to the wide variety of collective behaviours exhibited by groups of animals, bacteria and even individual cells. More generally, the term swarming describes the behaviour of an aggregate of agents (not necessarily biological) of similar size and shape which exhibit some emergent property such as directed migration or group cohesion. In this paper we review various individual-based models of collective behaviour and discuss their merits and drawbacks. We further analyse some one-dimensional models in the context of locust swarming. In specific models, in both one and two dimensions, we demonstrate how varying parameters relating to how much attention individuals pay to their neighbours can dramatically change the behaviour of the group. We also introduce leader individuals to these models. Leader individuals have the ability to guide the swarm to a greater or lesser degree as we vary the parameters of the model. Finally, we consider evolutionary scenarios for models with leaders in which individuals are allowed to evolve the degree of influence neighbouring individuals have on their subsequent motion.

1 Introduction Often we need only look out of the window to see the principles of collective behaviour at work; in a flock of birds for example. Other collective behaviours which capture our imaginations are those of fish schools, locust plagues and bee swarms. Sometimes we almost forget that, as humans in crowded environments, we exhibit some of the most interesting collective behaviours of any species [**14**]. As well as the aesthetically pleasing aspect of watching a swarm in motion, studying collective behaviour has practical applications: understanding fish schooling can lead to more well developed fishing strategies; a knowledge of the way locusts interact and stay together in devastatingly large, coherent groups may shed light on potential strategies which

may be used to disrupt these groups and halt the swarm's destructive progress [**8, 72**]; studying how humans interact in crowds can be key in devising and implementing efficient and safe crowd control policies, which will help to avoid fatalities in possibly overcrowded situations such as football grounds, music concerts and shopping centres [**14**].

Over the course of the last twenty-five years several so called 'self-propelled particle' (SPP) models (a specific class of individual-based model (IBM)) have been introduced and analysed, in one dimension [**22**], in higher dimensions [**15, 19, 69**], both without and with leaders [**16, 41, 61**], with and without internally or externally generated noise and incorporating more biologically realistic interaction rules [**2, 15, 31, 39, 49, 60**]. These models attempt to replicate naturally observed phenomena from not only animal groups but other self-propelled interacting groups of individuals such as robots. Several attempts have been made to evolve the characteristics of such models [**63**] and to compare the evolved characteristics with those of actual animal swarms [**71**] in order to understand better the possible mechanisms by which these characteristics may have evolved. Many of the models have been found, at least qualitatively, to replicate some of the complex emergent behaviours exhibited by these groups.

In this paper we aim to give a brief overview of some of the above classes of SPP models or IBM. We will discuss a paradigmatic example of each in more detail, illustrating the highlights and pitfalls. For some of the most important and seminal models [**22, 69**] we derive novel results pertaining to phase transitions as parameters of the model vary (see Sections 2 and 3). We also consider the effects of evolving these SPP models in the presence of an 'informed individual' allowing parameters to change in subsequent generations of particles in order to find stable parameter values in a manner analogous to [**71**]. The overview we give is by no means an exhaustive review of SPP models or IBMs. We have selected a variety of archetypal models which cover a wide range of biological situations and modelling regimes but there are, of course, several areas which we will not have space to discuss. These include traffic [**36, 37**] and human crowd modelling [**38, 68**] to name but two.

The remainder of this paper is structured as follows. In Section 2 we discuss the most basic SPP model, formulated in one dimension by Czirók *et al.* [**22**]. After discussing alterations to the model we go on to consider suitable application areas in the specific context of locust swarming. We present higher dimensional models in Section 3. We begin by discussing the simple model proposed by Vicsek *et al.* [**69**] and show that complex emergent behaviours are possible given even the sim-

plest of interaction rules in two dimensions. We consider alterations to the model and summarise the analysis of the potential convergence of the individuals in the model to a common direction. We also consider the archetypal model of [**19**] in both a discrete and continuous formulation and the biologically motivated model of [**15**]. In Section 4 we introduce the concept of leadership in SPP models. We analyse three different types of leadership in the three paradigmatic models previously introduced and discuss possible effects these leader particles can have upon the behaviour of the group as a whole. Finally we discuss the concept of evolving SPP models in order to test theories about the selective pressures acting on individuals which give rise to specific behavioural types. We combine the work of previous sections by introducing evolvable parameters to SPP models in one and two dimensions, studying how the parameters vary as a result of each individual's ability to follow an informed leader. We conclude with a short discussion of the current status of SPP modelling and possible avenues for further research.

2 One-dimensional models The simplest case of a SPP model is the Czirók model in one dimension. The system studied by Czirók *et al.* [**22**], for example, consists of N autonomous agents on a line of length L with periodic boundary conditions. The i^{th} particle, with lateral position x_i moves in one direction or the other along the line with variable speed u_i. The velocity of each agent is updated by considering an average of its own velocity and those of its 'neighbours'. The neighbours of agent i are those particles which lie within a distance r of agent i's position, x_i. The interaction radius, r, is a predefined constant which, in the original model, was chosen to be the same for each particle in the simulation. We use $\mathcal{J}_i^r$ to denote the set of neighbours of particle i, excluding itself, and $n_i(t) = |\mathcal{J}_i^r|$ is the number of these neighbours. The particles are initialised with a random initial position, $x_i \in [0, L)$, and a random initial velocity, $u_i \in \{-1, 1\}$. Their positions and nondimensionalised velocities are evolved using the following two rules, identical for each particle, $i = 1, 2, \ldots, N$,

$$x_i(t+1) = x_i(t) + v_0 u_i(t), \tag{2.1}$$

$$u_i(t+1) = \{G(\langle u(t)\rangle_i)\} + \Delta Q_i, \tag{2.2}$$

and

$$\langle u \rangle_i = \frac{1}{n_i + 1} \left(u_i + \sum_{j \in \mathcal{J}_i^r} u_j \right). \tag{2.3}$$

$\langle u \rangle_i$ is the mean of the nondimensionalised velocity of the particles in the interaction radius of particle i and v_0 is a velocity scaling necessitated by choosing a specific time-step. The role of the function G is to scale the average nondimensional velocity measured by each particle towards unity. A generalisation of the odd function, G, representing this velocity adjustment, is as follows:

$$G(z) = \frac{z + \beta \operatorname{sign}(z)}{1 + \beta}, \tag{2.4}$$

where

$$\operatorname{sign}(z) = \begin{cases} 1, & \text{for } z > 0; \\ 0, & \text{for } z = 0; \\ -1, & \text{for } z < 0, \end{cases}$$

and in [**22**] β is taken to be 1. ΔQ_i represents noise drawn from a uniform distribution with support $[-\eta/2, \eta/2]$ where η will be referred to as the amplitude of the noise. In this, original, statement of the Czirók model (or adapted versions thereof) the time-step is simply taken to be unity, which avoids the necessity of scaling the noise with the time-step, but leaves the model with one fewer degree of freedom. When recapitulating the models we prefer to state them with generality and consistency, which necessitates, in some cases, the introduction of an arbitrary time-step, Δt:

$$x_i(t + \Delta t) = x_i(t) + \Delta t v_i(t) \tag{2.5}$$

$$v_i(t + \Delta t) = v_i(t) + \Delta t \left\{ G(\langle v(t) \rangle_i) - v_i(t) \right\} + \Delta Q_i. \tag{2.6}$$

Note that v_i now represents the dimensional velocity of particle i. The noise parameter η should scale with the square root of the time-step, Δt. We note that the choice of distribution from which the noise is drawn is perhaps less than optimal. Mathematically, at least, it would make more sense to consider normal distributions, especially when considering ensemble properties of the group, such as the average velocity. In order to obtain an update equation for the average velocity of the group, we must sum equation (2.6) over all individuals. Normal noise

has the convenient property that the sum of two (or more) normal distributions is still normal, whereas the sum of two uniform distributions gives the nonstandard, nondifferentiable triangle distribution. We must sum a large number of uniform random variables in order to be able to invoke the Central Limit Theorem and obtain an accessible (normal) distribution. It is pertinent to suggest that noise should be chosen with a specific context in mind. For example, the appropriate noise could be chosen in line with the experimental situation we are trying to model. Often, however, in biological situations it is difficult to determine the nature of the noise and a simple, mathematically tractable, distribution is used to approximate the actual noise distribution.

Figure 1 displays the typical behaviour of particles in the Czirók model over a short time period. Alignment is reached in the absence of underlying or external group coordination although individual particles can sporadically break away from the group. Once the group has formed, its direction is not fixed in time. Given sufficient random particle turning events, due to noise, the direction of the group may be reversed (see Figure 1 (c)).

Czirók *et al.* [**22**] present a continuum analogue of their discrete, one-dimensional model:

$$\frac{\partial U}{\partial t} = f(U) + \mu^2 \frac{\partial^2 u}{\partial x^2} + \frac{2\mu^2}{\rho} \frac{\partial U}{\partial x} \frac{\partial \rho}{\partial x} + \eta, \tag{2.7}$$

$$\frac{\partial \rho}{\partial t} = -v_0 \frac{\partial (\rho U)}{\partial x} + D \frac{\partial^2 \rho}{\partial x^2}, \tag{2.8}$$

where $U(x,t)$ is the coarse-grained dimensionless velocity field and $\rho(x,t)$ is the coarse-grained density. The self-propulsion term, $f(U)$, is an antisymmetric function with $f(U) > 0$ for $0 < U < 1$ and $f(U) < 0$ for $U > 1$. The noise, η, has zero mean and a density-dependent standard deviation: $\eta^2 = \sigma^2/\rho\tau^2$. For a more thorough discussion of the derivation and analysis of this continuum limit, see [**22**].

Using this continuum analogue Czirók and Viscek [**21**] demonstrate that, in the zero noise case (i.e., $\eta = 0$), the two steady states of the system, where all the particles move in the same direction (either positive or negative), are absolutely stable to small perturbations and that it is plausible to assume that the behaviour of the system converges to those steady states. However, since the model without noise is physically uninteresting in one dimension we will not discuss it further and will concentrate, subsequently, on finite, nonzero noise.

In the original formulation of the model ($\Delta t = 1$) it is possible to adapt the update rule (2.2) so the importance that particle i places

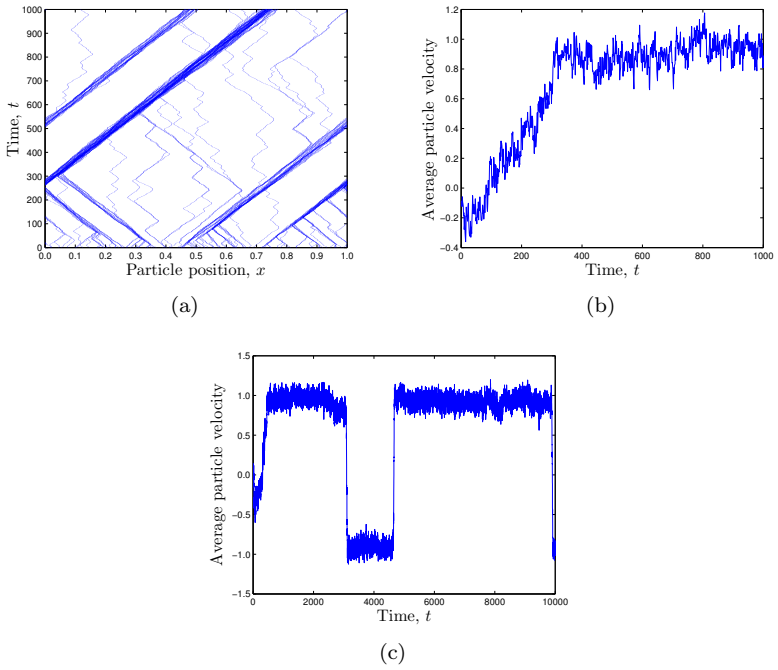


FIGURE 1: Panel (a) shows the typical evolution of the particles' positions over 1000 time-steps ($\Delta t = 1$) using the Czirók model (see equations (2.1) and (2.2)). Panel (b) shows the corresponding average nondimensionalised velocity of the particles over the same time period. The high average velocity from $t \approx 400$ onwards in (b) corresponds to movement in a common direction of most of the particles seen in (a). Panel (c) shows a longer time course for the same model (10,000 time-steps). The ensemble directional switches at $t \approx 3000$ and $t \approx 5000$ are characteristic of the model. Parameter values are as follows: number of individuals, $N = 100$; uniform size of interaction radius, $r = 0.005$; normalised domain length, $L = 1$; particle speed, $v_0 = 0.001$ and noise parameter, $\eta = 2$.

on its own, current velocity is variable; from complete self-importance, ignoring all the particles around it, to a complete disregard for its own previous velocity. Introducing a weighting parameter, α, the equation for the mean velocity of the surrounding particles used in the velocity

update equation becomes

$$\langle u \rangle_i = \frac{1}{n_i(t) + \alpha} \left(\alpha u_i + \sum_{j \in \mathcal{J}_i^r} u_j \right). \tag{2.9}$$

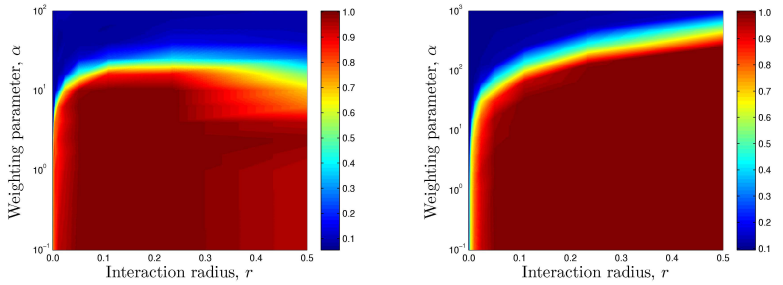
Choosing $\alpha = \infty$ implies complete independence from other particles whereas choosing $\alpha = 0$ implies a particle's next velocity is independent of its previous velocity. Choosing $\alpha = 1$ returns us to an equivalent formulation of equation (2.3), where the velocities of all particles in the near neighbourhood, including particle i itself, are given equal weighting. We must be careful when choosing $\alpha = 0$ since it is possible that a particle may find itself in a position where it has no nearest neighbours. In this case the velocity update rule is undefined. If this happens, since the particle has no neighbours from which to draw influence, it continues with its current velocity.

Figure 2 shows the variation of group alignment with the weighting parameter and the interaction radius. By alignment at each time-step we refer to the normalised sum of the absolute value of the mean velocities of all particles,

$$\frac{1}{N} \left| \sum_{i=1}^{N} \frac{u_i}{u_{\max}} \right|, \tag{2.10}$$

where $u_{\max}$ is the maximum, over all particles, of the absolute value of their velocities. This measure gives us a sense of the cohesion of the group. The group alignment for a simulation was taken to be the average of the values of the alignments during the second half of the simulation. We consider only the second half of the simulation in order to allow the particles to 'warm up' (to allow aggregation to occur or otherwise) so that initial conditions do not have too significant an effect on the measured alignment. In general we see reduced cohesion when particles pay little attention to each other and are more focussed on their own velocities (i.e., $\alpha \gg 1$ or $r = 0$) and larger cohesion when particles are more aware of each other, either by giving more weight to other particles' velocities (reduced α) or simply by considering more particles (increased r). The addition of noise to the model has the effect of smoothing the transition between ordered and unordered states and changing the values of α and r at which this transition occurs.

In one dimension there is a limit to the complexity of an SPP model. Eftimie *et al.* [**24, 25**] introduce a continuum version of an IBM [**51**]. By



(a) (b)

FIGURE 2: Alignment averaged over the second half of 50 simulations of the weighted Czirók model for varying values of the interaction radius, r, and the weighting parameter, α. The simulations in (a) were run in the absence of noise (i.e., $\eta = 0$) and those in (b) with noise parameter $\eta = 2$. Alignment values near unity indicate that the individuals moved in the same direction for the majority of the second half of the simulation. Alignment values near zero indicate that individuals travelled in different directions in a disordered manner for the majority of the simulation. In both the presence and absence of noise there is reduced group alignment when particles ignore each other and higher alignment when particles interact strongly with each other. All particles were initialised with a random position on the domain and a randomly chosen initial velocity of either v_0 or $-v_0$. The other parameters common to both simulations were $N = 100$, $T_{final} = 500$ (time for which the simulation was run), $\Delta t = 1$, $v_0 = 0.01$ and $L = 1$ (domain length). We have smoothed the alignment surface for ease of viewing. Data points are on an equally spaced 11 by 11 grid of interaction radii and weighting parameters.

assuming several different types of inter-particle interactions in the IBM (combinations of attraction, repulsion, and alignment) they derive corresponding systems of hyperbolic PDEs. They show that a variety of observed and, as yet, unobserved (in nature) behaviours result from these continuum models. Levine *et al.* [**47**] describe a one-dimensional model which incorporates attraction and repulsion forces, and leads rapidly to a steady state distribution of particles. In their original model (incorporating soft-core repulsions only) they find that, as the number of individuals in the group increases, the size of the group remains constant. They recognise that this is biologically unrealistic and, as such, introduce hard-core (large) repulsions which are much stronger than the corresponding attractive forces. This allows group size to expand lin-

early with an increase in members, with each individual maintaining a certain distance from its neighbours. This model, incorporating attraction and repulsion, does not rely on physically unrealistic periodic boundary conditions for cohesion as the Czirók model does. Groups can be held together on an infinite domain, which represents a major advantage. However, in one dimension this model lacks the interesting dynamic behaviour of the Czirók model since the attraction and repulsion forces quickly balance and particles are not able to pass through each other, leading to a steady state distribution of particles. The model also relies on global particle interaction which is often deemed physically unrealistic. Similar attraction-repulsion models in two dimensions [**29, 30, 31, 32, 49**] have more interesting behaviour and will be discussed further in Section 3.5.

2.1 Applications Adaptations of Czirók's original representation have been used to model collective motion in animals. Buhl *et al.* [**8**] used an adapted version of the Czirók model to describe the alignment of a number of locust nymphs marching in a ring-shaped arena. Despite the locusts having freedom to move in two dimensions, the annular nature of the arena means that this situation can be considered to be a one-dimensional domain with periodic boundary conditions, and therefore that the locusts' movements are amenable to representation using the Czirók model (see equations (2.1)–(2.3)). Buhl *et al.* [**8**] found that, as the density of locusts increased, they showed a dramatic shift from seemingly solitary, individual-based behaviour to highly cohesive group-level behaviour, identifying a critical density for the onset of collective motion in locusts. The locusts demonstrated a transition from random movement with no clear group orientation to almost ubiquitous motion in one direction around the ring-shaped arena, within a short range of densities. Crucially, Buhl *et al.* [**8**] also discovered a dynamic instability in locust density which allowed the locust groups to switch direction in the absence of external perturbation. This directional switching was found to be exhibited at densities typical of locusts in the field indicating that naturally formed locust groups in the wild can exhibit spontaneous directional switching. The Czirók model with noise also exhibits these directional switches (see Figure 1 (c)) and, as such, a weighted adaptation of the model was deemed appropriate to model the average velocity of the locusts.

The adapted version of the Czirók model used to represent the alignment of the locusts employs the same position update equation as in [**22**] (see equation (2.1)), but alters the weighting of the velocities in a

manner which is significantly different to the weighting described above. The velocity update equation (2.2) of the original model becomes

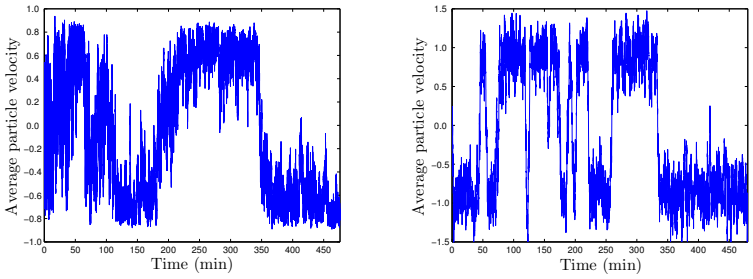
$$u_i(t+1) = \alpha u_i(t) + (1-\alpha)[G(\langle u(t)\rangle_i) - u_i(t)] + \Delta Q_i, \tag{2.11}$$

where the time-step, Δt, is explicitly taken to be 1 (which corresponds to approximately 3.6 seconds of experimental time) and $\langle u(t)\rangle_i$ is, as in the original formulation of the model, the mean velocity of all the particles in the interaction radius of particle i, including i itself. For each simulation all particles were initialised with random initial positions and nondimensional velocities, $u_i \in \{-1, 1\}$. Significantly there is no value of α that will return us to the original model[1] and, as such, this model represents a departure from Czirók's model. We will therefore refer to the model given by equation (2.11) as the Buhl model. Buhl *et al.* [**8**] claim that, for locusts, α corresponds to an inertia-like term, when a locust is considered in the absence of near neighbours.

Interestingly, although Buhl *et al.* [**8**] chose to model the movement of locusts using a Czirók-type model to describe velocity of the particles, the actual experimental data they extracted were for instantaneous alignment, which they referred to as the average orientation for all moving locusts. They found, at least qualitatively, that the switching behaviour of the alignment of the moving locusts was replicated by the average velocity of the particles in their revised model. The interaction radius of the particles (assumed to be the same for each individual) was determined from detailed analysis of the behaviour of locusts in the experimental arena. Other parameter values, such as the underlying speed of the locusts, v_0, and the domain length, L, were also determined experimentally, whereas the magnitude of the noise term, η, was chosen by trial and error in order to replicate qualitatively the switching behaviour of the locusts.

Buhl's groundwork on the modelling of locust behaviour was built upon by Yates *et al.* [**72**] who attempted to quantify the comparison between experimental data and the model. They were able to implement this quantitative comparison by introducing a 'coarse-grained' variable which avoids the 'fine' details of the behaviour of individual locusts. The coarse variable found to represent the dynamics of the model system appropriately was the average particle velocity, U. This is the same quantity compared qualitatively to the locusts' alignment in [**8**]. This is convenient as it is an easy quantity to calculate given the velocities of the

[1]However, choosing $\alpha = 1/2$ will return us to the more generalised formulation of the velocity update equation (2.6) with $\Delta t = 1/2$.



(a) (b)

FIGURE 3: Panel (a) exhibits a sample time course of alignment of 30 locusts from the experiments described in [**8**]. The switching behaviour of the group is clearly evident. Panel (b) displays the average velocity of 30 individuals simulated according to the Buhl model with experimentally derived parameter values: $L = 36$, $r = 2$, $v_0 = 1$, $\alpha = 0.66$, $\eta = 2$. Qualitatively there appears to be a good comparison between the switching behaviour exhibited by the data and that replicated by the model. First published in Proc. Natl. Acad. Sci. USA [**72**].

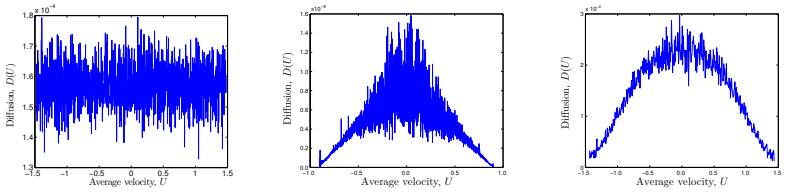
individual locusts. In special cases (i.e., global interaction, $r = L/2$) it is possible to derive analytically a stochastic differential equation (SDE) describing the evolution of the coarse-grained variable, U, from the IBM. Yates *et al.* [**72**] estimated the diffusion coefficient from experimental data and, following a method from [**27**], used this to revise the IBM in order to replicate the coefficient of the SDE describing the coarse variable, U. It was necessary to change the way in which individuals reacted to locally ordered and disordered situations. When individuals found themselves in a more ordered state (i.e., a high value of $|\langle u \rangle_i|$, indicating the locusts in the interaction radius of locust i were travelling mainly in the same direction) it was hypothesised that individuals tended to decrease the randomness of their motion in order to stay aligned with their cohort. However, when a locust found itself in a disordered situation, with the absolute value of the local average velocity being low, it was supposed that the individual increased the random element of its movement in order to bring the group back into an ordered phase.

Consequently, the velocity update equation (2.11) of the IBM was altered to reflect these hypotheses. Instead of assuming the magnitude of the noise term, η, to be constant Yates *et al.* [**72**] assumed it to be a quadratic function of the local average velocity:

$$\eta(\langle u\rangle_i) = \frac{3}{2}\left\{1 - \left(\frac{\langle u\rangle_i}{|\langle u\rangle_i|_{\max}}\right)^2\right\}, \tag{2.12}$$

where $|\langle u\rangle_i|_{\max}$ is the maximum absolute value of the local average velocity. This ensures that the function η, representing the magnitude of the noise term, is always positive. This quadratic dependency of η on $\langle u\rangle_i$ ensured that when the absolute value of local alignment measured by an individual was high the randomness the individual locust introduced to the system remained relatively low and when an individual found its neighbours in a locally disordered state, it correspondingly increased the randomness of its motion.

The model with the revised noise term (2.12) showed, at least qualitatively, a diffusion coefficient which replicated that of the experimental locust data much more accurately than the original Buhl model. The results indicated that individuals move more randomly in locust groups with low alignment, which appears to enable the group to find (and remain in) a highly aligned state more easily. Yates *et al.* [**72**] hypothesised that the revised functional form of their position update equation may be useful in the formulation of other SPP models characterising collective animal behaviour.



(a) (b) (c)

FIGURE 4: Diffusion coefficients of the SDEs (corresponding to the behaviour of the original and revised Buhl models and the experimental data) as a function of average velocity/alignment found by estimation from time series data [**72**] for 30 locusts/particles. Panel (a) represents the diffusion coefficient as a function of alignment for the Buhl model (see equation (2.11)). (b) The diffusion coefficient of an SDE supposed to underlie the experimental data. (c) The diffusion coefficient of the SDE underlying the model with adapted noise magnitude (see equation (2.12)). First published in Proc. Natl. Acad. Sci. USA [**72**].

3 Higher dimensional models

3.1 The Vicsek model The model originally described by Vicsek *et al.* [**69**] to represent flocking in noisy environments is often considered to be the two-dimensional analogue of the Czirók model described in the previous section. Here we initially consider a version of the model in the absence of noise. The model (which we shall henceforth refer to as Vicsek's model) assumes that each particle has a velocity with a constant magnitude and that the directions (or headings) of the particles change at each time-step. In the discrete, deterministic, two-dimensional implementation of this model the direction, θ_i, of particle i ($i = 1, \ldots, N$) is updated as follows:

$$\theta_i(t+1) = \arg\left\{ \exp\left(\mathbb{i}\theta_i(t)\right) + \sum_{j \in \mathcal{J}_i^r} \exp\left(\mathbb{i}\theta_j(t)\right) \right\}, \tag{3.1}$$

where[2] the right-hand side in the above update equation can be summarised as the 'average' direction of the particles (excluding particle i) within a circle of radius $r > 0$ (the interaction radius), i.e., $\mathcal{J}_i^r$ represents the set of all the particles, j, such that $\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2 < r^2$, excluding particle i itself. Again we choose to recapitulate the model in a time-step independent manner as follows:

$$\theta_i(t+\Delta t) = \theta_i(t) + \Delta t \Bigg[\arg\left\{ \exp\left(\mathbb{i}\theta_i(t)\right) + \sum_{j \in \mathcal{J}_i^r} \exp\left(\mathbb{i}\theta_j(t)\right) \right\} - \theta_i(t) \Bigg]. \tag{3.2}$$

Often, when simulating flocking models, periodic boundary conditions are used on square domains [**21, 69**]. Taking r to be small in comparison to the domain length gives the 'local-interaction' model. In the 'global-interaction' model all particles interact with all other particles at every time-step.

The position update equation is analogous to that of the Czirók model (see equation (2.5)), but now our position is a vector rather than a scalar:

$$\boldsymbol{x}_i(t+\Delta t) = \boldsymbol{x}_i(t) + \Delta t \boldsymbol{v}_i, \tag{3.3}$$

[2]To distinguish from particle index, i, we have labelled $\sqrt{-1}$ as $\mathbb{i}$.

where $\boldsymbol{v}_i = \cos(\theta_i)\boldsymbol{i} + \sin(\theta_i)\boldsymbol{j}$ and $\boldsymbol{i}$ and $\boldsymbol{j}$ are unit vectors in the x and y directions, respectively. In order to simulate biological situations we add noise to the revised direction update equation (3.2):

$$\theta_i(t+\Delta t) = \theta_i(t) + \Delta t\Bigg[\arg\Bigg\{\exp\left(\mathtt{i}\theta_i(t)\right) + \sum_{j\in\mathcal{J}_i^r}\exp\left(\mathtt{i}\theta_j(t)\right)\Bigg\} - \theta_i(t)\Bigg] + \Delta\theta_i, \tag{3.4}$$

where $\Delta\theta_i$ is a random number chosen with uniform probability from the region $[-\eta/2, \eta/2]$ where, as in the Czirók model, η is referred to as the magnitude of the noise and scales with $\sqrt{\Delta t}$. Choosing $\eta = 2\pi$ represents random reorientation of each particle, ignoring its original velocity and those of its neighbours. With such a large noise term we would not expect to see any coordinated motion.

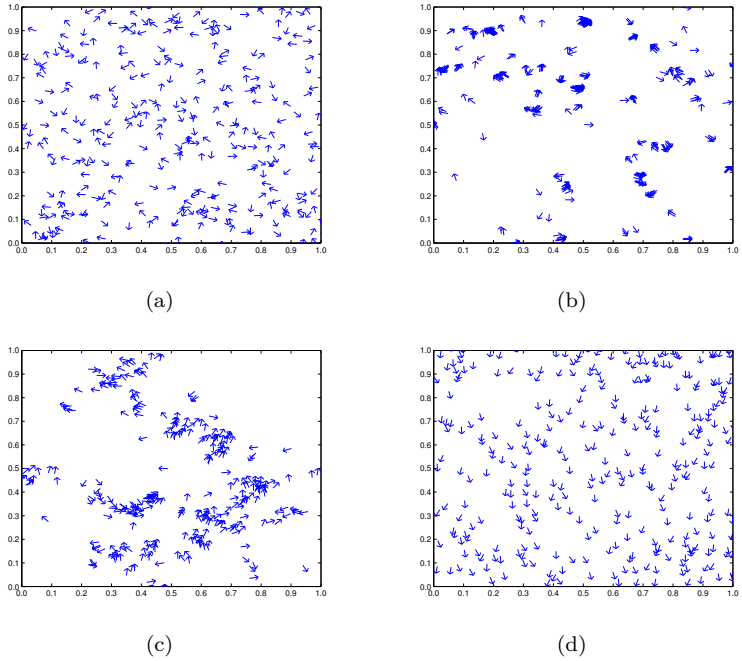
It should be noted, quite apart from any differences of opinion as to the statistical distribution of the noise, the precise step in which the noise is implemented is also a topic of debate. Vicsek *et al.* [**69**] choose to add the noise term to the average of the neighbours' perfectly measured headings, implying that the noise is as a result of the individuals being unable to implement the measured velocity accurately ('angular noise'). Grégoire and Chaté [**34**] make a strong argument that "errors are made when estimating interactions, for example, because of a noisy environment." Consequently they propose that the heading update rule should in fact be

$$\theta_i(t+\Delta t) = \theta_i(t) + \Delta t\Bigg[\arg\Bigg\{\eta\, n_i(t)\exp\left(\mathtt{i}\Delta\theta_i\right) + \sum_{j\in\mathcal{J}_i^r}\exp\left(\mathtt{i}\theta_j(t)\right) + \exp\left(\mathtt{i}\theta_i(t)\right)\Bigg\} - \theta_i(t)\Bigg], \tag{3.5}$$

where $\Delta\theta_i$ is now delta-correlated white noise ($\Delta\theta_i \in [-\pi, \pi]$) and η again controls the magnitude of the noise term.

Vicsek *et al.* [**69**] demonstrate a series of interesting flock-like behaviours (see Figure 5) and show that their models can give rise to a system where all agents eventually move in the same direction (see Figure 5 (d)).

By varying the interaction radius and noise parameter, whilst keeping the number of individuals constant, we elicit a wide variety of behaviours



(a) (b)

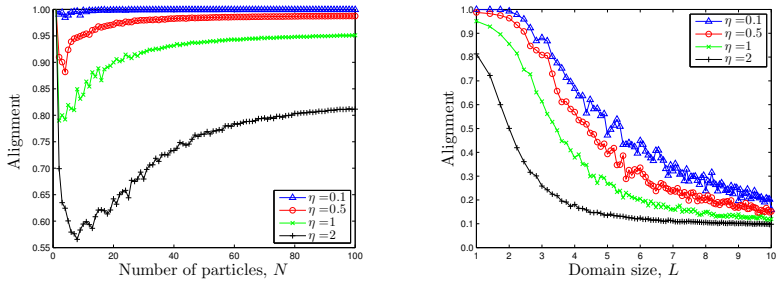
(c) (d)

FIGURE 5: Typical behaviours of the Vicsek model simulated using the original model formulation (see equations (3.1) and (3.3)) with $\Delta t = 1$ and noise incorporated. The directional alignment of all individuals is displayed for varying values of interaction radius and noise. Noise is implemented as angular noise, as in the original Vicsek model. In each case the number of particles is $N = 300$, the domain is normalised ($L = 1$) and the particles move with common speed $v_0 = 0.01$. Panel (a) shows a typical initial configuration of the particles at random positions with random headings. Panel (b) demonstrates the formation of small compact groups of individuals who travel together in random directions in the case where the interaction radius is small and noise is also small: $r = 0.01$, $\eta = 0.1$. Panel (c) shows the formation of larger, better aligned/coherent groups as a consequence of an increased interaction radius at the same noise level: $r = 0.03$, $\eta = 0.1$. Panel (d) shows the formation of a single coherent group encompassing all individuals as a result of a further increase in interaction radius which more than compensates for an increase in the noise term: $r = 0.1$, $\eta = 1$.

from the Vicsek model: Figure 5 demonstrates just some of these behaviours. Clearly the model exhibits several different types of behaviour as the parameters vary: from completely random uncorrelated behaviour of individuals (not shown, often known as 'gas' phase), through the formation of small cohesive groups with an absence of global correlation (see Figure 5 (b) and (c) often known as 'liquid' phase), to a single group encompassing all individuals travelling in the same direction (see Figure 5(d), 'solid' phase).

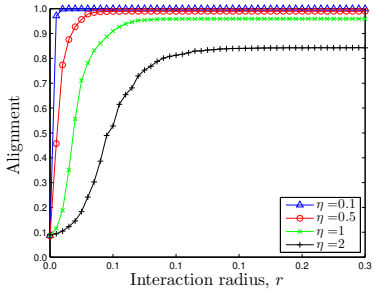
In the classical sense a phase transition refers to the transformation of a thermodynamic system from one 'phase' or 'state' to another (e.g., liquid to solid (evaporation) or solid to gas (sublimation)). As a result of a change in some external condition (temperature, pressure or volume in the classical definition) certain properties of a medium change (often in a discontinuous manner). Analogously, in a system of self-driven particles, when a parameter of the system changes causing the particles to alter their behaviour from a disordered to an ordered state (or *vice versa*) this is known as a phase transition. The interaction-radius-dependent phase transition described above is reminiscent of the density-dependent phase transitions detailed by Vicsek *et al.* [**69**]. In order to elicit their density-dependent phase transition they reduced the size of the domain whilst keeping particle numbers constant, leading to an increase in density. This process is essentially analogous to increasing the interaction radius. In Figure 6 we demonstrate phase transitions as a function of interaction radius and density (by varying particle number and domain size).

In Figure 6 we see that the density-dependent phase transition arising from altering the number of particles (see Figure 6 (a)) is less sharp than that caused by altering the domain size with a constant number of particles (see Figure 6 (b)). In the first case, low density implies low particle numbers, which are unlikely to have low alignment values. In the extreme, $N = 1$, the alignment is, by definition, equal to its maximum, 1. In a two-particle system ($N = 2$) in order for the alignment to be close to zero, both particles would have to move in opposite directions for the duration of each simulation, which is extremely unlikely. The same low particle density with a large fixed number of individuals on a correspondingly large domain gives a much higher probability of low alignment due to the eventual formation of small clusters of individuals often travelling in opposite directions to each other (see Figure 5 (b)). The phase transition exhibited in Figure 6 (c) as a result of varying the interaction radius is sharp, indicating some critical interaction radius, r_c, above which group motion is dominant and below which particles act as individuals. Vicsek *et al.* [**69**] interpret the existence of such critical



(a)

(b)



(c)

FIGURE 6: Phase transitions exhibited by the Vicsek model as a function of density and interaction radius over a range of values of the noise parameter, η. Panels (a) and (b) show the phase transition caused by changing the density. In (a) the density is changed by varying the number of individuals between $N = 1$ and $N = 100$ with constant domain length, $L = 1$, whereas in (b) the number of individuals remains constant, $N = 100$, but the length of the domain changes between $L = 1$ and $L = 10$, giving the same range of densities as in (a). In both cases the density varies from 1 individual per unit area to 100 individuals per unit area and all particles have a constant interaction radius, $r = 0.2$. Panel (c) exhibits the phase transition caused by varying the interaction radius between $r = 0$ and $r = 0.5$ with a constant number of individuals, $N = 100$, and constant domain size, $L = 1$. Each simulation is run for a range of noise parameters from $\eta = 0.1$ to $\eta = 2$. Alignments are averaged over the second half of a 500 time-step simulation and over 50 repeats.

thresholds in density as evidence that, in the limit of large particle numbers, the model exhibits a kinetic phase transition analogous to phase transitions found in equilibrium systems.

As in the case of the Czirók model, it is possible to adapt the update rule (3.1) so that the importance particle i places on its own current velocity is variable; from complete self-importance, ignoring all neighbouring particles, to complete disregard for its own previous direction. Reintroducing the weighting parameter, α, the direction update equation (in the absence of noise), (3.1), becomes

$$\theta_i(t+1) = \left[\arg \left\{ \frac{1}{n_i(t)+\alpha} \left(\alpha \exp(\mathrm{i}\theta_i(t)) + \sum_{j \in \mathcal{J}_i^r} \exp\left(\mathrm{i}\theta_j(t)\right) \right) \right\} \right]. \tag{3.6}$$

Figure 7 shows the variation in group alignment with weighting parameter and interaction radius. By alignment at each time-step we refer to the normalised cumulative Euclidean distance travelled by the particles,

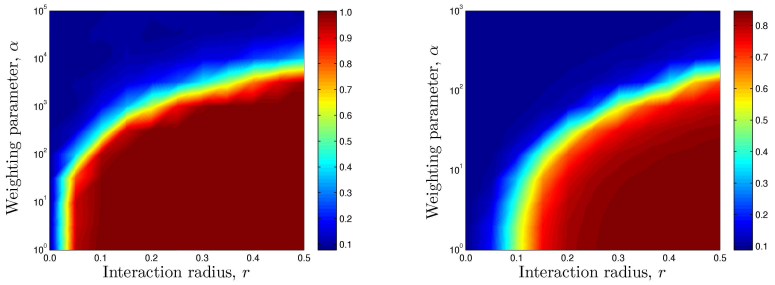
$$\frac{1}{N} \left| \sum_{i=1}^{N} \boldsymbol{v}_i \right| = \frac{1}{N} \sqrt{\left(\mathrm{Re} \sum_{i=1}^{N} \exp(\mathrm{i}\theta_i) \right)^2 + \left(\mathrm{Im} \sum_{i=1}^{N} \exp(\mathrm{i}\theta_i) \right)^2}, \tag{3.7}$$

which gives us a sense of the cohesion of the group. As in the one-dimensional Czirók model we see reduced cohesion when particles pay little attention to each other and are more focussed on their own velocities (i.e., $\alpha \gg 1$ or $r = 0$) and larger cohesion when particles are more aware of each other as a result of either giving more weight to other particles' velocities (reduced α) or simply through considering more particles (increased r). The addition of noise to the model again has the effect of smoothing the transition between ordered and disordered states but means that the maximum alignment achieved is never as high as the maximum alignment in the noiseless model.

The three-dimensional analogue of this model is given by Czirók and Vicsek [**21**] as

$$\boldsymbol{v}_i(t+\Delta t) = \boldsymbol{v}_i(t) + \Delta t \left(\boldsymbol{N}(\langle \boldsymbol{v}(t) \rangle) - \boldsymbol{v}_i(t) \right), \tag{3.8}$$

where $\boldsymbol{N}(\boldsymbol{v}) = \boldsymbol{v}/|\boldsymbol{v}|$ and the position update rule is as in equation (3.3). If noise were to be reintroduced to the model we would simply add ΔQ,



(a) (b)

FIGURE 7: Alignment averaged over the second half of 50 simulations of the weighted Vicsek model for varying values of the interaction radius, r, and the weighting parameter, α. The simulations in (a) were run in the absence of noise (i.e., $\eta = 0$) and those in (b) with noise parameter $\eta = 2$. Alignment values near unity indicate that the individuals travelled in the same direction as each other for the majority of the simulation. Alignment values near zero indicate that individuals travelled in different directions in a disordered manner for the majority of the simulation. The other parameters common to both simulations were $N = 100$, $T_{final} = 500$ (time for which the simulation was run), $\Delta t = 1$, $v_0 = 0.01$ and $L = 1$ (domain length). All individuals were initialised with random positions and direction. We have smoothed the alignment surface for ease of viewing. Data points are on an equally spaced 11 by 11 grid of interaction radii and weighting parameters.

uniformly distributed noise in a sphere of radius η, to equation (3.8), where η scales with $\sqrt{\Delta t}$ (in analogy to the one- and two-dimensional models in [**22**] and [**69**], respectively).

Chaté *et al.* [**11, 12**] consider the effect of changing particle interactions and orientations from polar (i.e., orientation and direction vector pointing in one specific direction) to apolar (orientation vector with no preferred direction). At every time-step each particle moves forwards or backwards (with equal probability) with a newly calculated orientation. Using this model Chaté *et al.* [**12**] find that "giant density fluctuations" occur in the ordered phase in contrast to the original Vicsek model in which high density high order bands occur interspersed with low density low order bands which do not fluctuate hugely. In the new model only quasi-long-range order is observed in 2 D as opposed to true long-range order in the original model. Chaté *et al.* [**12**] also suggest and investigate the addition of cohesion to the Vicsek model (see Section 3.6) and

the consideration of the role of the ambient fluid in which the particles are moving. Both additions serve to make the Vicsek model more biologically realistic.

Peruani *et al.* [**56**] also consider a variation on the Vicsek model in which the particles are rod-shaped. The rods are assumed to exhibit volume exclusion, but in a "soft" manner which means that a certain degree of overlap is allowed. This enables the packing fraction to be controlled. They find that a sufficiently large packing fraction, ϕ, and aspect ratio of the cells, κ, are necessary for clustering. They derive a mean field approximation and find an analytical relationship between κ and ϕ which characterises the transition from nonclustered to clustered particles. Individual-based simulations correctly characterised the transition curve in the $(\phi - \kappa)$ parameter space plane as predicted by the mean field model, when correctly parameterised. This model should be considered a minimal model for collective phenomena in actively moving rod-shaped cellular entities.

We will focus next on analysing the two-dimensional version of the original polar, point-particle Vicsek model.

3.2 Analysis of the Vicsek model Jadbabaie *et al.* [**41**] consider a version of the Vicsek model which is amenable to mathematical analysis using adjacency matrices, but which by their own admission can introduce inconsistencies to the model which the original Vicsek model does not have. Here, by "adjacency matrix," we refer to a matrix which has nonzero entries if and only if two particles are sufficiently close to interact or, in graph theoretic terms, if and only if the two nodes (particles) of the graph representing particle interactions are connected (sufficiently close). Jadbabaie *et al.* [**41**] alter the heading update equation from equation (3.2) to

$$\theta_i(t+1) = \frac{1}{n_i(t)+1}\left(\theta_i(t) + \sum_{j\in\mathcal{J}_i^r} \theta_j(t)\right), \tag{3.9}$$

where, we recall, $n_i(t) = |\mathcal{J}_i^r|$ is the number of particles in the interaction radius of particle i *excluding* itself. For example, using this new heading update rule, three mutually and exclusively interacting particles each with headings, $\theta_i = 2\pi/3$ for $i = 1, 2, 3$, would inexplicably change direction at the next time-step to have a heading of zero instead of retaining their original heading, as would be expected. Generally this type of behaviour is unrealistic, but in the limit of large particle numbers we will assume this effect diminishes as the occurrences of events like that described above becomes less likely.

Jadbabaie *et al.* [**41**] analyse the relationships between neighbours using a simple, undirected, time-dependent graph with vertex set $\{1, 2, \ldots, N\}$. An edge connects the vertices corresponding to particles i and j if and only if i and j are neighbours (using the previous space-dependent definition of neighbours). The adapted heading update rule of equation (3.9) can be rewritten in the matrix form

$$\boldsymbol{\theta}(t+1) = F_p \boldsymbol{\theta}(t), \tag{3.10}$$

where

$$F_p = (1 + D_p)^{-1}(A_p + I). \tag{3.11}$$

Here, A_p is the symmetric adjacency matrix of the undirected graph G_p describing interactions between the particles and D_p is the diagonal matrix whose i^{th} diagonal element corresponds to the degree of vertex i. The subscript p is a member of $\mathcal{P}$, a suitably defined indexing set of all possible simple graphs $\mathbb{G}$ on N vertices. The heading state vector $\boldsymbol{\theta}$ is given by $[\theta_1, \theta_2, \ldots, \theta_N]'$. Clearly the adjacency matrix, A_p, of graph $G_p \in \mathbb{G}$, representing the neighbour relationships of each agent, changes over time and, as such, a complete description would include a model of how the index $p(t)$ varies with time as a function of the positions of the N agents. Such a model, while essential for simulation purposes, can be difficult to account for when investigating convergence and, as such, can be bypassed by assuming that the indices $p(t)$ all come from a suitably defined subset (of $\mathcal{P}$) of interest.

Jadbabaie *et al.* [**41**] point out that naturally there will be situations for which the headings of the N agents will not converge to a common heading, but attempt to demonstrate that for a certain class of indices, given appropriate initial conditions (i.e., agents are not too spread out and interaction radii are not too small), convergence to a common heading, θ_{ss}, such that $\lim_{t\to\infty} \boldsymbol{\theta}(t) = \theta_{ss}\mathbf{1}$, will occur. Here $\mathbf{1} = [1, 1, \ldots, 1]'$ denotes the $1 \times N$ column vector of ones. A trivial case for which this is true is $r \gg 1$. In this case we can assume that all agents will remain neighbours of all others for all time and hence that index $p(t)$ will remain on a fixed trajectory in $\mathcal{P}$ such that G_p is always a complete graph. In this case convergence to a common heading is easy to establish as the heading update equation (3.9) reduces to a linear, time-invariant, discrete system.

Jadbabaie *et al.* [**41**] prove a central theorem relating to the convergence of the more interesting case of an intermediate value of the

interaction radius, r, which is neither too small to inhibit convergence, nor too large to ensure it with certainty. In this case $G_{p(t)}$ is not necessarily complete or even connected for any single time $t \geq 0$, but at the same time no proper set of vertices of $G_{p(t)}$ is isolated from the rest for all time. A set of graphs is said to be *jointly connected* if the (graph theoretic) union of its members is connected. Consequently a set of consecutive graphs (each representing the adjacency of the particle system at a time point in the interval $[t, \tau]$, where $t < \tau$) are said to be *linked together* across that interval if they are jointly connected.

Jadbabaie's central theorem states that, when all agents are linked together over contiguous intervals of arbitrary but finite (and uniformly bounded) length, convergence of the agent's headings to a common heading, θ_{ss}, is guaranteed, where θ_{ss} depends only on the initial headings and the intermediate indices $p(t)$.

The proof of eventual convergence of the N-agent system to a common heading under the specific conditions given is useful in so far as it provides us with insight into the dynamic situation which will give rise to collective behaviour of the individual agents. However, these results should be interpreted with caution. Firstly, as we have already mentioned, the model considered in [**41**] is not the same as the Vicsek model and, due to the possibility of erratic unpredictable changes in direction, is unlikely to be a realistic or desirable model of either biological or robotic agents. The second point to make is that, although conditions for convergence are given, they are dynamic conditions which depend upon the state of the system at each update, which makes predicting the model behaviour, given only the initial conditions, difficult. It should be noted, however, that the dynamic conditions outlined above are far more likely to be effectively satisfied if N is large and we consider individuals on a finite domain with periodic boundary conditions. On a sufficiently small domain this last condition ensures that no particle will stray too far from other members of the group and, hence, that the likelihood of finite sequences of states in which the corresponding set of adjacency graphs in each sequence are jointly connected increases.

3.3 The Cucker-Smale model Cucker and Smale [**19**] introduce another model for representing the arrival at consensus decisions in the absence of central direction. They discuss their model in the specific context of animal flocking in physical space $\mathbb{R}^3$ although they point out that their approach can be used to model systems as diverse as the emergence of a common price system in a financial market and of common languages in primitive societies.

Their main goal when considering the model in the context of flocking is to show that, given certain types of communication between individuals, the group converges to a state in which all individuals move with the same velocity. Their analysis enables straightforward identification of parameter regimes in which this behaviour is a necessary outcome of the model's evolution. In other parameter regimes it transpires that Cucker-Smale (CS) convergence to a uniform state or otherwise is dependent only on the initial state of the system. This is in direct contrast to the strict dynamic requirements for convergence of the Vicsek model outlined in [**41**].

The CS model is based on the following hypothesis on the individual-level behaviour: each individual adjusts its velocity between time-steps by adding to it a weighted average of the differences between its velocity and those of the other individuals. The weighting function is postulated to be some nonincreasing function of the distance between individuals. The discrete version of the velocity update rule for individual i at time t, given N individuals, is

$$\boldsymbol{v}_i(t+\Delta t) = \boldsymbol{v}_i(t) + \frac{\lambda \Delta t}{N} \sum_{j=1}^{N} a_{ij}(\boldsymbol{v}_j(t) - \boldsymbol{v}_i(t)), \tag{3.12}$$

where

$$a_{ij} = \zeta(||\boldsymbol{x}_i - \boldsymbol{x}_j||^2), \tag{3.13}$$

is the weighting function and $\zeta : \mathbb{R}^+ \to \mathbb{R}^+$ describes the distance-dependent weighting function. λ parameterises the strength of the interactions and Δt is the time-step between updates of the position and velocity. Cucker and Smale [**19**] choose

$$\zeta(y) = \frac{K}{(\sigma^2 + y)^\beta}, \tag{3.14}$$

for $K > 0$ and $\beta \geq 0$, and incorporate the individual pair weightings $\{a_{ij}\}$ into an adjacency matrix, A.

Cucker and Smale [**19**] also consider the continuum limit of their model. Let D be the $N \times N$ diagonal matrix, $D = \mathrm{diag}(d_1, d_2, \ldots, d_N)$ such that $d_i = \sum_{j=1}^{N} a_{ij}$ is the sum of all the weights corresponding to individual i. Defining the graph Laplacian $L = D - A$ to be the difference between the two matrices, the velocity update rule for particle i is reformulated more succinctly as

$$\boldsymbol{v}_i(t+\Delta t) = \boldsymbol{v}_i(t) - \Delta t[L\boldsymbol{v}(t)]_i, \tag{3.15}$$

and the corresponding position update rule is as in equation (3.2). The continuous time version of the model is formulated naturally in the limit, as $\Delta t \to 0$, as a system of differential equations.

$$\boldsymbol{x}' = \boldsymbol{v}, \tag{3.16}$$

$$\boldsymbol{v}' = -L\boldsymbol{v}. \tag{3.17}$$

For $\beta < 1/2$ decay of the function ζ is relatively slow, which guarantees convergence of the discrete model independent of initial conditions. Cucker and Smale [**19**] show that this unconditional flocking behaviour occurs exponentially quickly. However, in the $\beta \geq 1/2$ case the decay of ζ is moderately fast and convergence is only guaranteed under specific initial conditions [**19**]. In the continuum model a similar convergence result is achieved.

In order to prove the above, the authors make use of some properties of the graph Laplacian, L, of the dynamic network. They decompose the state space $(\mathbb{R}^3)^N$ into two subspaces Δ and its orthogonal complement $\Delta^\perp$ defined by the action of L i.e., $L(\Delta) = 0$ and $L(\Delta^\perp) \subseteq \Delta^\perp$ where

$$\Delta = \{(\boldsymbol{u}, \boldsymbol{u}, \ldots, \boldsymbol{u}) | \boldsymbol{u} \in \mathbb{R}^3\}. \tag{3.18}$$

Each point $\boldsymbol{x} \in (\mathbb{R}^3)^N$ decomposes in a unique manner as $\boldsymbol{x}_\Delta + \boldsymbol{x}_\perp$ where $\boldsymbol{x}_\Delta \in \Delta$ and $\boldsymbol{x}_\perp \in \Delta^\perp$. The velocities $\boldsymbol{u}$ can be decomposed in a similar manner. Denote the mean velocity of the flock by

$$\bar{\boldsymbol{v}} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{v}_i. \tag{3.19}$$

Then $\boldsymbol{v}_\Delta = (\bar{\boldsymbol{v}}, \bar{\boldsymbol{v}}, \ldots, \bar{\boldsymbol{v}})$ and, necessarily, $\boldsymbol{v}_\perp = (\boldsymbol{v}_1 - \bar{\boldsymbol{v}}, \boldsymbol{v}_2 - \bar{\boldsymbol{v}}, \ldots, \boldsymbol{v}_N - \bar{\boldsymbol{v}})$. Using this decomposition we can view the evolution of the particles' velocities as the evolution of their mean velocity, $\boldsymbol{v}_\Delta = \bar{\boldsymbol{v}}(t)$, and the evolution of the differences between individuals' velocities and the mean velocity, $\boldsymbol{v}_\perp = (\boldsymbol{v}_1 - \bar{\boldsymbol{v}}, \boldsymbol{v}_2 - \bar{\boldsymbol{v}}, \ldots, \boldsymbol{v}_N - \bar{\boldsymbol{v}})$. Cucker and Smale [**19**] are able to show that the projections over $\Delta^\perp$ of the solutions of equations (3.15) and (3.2) are the solutions of the restriction of (3.15) and (3.2) to $\Delta^\perp$, respectively. This is of crucial importance since it is the relative differences in the positions and the velocities of the individuals $\boldsymbol{x}_i - \boldsymbol{x}_j$ and $\boldsymbol{v}_i - \boldsymbol{v}_j$, respectively, that we are interested in, rather than the difference in their positions from a fixed point in space or an arbitrarily

chosen velocity. Clearly, convergence to a common velocity is a feature of the evolution of $\boldsymbol{v}_\perp$ only. Hence, [**19**] study the two quantities

$$\Gamma = \frac{1}{2}\sum_{i\neq j} \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2, \tag{3.20}$$

and

$$\Lambda = \frac{1}{2}\sum_{i\neq j} \|\boldsymbol{v}_i - \boldsymbol{v}_j\|^2, \tag{3.21}$$

and attempt to bound Λ and hence demonstrate the decay of Γ to zero.

Through a series of intricate propositions Cucker and Smale [**19**] prove convergence of the velocities through the decay of $\Gamma(t)$ as $t \to \infty$, and demonstrate boundedness of the differences in positions such that $\Lambda(t) \leq B_0 \ \forall \ t \in \mathbb{R}^+$. In some parameter regimes convergence is found to be unconditional, but in others there are possible dependencies on the initial conditions.

Cucker and Smale [**19**] achieve the remarkable result of exponentially fast convergence in time to the equilibrium dynamics in the absence of any restrictions on the initial conditions in the case $\beta < 1/2$. Even in the case $\beta \geq 1/2$, exponentially fast convergence is proved given that certain conditions on the initial state of the system are satisfied.

3.4 Further analysis of the Cucker-Smale model Ha and Tadmor [**35**] analyse a Vlasov-type[3] kinetic version of the CS model. They use this model to demonstrate the existence of time-asymptotic flocking behaviour for arbitrary initial data of compact support. They also derive a hydrodynamic description of flocking based on their adapted CS Vlasov-type kinetic model and again prove flocking behaviour. They are only able to show that the fluctuations in the energy of the system decay in sub-exponentially fast time. They also require stronger long-range interactions in order to achieve unconditional flocking in their continuum model than were necessary for the analysis of the original discrete CS model.

Carrillo *et al.* [**10**] build on the results of Ha and Tadmor [**35**]. They motivate their continuum description by arguing that instead of simulating the behaviour of a large number of individuals, which may be

[3]A Vlasov system is a system of nonlinear integro-differential equations usually used to describe the dynamics of plasma consisting of charged particles with long-range interaction.

computationally expensive since all individuals interact at each time-step, a mesoscopic partial differential equation (PDE) could be employed to describe the collective behaviour of the flock in terms of its density, $f = f(\boldsymbol{x}, \boldsymbol{v}, t)$, at position $\boldsymbol{x} \in \mathbb{R}^d$, with velocity $\boldsymbol{v} \in \mathbb{R}^d$ at time t, where d is the number of spatial dimensions under consideration. In the absence of other effects the change in density depends essentially on velocity changes arising from binary interactions between individuals. As such, Carrillo *et al.* [**10**] begin by invoking standard 'collisional kinetic theory' to show that their simplified version of the model, for purely binary interactions, leads to the following Boltzmann-type integro-differential equation (IDE):

$$\left(\frac{\partial f}{\partial t} + \boldsymbol{v} \cdot \nabla_x f\right)(\boldsymbol{x}, \boldsymbol{v}, t) = P(f, f)(\boldsymbol{x}, \boldsymbol{v}, t), \tag{3.22}$$

where, as is usual with Boltzmann-type equations, the function P represents the change in particle density due to collisions. In the absence of collisions, this term is zero. Carrillo *et al.* [**10**] use a modified version of the Boltzmann collision operator first postulated by Povzner [**58**] which considers a so-called 'smearing process' for pairwise collisions:

$$P(f, f)(\boldsymbol{x}, \boldsymbol{v}) = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} B(\boldsymbol{x} - \boldsymbol{y}, \boldsymbol{v} - \boldsymbol{w})(f(\boldsymbol{x}, \boldsymbol{v}_*) f(\boldsymbol{y}, \boldsymbol{w}_*) - f(\boldsymbol{x}, \boldsymbol{v}) f(\boldsymbol{y}, \boldsymbol{w}))\, d\boldsymbol{w}\, d\boldsymbol{y}, \tag{3.23}$$

where $\boldsymbol{v}_*$ and $\boldsymbol{w}_*$ are shorthand for $\boldsymbol{v}(t+\Delta t)$ and $\boldsymbol{w}(t+\Delta t)$, respectively, and the function B is known as the collision kernel.

Although the Boltzmann equation (3.22) describes in detail the time evolution of the population density of the flock, a useful description of the phenomenology of flocking is typically given by considering the asymptotically large-time behaviour of the solution. Carrillo *et al.* [**10**] consider what is known as the 'grazing collision' limit, in which the strength of the binary interactions, γ, is small. In order that the effect of the collision integral, P, (given by equation (3.23)) is not diminished when $\gamma \ll 1$ it is imposed that the collision frequency, σ, is increased so that the product $\sigma\gamma$ remains constant. This is the embodiment of the *grazing collision* limit. In this limit the Boltzmann collision operator, P, can be approximated by a dissipative divergence operator, which is easier to analyse. Whereas Ha and Tadmor [**35**] show that the fluctuations in energy are a Lyapunov functional vanishing in sub-exponentially fast

time, using this approximated Boltzmann equation, Carrillo *et al.* [**10**] are able to show that kinetic energy for the continuum interpretation of the CS model vanishes at an exponential rate. This enables them to provide an unconditional flocking theorem for the continuum model with interaction strength estimates that are valid for particle models.

In contrast to the Vicsek model, the CS model is formulated inherently without noise. In a follow up paper, Cucker and Mordecki [**18**] introduce noise to the CS model. They are able to show, using similar arguments to those given in the deterministic formulation of the model, that flocking occurs in a finite time with a certain confidence. As in [**19**] they demonstrate that convergence can be established conditional on the initial conditions only and that for some parameter regimes, even these restrictions are unnecessary. The assumption on the initial conditions (when necessary) can be stated informally as the positions and velocities of the flock not being too dissimilar. Due to the noise in the model, when the velocities of the particles become sufficiently similar, the contribution to the velocity update equation from the noise term will be of approximately the same order as that from a particle's neighbours and, as such, perfect alignment is not possible. By convergence Cucker and Mordecki [**18**] refer to a property called 'nearly alignment' which they define to be a formal measure of (dis)similarity. The dissimilarity of position and velocity are the quantities previously defined as Γ and Λ, respectively. The dissimilarity of $\boldsymbol{x} \in (\mathbb{R}^3)^N$ is in some sense a measure of the dispersion of the group related to the concept of the group's 'diameter'. The flock is said to be 'nearly-aligned' or 'ν-nearly-aligned' when $\Lambda \leq \nu$ for $\nu \geq 0$. Cucker and Mordecki [**18**] demonstrate that ν-nearly-alignment is reached quickly with a certain probability (a lower bound which is given in terms of the initial similarity of positions, velocities, the variance of the distribution from which the noise is drawn and other parameters of the model).

3.5 Attraction-repulsion models Purely alignment-based models such as those of Czirók and Vicsek can never maintain true cohesion of their individual constituents in the absence of biologically unrealistic periodic boundary conditions. In an infinite domain, particles evolving under these types of models will eventually fly apart [**12**]. To avoid this it is necessary to add some sort of cohesion or mutual attraction between particles, the magnitude of which is often a function of the distance separating two particles. In order to avoid collapse of the particles to a point due to mutual attraction it is also often necessary to include a mutual repulsion between particles of a different range, size and, of

course, acting in the opposite direction to the mutual attraction.

Attraction and repulsion forces corresponding to collective motion have been observed in nature [**13, 17, 26, 54**] and one of the earliest mathematical models was formulated by Breder [**6**]. By considering a Lagrangian approach, ignoring the overall motion of the group in preference to a detailed consideration of the spacing of individuals, Mogilner [**49**] investigate the effect of attraction/repulsion parameters on inter-individual distances. They give conditions on the relative strength and range of the repulsion and attraction forces showing that unless the range of attraction is larger than that of repulsion and the strength of repulsion is stronger than the strength of attraction a stable, well-spaced group cannot form. They also demonstrate that isolated pairwise considerations are not sufficient to determine the magnitude of the spacing between individuals in a large, stable group, which is, in fact, smaller than in the isolated case.

Further work on the study of attraction-repulsion IBMs has been carried out by Gazi and Passino [**29, 30, 31, 32**]. Their work focusses on an individual-based continuum model (in both space and time) with long range attraction and short range repulsion. The update rule for the velocity of particle i in $\mathbb{R}^n$ is

$$\boldsymbol{v}_i = \sum_{j=1, j\neq i}^{N} g(\boldsymbol{x}^i - \boldsymbol{x}^j), \tag{3.24}$$

where N is the total number of particles in the system and $g(\boldsymbol{y})$ is some attraction-repulsion function dependent on the distance between particles, $\boldsymbol{y} = \boldsymbol{x}^i - \boldsymbol{x}^j$. The odd attraction-repulsion function, g, can be broken down into its component parts, g_a (corresponding to attraction) and g_r (corresponding to repulsion) and takes the general form

$$g(\boldsymbol{y}) = -\boldsymbol{y}[g_a(\|\boldsymbol{y}\|) - g_r(\|\boldsymbol{y}\|)]. \tag{3.25}$$

The position update rule is the standard one given by equation (3.16) for the continuous time Cucker-Smale model.

Using this model Gazi and Passino [**29**] show that if members of the swarm are sufficiently far away from other individuals they do not feel any repulsion force, only attraction, and correspondingly move towards the centre of mass of the group. They further formalise this idea by bounding the members of the swarm in an n-dimensional hyperball and giving an upper bound on the radius in terms of the parameters of the model. This upper bound increases as the range and strength of

repulsion increases but decreases as the strength of attraction increases, as expected. This model predicts that, for large N, as N increases the size of the group remains bounded, implying that the density of the group increases, which is biologically unrealistic.

Gazi and Passino [**31**] improve on their original model by considering a larger general class of interaction attraction-repulsion forces. They allow the repulsion force to become unbounded so that two particles do not occupy the same position. This addition, in combination with a "safe distance" for each particle in which the repulsion force becomes unbounded (to account for the finite, noninfinitesimal, size of individuals, leads to swarm expansion with constant density as the number of individuals increases.

Gazi and Passino [**32**] combine their IBM with an attractant-repellant gradient, which, in general, individuals attempt to climb down. Their models illustrate a basic foraging principle: the existence of a balance between the desire to stay with the swarm and the desire to interact with the environment.

3.6 Models with multiple interaction radii Although some steps have been taken to ensure they are biologically relevant, the models discussed in the previous sections (Sections 3.3–3.5) allow global interaction, which is, of course, unrealistic in large groups. A more realistic model would take into account the finite size of an individual's senses and build in a finite interaction zone for each individual. The concept of multiple interaction radii has been considered by a number of authors [**2, 15, 39, 60**]. Within each interaction radius, or shell, individuals respond to each other differently. The inner-most region is usually reserved for repulsive reactions, corresponding to the idea that individuals dislike overcrowding and the associated problems it causes (reduced perception of the surroundings, for example, in a biological context). The radius of repulsion can also represent the exclusion principle, assuming that particles have solid physical volumes. The next shell usually corresponds to the zone in which individuals tend to align with one another and the outer shell (or possibly the whole of space outside the inner two regions) corresponds to the region of attraction between individuals which ensures that they do not disperse too widely, given that the innermost region will cause them, in general, to disperse. The order of these shells is not necessarily as described above, nor are the shells necessarily mutually exclusive. A schematic view of the nature of the interactions in each zone is given in Figure 8.

Multiple interaction radii models bring biological realism to SPP mod-

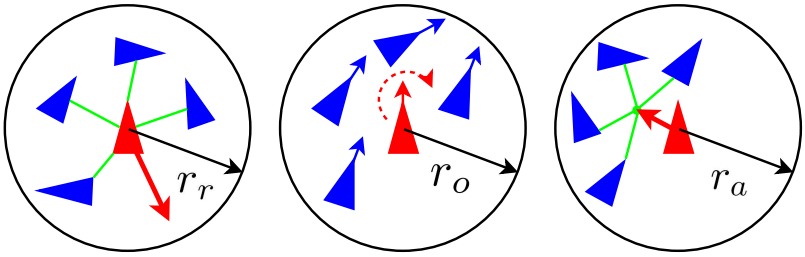


FIGURE 8: Schematic views of the nature of the interactions in each of the three interaction radii used in the models of [**2, 60, 39, 15**]. Panel (a) demonstrates the mutual repulsion at close range found in the inner-most shell with 'radius of repulsion', r_r. In the majority of models if there are particles within the zone of repulsion, then this interaction occurs to the exclusion of the other two. Panel (b) shows the central particle aligning itself with the other particles inside the 'zone of orientation' with radius r_o. Panel (c) exhibits the mutual attraction of the particles in the outermost shell with 'radius of attraction', r_a. In panels (a) and (c) the thicker arrow emanating from the central particle indicates its desired direction of motion. In panel (b) the central particle's desired realignment is indicated via the curved, dashed arrow. For the purposes of clarity in (b) and (c) it has been assumed that the inner radii are of zero size. In reality, these interaction zones will look more like concentric shells.

els that single interaction radius models [**22, 69**] lack. For example, in the most simple models, particles neither avoid collisions nor exhibit attraction towards other particles, which makes the formation of stable, bounded groups in the presence of noise impossible. These simple models cannot therefore replicate the well-defined animal groups seen in nature. Importantly, these revised models show cohesive collective behaviour and group formation in the absence of periodic boundary conditions which are necessary for group formation and alignment in simpler models. Recently, experimental evidence for such multiple interaction radii models has been discovered. Tien *et al.* [**66**] studied the movements of individual fish 'in the field' in both normal unperturbed situations and in the presence of a simulated predator. In the neutral situation both attraction and repulsion zones are evident and in the case of increased fear levels of the fish, induced by the presence of a 'predator', the nearest neighbour distance and the size of these interaction zones were found to

decrease.

The most widely used multiple interaction radii model, in the context of animal migration and collective motion, is that of Couzin *et al.* [**15**]. They assume that the behaviour of individuals results from local repulsion, alignment and attraction based upon the relative position and alignment of individuals with each other. Time is discretised and the position update rule is as in equation (3.3) for the Vicsek model. The velocity update rule is more complicated owing to the increased complexity of interactions considered between individuals.

An individual's highest priority is chosen to be the avoidance of all other individuals, which simulates desire to maintain personal space as well as reducing the probability that two particles are at the same position at the same time [**42**]. Particle i, therefore, turns away from neighbours, j, within the inner-most interaction radius, r_r, (see Figure 8 (a) and Figure 9 (a)) choosing its desired direction of motion at the next time-step, $\boldsymbol{u}_i^r(t+1)$, as follows,

$$\boldsymbol{u}_i^r(t+1) = -\sum_{j\in\mathcal{J}_i^{r_r}} \frac{\boldsymbol{x}_j(t)-\boldsymbol{x}_i(t)}{|\boldsymbol{x}_j(t)-\boldsymbol{x}_i(t)|}. \tag{3.26}$$

If no neighbours are detected within the inner region then individuals tend to be attracted towards each other (in order to avoid isolation) and to align themselves with neighbouring individuals [**52, 54**]. In two dimensions these zones of interaction are circular with the possible exception of a 'blind zone' behind each individual in which neighbours and other external stimuli are undetectable. The blind zone is, in two dimensions, a sector of the circle with interior angle $2\pi-\psi$ radians, where ψ is known as the viewing angle (see Figure 9(b)). Analogously in three dimensions the interaction radii are spheres and the blind zone is represented by a cone with the same interior angle, ψ. Whether the zones of interaction have blind zones depends on the situation the model is designed to simulate. In some biological situations, such as modelling fish schooling, a blind zone is appropriate [**71**] but often full circles/spheres are used since physical interactions can be detected independently of visual interactions.[4]

[4]Eftimie *et al.* [**24**] introduce an adaptation of the traditional IBM in which the relative orientation of individuals has an effect on how they receive and transmit information to each other. They present five different one-dimensional submodels in which individuals receive different information about attraction/repulsion or alignment depending on their orientation and the orientation of their near neighbours. These sub-models are then related to hyperbolic PDEs which are solved with periodic boundary conditions. Several of the emergent patterns from these models are novel and relate to previously unmodelled collective behaviour in animals.

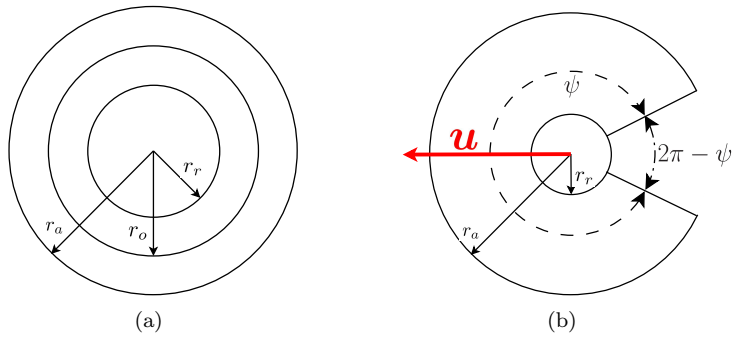


FIGURE 9: Schematic views of interaction zones in two dimensions used in the models in [**2, 15, 39, 60**], amongst others. Panel (a) demonstrates the idea of three mutually exclusive interaction radii. The circle defined by r_r represents the zone of repulsion. The shell defined by r_o represents the zone of orientation and that defined by r_a the zone of attraction. Although the zone of repulsion is generally taken to be the most important, this is not always necessarily the case. Neither are the zones necessarily mutually exclusive or ordered in this manner. Panel (b) represents the actual implementation of the interaction zones. To represent the blind spot exhibited in some species a segment of the outer interaction radii to the rear of the individual (in the opposite direction to an individual's velocity for forward facing individuals) is removed. In practise animals do not always have blind spots or are able to sense movement in a visual blind-spot using sense organs other than sight. For this reason the radius of repulsion, r_r, (often modelling the volume exclusion principle) is rarely modelled with a blind-spot.

Letting $\mathcal{J}_i^{r_o}$ denote the set of neighbours of particle i in the zone of orientation ($r_r \leq |\boldsymbol{x}_j - \boldsymbol{x}_i| \leq r_o$), the models suppose individual i attempts to align its heading with these neighbours,

$$\boldsymbol{u}_i^o(t+1) = \sum_{j \in \mathcal{J}_i^{r_o}} \frac{\boldsymbol{u_j}}{|\boldsymbol{u_j}|}, \tag{3.27}$$

and attempts to reposition itself towards particles $j \in \mathcal{J}_i^{r_a}$, its neighbours in the zone of attraction

$$\boldsymbol{u}_i^a(t+1) = \sum_{j \in \mathcal{J}_i^{r_a}} \frac{\boldsymbol{x}_j(t) - \boldsymbol{x}_i(t)}{|\boldsymbol{x}_j(t) - \boldsymbol{x}_i(t)|}. \tag{3.28}$$

If neighbours are found in only one of the two outer zones of interaction then the desired direction of movement is given by

$$\boldsymbol{u}_i(t+1) = \boldsymbol{u}_i^{o,a}(t+1), \tag{3.29}$$

depending in which zone neighbours are found [**15**]. If neighbours are found in both orientation and interaction zones then the desired direction of movement is an average of the two preferred directions,

$$\boldsymbol{u}_i(t+1) = \frac{1}{2}[\boldsymbol{u}_i^o(t+1) + \boldsymbol{u}_i^a(t+1)]. \tag{3.30}$$

If no neighbours are detected or social forces result in a zero desired direction vector then the particle continues with its original direction [**15**].

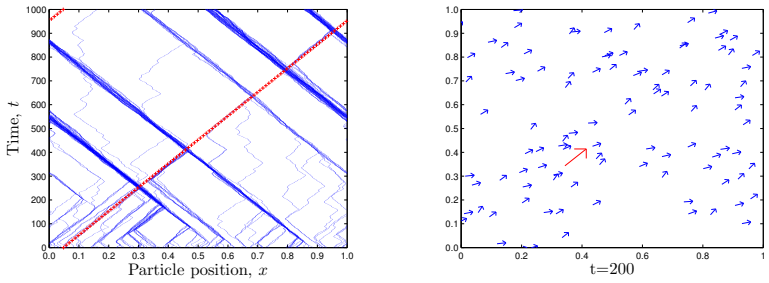
In order to simulate the effects of a noisy environment and the fact that decision making in animals is subject to stochastic effects, a random number drawn from a (spherically wrapped in three dimensions) Gaussian distribution with standard deviation σ is added to the desired direction vector. The particles simultaneously turn as far towards their desired direction vector as is permitted by their predefined 'maximum turning angle', $\phi\Delta t$, (where the value of ϕ is often motivated by experimentally observed turning rates) and move in this direction, $\boldsymbol{u}_i$, as in equation (3.3). A time-step independent reformulation of this model is possible, but technical, and hence we omit it.

Couzin *et al.* [**15**] showed this model to be capable of producing some complex and, at least qualitatively, biologically realistic behaviour. Four distinct behaviours were discovered to be inherent to the model in different parameter regimes. Two important parameters which seemed to correspond to alterations in the qualitative behaviour were variation in the radius of orientation, Δr_o, and variation in the radius of attraction, Δr_a. For low values of these variational parameters group cohesion was low and individuals tended to disperse more often than forming groups. For low Δr_o but high Δr_a a 'swarming' behaviour was exhibited with good group cohesion, but low alignment and low angular momentum about the group's centroid. For slightly larger values of Δr_o a 'torus' behaviour was exhibited in which angular momentum was high, but overall alignment was low as the individuals rotate perpetually around a hollow core. This behaviour is reminiscent of natural schools of fish [**57**]. A third behaviour type exhibited for intermediate values of Δr_o was the 'dynamic parallel' group, whose alignment was high but whose angular momentum was low. Individuals are easily able to interchange positions,

allowing the density and form of the group to fluctuate, presenting behaviour similar to that of flocks of birds or schools of fish. A fourth behavioural type is exhibited when Δr_o and Δr_a are both large. The group is highly aligned and interchanges of positions between individuals are infrequent leading to consistent group density and shape.

4 Models with leaders The concept of leadership in IBMs can be interpreted in a variety of different ways. Some models take it to be a strict hierarchical ordering of individuals, as in a military situation, where lower-ranked individuals take their lead from higher-ranked individuals who, in turn, follow the behaviour of even higher-ranked individuals. The ordering culminates with a single individual who is the *overall* leader. This type of leadership is more usually associated with active or intrinsic leadership exhibited by governmental and military hierarchies or social animals such as bees [**16, 61, 64**], for example, where leadership structures can clearly be identified and are known to all individuals. Another type of leadership is passive or transient leadership, such as that exhibited in disturbed bird flocks or traffic jams, when individuals tend to react mainly to near neighbours irrespective of their putative leadership qualities. For example, if a bird senses the approach of a predator, by taking flight or calling to other birds who may mimic this behaviour, the first bird becomes a *virtual* leader with a transient leadership structure developing rapidly as the call or movement spreads throughout the flock. There is nothing preordained or distinguishing about this virtual leader other than that it was the first to spot the potential predator. In dynamic, crowded situations the range over which an individual can sense is reduced and individuals may lose the capacity for individual recognition [**14, 42**]. In these situations the effectiveness of a hierarchical leadership system will be vastly reduced and passive leadership may become more appropriate.

4.1 Leadership in the Vicsek model Jadbabaie *et al.* [**41**] introduce the concept of a transient leader particle as an individual labelled with index 0 in a group of N other identical individuals. The leader differs from the others in that it has a constant heading, θ_0. The leader moves with the same speed as the other particles but ignores them completely, never wavering from its predefined heading. The other individuals in the model update their headings as before (see equation (3.1) for Vicsek's update rule and equation (3.9) for the rule which [**41**] use to prove convergence in their analysis) considering the leader particle as any other in their cohort. This leads to the revised heading update rule



(a) (b)

FIGURE 10: Panel (a) shows a typical realisation of particle positions over 1000 time-steps ($\Delta t = 1$) using the Czirók model in the presence of a single leader. The leader (red crosses) ignores all other particles and does not deviate from his chosen direction. The leader is capable of changing the direction of small groups of particles (see $t \approx 300$) or changing the direction of a small number of particles in larger well established groups (see $t \approx 750$) but will very rarely cause the direction of a large well established group to change by passing though it (see $t \approx 900$). Parameter values are as follows: number of individuals, $N = 100$; uniform size of interaction radius, $r = 0.005$; normalised domain length, $L = 1$; particle speed, $v_0 = 0.001$ and noise parameter, $\eta = 2$. Panel (b) shows a typical snapshot of the Vicsek model in the presence of a leader after 200 time-steps. The particles have had time to 'warm up' and are all roughly aligned with the direction of the leader (large arrow positioned at approximately $(0.4, 0.4)$). Parameter values are as follows: number of individuals, $N = 100$; uniform size of interaction radius, $r = 0.5$; normalised domain length, $L = 1$; particle speed, $v_0 = 0.01$ and noise parameter, $\eta = 1$.

for particles $i = 1, 2, \ldots, N$:

$$\theta_i(t+1) = \frac{1}{n_i(t) + 1 + b_i(t)} \left(\theta_i(t) + b_i(t)\theta_0 + \sum_{j \in \mathcal{J}_i^r} \theta_j(t) \right), \tag{4.1}$$

where $b_i(t)$ is defined to be one whenever the leader particle is in the neighbourhood of particle i and zero otherwise. Using similar notation as in Section 3.2 (where the leaderless Vicsek model was analysed) we can vectorise equation (4.1) (in analogy to equation (3.10) for the leaderless model) as follows:

$$\boldsymbol{\theta}(t+1) = (I + D_p + B_p)^{-1} \big[(I + A_p)\boldsymbol{\theta}(t) + B_p \mathbf{1}\theta_0 \big], \tag{4.2}$$

where p is the index of the graph $\tilde{G}_p \in \tilde{\mathbb{G}}$ (representing the neighbour relationships between all $N+1$ particles) drawn from a suitable set of indices, $\tilde{P}$, of all the possible graphs on $N+1$ vertices. B_p is the $N \times N$ diagonal matrix whose i^{th} diagonal element is one if node i is connected to the leader and zero otherwise. $\tilde{\mathbb{G}}$ is the set of all possible graphs on $N+1$ nodes.

The analysis carried out by Jadbabaie *et al.* [**41**] proceeds in a similar manner to the analysis in the leaderless case. However, in order to show that the alignments of the particles in the group tend towards the leader's alignment, it is necessary that each particle is *linked to the leader* sufficiently frequently rather than that the particles are simply *linked together* as was previously required for convergence to a common heading. Clearly, if the particles are to converge to a common heading they must be linked together, but we say the N agents are linked to the leader across a finite time interval $[t, \tau]$, (where, as before, $t < \tau$) if the collection of graphs $\{\tilde{G}_{p(t)}, \tilde{G}_{p(t+\Delta t)}, \ldots, \tilde{G}_{p(\tau)}\}$ encountered along the interval is jointly connected. In other words, the N nonleader agents are leader-linked across a time interval when the $(N+1)$-member group (including the leader) is linked together across the time interval. It should be noted that for the N-particle group to be leader-linked it does not necessarily have to be linked together and neither does the N-member group being linked together imply that the individuals are linked to the leader.

Jadbabaie *et al.* [**41**] use properties of stochasticity of the heading update matrix and sub-multiplicativity of the induced infinity norm on nonnegative square matrices, amongst other useful properties, as well as arguments similar to those used to prove the convergence to a common heading in the leaderless case, to prove the following remarkable result: if there exists an infinite sequence of contiguous, nonempty, bounded time intervals $[t_i, t_{i+1})$, $i \geq 0$, starting at $t_0 = 0$, with the property that across each such interval, the N-particle group of 'followers' is linked to the leader, then

$$\lim_{t \to \infty} \boldsymbol{\theta}(t) = \theta_0 \mathbf{1}, \tag{4.3}$$

where θ_0 is the constant heading of the leader particle (see Figure 10 (b) for an example of the particle alignments of the Vicsek model in the presence of a leader after a relatively short amount of time).

In order to prove this, Jadbabaie *et al.* [**41**] rewrite the system in terms of $\varepsilon(t) = \boldsymbol{\theta}(t) - \theta_0 \mathbf{1}$ so that the update equation (4.2) can be

restated as

$$\varepsilon(t+1) = F_p \varepsilon(t), \tag{4.4}$$

for

$$F_p = (I + D_p + B_p)^{-1}(I + A_p). \tag{4.5}$$

They are then able to show that the property of the group being leader-linked over each contiguous interval $[t_i, t_{i+1})$ implies that the matrix infinity norm of the product of the update matrices, $\|F_{p(t_{i+1})}F_{p(t_{i+1}-\Delta t)}\cdots F_{p(t_i)}\|$, encountered during that interval is bounded by a constant $\lambda < 1$. As such, it is possible to show that the product of infinitely many update matrices $F_p(t)$ satisfying the above conditions will tend to zero and hence $\varepsilon(t) \to 0$ proving convergence of all the 'follower' particles' headings to that of the leader.

4.2 Hierarchical leadership in the Cucker-Smale model Shen [**61**] studies the CS model in the context of hierarchical leadership (HL) and aims to prove similar convergence results about the flock given specific parameter values and possible restrictions on the initial conditions of the flock. A flock of size $N+1$ with adjacency matrix, $A = (a_{ij})$, (as previously defined) is said to be under HL if the agents can be labelled $0, \ldots, N$ such that the following two properties hold:

1. $a_{ij} \neq 0$ implies that agent i is led by agent j and this is only the case if $j < i$;
2. if the *leader set* of each agent i is defined as

$$\mathcal{L}(i) = \{j \,|\, a_{ij} > 0\}, \tag{4.6}$$

then for any $i > 0$ the leader set, $\mathcal{L}(i)$, is nonempty.

Condition 1 implies that the adjacency matrix of the flock is lower triangular. Shen [**61**] goes further than this by proving that a flock of size $N+1$ is under HL if and only if, after some ordered labelling, the adjacency matrix is lower triangular and for any row $i > 0$, there exists at least one positive off-diagonal entry. Figure 11 shows the leadership structures of four simple flocks. Two are HL flocks and two are non-HL flocks, each violating a different condition of the definition of hierarchical leadership.

The framework for the proof of convergence of the leaderless CS model Cucker and Smale [**19**] relies heavily on the existence of a special inner

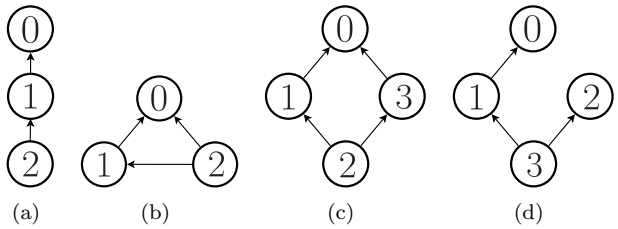


FIGURE 11: The arrows $i \to j$ indicate that agent i is led by agent j. (a) and (b) are examples of HL flocks, whereas (c) and (d) are non-HL flocks. Flock (c) violates condition 1 of the above definition of HL in that $a_{23} \neq 0$. Flock (d) violates condition 2 in that the leader set of agent 2 is empty.

product $\langle \cdot, \cdot \rangle_Q$, which in turn relies on the symmetry assumption of the CS model. Shen [**61**] proves that such inner products could fail to exist for nonsymmetric systems such as HL-flocks. Even when such inner products exist in nonsymmetric systems it is likely that they are time-dependent and therefore unsuitable for proving *a priori* convergence results.

Despite this general lack of compatible inner products for an HL flock, Shen [**61**] shows, under a set of conditions similar to those given by Cucker and Smale [**19**], for an HL flock of size $N+1$ with sufficiently small discrete time-step, Δt, between updates of position and velocity, that the velocities of the flock converge at rate $O(\rho_{\Delta t}^t t^{N-1})$ where $\rho_{\Delta t} \in (0,1)$ depends on Δt, system parameters and the initial configuration of the flock. The CS-type conditions are such that whether or not convergence of the system is dependent on initial conditions depends on the magnitude of the parameter β, where β is the exponent of the denominator of the interaction weighting function ζ (see equation (3.14)). Specifically, if $\beta < \beta_c$ then the convergence is independent of initial conditions, where instead of $\beta_c = 1/2$, as in the CS model, $\beta_c = 1/(2N)$. Shen [**61**] also proves a result for CS-type flocks under HL in continuous time; for $\beta < 1/2$ there exists some $B > 0$ such that the velocities of the flock converge at an exponential rate $O(e^{-Bt})$, where the constant B depends on the system parameters and initial conditions of the flock.

4.3 Inherent leadership Animals such as locusts, fish or birds often move in groups which are densely crowded. This limits the range over which individuals can sense each other and reduces their capacity for individual recognition. As such, it is often realistic, when modelling

biological individuals, to allow the alteration of an individual's behaviour based only on the behaviour of its neighbours and to implement 'leader' individuals with no external characteristics different from those of the 'followers'. The only difference between leader particles and followers should be in the preferential movement of the leaders in a predefined direction.

Couzin *et al.* [**16**] adapt their original model [**15**] in order to deal effectively with biologically motivated model systems incorporating informed individuals. In the new model the two outer interaction radii, corresponding to orientation and alignment zones, are incorporated into one radius and, as such, the contribution to the desired direction $\boldsymbol{u_i}$ from the individuals in this radius, providing there are no neighbours in the zone of repulsion, is the summation of the terms, $\boldsymbol{u}_i^o(t+\Delta t)$ and $\boldsymbol{u}_i^a(t+\Delta t)$, respectively, relating to attraction and orientation (see Section 3.6). In either case the resultant direction vector, $\boldsymbol{u}_i(t+\Delta t)$, from interactions is subsequently normalised, $\hat{\boldsymbol{u}}_i(t+\Delta t) = \boldsymbol{u}_i(t+\Delta t)/|\boldsymbol{u}_i(t+\Delta t)|$. In a further alteration to the original model no 'blind zone' is considered.

The influence of informed individuals is naturally incorporated into the model by further altering the desired direction vector of a predefined proportion of the individuals, p, so that they tend to follow a preferred direction, $\boldsymbol{g}_i$, for individual i. The degree to which the individuals adhere to this direction as opposed to simply following the desired direction of movement $\hat{\boldsymbol{u}}_i(t+\Delta t)$ is altered using the weighting parameter, ω:

$$\boldsymbol{u}_i'(t+\Delta t) = \frac{\hat{\boldsymbol{u}}_i(t+\Delta t) + \omega \boldsymbol{g}_i}{|\hat{\boldsymbol{u}}_i(t+\Delta t) + \omega \boldsymbol{g}_i|}, \tag{4.7}$$

where $\boldsymbol{u}_i'(t+\Delta t)$ becomes the desired direction vector of the informed particle.

Couzin *et al.* [**16**] measured how well the group followed the preferred direction of migration of the informed individuals. They found that the accuracy increased asymptotically as the proportion of individuals increased, and that for larger groups a smaller proportion of individuals was needed to guide the group to the same level of accuracy.

The authors simulate the possibility that there exists more than one group of informed individuals with differing preferential directions. They found that, when the numbers of informed individuals in the two groups were equal, an average direction is taken by the group until the angle between the preferential directions is increased beyond a certain limit, after which the group randomly selected one of the two preferred directions with equal probability. When the size of one of the informed sets

is even marginally larger than that of the other set, the group selects the preferred direction of the larger informed set (provided there existed an appropriate difference in the preferences).

Positive feedback on the weighting parameter, ω, which causes it to increase for those individuals following their preferred direction to a good degree of accuracy and decrease otherwise, is introduced. This reduces the necessary difference in preferred direction of the two groups above which the average direction is abandoned and a consensus decision on which of the two preferred directions to follow is reached. Using feedback on the weighting parameter, Couzin *et al.* [**16**] were able to tune the necessary difference in preferred direction for a consensus decision to be made. This feedback mechanism also makes it possible for a group to discriminate between the quality of information of two sets of informed individuals. Individuals in the set with (even slightly) poorer quality of information are less likely to head in their preferred direction and hence the weighting given to their preferred direction will be decreased as part of the feedback loop, until the set of poorly informed individuals give no weight to their preferred direction and are in effect uninformed. A consensus decision is easily reached in this case.

5 Evolutionary self-propelled particle models The evolution of collective behaviour is becoming an increasingly important area of research. Several advantages and disadvantages have been postulated and argued for qualitatively as the evolutionary basis of group formation. In order for collective behaviour to evolve the advantages to group formation must outweigh the disadvantages. Group behaviour often provides advantages that individuals could not hope to achieve acting alone: it can often mean the difference between life and death for the individuals involved. For example, warm-blooded animals such as emperor penguins huddle together in groups sharing body heat and reducing the need for energy expenditure [**59**]. On a less dramatic scale, schooling fish and flocking birds are thought to gain a dynamic or hydrodynamic advantage by moving together in a group [**4, 7, 70**]. Their collective motion reduces the average resistance force on each animal which therefore leads to a smaller expenditure of energy for each individual. By swarming, groups of animals, such as locusts, act as a multicellular organism, increasing their collective zone of awareness in comparison to that of an individual allowing them to forage more efficiently.

The disadvantages of group formation can, however, be severe. Disease spreads quickly in crowded environments, often debilitating every

member of the group. Despite the foraging ability of the group increasing, competition for the available resources also increases meaning that some individuals may, in fact, find it harder to obtain enough food [**44**].

Even the validity of these simple pros and cons have been much debated [**1, 53**], even more so their importance in an evolutionary scenario. However, since groups are evident in many animal species, it is clear that, whatever the advantages of the formation of these groups are, they outweigh the disadvantages.

In recent years it has become easier to conduct increasingly complex computational evolutionary simulations on relatively modest hardware. This has led to an increase in algorithms designed to simulate various evolutionary scenarios testing quantitatively the possible hypotheses on the advantages and disadvantages of group formation.

Spector *et al.* [**63**] describe numerical experiments on the emergence of collective behaviour in populations of flying agents incorporating multiple species and goal orientation as evolutionary factors. They find the emergence of collective behaviour akin to that of multicellular organisms. In a second set of experiments the behaviour of the agents was observed to exhibit altruistic food-sharing behaviours.

Wood and Ackland [**71**] introduced the first multi-parameter SPP evolution algorithms. The model they consider is based on a two-dimensional version of the models described by Couzin *et al.* [**15**], given in Section 3.6. The parameters of the flock are allowed to evolve via a caricature of natural selection, with each individual in the new generation having different characteristics to each other, but sharing similar characteristics to a single haploid parent in the previous generation (i.e., asexual reproduction with mutation). Wood and Ackland's model imposes two additional restrictions on the parameter regimes over the Couzin model. These restrictions ensure that all possible parameter values are not simply maximised. Indeed parameters such as speed and maximum allowed turning angle are traded off against each other. Sighting area parameter, $A_s = \psi r_a^2$ (where ψ, as before, is the viewing angle) is kept constant to simulate a trade-off between long-range and peripheral vision. Similarly the possible movement 'area' parameter, $A_m = 2\theta \boldsymbol{v}_i^2$, is kept constant to simulate faster moving individuals being less free to change their direction[5].

In addition to the Couzin model, Wood and Ackland [**71**] introduce stimuli such as a food source or a predator. The response to these stimuli in each generation of individuals is a basis for the fitness upon which

[5]We suggest that, in order to keep the model time-step independent, the movement 'area' be formulated with the units of area: $A_m = 2\theta(\boldsymbol{v}_i \Delta t)^2$.

the individuals are selected to pass on their characteristics to the next generation. Therefore, in addition to the desired direction of movement due to social interactions, $\boldsymbol{u}_i$, a stimulus-related movement vector, $\Omega_i \hat{\boldsymbol{v}}$, is added into an individual's velocity vector, where the script letters f and p relate to stimuli in response to food or predation and $\hat{\boldsymbol{v}}$ is the unit vector towards food or away from the predator, respectively. Ω_i is an evolvable weighting parameter which determines to what degree individuals attempt to attain the goals of predator avoidance or food retrieval. Other evolvable parameters of the model are the orientation radius, r_o, such that $r_r < r_o < r_a$ and the magnitude of the noise for each particle, η_i.

In the terminology of [**71**] one *run* of the algorithm is the simulation of a predefined number of generations or *instances*. Between each instance a new generation of individuals is generated. The new generation is largely dependent on the previous generation, with the fitter individuals of the previous generation having a higher probability of passing on their characteristics to the next generation. Mutation is also implemented by adding a random number, chosen from a uniform distribution in the range $[-m, m]$, to the magnitude of the interaction radii of each particle in the new generation. Here m is a fixed value and precautions are taken to ensure parameters do not take unrealistic values (i.e., negative interaction radii, etc.).

After evolving their individuals for sufficiently many generations, Wood and Ackland [**71**] found that very different flocks evolved given different types of stimuli. The predation stimulus necessitated group formation, whereas food searching did not, although group formation was possible. In almost all cases a small viewing angle was strongly selected for.

Two sub-types of behaviour emerged from the predation model: slow moving, milling behaviour and fast moving, dynamic groups. The milling behaviour allows individuals to interact strongly with each other and to respond rapidly to the presence of a predator. These types of groups evolve when the mutation rate is high and the viewing angle of the predator is high (a smaller more compact group does not allow the predator to take advantage of his large viewing angle). Alternatively, with low mutation rates and predators with a small viewing angle, fast moving dynamic groups evolve. Their response to a predator is to fan out and take advantage of the predator's low viewing angle.

In flocks selected for their ability to forage, group evolution was highly dependent on the type of food stimulus on offer. Foraging seemed also to promote strong selection on the speed and the attraction radius (and

hence demote turning efficiency and viewing angle), but not necessarily on the orientation radius. The distribution of single food packets did not lead to group formation. The first individual to arrive ate the whole packet and there was no advantage to group foraging. However, when multiple food packets were distributed, group behaviour became more efficient as the group was more likely to find food than a lone individual with a small viewing angle. Several of the behaviours to which the system evolved in the models of Wood and Ackland [**71**] are reminiscent of those found in nature [**5, 40, 43**].

5.1 Evolution of interaction radii for particles following a leader In groups of birds, fish and insects crowding is thought to restrict an animal's capacity for individual recognition as well as its field of vision and other external signal sensing abilities [**14, 42**]. Hence an individual's foraging efficiency and ability to avoid predation are likely to depend on its ability to follow an individual informed about the location of food or the appearance of a predator. Here we combine the ideas of intrinsic leadership and evolution in order to evolve a single characteristic of the individuals in Czirók-type (one dimension) and Vicsek-type (two dimensions) models. Instead of basing the relative fitness of a particle on its ability to forage or avoid predation (which, in a crowded environment, is often a random process depending on the position of the food source, appearance of the predator and the position of the individual in the group) we focus on the ability of individuals to follow an informed leader particle. We assume that a leader particle has seen a stimulus and, ignoring the influence of its neighbours, heads directly towards the food source or away from the predator. This is equivalent to the implementation of informed individuals by Jadbabaie *et al.* [**41**] or by Couzin *et al.* [**16**] where the weighting parameter, ω, has been allowed to tend to infinity (see equation (4.7)).

We largely follow the evolution algorithm of Wood and Ackland [**71**], with a few subtle exceptions. Since we now base fitness of the individuals on their ability to follow an unmarked, but informed individual, we must alter the way in which the fitness is calculated. We find an auxiliary vector, $\boldsymbol{e}$, such that

$$e_i = |X_i - X_0| \quad \text{for } i = 1, \ldots, N, \tag{5.1}$$

where X_i represents the distance travelled by the i^{th} particle over the period ($j = J_0, \ldots, T_{final}$) and, as before, 0 is the index of the leader

particle. In one dimension this can be written simply as

$$e_i = \left| \sum_{j=J_0}^{T_{final}} (v_i(t_j) - v_0(t_j)) \right|. \tag{5.2}$$

In two dimensions the formulation is slightly more complicated:

$$e_i = \left| \sum_{j=J_0}^{T_{final}} (\exp(\mathrm{i}\theta_i(t_j)) - \exp(\mathrm{i}\theta_0(t_j))) \right|. \tag{5.3}$$

We use $\boldsymbol{e}$ to calculate the fitness vector, $\boldsymbol{f}$,

$$f_i = \max_{j=1,\ldots,N}\{e_j\} - e_i \quad \text{for } i = 1, \ldots, N. \tag{5.4}$$

We also normalise the fitness vector to find the probability vector, $\boldsymbol{p}$, that each individual will pass on its characteristics to an individual in the next generation in a slightly different manner to Wood and Ackland [**71**]:

$$\boldsymbol{p} = \frac{1 - \kappa + \kappa \boldsymbol{f}}{N(1-\kappa) + \kappa \sum_{i=1}^{N} f_i}, \tag{5.5}$$

where $\kappa \in [0, 1]$ represents the weighting given to the fitness. Clearly $\kappa = 1$ represents strong selection and $\kappa \ll 1$ represents weak selection. Here we choose $\kappa = 1$ as we have found that weaker selection gives similar results, but requires more generations to achieve a quasi-steady-state distribution.

We choose to evolve a single characteristic of the individuals between generations so as to see clearly the effect of evolution based on fitnesses corresponding to an individual's ability to follow the leader in the absence of variation of other characteristics. We are interested in the evolution of the interaction radius. Between generations we assess the fitness of each individual and pass the interaction radius of individual i to each member of the new generation with probability p_i, for $i = 1, \ldots, N$. Once individual i, in the new generation, is assigned a value, r_i, for its interaction radius (from those in the previous generation) a random

number, ξ_i, drawn from a uniform distribution in $[-m, m]$, is used to mutate the interaction radius as follows:

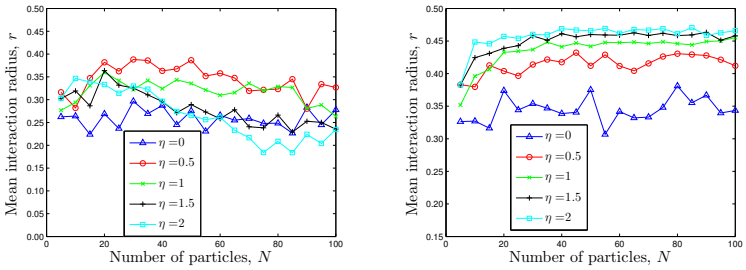
$$r_i = \begin{cases} 0, & \text{for } r_i + \xi_i < 0; \\ 0.5, & \text{for } r_i + \xi_i > 0.5; \\ r_i + \xi_i, & \text{otherwise,} \end{cases}$$

which ensures the interaction radius does not become unrealistically large or small.

In order to implement this evolutionary method all particles are initialised with random positions, velocities and interaction radii (between 0 and 0.5) and are allowed to move for T_{final} time-steps. This is known as one 'run'. At the end of this run, fitnesses are calculated and a new generation of individuals evolved. This evolutionary process occurs for M generations. This is known as one 'simulation'. We repeat the simulations K times and take an average of the final evolved radii.

Figure 12 shows the results of implementing the above evolution algorithm in order to evolve the interaction radii for the one-dimensional Czirók and the two-dimensional Vicsek models. In Figure 12 (a) a larger noise value tends to lead to the evolution of a smaller interaction radius, especially for large numbers of particles. The opposite is seen in Figure 12 (b). This suggests that, in two dimensions, increased noise necessitates a larger interaction radius to be evolved, since a set of particles with smaller interaction radii will not be able to form as coherent a group as a set with larger interaction radii and will hence be impeded when following the leader. The fact that a smaller interaction radius is evolved in low noise conditions hints that there is a cost to evolving a large interaction radius. It could be that, in a group of particles with only moderately large or very large interaction radii, all individuals benefit from forming a group heading in the direction of the leader, but individuals with only moderately large interaction radii see fewer non-leader particles in their interaction zones and hence follow the leader to a greater degree when the leader is in their interaction zone, allowing them to gain an advantage. This advantage is only made possible by low noise conditions. In high noise conditions, unless all particles have large interaction radii, group formation may be affected and those particles with very large interaction radii may gain an advantage.

The evolutionary algorithm we have implemented above is far from perfect. Facets such as periodic boundary conditions, homogeneous square domains and a single interaction radius make the models themselves unrealistic. We therefore hesitate to draw strong conclusions from



(a) (b)

FIGURE 12: Panel (a) exhibits the results of evolving a set of N SPPs in one dimension in the presence of an informed individual using the Czirók model and the evolution algorithm described above for varying values of the noise parameter η. Panel (b) exhibits similar evolution results for a set of N SPPs which move according to the Vicsek model in two dimensions. For both models the number of particles, excluding the leader particle, ranges from $N = 5, 10, \ldots, 100$ and the noise parameter takes the values $\eta = 0, 0.5, \ldots, 2$. Evolutionary parameter values were as follows: number of time-steps in one run, $T_{final} = 1000$, number of generations per simulation, $M = 100$, number of repeats per simulation, $K = 50$, and mutation rate, $m = 0.005$. The parameters for each run were, $L = 1$, $\Delta t = 1$ and $v_0 = 0.1$ for both the Czirók and Vicsek models. In both cases we have allowed the particles to 'warm up' and calculated fitnesses based on how well the particles followed the leader in the second half of the simulation (i.e., $J_0 = 501$ in equations (5.2) and (5.3)).

these results, but note that the degree to which particles follow a leader is a good candidate for fitness in evolutionary SPP models.

6 Discussion The continuing study of emergent collective behaviour in particle models will undoubtedly have a significant impact on our understanding of biological systems on many levels: from symbiosis in microbes, through swarming and flocking behaviours in insects and larger animals to the organisation of human societies. Increasing numbers of processes discovered in microbiology require IBMs in order to represent them realistically. Biofilm formation and degradation [**45, 46**], wound healing [**9, 62**], tumour growth [**23, 33**], and other important cell migration problems observed in a variety of areas of developmental biology [**3, 50**], are amenable to particle-based modelling.

On a smaller scale, IBMs are also being used in a subcellular context to model protein organisation [**28**], molecular motility and signal transduction reactions [**65**] amongst other processes [**67**].

Possible application areas of the IBMs described in this work are not solely restricted to biological systems. They encompass the robotics of multiagent systems [**48**], language evolution [**20**] as well as more unforeseen circumstances such as space-flight control for the European space agency mission DARWIN [**55**] and the modelling of traffic flow in both vehicles and humans [**36, 37**].

In the past, multi-particle systems exhibiting collective behaviour have been modelled almost exclusively using partial differential equations, deriving mean field descriptions on a macroscopic scale. These coarse, whole-population representations of multi-agent systems naturally engender a loss of information, specifically about the individual agents which make up the model. This has made evolutionary models of multi-agent systems and models incorporating individual leadership difficult to formulate. With the steady increase in computational power has come the ability to simulate complex multi-agent systems on even modest desktop machines and with it the rise in verifiable IBMs for emergent, collective behaviour. Such models have several distinct advantages over population-based model: (i) they are bottom-up approaches which describe the whole system by establishing procedural rules which individuals obey. These rules are often drawn from observations of individuals, rather than inferring individual behaviour from the mean-field model which best describes the behaviour of the whole group, allowing more realistic modelling assumptions; (ii) they allow for the introduction of individual variability (leaders for example) and individual-scale randomness which enables models to mimic the variability of real systems; and (iii) they allow for the independent evolution of individuals to their environment which implies that group evolution arises naturally from the dynamics governing the evolution of the individuals. There are, of course, several disadvantages inherent to multi-agent models: (i) they may lack the clarity of continuum models and easily become extremely complicated, with large numbers of free parameters, making their analysis and importantly their communication difficult; (ii) with increased complexity and variability IBMs quickly erode the advances in computational power made in recent years.

Although in this work we have barely scratched the surface of the huge variety of particulate models, we have demonstrated that there exists a wide variety of models which are applicable in many diverse situations. There is work to be done on unifying these IBMs with previously exist-

ing and novel mean field models in order to create a framework which combines the advantages of the two different types of models. Of course not all models will have a corresponding mean field description in all cases but the advantage of finding such a connection where it exists will be tangible.

Acknowledgements This work was partially supported by the Engineering and Physical Sciences Research Council via the Systems Biology Doctoral Training Centre, University of Oxford and St. Catherine's College, University of Oxford (C.A.Y.). R.E.B. would like to thank RCUK for a fellowship in Math. Biol. and St Hugh's College, University of Oxford for a Junior Research Fellowship. This publication is partially based on work supported by Award No. KUK-C1-013-04, made by King Abdullah University of Science and Technology (KAUST). R.E. would also like to thank Somerville College, University of Oxford for Fulford Junior Research Fellowship. The research leading to these results has received funding from the European Research Council under the *European Community's* Seventh Framework Programme (*FP7/2007–2013*)/ERC *grant agreement* No. 239870. P.K.M. was partially supported by a Royal Society-Wolfson Research Merit Award.

REFERENCES

1. M. V. Abrahams and P. W. Colgan, *Fish schools and their hydrodynamic function: a reanalysis*, Environ. Biol. Fish. **20**(1) (1987), 79–80.
2. I. Aoki, *A simulation study on the schooling mechanism in fish*, Bulletin of the Japanese Society of Scientific Fisheries (Japan), (1982).
3. R. E. Baker, C. A. Yates and R. Erban, *From microscopic to macroscopic descriptions of cell migration on growing domains*, Bull. Math. Biol. **72**(3) (2010), 719–762.
4. V. V. Belyayev and G. V. Zuyev, *Hydrodynamic hypothesis of school formation in fishes*, Problems of Ichthyology **9** (1969), 578–584.
5. C. W. Benkman, *Flock size, food dispersion, and the feeding behavior of crossbills*, Behav. Ecol. Sociobiol. **23**(3) (1988), 167–175.
6. C. M. Breder, *Equations descriptive of fish schools and other animal aggregations*, Ecology, **35**(3) (1954), 361–370.
7. C. M. Breder, *Vortices and fish schools*, Zoologica.-New York, **50** (1965), 97–114.
8. J. Buhl, D. J. T. Sumpter, I. D. Couzin, J. Hale, E. Despland, E. Miller and S. J. Simpson, *From disorder to order in marching locusts*, Science **312**(5778) (2006), 1402.
9. A. Q. Cai, K. A. Landman and B. D. Hughes, *Multi-scale modeling of a wound-healing cell migration assay*, J. Theor. Biol. **245**(3) (2007), 576–594.

10. J. A. Carrillo, M. Fornasier, J. Rosado and G. Toscani, *Asymptotic flocking dynamics for the kinetic Cucker-Smale model*, (2009).
11. H. Chaté, F. Ginelli and R. Montagne, *Simple model for active nematics: Quasi-long-range order and giant fluctuations*, Phys. Rev. Lett. **96**(18) (2006), 180602.
12. H. Chaté, F. Ginelli, G. Grégoire, F. Peruani and F. Raynaud, *Modeling collective motion: variations on the Vicsek model*, Eur. Phys. J. B **64**(3-4) (2008), 451–456.
13. P. J. Conder, *Individual distance*, Ibis. **91**(4) (1949), 649–655.
14. I. D. Couzin and J. Krause, *Self-organization and collective behavior in vertebrates*, Adv. Stud. Behav. **32** (2003), 1–75.
15. 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), 1–11.
16. I. D. Couzin, J. Krause, N. R. Franks and S. A. Levin, *Effective leadership and decision-making in animal groups on the move*, Nature **433**(7025) (2005), 513–516.
17. J. H. Crook, *The basis of flock organization in birds*, in: Current Problems in Animals Behaviour, Cambridge University Press, Cambridge (1961), 125–149.
18. F. Cucker and E. Mordecki, *Flocking in noisy environments*, J. Math. Pure. Appl. (2007).
19. F. Cucker and S. Smale. *Emergent behavior in flocks*, IEEE. T. Automat. Contr. **52**(5) (2007), 852–862.
20. F. Cucker, S. Smale and D. X. Zhou, *Modeling language evolution*, Found. Comput. Math. **4**(3) (2004), 315–343.
21. A. Czirók and T. Vicsek, *Statistical mechanics of biocomplexity*, **527**, Lecture Notes in Physics, Ch. 29, 152–164, Collective Motion, Springer, Berlin/Heidelberg, 1999.
22. A. Czirók, A. L. Barabási and T. Vicsek, *Collective motion of self-propelled particles: kinetic phase transition in one dimension*, Phys. Rev. Lett. **82**(1) (1999), 209–212.
23. D. Drasdo and S. Hohme, *Individual-based approaches to birth and death in avascu1ar tumors*, Math. Comput. Modelling **37**(11) (2003), 1163–1176.
24. R. Eftimie, G. de Vries and M. A. Lewis, *Complex spatial group patterns result from different animal communication mechanisms*, P. Natl. Acad. Sci. USA **104**(17) (2007), 6974.
25. R. Eftimie, G. de Vries, M. A. Lewis and F. Lutscher, *Modeling group formation and activity patterns in self-organizing collectives of individuals*, Bull. Math. Biol. **69**(5) (2007), 1537–1565.
26. J. T. Emlen, *Flocking behavior in birds*, The Auk **69**(2) (1952), 160–170.
27. R. Erban, I. G. Kevrekidis, D. Adalsteinsson and T. C. Elston, *Gene regulatory networks: a coarse-grained, equation-free approach to multiscale computation*, J. Chem. Phys. **124**(8) 084106 (2006), 1–17.
28. M. J. Fisher, R. C. Paton and K. Matsuno, *Intracellular signalling proteins as 'smart' agents in parallel distributed processes*, Biosystems **50**(3) (1999), 159–171.
29. V. Gazi and K. Passino, *Stability analysis of swarms*, IEEE. T. Automat. Contr. **48**(4) (2003), 692–697.
30. V. Gazi and K. M. Passino, *Stability analysis of swarms in an environment with an attractant/repellent profile*, Proc. Amer. Cont. Conf. **3** (2002), 1819–1824.
31. V. Gazi and K. M. Passino, *A class of attractions/repulsion functions for stable swarm aggregations*, Int. J. Control. **77**(18) (2004), 1567–1579.
32. V. Gazi and K. M. Passino, *Stability analysis of social foraging swarms*, IEEE. T. Syst. Man. Cy. B **34**(1) (2004), 539–557.

33. P. Gómez-Mourelo, E. Sánchez, L. Casasús and G. F. Webb, *A fully continuous individual-based model of tumor cell evolution*, Comptes rendus-Biologies **331**(11) (2008), 823–836.
34. G. Grégoire and H. Chaté, *Onset of collective and cohesive motion*, Phys. Rev. Lett. **92**(2) (2004), 025702.
35. S. Y. Ha and E. Tadmor, *From particle to kinetic and hydrodynamic descriptions of flocking*, Kinetic Related Models **1**(3) (2008), 415–435.
36. D. Helbing, *Traffic and related self-driven many-particle systems*, Rev. Mod. Phys. **73**(4) (2001), 1067–1141.
37. D. Helbing and B. Tilch, *Generalized force model of traffic dynamics*, Phys. Rev. E **58**(1) (1998), 133–138.
38. D. Helbing, A. Johansson and H. Z. Al-Abideen, *Dynamics of crowd disasters: An empirical study*, Phys. Rev. E **75**(4) (2007), 46109.
39. A. Huth and C. Wissel, *The simulation of the movement of fish schools*, J. Theor. Biol. **156**(3) (1992), 365–385.
40. Y. Inada and K. Kawachi, *Order and flexibility in the motion of fish schools*, J. Theor. Biol. **214**(3) (2002), 371–387.
41. A. Jadbabaie, J. Lin and A. S. Morse, *Coordination of groups of mobile autonomous agents using nearest neighbor rules*, IEEE T. Automat. Contr. **48**(6) (2003), 988–1001.
42. J. Krause and G. D. Ruxton, *Living in Groups*, Oxford University Press, USA, 2002.
43. C. J. Krebs, *Ecology: The Experimental Analysis of Distribution and Abundance*, HarperCollins, 1978.
44. J. R. Krebs, N. B. Davies and J. Parr, *An Introduction to Behavioural Ecology*, Wiley-Blackwell, 1993.
45. J. U. Kreft, G. Booth and J. W. T. Wimpenny, *BacSim, a simulator for individual-based modelling of bacterial colony growth*, Microbiology, **144**(12) (1998), 3275.
46. J. U. Kreft, C. Picioreanu, J. W. T. Wimpenny and M. van Loosdrecht, *Individual-based modelling of biofilms*, Microbiology, **147**(11) (2001), 2897.
47. H. Levine, W. J. Rappel and I. Cohen, *Self-organization in systems of self-propelled particles*, Phys. Rev. E **63**(1) (2000), 017101, 1–4.
48. M. J. Mataric, *Issues and approaches in the design of collective autonomous agents*, Robot. Auton. Syst. **16**(2) (1995), 321–332.
49. A. Mogilner, L. Edelstein-Keshet, L. Bent and A. Spiros, *Mutual interactions, potentials, and individual distance in a social aggregation*, J. Math. Biol. **47**(4) (2003), 353–389.
50. H. G. Othmer and A. Stevens, *Aggregation, blowup, and collapse: The ABC's of taxis in reinforced random walks*, SIAM J. Appl. Math. **57**(4) (1997), 1044–1081.
51. H. G. Othmer, S. R. Dunbar and W. Alt, *Models of dispersal in biological systems*, J. Math. Biol. **26** (1988), 263–298.
52. B. L. Partridge, *The structure and function of fish schools*, Sci. Amer. **246**(6) (1982), 114–123.
53. B. L. Partridge and T. J. Pitcher, *Evidence against a hydrodynamic function for fish schools*, Nature **279** (1979), 418–419.
54. B. L. Partridge and T. J. Pitcher, *The sensory basis of fish schools: relative roles of lateral line and vision*, J. Comp. Physiol. A **135**(4) (1980), 315–325.
55. L. Perea, G. Gómez and P. Elosegui, *Extension of the Cucker-Smale control law to space flight formations*, J. Guid. Control. Dynam. **32**(2) (2009).
56. F. Peruani, A. Deutsch and M. Bär, *Nonequilibrium clustering of self-propelled rods*, Phys. Rev. E **74**(3) 30904, (2006).
57. T. J. Pitcher. *Behaviour of Teleost Fishes*, Springer, 1993.

58. A. Y. Povzner. *On the Boltzmann equation in the kinetic theory of gases*, Matematicheskii Sbornik **100**(1) (1962), 65–86.
59. J. Prévost, *Écologie du manchot empereur Aptenodytes forsteri Gray*, 1961.
60. C. W. Reynolds, *Flocks, Herds, and Schools: A Distributed Behavioral Model (Proceedings of SIGGRAPH'87)*, Comput. Graphics.-Us. **21**(4) (1987), 25–34.
61. J. Shen, *Cucker-Smale flocking under hierarchical leadership*, SIAM J. Appl. Math. **68**(3) (2008), 694–719.
62. M. J. Simpson, K. A. Landman and B. D. Hughes, *Distinguishing between directed and undirected cell motility within an invading cell population*, Bull. Math. Biol. **71**(4) (2009), 781–799.
63. L. Spector, J. Klein, C. Perry and M. Feinstein, *Emergence of collective behavior in evolving populations of flying agents*, GPEM **6**(1) (2005), 111–125.
64. D. J. T. Sumpter, *The principles of collective animal behaviour*, Philosophical Transactions of the Royal Society B: Biological Sciences **361**(1465) (2006), 5.
65. K. Takahashi, S. N. V. Arjunan and M. Tomita, *Space in systems biology of signaling pathways–towards intracellular molecular crowding in silico*, Febs. Lett. **579**(8) (2005), 1783–1788.
66. J. H. Tien, S. A. Levin and D. I. Rubenstein, *Dynamics of fish shoals: identifying key decision rules*, Evol. Ecol. Res. **6**(4) (2004), 555–565.
67. D. P. Tolle and N. Le Novere, *Particle-based stochastic simulation in systems biology*, Current Bioinformatics **1**(3) (2006), 315–320.
68. A. Turner and A. Penn, *Encoding natural movement as an agent-based system: an investigation into human pedestrian behaviour in the built environment*, Environ. Plann. B **29**(4) (2002), 473–490.
69. T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen and O. Shochet, *Novel type of phase transition in a system of self-driven particles*, Phys. Rev. Lett. **75**(6) (1995), 1226–1229.
70. D. Weihs, *Hydromechanics of fish schooling*, Nature (1973).
71. A. J. Wood and G. J. Ackland, *Evolving the selfish herd: emergence of distinct aggregating strategies in an individual-based model*, P. Roy. Soc. B-Biol. Sci. **274**(1618) (2007), 1637.
72. C. A. Yates, R. Erban, C. Escudero, I. D. Couzin, J. Buhl, I. G. Kevrekidis, P. K. Maini and D. J. T. Sumpter, *Inherent noise can facilitate coherence in collective swarm motion*, P. Natl. Acad. Sci. USA **106**(14) (2009), 5464–5469.

Corresponding author: C. A. Yates
Centre for Mathematical Biology, Mathematical Institute,
University of Oxford, 24-29 St Giles',
Oxford, OX1 3LB, UK.
E-mail address: yatesc@maths.ox.ac.uk

Centre for Mathematical Biology, Mathematical Institute,
University of Oxford, 24-29 St Giles',
Oxford, OX1 3LB, UK.

Oxford Centre for Collaborative Applied Mathematics,
Mathematical Institute,
University of Oxford, 24-29 St Giles',
Oxford, OX1 3LB, UK.

Centre for Mathematical Biology, Mathematical Institute,
University of Oxford, 24-29 St Giles',
Oxford, OX1 3LB, UK.
and
Oxford Centre for Integrative Systems Biology,
Department of Biochemistry, University of Oxford,
South Parks Road, Oxford, OX1 3QU.