Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

doi:10.3934/xx.xx.xx.xx

pp. **X–XX**

# NON-LOCAL KINETIC AND MACROSCOPIC MODELS FOR SELF-ORGANISED ANIMAL AGGREGATIONS

José A. Carrillo

Department of Mathematics, Imperial College London
London SW7 2AZ, United Kingdom

Raluca Eftimie

Division of Mathematics, University of Dundee
Dundee, DD1 4HN, United Kingdom

Franca Hoffmann

University of Cambridge, Centre for Mathematical Sciences
Wilberforce Road, Cambridge CB3 0WA, United Kingdom

(Communicated by the associate editor name)

Abstract. The last two decades have seen a surge in kinetic and macroscopic models derived to investigate the multi-scale aspects of self-organised biological aggregations. Because the individual-level details incorporated into the kinetic models (e.g., individual speeds and turning rates) make them somewhat difficult to investigate, one is interested in transforming these models into simpler macroscopic models, by using various scaling techniques that are imposed by the biological assumptions of the models. However, not many studies investigate how the dynamics of the initial models are preserved via these scalings. Here, we consider two scaling approaches (parabolic and grazing collision limits) that can be used to reduce a class of non-local 1D and 2D models for biological aggregations to simpler models existent in the literature. Then, we investigate how some of the spatio-temporal patterns exhibited by the original kinetic models are preserved via these scalings. To this end, we focus on the parabolic scaling for non-local 1D models and apply asymptotic preserving numerical methods, which allow us to analyse changes in the patterns as the scaling coefficient $\epsilon$ is varied from $\epsilon = 1$ (for 1D transport models) to $\epsilon = 0$ (for 1D parabolic models). We show that some patterns (describing stationary aggregations) are preserved in the limit $\epsilon \to 0$, while other patterns (describing moving aggregations) are lost. To understand the loss of these patterns, we construct bifurcation diagrams.

1. **Introduction.** Over the past 10-20 years a multitude of kinetic and macroscopic models have been introduced to investigate the formation and movement of various biological aggregations: from cells [5, 1] and bacteria [56] to flocks of birds, schools of fish and even human aggregations (see, for example, [60, 18, 55, 19, 29, 8, 25] and the references therein). The use of kinetic or macroscopic approaches is generally dictated by the problem under investigation: (i) kinetic (transport) models focus on

2010 *Mathematics Subject Classification.* Primary: 35L40, 35L65, 35K55, 92B99; Secondary: 82C40, 65M06,

*Key words and phrases.* self-organised aggregations; kinetic models; macroscopic models; non-local interactions; asymptotic preserving methods.

changes in the density distribution of individuals that have a certain spatial position, speed and movement direction (or are in some activity state [7]); (ii) macroscopic models focus on changes in the averaged total density of individuals [20, 31].

Generally, these kinetic and macroscopic models assume that individuals, particles, or cells can organise themselves in the absence of a leader. The factors that lead to the formation of self-organised aggregations are the interactions among individuals as a result of various social forces: repulsion from nearby neighbours, attraction to far-away neighbours (or to roosting areas [23]) and alignment/orientation with neighbours positioned at intermediate distances. These interaction forces are usually assumed to act on different spatial ranges, depending on the communication mechanisms used by individuals; e.g., via acoustic long-range signals, or via chemical/visual short-range signals. The non-locality of the attractive and alignment/orientation interactions is supported by radar tracking observations of flocks of migratory birds, which can move with the same speed and in the same direction despite the fact that individuals are 200-300 meters apart from each other [48]. For the repulsive forces some models consider non-local effects generated by decaying interactions with neighbours positioned further and further away [32], while other models consider only local effects [59]. In the case of continuous mesoscopic and macroscopic models, the non-local interactions are modelled by interaction kernels (see Figure 1 for 2D and 1D kernels). The most common choices for these kernels are Morse potential-type kernels [20, 18, 19, 22] (see Figure 1(b)) and Gaussian kernels [33, 32, 31, 49] (see Figure 1(c)).

Due to their complex structure, kinetic models are difficult to investigate. Although progress has been made in recent years, mainly regarding the existence and stability of various types of solutions and the analytic asymptotic methods that allow transitions from kinetic (mesoscopic) to macroscopic models (see, for example, [44, 54, 6, 18, 17, 29, 28, 11, 9, 41] and the references therein), it is still difficult to study analytically and numerically the spatial and spatio-temporal aggregation patterns exhibited by the kinetic models. For example, there are very few studies that investigate the types of spatiotemporal patterns obtained with 2D and 3D kinetic models (see the review in [31]). Moreover, the presence of non-local interaction terms increases the complexity of the models, leading to a larger variety of patterns that are more difficult to be analysed. While numerical and analytical studies have been conducted to investigate the patterns in 1D non-local models [32, 34, 14], such an investigation is still difficult in the 2D non-local case (see [36]).

The first goal of this article is to start with a class of 1D and 2D non-local kinetic models for self-organised aggregations that incorporate all three social interactions, and to show, through different parabolic scaling approaches, that these models can be reduced to known non-local parabolic models for swarming; see Figure 2 for a diagram illustrating this approach. For the 1D case, similar analytical scalings have been done in the context of bacterial chemotaxis [58] and for the kinetic model (1) for individuals moving along a line [35].

The next aim is to investigate the numerical preservation of patterns between the mesoscopic and macroscopic scales. We use asymptotic preserving numerical methods [46, 47, 21, 24], to obtain a better understanding of what happens with the 1D patterns via the parabolic scaling. With the help of these methods, we investigate numerically the preservation of stationary aggregations (that arise via steady-state bifurcations) and moving aggregations (that arise via Hopf bifurcations), as the scaling parameter $\varepsilon$ is varied from large positive values ($\varepsilon = 1$) corresponding to

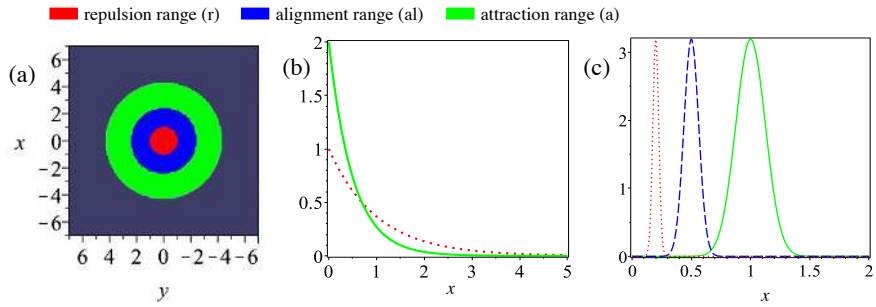


FIGURE 1. 2D and 1D spatial kernels for social interactions. (a) 2D: Attractive ($K_a$), repulsive ($K_r$) and alignment ($K_{al}$) kernels described by equation (21); (b) 1D: Morse-type kernels: $K_{r,a}(x) = e^{-|x|/s_{r,a}}$. (c) 1D: Translated Gaussian kernels $K_j$ as defined in (3) with $j = r, al, a$.

the kinetic models to zero values corresponding to the limiting parabolic models. To visualise the transitions between different patterns as $\varepsilon \to 0$, we construct bifurcation diagrams for the amplitude of the solutions. For the 2D kinetic models, we focus on two analytical scalings that lead to two different nonlocal parabolic models. Our final target is to show the reader that by considering such scaling approaches, we may lose certain aspects of the model dynamics - as emphasised by the numerical simulations in the 1D case.

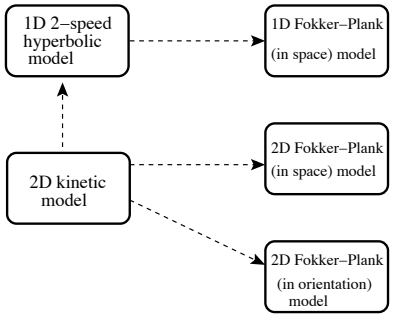


FIGURE 2. Schematic diagram of the scaling and reductionist approach taken here.

The article is structured as follows. Section 2 contains a detailed description of the 1D non-local models for animal aggregations, followed by the parabolic scaling of these models. We also investigate analytically the steady states of the kinetic and corresponding parabolic models. Section 3 contains a description of the 2D nonlocal models, followed by a parabolic limit and a "grazing collision" limit, which lead to different types of macroscopic models of parabolic type. Section 4 focuses on asymptotic preserving methods for 1D models, and shows the spatial and spatio-temporal patterns obtained with the parabolic and kinetic models, for some specific parameter values. Here, we come back to the steady states of the 1D kinetic and parabolic models, and investigate them numerically. We conclude in Section 5 with a summary and discussion of the results.

2. **Description of 1D models.** The following one-dimensional model was introduced in [33, 32] to describe the movement of the densities of left-moving ($u^-$) and right-moving ($u^+$) individuals that interact with conspecifics via social interactions:

$$\frac{\partial u^+}{\partial t} + \gamma \frac{\partial u^+}{\partial x} = -u^+\lambda^+[u^+,u^-] + u^-\lambda^-[u^+,u^-], \tag{1a}$$

$$\frac{\partial u^-}{\partial t} - \gamma \frac{\partial u^-}{\partial x} = u^+\lambda^+[u^+,u^-] - u^-\lambda^-[u^+,u^-], \tag{1b}$$

$$u^\pm(x,0) = u_0^\pm(x). \tag{1c}$$

Here $\gamma$ is the constant speed and $\lambda^+$ ($\lambda^-$) is the rate at which right-moving (left-moving) individuals turn left (right). Since the rates $\lambda^\pm$ are related to the probability of turning (see the derivation of model (1) in [31]), they are positive functions defined as:

$$\begin{aligned}\lambda^\pm[u^+,u^-] =& \lambda_1 + \lambda_2 f(y_N[u^+,u^-]) + \lambda_3 f(y_D^\pm[u^+,u^-]) \\ =& \lambda_1 + \lambda_3\Big(\lambda_2^0 f(y_N[u^+,u^-]) + f(y_D^\pm[u^+,u^-])\Big),\end{aligned} \tag{2}$$

where we denote by $u = u^+ + u^-$ the total population density. In this paper, we generalise the turning rates in [33, 32, 31] and assume that:

- individuals can turn randomly at a constant rate approximated by $\lambda_1$ [33];
- individuals can turn randomly in response to the perception of individuals inside any of the repulsive/attractive/alignment ranges (and independent of the movement direction of their neighbours). These *non-directed interactions* with neighbours are described by the term $y_N[u^+,u^-]$ with turning rate $\lambda_2$.
- individuals can turn in response to interactions with neighbours positioned within the repulsive ($r$), attractive ($a$) and alignment ($al$) zones, respectively (see Figure 1(a)) [33]. This turning is *directed towards or away* from neighbours, depending on the type of interaction (attractive or repulsive). For alignment interactions, individuals turn to move in the same direction as their neighbours. The non-local *directed interactions* with neighbours are described by terms $y_D^\pm[u^+,u^-]$ with turning rate $\lambda_3$.

If $\lambda_3 \neq 0$, we denote by $\lambda_2^0$ the quotient of the turning rates, $\lambda_2/\lambda_3$. This choice of notation is motivated by the corresponding 2D model (Section 3), the connection between the 1D model (1) and the 2D model (18) will be made clearer in Remarks 3, 4, 5 and 6. The turning function $f(\cdot)$ is a non-negative, increasing, bounded functional of the interactions with neighbours. An example of such function is $f(Y) = 0.5 + 0.5\tanh(Y - y_0)$ (see [32]), where $y_0$ is chosen such that when $Y = 0$ (i.e., no neighbours around), then $f(0) \approx 0$ and the turning is mainly random.

To model the long-distance social interactions that lead to turning behaviours, we define the interaction kernels in 1D, see Figure 1, as decreasing functions of the distance between the reference position $x$ (of the population density) and the mid of the interaction ranges $s_j$, $j = r, al, a$,

$$K_j(x) = \frac{1}{\sqrt{2\pi m_j^2}} e^{-(x-s_j)^2/(2m_j^2)}, \tag{3}$$

for $x > 0$ and zero otherwise, with $j = r, al, a$ denoting short-range repulsion ($K_r$), medium-range alignment ($K_{al}$) and long-range attraction ($K_a$) interaction kernels. Here, $m_j = s_j/8$ controls the width of the interaction range $j$.

For the non-directed density-dependent turning we define the turning kernel, $K^N(x) = \hat{K}^N(x) + \hat{K}^N(-x)$ with $\hat{K}^N = q_r K_r + q_{al} K_{al} + q_a K_a$ obtained by superimposing the kernels $K_j$, $j = r, al, a$. Here $q_r$, $q_{al}$ and $q_a$ represent the magnitudes of the repulsive, alignment and attractive social interactions. Note that in [32], $\lambda_2^0 = 0$ and the density-dependent non-directed turning term does not exist. However, in 2D, this term appears naturally when we incorporate random turning behaviour (as discussed in Section 3). With these notations we may define

$$y_N[u] = K^N * u, \quad \text{with } u = u^+ + u^-,$$

for the *non-directed* turning mechanisms. We assume here that individuals turn randomly whenever they perceive other neighbours around (within the repulsive, alignment and attractive ranges).

For the *directed* density-dependent turning, we define

$$y_D^{\pm}[u^+, u^-] = y_r^{\pm}[u^+, u^-] - y_a^{\pm}[u^+, u^-] + y_{al}^{\pm}[u^+, u^-]. \tag{4}$$

Here, $y_j^{\pm}[u^+, u^-]$, $j = r, al, a$, describe the directed turning in response to neighbours within the repulsive ($r$), alignment ($al$) and attractive ($a$) social ranges (as in [33]). As we will explain shortly, the direction of the turning will be given by incorporating movement direction towards or away conspecifics. For this reason, $y_a^{\pm}$ and $y_r^{\pm}$ enter equation (4) with opposite signs.

The density-dependent turnings depend greatly on how individuals communicate with each other, namely whether they can emit (perceive) signals to (from) *all* or *some* of their neighbours. Two particular situations, described by models called M2 and M4 as in [32] (see Figure 3) are considered:

- *Model M2:* Individuals communicate via omni-directional communication signals, and thus they can perceive *all* their neighbours positioned around them within all social interaction ranges. For instance, the majority of mammals communicate via a combination of visual, chemical and auditory signals, which allows them to receive/send information from/to all their neighbours. With this assumption (see Figure 3(a)), the terms $y_{r,a,al}^{\pm}$ are defined as follows:

$$y_{r,a}^{\pm}[u^+, u^-] = q_{r,a} \int_0^\infty K_{r,a}(s)\big(u(x \pm s) - u(x \mp s)\big) ds, \tag{5a}$$

$$y_{al}^{\pm}[u^+, u^-] = q_{al} \int_0^\infty K_{al}(s)\big(u^{\mp}(x \mp s) + u^{\mp}(x \pm s) - u^{\pm}(x \mp s) - u^{\pm}(x \pm s)\big) ds. \tag{5b}$$

Here, $q_j$ describe the magnitudes of the social interactions associated to the interaction kernels defined in (3). To understand the effect of these terms on the turning rates, let us focus on $y_r^+$, for example. If $u(x+s) > u(x-s)$, then $y_r^+$ enters $\lambda^+$ with positive sign, suggesting a higher likelihood of turning, to avoid collision with neighbours ahead at $x + s$. If, on the other hand, $u(x + s) < u(x - s)$, then $y_r^+$ enters $\lambda^+$ with a negative sign, suggesting a lower likelihood of turning. In this case, the individuals at $x$ will keep moving in the same direction, to avoid collision with neighbours behind at $x - s$. Note that the directionality of neighbours influences only the alignment interactions (the attractive and repulsive interactions being defined in terms of the total density $u$). Also, for this particular model, the random density-dependent

terms are given by

$$y_N[u] = \int_0^\infty \hat{K}^N(s)\big(u(x+s) + u(x-s)\big)ds. \tag{6}$$

- *Model M4:* Individuals communicate via unidirectional communication signals, and thus they can perceive only those neighbours moving towards them. For example, birds communicate via directional sound signals, and to ensure an effective transmission of their signals they orient themselves towards their targeted receivers [12]. With this assumption (see Figure 3(b)), the terms $y^{\pm}_{r,a,al}$ are defined as follows:

$$y^{\pm}_{r,a,al}[u^+, u^-] = q_{r,a,al}\int_0^\infty K_{r,a,al}(s)\big(u^{\mp}(x \pm s) - u^{\pm}(x \mp s)\big)ds. \tag{7}$$

Here, the directionality of neighbours influences all three social interactions. Moreover, for this model, the random density-dependent terms are given by

$$y_N[u^+, u^-] = \int_0^\infty \hat{K}^N(s)\big(u^-(x+s) + u^+(x-s)\big)ds. \tag{8}$$

In this equation, we assume that individuals turn randomly in response to $u^-$ and $u^+$ individuals (i.e., in (8) we add all perceived individuals; this is in contrast to equation (7), where we subtract individuals positioned ahead from individuals positioned behind, to impose directionality in the turning behaviour). Note that in (8), $y_N$ does not depend anymore on $u = u^+ + u^-$ (as in (6)), since the individuals at $x$ cannot perceive all their neighbours at $x \pm s$.

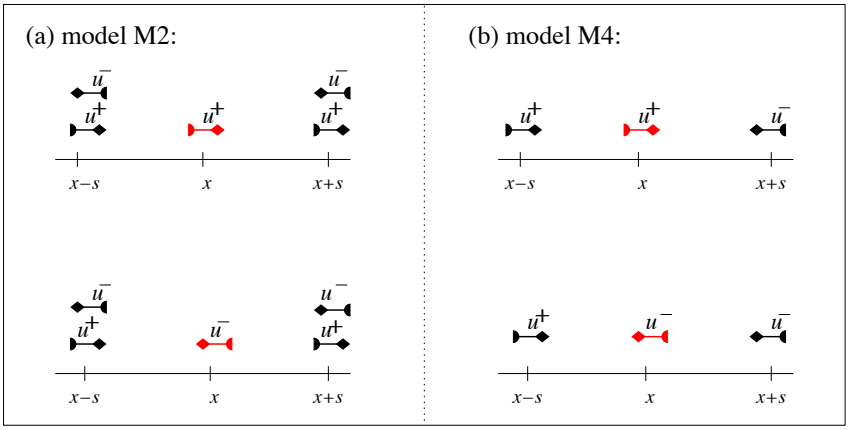


FIGURE 3. Diagram describing the mechanisms through which a reference individual positioned at $x$ (right-moving – top; left-moving – bottom) perceives its neighbours positioned at $x - s$ and $x + s$. The reference individual can perceive (a) all its neighbours (model M2 in [32]); (b) only its neighbours moving towards it (model M4 in [32]).

We focus on these two particular models because: (i) the model (1)+(2)+(5)+(6) assuming $\lambda_1 = 0$ has been generalised to 2D; (ii) the model (1)+(2)+(7)+(8) assuming $\lambda_2 = 0$ has been investigated analytically and numerically, and showed that

it can exhibit Hopf bifurcations (even when $q_{al} = 0$), which give rise to spatio-temporal patterns such as rotating waves and modulated rotating waves [14]. In contrast, model (1)+(2)+(5)+(6) with $\lambda_2 = 0$ does not seem to exhibit rotating waves when $q_{al} = 0$, see [32].

To complete the description of the model, we need to specify the domain size and the boundary conditions. Throughout most of this article, we will consider an infinite domain. However, for the purpose of numerical simulations, in Sections 2.2 and 4 we will consider a finite domain of length $L$ (i.e., $[0, L]$) with periodic boundary conditions: $u^+(L,t) = u^+(0,t), \quad u^-(0,t) = u^-(L,t)$. This assumption will also require wrap-around conditions for the kernels describing the nonlocal social interactions, see Section 4. For large $L$, this assumption approximates the dynamics on an infinite domain.

In the following, we show how this hyperbolic 2-velocity model can be reduced to a parabolic equation by considering suitable scalings, which depend on the biological phenomena and biological assumptions. Of course, to be useful in practice, these parameters have to be calibrated and adapted to particular species as in [42, 43]. The scaling arguments are classically obtained by writing a dimensionless formulation of the problem. We refer to [57] in bacterial chemotaxis and [2] in semiconductor modelling for a detailed description. After this dimensionless rescaling, we typically end up with two different time scales whose balance determines our small parameter: the drift time and the diffusion time.

We start in Subsection 2.1 with a parabolic scaling, which describes the situation where the drift time of a population is much smaller than its diffusion time. To this end, we discuss two separate cases (i.e., social interactions described by nonlinear or linear functions $f(y)$ in (2)), which lead to two different parabolic equations.

2.1. **Parabolic limit for non-linear interactions.** Next, we focus only on model M2 (i.e., equations (1)+(2)+(5)+(6)), since the results for model M4 are similar. The scaling argument applied in [44] transforms the hyperbolic system (1) into a parabolic equation. One can scale the space and time variables ($x = x^*/\varepsilon$, $t = t^*/\varepsilon^2$, with $\varepsilon \ll 1$), or can scale the speed ($\gamma$) and the turning rates ($\lambda_{1,2,3}$). In both cases, we consider the rescaled interaction kernels $K_j^*(x^*) = \frac{1}{\varepsilon} K_j(\frac{x^*}{\varepsilon})$ in the expressions for $y_j^{\pm}$, $j = r, al, a$. Here, we scale the time and space variables to be consistent with the approach in Section 3.1. As mentioned above, the scaling parameter $\varepsilon$ depends on the biological problem modelled. For example, in [44] the authors connect $\varepsilon$ to the ratio of the drift ($\tau_{drift}$) and diffusion ($\tau_{diff}$) times observed in bacteria such as *E. coli*, where $\tau_{drift} \approx 100$ seconds and $\tau_{diff} \approx 10^4$ seconds, and thus $\varepsilon \approx O(10^{-2})$. Similar scaling arguments are used in [57, Appendix] to analyse the ability of parabolic scalings to describe travelling pulses.

To perform the scaling, let us re-write model (1) in terms of the total density $u(x,t)$ and the flux $v(x,t) = \gamma(u^+(x,t) - u^-(x,t))$ of individuals (see also [44, 45]):

$$\varepsilon^2 \frac{\partial u}{\partial t} + \varepsilon \frac{\partial v}{\partial x} = 0, \tag{9a}$$

$$\varepsilon^2 \frac{\partial v}{\partial t} + \varepsilon\gamma^2 \frac{\partial u}{\partial x} = \gamma u\big(\lambda^-[u,v] - \lambda^+[u,v]\big) - v\big(\lambda^+[u,v] + \lambda^-[u,v]\big), \tag{9b}$$

with initial conditions $u(x,0) = u_0(x)$, $v(x,0) = v_0(x)$. For clarity, here we dropped the "$*$" from the rescaled space ($x^*$) and time ($t^*$) variables. In addition, we assume that individuals have a reduced perception of the surrounding neighbours for small

values of $\varepsilon$, [30]:

$$f_\varepsilon\Big(y_D^\pm[u,v]\Big) = \varepsilon f\Big(y_D^\pm[u, \int_{\frac{x^*}{\varepsilon}} \varepsilon \frac{\partial u}{\partial t^*}]\Big), \quad f_\varepsilon\Big(y_N[u]\Big) = \varepsilon f\Big(y_N[u]\Big), \tag{10}$$

where $f$ enters the turning functions $\lambda^\pm$ (2):

$$\begin{aligned}\lambda^+[\cdot] + \lambda^-[\cdot] &= 2\lambda_1 + 2\lambda_2\, \varepsilon f(y_N[\cdot]) + \varepsilon\, \lambda_3 \Big(f(y_D^+[\cdot]) + f(y_D^-[\cdot])\Big),\\ \lambda^-[\cdot] - \lambda^+[\cdot] &= \lambda_3 \varepsilon \Big(f(y_D^-[\cdot]) - f(y_D^+[\cdot])\Big).\end{aligned}$$

By eliminating $v = \varepsilon \int_x \frac{\partial u}{\partial t}$ from equations (9), and taking the limit $\varepsilon \to 0$, we obtain the following parabolic equation

$$\frac{\partial u}{\partial t} = \frac{\gamma^2}{2\lambda_1} \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right) - \frac{\lambda_3 \gamma}{2\lambda_1} \frac{\partial}{\partial x}\left(\left(f(y_D^-[u]) - f(y_D^+[u])\right) u\right). \tag{11}$$

We note here that the non-local terms $f(y_D^\pm[u])$ now depend only on the repulsive and attractive interactions. The reason for this is that the alignment interactions are defined in terms of $u^\pm = (u \pm \frac{1}{\gamma} v)/2 = 0.5(u \pm \frac{1}{\gamma} \int_{x/\varepsilon} \varepsilon^2 \partial u/\partial t)$. As $\varepsilon \to 0$, the $u$ terms in (5) cancel out, and the integrals approach zero. Equation (11) can be re-written as

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(D_0 \frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial x}\Big(B_0\, u\, V(u)\Big), \tag{12}$$

with diffusion rate $D_0 = \gamma^2/(2\lambda_1)$ and drift rate $B_0 = \lambda_3\gamma/(2\lambda_1)$. The velocity $V(u)$ depends on the communication mechanism incorporated. For example, for model M2 we have $y_D^\pm[u] = \pm K * u$, and so the velocity is given by

$$V[u] = f\big(-K * u\big) - f\big(K * u\big)$$

where we define

$$K * u = \bar{K}^+ * u - \bar{K}^- * u, \quad \bar{K}^\pm * u = \int_0^\infty \bar{K}(s) u(x \pm s) ds,$$
$$\bar{K} = q_r K_r - q_a K_a. \tag{13}$$

For model M4, we have $y_D^\pm[u] = \pm 0.5 K * u$, and so the velocity is quite similar: $V[u] = f\big(-0.5K*u\big) - f\big(0.5K*u\big)$, the factor 0.5 appearing from $u^\pm = 0.5(u \pm \frac{1}{\gamma} v)$.

**Remark 1.** We observe that the random density-dependent turning $f(y_N[u])$ does not appear in this parabolic limit. This is the result of the scaling assumptions (10).

**Remark 2.** Here, the turning functions $f(\cdot)$ were chosen to be bounded, since individuals cannot turn infinitely fast when subject to very strong interactions with neighbours [32, 34]. However, for simplicity, many models consider linear functions: $f(z) = z$ (see, for example, [49, 51, 36]). The choice of having bounded or non-bounded turning functions $f(\cdot)$ has further implications on the models. In particular, for linear functions, the argument $y_D^\pm = y_r^\pm - y_a^\pm + y_{al}^\pm$ can be either positive or negative (depending on the magnitudes of the social interactions), with $y_D^+ = -y_D^-$. For very small constant and non-directional turning rates ($\lambda_1, \lambda_2 \approx 0$), this can lead to $\lambda^+ < 0$ and $\lambda^- > 0$, or vice versa. Now the $u^+\lambda^+$ terms add to the $u^-\lambda^-$ terms, causing both $u^+$ and $u^-$ populations to decide very fast to move in the same direction (in fact, one of the populations is reinforced to keep its moving direction). This is different from the case with bounded turning functions, where if $y_D^+ = -y_D^- \ll 0$, then $0 < \lambda^+ \approx \lambda_1 + \lambda_2 f\left(y_N[u^+, u^-]\right) < \lambda^-$. So if $\lambda_1, \lambda_2 \approx 0$,

then $u^+\lambda^+ \approx 0$ and hence population $u^+$ is not reinforced to keep its movement direction.

Because the 2D kinetic model that we will investigate in Section 3 assumes $f$ to be a linear function, with a very weak directed turning behaviour ($\varepsilon\lambda_3$), we now consider the case $f(y_N[u]) = y_N[u] = K^N * u$ and $f(y_D^{\pm}[u]) = \varepsilon y_D^{\pm}[u]$, and so the turning rates can be written as

$$\lambda^{\pm}[u^+, u^-] = \lambda_1 + \lambda_2 K^N * u + \varepsilon \lambda_3 y_D^{\pm}[u] \tag{14}$$

By taking the limit $\varepsilon \to 0$ in (9), we obtain the following parabolic equation with density-dependent coefficients:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(D[u]\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial x}\Big(B[u]u\big(y_D^-[u] - y_D^+[u]\big)\Big), \tag{15a}$$

$$D[u] = \frac{\gamma^2}{2(\lambda_1 + \lambda_2 K^N * u)} \quad \text{and} \quad B[u] = \frac{\lambda_3\gamma}{2(\lambda_1 + \lambda_2 K^N * u)}. \tag{15b}$$

This expression is similar to the asymptotic parabolic equation (30) for the 2D model. We will return to this aspect in Section 3.1.

2.2. **The preservation of steady states and their stability as $\varepsilon \to 0$.** The spatially homogeneous steady states describe the situation where individuals are evenly spread over the whole domain. In the following we investigate how these steady states and their linear stability are preserved in the parabolic limit. To this end, we focus on the more general case of non-linear social interactions (the case with linear interactions is similar). For simplicity we assume here that $\lambda_2 = 0$ and $q_{al} = 0$. To calculate these spatially homogeneous states we need to define $A = \int_0^L (u^+ + u^-)dx$ the total population density. For simplicity, throughout this paper we assume that $A = 2$; similar results can be obtained for different values of $A$.

Figure 4(a) shows the number and magnitude of the steady states $u^*$ displayed by (9)-(10) with communication mechanism M4, for different values of $\varepsilon$, as one varies the difference in the magnitude of the repulsive and attractive social interactions, $q_r - q_a$. For medium $\varepsilon$, the model can display up to 5 different steady states: one "unpolarised" state $(u^+, u^-) = (u^*, u^*) = (A/2, A/2)$ (where half of the individuals are facing left and half are facing right), and two or four "polarised" states $(u^*, A - u^*), (A - u^*, u^*)$ characterised by $u^* < A/2$ or $u^* > A/2$. Two of these "polarised" states exist only in a very narrow parameter range: e.g., for $\varepsilon = 1$, they exist when $q_r - q_a \in (2, 3.7)$. The other two "polarised" states exist for any $q_r - q_a > 2$. For a calculation of the threshold values of $q_r - q_a$ that ensure the existence of 3 or 5 steady states see [33]. As $\varepsilon$ decreases, the magnitude of the polarised states decreases (i.e., the differences between the number of individuals facing right and those facing left are decreasing). Moreover, for small $\varepsilon$, these polarised states appear only when repulsion becomes much stronger than attraction (i.e., $q_r - q_a \gg 10$). When $\varepsilon = 0$ there is only one steady state $u^* = A/2$. Since this state exists for all $\varepsilon \geq 0$, from now on we will focus our attention only on it. Note that for $q_{al} = 0$ and for the communication mechanism M2 (not shown here), the nonlocal attractive-repulsive terms vanish, and there is only one steady state, $u^* = A/2 = 1$, which does not depend on $\varepsilon$.

Models (1) and (9) do exhibit a large variety of local bifurcations: codimension-1 Steady-state and Hopf bifurcations [34] as well as codimension-2 Hopf/Hopf,

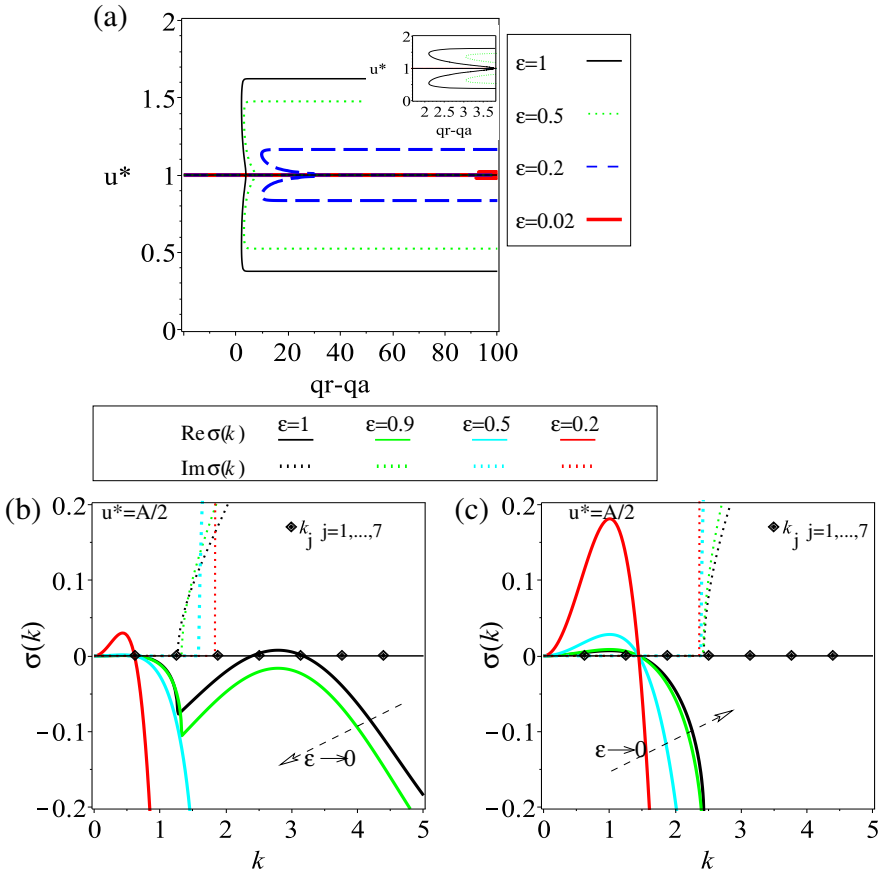


FIGURE 4. (a) Spatially homogeneous steady states $u^*$ for model (9) with communication signals (7) and (8) (communication mechanism M4), for different values of $\varepsilon$. The small inset figure shows the 5 possible steady states occurring for $\varepsilon = 1$ and $q_r - q_a \in (2, 3.7)$ (see the black continuous curve); (b) Dispersion relation $\sigma(k_j)$ for M4 (given by (16)), showing the stability of the spatially homogeneous steady state $u^* = A/2$, for different values of $\varepsilon$; (c) Dispersion relation $\sigma(k_j)$ for M2, for the stability of the spatially homogeneous steady state $u^* = A/2$, for different values of $\varepsilon$. The continuous curves describe Re $\sigma(k_j)$, while the dotted curves describe the Im $\sigma(k_j)$. The small diamond-shaped points show the discrete wavenumbers $k_j, j = 1, ..., 7$, with $k_j = 2\pi j/L$ (and thus $k_j \in (0, 5)$ for $j = 1, .., 7$ and $L = 10$). The parameter values are: (b)$q_a = 1.545$, $q_r = 2.779$; (c) $q_a = 1.5$, $q_r = 0.93$. The rest of parameters are: $q_{al} = 0$, $\lambda_1 = 0.2$, $\lambda_2 = 0$, $\lambda_3 = 0.9$, $A = 2$.

Hopf/Steady-state and Steady-state/Steady-state bifurcations [14]. Next we focus on the parameter region where two such bifurcations can occur. We choose a Hopf/steady-state bifurcation for M4 (Figure 4(b)) and a steady-state bifurcation for M2 (Figure 4(c)), and investigate what happens when $\varepsilon \to 0$. To identify the

parameter regions where these bifurcations occur, we consider a finite domain of length $L$, and investigate the growth of small perturbations of spatially homogeneous solutions. We assume $u^{\pm} \propto u^* + a_{\pm}\exp(\sigma t + ik_j x)$, with $k_j = 2\pi j/L, j \in \mathbb{N}^+$, the discrete wave-numbers, and $|a_{\pm}| \ll 1$. We substitute these solutions into the linearised system (9), and by imposing that the determinant of this system is zero, we obtain the following dispersion relation, which connects $\sigma$ (the growth/decay of the perturbations) with the wave-numbers $k_j$:

$$\varepsilon^2\sigma^2 + \sigma(2L_1^{\varepsilon} - R_2^{\varepsilon}\mathrm{Re}(\hat{K}^+)) + \gamma^2 k_j^2 - \gamma k_j R_2 \mathrm{Im}(\hat{K}^+) = 0, \tag{16}$$

where $L_1^{\varepsilon} = \lambda_1 + \varepsilon\lambda_3 f(0)$, $R_2^{\varepsilon} = 2\varepsilon u^*\lambda_3 f'(0)$, and $\hat{K}^+ = \mathrm{Re}(\hat{K}^+) + i\mathrm{Im}(\hat{K}^+)$ the Fourier transforms of $\bar{K}^+ * u$ described in equations (13). Note that the wave numbers $k_j$ that become unstable (i.e., for which $\mathrm{Re}(\sigma(k_j)) > 0$) determine, at least for a short time, the number of "peaks" $j$ that emerge in the spatial distribution of the density.

Figure 4(b) shows the stability of the spatially homogeneous steady state $u^* = A/2$, for model M4, as given by the dispersion relation (16). Even if the wave-numbers $k_j$ are discrete (see the diamond-shaped points on the x-axis of Figure 4(b)), we plot $\sigma(k_j)$, $j > 0$ as a continuous function of $k_j$ for clarity. To discuss what happens with a Hopf bifurcation as $\varepsilon \to 0$, we focus in Figure 4(b) on a parameter space where such a bifurcation occurs (i.e., where $\mathrm{Re}(\sigma(k_j)) = 0$ in (16)): $q_a = 1.545$, $q_r = 2.779$, $\lambda_1 = 0.2$, $\lambda_2 = 0$, $\lambda_3 = 0.9$ and $\varepsilon = 1$ (see also [15]). For these parameter values, three modes become unstable at the same time: a steady-state mode $k_1$ ($\mathrm{Im}(\sigma(k_1)) = 0$; associated with stationary patterns with 1 peak) and two Hopf modes $k_4$ and $k_5$ ($\mathrm{Im}(\sigma(k_{4,5})) > 0$; associated with travelling patterns with 4 or 5 peaks). As $\varepsilon \to 0$, the steady-state mode persists while the Hopf modes disappear (i.e., $0 < \mathrm{Re}(\sigma(k_1)) \ll 1$ and $\mathrm{Re}(\sigma(k_{4,5})) < 0$; see Figure 4(b).) This can be observed also from equation (16): as $\varepsilon \to 0$, we have $\sigma \in \mathbb{R}$. A similar investigation of the local stability of the spatially homogeneous steady states associated with the non-local parabolic equation (12) shows that this equation cannot have complex eigenvalues (i.e., $\mathrm{Im}(\sigma(k_j)) = 0$ for all $j > 0$), and thus cannot exhibit local Hopf bifurcations [16].

Figure 4(c) shows the stability of the spatially homogeneous steady state $u^* = A/2$, for model M2, as given by the dispersion relation $\sigma(k_j)$:

$$\varepsilon^2\sigma^2 + \sigma(2L_1^{\varepsilon}) + \gamma^2 k_j^2 - 2\gamma k_j R_2 \mathrm{Im}(\hat{K}^+) = 0. \tag{17}$$

For $q_a = 1.5$, $q_r = 0.93$, $\lambda_1 = 0.2$, $\lambda_2 = 0$, $\lambda_3 = 0.9$ and $\varepsilon = 1$, model M2 exhibits a steady-state bifurcation, i.e., $\mathrm{Re}(\sigma(k_j)) = \mathrm{Im}(\sigma(k_j)) = 0$ in (17). In particular, two steady-state modes are unstable at the same time: $k_1$ and $k_2$ (both associated with stationary patterns). As $\varepsilon \to 0$, the two modes remain unstable. Hence, we expect that the spatial patterns generated by these modes will persist as $\varepsilon \to 0$. We will return to this aspect in Section 4.4, when we will investigate numerically the mechanisms that lead to the disappearance of the Hopf modes and the persistence of the steady-state modes, as $\varepsilon \to 0$.

## 3. Description of 2D Models.

An attempt to generalise a specific case of the 1D model (1)-(2)-(5)-(6) to two dimensions was made by Fetecau [36]. The Boltzman-type model described in [36] incorporates the non-local social interactions in the

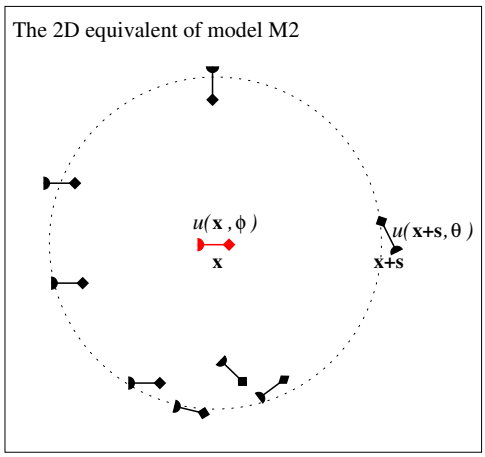


FIGURE 5. Caricature description of the M2 mechanism in 2D (where individuals can perceive *all* their neighbours within a certain interaction range). We assume that a reference individual is positioned at $\mathbf{x} = (x, y)$ and moves in direction $\phi$. Its neighbours are at various spatial positions $\mathbf{x} + \mathbf{s}$ within a certain interaction range (e.g., alignment range). The interaction ranges are described by the 2D kernels (21); see also Figure 1(a).

reorientation terms:

$$\frac{\partial u}{\partial t} + \gamma \mathbf{e}_\phi \cdot \nabla_{\mathbf{x}} u = -\lambda(\mathbf{x}, \phi) u + \int_{-\pi}^{\pi} T(\mathbf{x}, \phi', \phi) u(\mathbf{x}, \phi', t) d\phi'. \tag{18}$$

Here, $u(\mathbf{x}, \phi, t)$ is the total population density of individuals located at $\mathbf{x} = (x, y)$, moving at a constant speed $\gamma > 0$ in direction $\phi$. The term $\mathbf{e}_\phi = (\cos(\phi), \sin(\phi))$ gives the movement direction of individuals. The reorientation terms, $\lambda(\mathbf{x}, \phi)$ and $T(\mathbf{x}, \phi', \phi)$ depend on the non-local interactions with neighbours, which can be positioned in the repulsive, attractive, and alignment ranges depicted in Fig. 1(a). Thus, these terms have three components each, corresponding to the three social interactions:

$$T(\mathbf{x}, \phi', \phi) = T_{al}(\mathbf{x}, \phi', \phi) + T_a(\mathbf{x}, \phi', \phi) + T_r(\mathbf{x}, \phi', \phi).$$

In contrast to the model in [36], here we assume that the reorientation terms

$$\lambda_j(\mathbf{x}, \phi') = \int_{-\pi}^{\pi} T_j(\mathbf{x}, \phi', \phi) \mathrm{d}\phi, \quad j = r, a, al$$

have both a constant and a density-dependent component:

$$T_{al}(\mathbf{x}, \phi', \phi) = \frac{\eta_{al}}{2\pi} + \tag{19a}$$
$$\lambda_3 \, q_{al} \int_{-\pi}^{\pi} \int_{\mathbb{R}^2} K_{al}^d(\mathbf{x} - \mathbf{s}) K_{al}^o(\theta, \phi') \omega_{al}(\phi' - \phi, \phi' - \theta) u(\mathbf{s}, \theta, t) ds d\theta,$$
$$T_{r,a}(\mathbf{x}, \phi', \phi) = \frac{\eta_{r,a}}{2\pi} + \tag{19b}$$
$$\lambda_3 \, q_{r,a} \int_{-\pi}^{\pi} \int_{\mathbb{R}^2} K_{r,a}^d(\mathbf{x} - \mathbf{s}) K_{r,a}^o(\mathbf{s}, \mathbf{x}, \phi') \omega_{r,a}(\phi' - \phi, \phi' - \psi) u(\mathbf{s}, \theta, t) ds d\theta.$$

Therefore, the turning rate $\lambda(\mathbf{x},\phi) = \lambda_{al}(\mathbf{x},\phi) + \lambda_a(\mathbf{x},\phi) + \lambda_r(\mathbf{x},\phi)$ is defined by

$$\lambda = \lambda_1 + \lambda_3 \bar{\lambda}[u(\mathbf{x},\phi)], \tag{20}$$

with $\lambda_1 = \eta_r + \eta_{al} + \eta_a$ and $\bar{\lambda}[u(\mathbf{x},\phi)]$ being given as the integral over $\phi' \in [-\pi,\pi]$ of the sum of nonlocal terms in (19) with $\phi$ and $\phi'$ interchanged.

**Remark 3.** By defining the constant basic turning rate to be $\lambda_1 = \eta_r + \eta_{al} + \eta_a$, we generalised the model in [36] (where $\lambda_1 = 0$). Note that the turning rates here are linear functions of the non-local interactions with neighbours. This is in contrast to the more general non-linear turning function $f$ we considered in Section 2.1 for the 1D hyperbolic model. In what follows, we are interested in non-constant turning rates $\lambda_j(\mathbf{x},\phi')$, $j = r,a,al$, and so we will henceforth assume $\lambda_3 \neq 0$.

As in [36], $\lambda_j$, $j = r,a,al$, are defined in terms of both distance kernels and orientation kernels. The 2D distance kernels $K_j^d$, $j = r,a,al$ are given by

$$K_j^d(\mathbf{x}) = \frac{1}{A_j} e^{-(\sqrt{x^2+y^2}-d_j)/m_j^2}, \quad j = r,a,al, \tag{21}$$

where constants $A_j$ are chosen such that the kernels integrate to one. The orientation kernels $K_j^o$ measure the likelihood of turning in response to the movement direction of neighbours (for alignment interactions) or in response to the position of neighbours (for repulsive and attractive interactions):

$$\begin{aligned} K_{al}^o(\theta,\phi) &= \frac{1}{2\pi}(1-\cos(\phi-\theta)), \\ K_{r,a}^o(\mathbf{s},\mathbf{x},\phi) &= \frac{1}{2\pi}(1 \pm \cos(\phi-\psi)), \end{aligned}$$

where $\psi$ is the angle between the positive $x$-axis and the relative location $\mathbf{s}-\mathbf{x}$ of the neighbours at $\mathbf{s}$ with respect to the reference individual at $\mathbf{x}$. Finally, $\omega$ describes the tendency to turn from direction $\phi'$ to direction $\phi$, as a result of interactions with individuals moving in direction $\theta$:

$$\omega(\phi'-\phi,\phi'-\theta) = g(\phi'-\phi-R(\phi'-\theta)),$$

for some suitable choice of $g$. Note that in the case $\lambda_1 = 0$, the function $\omega$ describes the probability of re-orientation in the sense discussed in [38] and thus we require $\int \omega(\phi'-\phi,\phi'-\theta)d\phi = 1$. For example, $g$ could be a periodic function that integrates to one:

$$g(\theta) = \frac{1}{\sqrt{\pi}\sigma} \sum_{z\in\mathbb{Z}} e^{-(\frac{\theta+2\pi z}{\sigma})^2}, \quad \theta \in (-\pi,\pi),$$

with $\sigma$ a parameter measuring the uncertainty of turning (with small $\sigma$ leading to exact turning) [36, 37]. Another typical choice could be the von Mises distribution, as in Vicsek-type models [27].
On the other hand, when $\lambda_1 > 0$, then $g$ can be interpreted as a small reorientation perturbation from the random turning behaviour and so $\omega$ satisfies $\int \omega(\phi'-\phi,\phi'-\theta)d\phi = 0$ and therefore $g$ is required to be odd.

**Remark 4.** Fetecau [36] showed that by imposing the turning angle to have only two possible values $\phi = \pm\pi$, the 2D model (18) can be reduced to the 1D model (1) for a specific choice of turning rates $\lambda^{\pm}[u^+,u^-]$. More precisely, considering the more general turning operators (19a) and (19b), we recover (2) with $\lambda_1, \lambda_3 \geq$

0, $\lambda_2 = 0$ for a linear turning function $f(z) = z$, and with the communication mechanism

$$\begin{aligned} y_D^{\pm}[u^+, u^-] = &\frac{1}{\pi} q_{al} \int_{-\infty}^{\infty} K_{al}(x-s) \left(u^{\mp}(s,t)\right) \mathrm{d}s \\ &+ \frac{1}{\pi} q_a \int_{-\infty}^{x} K_a(x-s) \left(u^+(s,t) + u^-(s,t)\right) \mathrm{d}s \\ &+ \frac{1}{\pi} q_r \int_{x}^{\infty} K_r(x-s) \left(u^+(s,t) + u^-(s,t)\right) \mathrm{d}s. \end{aligned}$$

This is a similar turning behaviour to model M2 in [32], since individuals receive and emit omni-directional communication signals, but with the function $f$ linear. Moreover, as we will show in the next section, even if the 2D model (18) can be reduced to a special case of the 1D model (1) without $\lambda_2$ in (2), the parabolic scaling of the 2D model reduces to a special case of the parabolic scaling of the 1D model, which includes a $\lambda_2$ term for non-directed turning. As we will see shortly, this 2D parabolic scaling leads to the natural appearance of a non-directed interaction contribution, suggesting that there are more subtle differences between the 1D and 2D models.

The diffusion limit (i.e., $x = x^*/\varepsilon$, $t = t^*/\varepsilon^2$) of a transport model similar to (18), but with constant turning rates $\lambda$ was discussed in [44, 54]. In the following we consider the parabolic limit for model (18) with density-dependent turning rates.

3.1. **Parabolic Drift-Diffusion limit.** We focus on the case where individuals are only influenced slightly by the presence of neighbours, i.e., the turning mechanism can be assumed to be a small perturbation of a uniform turning probability. In this case, we will show that the Boltzmann-type equation (18) can be reduced to a drift-diffusion equation in the macroscopic regime.

We consider the scaling $t = t^*/\varepsilon^2$, $\mathbf{x} = \mathbf{x}^*/\varepsilon$, where $\varepsilon \ll 1$ is a small parameter. Since the velocity in the new variables is of order $1/\varepsilon$, then we make the scaling assumption that an individual's turning behaviour is only influenced slightly by the presence of neighbours:

$$T[u](\mathbf{x}, \phi', \phi) = \frac{\lambda_1}{2\pi} + \frac{\lambda_2}{2\pi} K^d * \rho(\mathbf{x}, t) + \varepsilon \lambda_3 B[u](\mathbf{x}, \phi', \phi), \tag{22}$$

with $\rho(\mathbf{x}, t) = \int_{-\pi}^{\pi} u(\mathbf{x}, \phi, t) \, \mathrm{d}\phi$, and where we define

$$K^d(\mathbf{x}) := q_{al} K_{al}^d(\mathbf{x}) + q_a K_a^d(\mathbf{x}) + q_r K_r^d(\mathbf{x})$$

to be the *social distance kernel*. As we have done in the 1D case, we have separated the non-directed and directed turning rates.

If $\lambda_3 \neq 0$, we factorise again the turning rate $\lambda_3$ corresponding to the directed interactions and write $\lambda_2^0 = \lambda_2/\lambda_3$ the quotient of turning rates. With this notation, $\bar{\lambda}[u(\mathbf{x}, \phi)]$ in (20) can be written as

$$\bar{\lambda}[u(\mathbf{x}, \phi)] = \lambda_2^0 K^d * u(\mathbf{x}, \phi, t) + \varepsilon \, y_D[u(\mathbf{x}, \phi, t)], \tag{23}$$

with $y_D[u] = \int B[u](\mathbf{x}, \phi', \phi) \mathrm{d}\phi'$. Note that the turning rate $\lambda$ given by (20)-(23) corresponds to the 1D turning rates (14) with this specific choice of $y_D[u]$. The scaling assumption (22) can be derived by introducing reduced perception of directionality of neighbours into the re-orientation function $\omega$ and into the orientation

kernels $K_j^o$,

$$g_j(\vartheta) = \lambda_2^0 + \varepsilon G_j(\vartheta),$$
$$K_{al}^o(\theta,\phi) = \frac{1}{2\pi}\left(1 - \varepsilon\,\cos\left(\phi-\theta\right)\right),$$
$$K_{r,a}^o(\mathbf{s},\mathbf{x},\phi) = \frac{1}{2\pi}\left(1 \pm \varepsilon\,\cos\left(\phi-\psi\right)\right),$$

where $G_j(\vartheta)$, $j = r,a,al$ are *signal response functions* to be chosen according to the biological context. Substituting these expressions into the reorientation terms (19), we define $\lambda_1 = \eta_{al} + \eta_r + \eta_a$ and we obtain (22) with a precise expression for the *social response function* $B[u]$.

If $\lambda_1 = 0$, we further have $\lambda_2 = \lambda_3/2\pi$ and $\int_{-\pi}^{\pi} G_j(\phi' - \phi - R(\phi'-\theta))\mathrm{d}\phi = 0$, $j = r,a,al$ as the probability to turn to any new angle is 1. In addition, we want the turning function $R(\vartheta)$ to be close to an unbiased turning mechanism. This can be expressed by taking $R(\vartheta) = \varepsilon\vartheta$, which indeed corresponds to weak interaction between individuals, [38]. We obtain $B[u] = B_{al}[u] + B_a[u] + B_r[u]$ with

$$\begin{aligned} B_{al}[u](\phi',\phi) =& \frac{1}{2\pi}\,q_{al}\,G_{al}(\phi'-\phi)\,K_{al}^d * \rho(\mathbf{x},t) \\ &- \frac{\lambda_2^0}{2\pi}\,q_{al}\int_{\mathbb{R}^2} K_{al}^d(\mathbf{x}-\mathbf{s})\int_{-\pi}^{\pi}\cos\left(\phi'-\theta\right)\,u(\mathbf{s},\theta,t)\,\mathrm{d}\theta\,\mathrm{d}\mathbf{s}, \end{aligned} \tag{24}$$

$$\begin{aligned} B_{r,a}[u](\phi',\phi) =& \frac{1}{2\pi}\,q_{r,a}\,G_{r,a}(\phi'-\phi)\,K_{r,a}^d * \rho(\mathbf{x},t) \\ &\pm \frac{\lambda_2^0}{2\pi}\,q_{r,a}\int_{\mathbb{R}^2} K_{r,a}^d(\mathbf{x}-\mathbf{s})\cos\left(\phi'-\psi\right)\,\rho(\mathbf{s},t)\,\mathrm{d}\mathbf{s}. \end{aligned} \tag{25}$$

**Remark 5.** Note that in 2D, $\lambda_2^0$ is introduced as the relative strength of non-directed and directed turning kernels. This is part of the scaling assumption in 2D, whereas in 1D, we introduced it as part of the model (1)-(2) before rescaling. Note that $\lambda_2^0 = 1/2\pi$ in Fetecau's model where no distinction is made between directed and non-directed turning.

Let us introduce

$$K_*^d(\mathbf{x}^*) = \frac{1}{\varepsilon}K^d\left(\frac{\mathbf{x}^*}{\varepsilon}\right), \quad B_*(\mathbf{x}^*,\phi',\phi) = \frac{1}{2\pi}B\left(\frac{\mathbf{x}^*}{\varepsilon},\phi',\phi\right).$$

Simplifying the notation by dropping $*$, system (18) writes in the new variables as

$$\begin{aligned} \varepsilon^2\partial_t u + \varepsilon\,\gamma\,\mathbf{e}_\phi\cdot\nabla_{\mathbf{x}}u =& \frac{1}{2\pi}\left(\lambda_1 + \lambda_2\,K^d * \rho\right)(\rho - 2\pi u) \\ &+ \varepsilon\,\lambda_3\,2\pi\int_{-\pi}^{\pi} B(\mathbf{x},\phi',\phi)\,u(\mathbf{x},\phi',t)\,\mathrm{d}\phi' \\ &- \varepsilon\,\lambda_3\,2\pi\;u(\mathbf{x},\phi,t)\int_{-\pi}^{\pi} B(\mathbf{x},\phi,\phi')\mathrm{d}\phi'\,. \end{aligned} \tag{26}$$

Using a Hilbert expansion approach, $u = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + ...$, and defining the macroscopic densities $\rho_i = \int_{-\pi}^{\pi} u_i\,\mathrm{d}\phi$ for $i \in \mathbb{N}_0$, we obtain at leading oder a relaxation towards a uniform angular distribution at each position:

$$u_0(\mathbf{x},\phi,t) = \rho_0(\mathbf{x},t)F(\phi), \tag{27}$$
$$F(\phi) = \frac{1}{2\pi}\mathbb{1}_{\phi\in(-\pi,\pi]}.$$

Integrating (26) with respect to the direction of motion $\phi$, we obtain the continuity equation

$$\partial_t \rho_0 + \gamma \int_{-\pi}^{\pi} \mathbf{e}_\phi \cdot \nabla_{\mathbf{x}} u_1 \mathrm{d}\phi = 0. \tag{28}$$

Comparing orders of $\varepsilon$ and using (27), we can derive an expression for $u_1$ in terms of $u_0, \rho_0, \rho_1$,

$$\begin{aligned} u_1 =& \frac{1}{2\pi}\rho_1 - \gamma \frac{\mathbf{e}_\phi \cdot \nabla_{\mathbf{x}} u_0}{\lambda_1 + \lambda_2 K^d * \rho_0} \\ &+ \rho_0 \frac{\lambda_3}{\lambda_1 + \lambda_2 K^d * \rho_0} \int_{-\pi}^{\pi} B[\rho_0](\mathbf{x}, \phi', \phi) - B[\rho_0](\mathbf{x}, \phi, \phi')\, \mathrm{d}\phi'. \end{aligned}$$

Substituting into (28), we arrive at a macroscopic drift-diffusion equation of the form

$$\partial_t \rho_0 = \nabla_{\mathbf{x}} \cdot \left( D[\rho_0] \nabla_{\mathbf{x}} \rho_0 - \rho_0 \mathbf{k}[\rho_0] \right),$$

where the macroscopic diffusion coefficient $D[\rho_0] = \gamma^2/(2(\lambda_1 + \lambda_2 K^d * \rho_0))$ and the social flux

$$\mathbf{k}[\rho_0] = \frac{\lambda_3 \gamma}{\lambda_1 + \lambda_2 K^d * \rho_0} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} (\mathbf{e}_\phi - \mathbf{e}_{\phi'}) B[\rho_0](\mathbf{x}, \phi', \phi) \mathrm{d}\phi' \mathrm{d}\phi \tag{29}$$

are both described in terms of microscopic quantities. In the context of collective behaviour of animal groups, we make two further assumptions:

(i) Individuals can process information in a similar manner for all three types of social interactions:

$$G_{al}(\vartheta) = G_r(\vartheta) = G_a(\vartheta) =: G(\vartheta) \quad \forall \vartheta.$$

(ii) Individuals have symmetric perception, in other words, they can process information equally well from left and right. Then the turning probability function $\omega$ is bisymmetric,

$$\omega(-\alpha, -\beta) = \omega(\alpha, \beta),$$

which implies symmetry of the signal response function $G$.

Under these assumptions, the first term of the social response functions $B_j[u]$ in (24) and (25) cancels when substituted into the social flux (29). The second term contains the factor $\lambda_2^0$ which cancels with $\lambda_3$ in (29), leaving us with a factor of $\lambda_2$ in the social flux. Using (27), we can simplify the social flux even further and obtain the drift-diffusion equation

$$\partial_t \rho = \nabla_{\mathbf{x}} \cdot \left( D_0[\rho] \nabla_{\mathbf{x}} \rho \right) - \nabla_{\mathbf{x}} \cdot \left( \rho\, \mathbf{k}[\rho] \right), \tag{30a}$$

$$D_0[\rho] = \frac{\gamma^2}{2(\lambda_1 + \lambda_2 K^d * \rho)}, \tag{30b}$$

$$\mathbf{k}[\rho](\mathbf{x}, t) = \frac{\lambda_2 \pi \gamma}{\lambda_1 + \lambda_2 K^d * \rho} \left( q_r K_r^d(\mathbf{x}) \frac{\mathbf{x}}{|\mathbf{x}|} - q_a K_a^d(\mathbf{x}) \frac{\mathbf{x}}{|\mathbf{x}|} \right) * \rho. \tag{30c}$$

For notational convenience, we dropped the zero in $\rho_0$. Note that this equation is similar to the 1D drift-diffusion equation (15) obtained via the parabolic limit for linear social interactions.

**Remark 6.** Integrating the 2D scaling assumption (22), we have

$$\lambda(\mathbf{x}, \phi') = \lambda_1 + \lambda_2 K^d * \rho(\mathbf{x}, t) + \varepsilon \lambda_3 \int B[u](\mathbf{x}, \phi', \phi) \mathrm{d}\phi,$$

which is a particular case of the 1D scaling assumption (14). More precisely, the 2D turning rate $\lambda(\mathbf{x}, \phi')$ corresponds to (2) on the projected velocity set $\{0, \pi\}$, with a linear turning function $f(z) = z$ and with the non-directed and directed communication mechanisms given by

$$\begin{aligned} y_N[u] =& K^d * \rho(\mathbf{x}, t), \\ y_D^{\pm}[u^+, u^-] =& \frac{G(0) + G(\pi)}{2} K^d * \rho(\mathbf{x}, t) \\ & \mp \lambda_2^0 \int_{\mathbb{R}} q_{al} K_{al}^d(\mathbf{x} - \mathbf{s}) \left(u^+(s_1, t) - u^-(s_1, t)\right) \mathrm{d}s_1 \\ & \mp \lambda_2^0 \int_{-\infty}^{x_1} \left(q_r K_r^d(\mathbf{x} - \mathbf{s}) - q_a K_a^d(\mathbf{x} - \mathbf{s})\right) \rho(\mathbf{s}, t) \mathrm{d}s_1 \\ & \pm \lambda_2^0 \int_{x_1}^{\infty} \left(q_r K_r^d(\mathbf{x} - \mathbf{s}) - q_a K_a^d(\mathbf{x} - \mathbf{s})\right) \rho(\mathbf{s}, t) \mathrm{d}s_1, \end{aligned} \tag{31}$$

where $\mathbf{x} = (x_1, 0)$, $\rho(\mathbf{x}, t) = u^+(x_1, t) + u^-(x_1, t) = u(x_1, t)$, and where we used assumptions (i) and (ii). Hence, model (1)-(14) with communication mechanism (31) corresponds exactly to the 2D non-local kinetic model (18)-(22)-(24)-(25). This means, for instance, that the macroscopic 2D model (30) reduces to the heat equation for $\lambda_2 = 0$, which is not the case in the parabolic limit (15) of the corresponding 1D hyperbolic model (1) with the turning rates given by (2). In fact, our 2D scaling assumption $g_j(\vartheta) = \lambda_2^0 + \varepsilon G_j(\vartheta)$, $j = al, r, a$, introduces the relative strength of directed and non-directed turning kernels into the expression of the social response function $B[u]$, which is responsible for the appearance of a factor $\lambda_2$ in the drift of the macroscopic 2D model (30).

**Remark 7.** For some particular choices of distance kernels, the limiting parabolic model (30) can be reduced to well known equations. Let us assume, for example, that the distance kernels are constant on the whole domain,

$$K_j^d(\mathbf{x}) = 1, \quad j = al, a, r. \tag{32}$$

This assumption corresponds to a setting in which individuals interact equally well with all other individuals present in the entire domain. This is true locally for example if we have many individuals packed in little space. Under assumption (32) together with $\lambda_1 = 0$, model (30) simplifies to

$$\partial_t \rho = \frac{C_0}{\lambda_2} \Delta\rho + C_1 \nabla . \left(\rho \int_{\mathbb{R}^2} \mathbf{e}_\psi \rho(\mathbf{s}) \, \mathrm{d}\mathbf{s}\right),$$

where

$$\mathbf{e}_\psi = \frac{\mathbf{s} - \mathbf{x}}{|\mathbf{s} - \mathbf{x}|},$$

and $C_0, C_1$ are constants depending only on $\gamma, q_{al}, q_a, q_r$ and the total mass $\int \rho \, \mathrm{d}\mathbf{x}$. If $q_a = q_r$, then the attraction and repulsion forces cancel out ($C_1 = 0$) and we obtain the heat equation. Let us henceforth assume $q_a \neq q_r$. Furthermore, we can write the social flux as

$$\mathbf{k}[\rho] = \nabla W * \rho, \tag{33}$$

where the interaction potential $W : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is given by $W(\mathbf{x}) = C_1 |\mathbf{x}|$. In fact, for the more general distance kernels (21) the social flux can also be written in the form (33), with the interaction potential $W$ behaving like $|\mathbf{x}|$ close to zero and decaying

exponentially fast as $|\mathbf{x}| \longrightarrow \infty$ (e.g. Morse potentials). Therefore, we recover the diffusive aggregation equation

$$\partial_t \rho = \Delta \rho + \nabla \cdot (\rho \, (\nabla W * \rho)) \,,$$

which models the behaviour of particles interacting through a pairwise potential while diffusing with Brownian motion. This type of equation has received a lot of attention in recent years because of its ubiquity in modelling aggregation processes, such as collective behaviour of animals [50, 52, 9, 26] and bacterial chemotaxis [10] (see also the references therein).

3.2. **Grazing collision limit.** In the following, we consider another type of scaling that leads to parabolic equations, by focusing on the case where individuals turn only a small angle upon interactions with neighbours. This is biologically realistic as, for example, many migratory birds follow favourable winds or magnetic fields [53] and social interactions with neighbours might not have a considerable impact on directional changes of individuals. The so-called *grazing collisions*, i.e. collisions with small deviation, correspond to this assumption. In this case, we show that the Boltzmann-type equation (18) can be reduced to a Fokker-Planck equation with non-local advective and diffusive terms in the orientation space.

For simplicity, the 2D kinetic model (18) can be re-written as

$$\frac{\partial u}{\partial t} + \gamma e_\phi \nabla_x u = -Q^-[u] + Q^+[u,u],$$

with

$$Q^-[u] = Q_r^-[u] + Q_a^-[u] + Q_{al}^-[u], \quad Q^+[u,u] = Q_r^+[u,u] + Q_a^+[u,u] + Q_{al}^+[u,u],$$
$$Q_j^-[u] = \lambda_j(x,\phi)u, \quad Q_j^+[u,u] = \int_{-\pi}^{\pi} T_j(x,\phi',\phi)u(x,\phi',t)d\phi', \quad \text{for } j = r, al, a.$$

Let us focus for now only on the alignment interactions; the analysis of attraction and repulsion interactions is similar. The grazing collision assumption suggests that we can rescale the probability of re-orientation as follows:

$$\omega_{al}^\varepsilon(\phi - \phi', \phi - \theta) = \frac{1}{\varepsilon} g_\varepsilon\Big(\frac{\phi - \phi' - \varepsilon R(\phi - \theta)}{\varepsilon}\Big).$$

Here, the parameter $\varepsilon$ is related to the small re-orientation angle following interactions with neighbours moving in direction $\theta$. If we denote by $\varepsilon\beta = \phi - \phi' - \varepsilon R(\phi - \theta)$, then since $\omega_\varepsilon$ integrates to 1, we obtain:

$$1 = \int_{-\pi}^{\pi} \omega_{al}^\varepsilon(\phi - \phi', \phi - \theta)d\phi' = \int_{-\pi + \phi - R(\phi-\theta)}^{\pi + \phi - R(\phi-\theta)} g_\varepsilon(\beta)d\beta = \int_{-\pi}^{\pi} g_\varepsilon(\beta)d\beta,$$

by periodicity of $g_\varepsilon$.

Generally, when an interaction kernel in the Boltzmann equation presents a singularity point, the troubles are avoided by considering a weak formulation of the Boltzmann operator [40, 20]. Expanding $Q_{al}[u] := -Q_{al}^-[u] + Q_{al}^+[u,u]$, we obtain

for all $\psi \in C_c^\infty([-\pi,\pi])$,

$$\int_{-\pi}^{\pi} Q_{al}[u]\psi(\phi)d\phi = \eta_{al}\int_{-\pi}^{\pi}\left(\frac{1}{2\pi}\rho(x,t) - u(x,\phi,t)\right)\psi(\phi)d\phi + \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{\mathbb{R}^2}\lambda_3 q_{al}K_{al}^d(x-s)K_{al}^0(\theta,\phi)u(x,\phi,t)u(s,\theta,t)\cdot \int_{-\pi}^{\pi}\omega_{al}^\varepsilon(\phi-\phi',\phi-\theta)\big[\psi(\phi')-\psi(\phi)\big]d\phi' ds d\theta d\phi. \tag{35}$$

By substituting $\phi' = \phi - \varepsilon\beta - \varepsilon R(\phi-\theta)$ into the $\psi(\phi')$ term in (35), and then expanding in Taylor series about $\phi$ we obtain:

$$\int_{-\pi}^{\pi}\omega_{al}^\varepsilon(\phi-\phi',\phi-\theta)\Big[\psi(\phi')-\psi(\phi)\Big]d\phi' \approx \int_{-\pi}^{\pi} g_\varepsilon(\beta)\Big[\big(-\varepsilon\beta-\varepsilon R(\phi-\theta)\big)\frac{\partial\psi}{\partial\phi} + \frac{\varepsilon^2}{2}\big(\beta+R(\phi-\theta)\big)^2\frac{\partial^2\psi}{\partial\phi^2}\Big]d\beta.$$

Equation (35) can thus be approximated by

$$\int_{-\pi}^{\pi} Q_{al}[u]\psi(\phi)d\phi = \eta_{al}\int_{-\pi}^{\pi}\left(\frac{1}{2\pi}\rho(x,t) - u(x,\phi,t)\right)\psi(\phi)d\phi - \int_{-\pi}^{\pi}\frac{\partial}{\partial\phi}\Big[u(x,\phi,t)C_{al}^\varepsilon[u,x,\phi]\Big]\psi(\phi)d\phi + \int_{-\pi}^{\pi}\frac{\partial^2}{\partial\phi^2}\Big[u(x,\phi,t)D_{al}^\varepsilon[u,x,\phi]\Big]\psi(\phi)d\phi,$$

with the definitions

$$C_{al}^\varepsilon[u,x,\phi] = \int_{-\pi}^{\pi}\int_{\mathbb{R}^2}\lambda_3 q_{al}K_{al}^d(x-s)K_{al}^0(\theta,\phi)A_{al}^\varepsilon(\phi-\theta)u(s,\theta,t)d\theta ds,$$
$$D_{al}^\varepsilon[u,x,\phi] = \int_{-\pi}^{\pi}\int_{\mathbb{R}^2}\lambda_3 q_{al}K_{al}^d(x-s)K_{al}^0(\theta,\phi)B_{al}^\varepsilon(\phi-\theta)u(s,\theta,t)d\theta ds,$$

where

$$A_{al}^\varepsilon(\phi-\theta) = -\varepsilon\big(M_1(\varepsilon) + M_0(\varepsilon)R(\phi-\theta)\big),$$
$$B_{al}^\varepsilon(\phi-\theta) = \frac{\varepsilon^2}{2}\big(M_2(\varepsilon) + 2M_1(\varepsilon)R(\phi-\theta) + M_0(\varepsilon)R(\phi-\theta)^2\big),$$

and $M_n(\varepsilon) = \int_{-\pi}^{\pi}\beta^n g_\varepsilon(\beta)d\beta$, $n = 0,1,2$, denote the moment generating functions of $g_\varepsilon(\beta)$. In a similar manner we can approximate the attractive and repulsive non-local terms:

$$\int_{-\pi}^{\pi} Q_{r,a}[u]\psi(\phi)d\phi = \eta_{r,a}\int_{-\pi}^{\pi}\left(\frac{1}{2\pi}\rho(x,t) - u(x,\phi,t)\right)\psi(\phi)d\phi - \int_{-\pi}^{\pi}\frac{\partial}{\partial\phi}\Big(u(x,\phi,t)C_{r,a}^\varepsilon[u,x,\phi]\Big)\psi(\phi)d\phi + \int_{-\pi}^{\pi}\frac{\partial^2}{\partial\phi^2}\Big(u(x,\phi,t)D_{r,a}^\varepsilon[u,x,\phi]\Big)\psi(\phi)d\phi,$$

where

$$C^{\varepsilon}_{r,a}[u,x,\phi] = \int_{-\pi}^{\pi}\int_{\mathbb{R}^2} \lambda_3 q_{r,a} K^d_{r,a}(x-s) K^0_{r,a}(s,x,\phi) A^{\varepsilon}_{r,a}(s,x,\phi) u(s,\theta,t) ds d\theta,$$
$$D^{\varepsilon}_{r,a}[u,x,\phi] = \int_{-\pi}^{\pi}\int_{\mathbb{R}^2} \lambda_3 q_{r,a} K^d_{r,a}(x-s) K^0_{r,a}(s,x,\phi) B^{\varepsilon}_{r,a}(s,x,\phi) u(s,\theta,t) ds d\theta,$$
$$A^{\varepsilon}_{r,a}(s,x,\phi) = -\varepsilon(M_1(\varepsilon)M_0(\varepsilon)R(\phi-\psi_s)),$$
$$B^{\varepsilon}_{r,a}(s,x,\phi) = \frac{\varepsilon^2}{2}\Big[M_2(\varepsilon) + 2M_1(\varepsilon)R(\phi-\psi_s) + M_0(\varepsilon)R(\phi-\psi_s)^2\Big].$$

Therefore, the kinetic model (18) in the strong formulation can be approximated (when individuals turn only by a small angle upon interactions with their neighbours) by the following Fokker-Planck model that contains all three social interactions:

$$\frac{\partial u}{\partial t} + \gamma e_\phi \cdot \nabla_x u = \lambda_1 \left(\frac{1}{2\pi}\rho(x,t) - u(x,\phi,t)\right) + \frac{\partial}{\partial \phi}\Big[-uC^{\varepsilon}[u,x,\phi] + \frac{\partial}{\partial\phi}(uD^{\varepsilon}[u,x,\phi])\Big], \tag{36}$$

with $\lambda_1 = \eta_a + \eta_{al} + \eta_r$ and

$$C^{\varepsilon}[u,x,\phi] = C^{\varepsilon}_{al}[u,x,\phi] + C^{\varepsilon}_{r}[u,x,\phi] + C^{\varepsilon}_{a}[u,x,\phi],$$
$$D^{\varepsilon}[u,x,\phi] = D^{\varepsilon}_{al}[u,x,\phi] + D^{\varepsilon}_{r}[u,x,\phi] + D^{\varepsilon}_{a}[u,x,\phi].$$

While non-local 2D Fokker-Planck models have been introduced in the past years in connection to self-organised aggregations, the majority of these models consider local diffusion [28, 3]. If we neglect the $\varepsilon^2$ terms (i.e., $B^{\varepsilon} \approx 0$) and assume $\lambda_1 = 0$, equation (36) reduces to a Vlasov-type flocking equation:

$$\frac{\partial u}{\partial t} + \gamma e_\phi \cdot \nabla_x u + \frac{\partial}{\partial \phi}\Big[uC^{\varepsilon}[u,x,\phi]\Big] = 0.$$

These type of models have been previously derived from individual-based models (Vicsek or Cucker-Smale models) with or without noise [28, 41, 20].

## 4. Asymptotic Preserving Methods for 1D models.

The kind of diffusion asymptotics we employed in the previous sections have been numerically investigated in [21] using so-called asymptotic preserving (AP) schemes. The AP methods, which improve the scheme already proposed in [39], are a fully explicit variation of the methods introduced in [46, 47]. They are a powerful tool to investigate how patterns are preserved in the parabolic limit by providing numerical schemes for all intermediate models of a scaling process given some scaling parameter $\varepsilon > 0$, and naturally produce a suitable numerical method for the limiting model as $\varepsilon \to 0$. Here, we apply these schemes only to the 1D models introduced in Section 2, since the numerics become much more complex in two dimensions. Taking advantage of our understanding of the limit process, we base our scheme on a splitting strategy with a convective-like step involving the transport part of the operator and an explicitly solvable ODE step containing stiff sources (see Section 4.2).

4.1. **Odd and Even Parity.** We consider the 1D kinetic model (1) written as an odd-even decomposition,

$$\begin{cases} \partial_t r + \gamma \partial_x j &= 0, \\ \partial_t j + \gamma \partial_x r &= -2\lambda^+[r,j](r+j) + 2\lambda^-[r,j](r-j), \end{cases}$$

with the equilibrium part (macro part/even part) $r$ and the non-equilibrium part (micro part/odd part) $j$ given by

$$r(x,t) = \frac{1}{2}\left(u^+(x,t) + u^-(x,t)\right), \quad j(x,t) = \frac{1}{2}\left(u^+(x,t) - u^-(x,t)\right).$$

Under scaling assumption (10) for (2), this model reads in the new variables $x = \tilde{x}/\varepsilon, t = \tilde{t}/\varepsilon^2$ as follows:

$$\begin{aligned} \varepsilon\partial_{\tilde{t}}\tilde{r} + \gamma\partial_{\tilde{x}}\tilde{j} &= 0 \\ \varepsilon\partial_{\tilde{t}}\tilde{j} + \gamma\partial_{\tilde{x}}\tilde{r} &= \tilde{r}\lambda_3(f[\tilde{y}^-] - f[\tilde{y}^+]) \\ &\quad - \frac{1}{\varepsilon}\tilde{j}\left(2\lambda_1 + 4\varepsilon\lambda_2 f\left(\tilde{K}^N * \tilde{r}\right) + \varepsilon\lambda_3(f[\tilde{y}^+] + f[\tilde{y}^-])\right), \end{aligned}$$

where $\tilde{K}^N(\tilde{x}) = \frac{1}{\varepsilon}K^N(\frac{\tilde{x}}{\varepsilon})$. Rearranging the terms and dropping "$\sim$" for notational convenience, we obtain for $r$ and $J := \frac{1}{\varepsilon}j$:

$$\begin{cases} \partial_t r + \gamma\partial_x J = 0 \\ \partial_t J + \gamma\partial_x r = \dfrac{1}{\varepsilon^2} r\lambda_3(f[y^-] - f[y^+]) + \left(1 - \dfrac{1}{\varepsilon^2}\right)\gamma\partial_x r \\ \qquad\qquad - \dfrac{1}{\varepsilon^2} J\left(2\lambda_1 + 4\varepsilon\lambda_2 f\left(K^N * r\right) + \varepsilon\lambda_3(f[y^+] + f[y^-])\right). \end{cases} \tag{37}$$

4.2. **Operator Splitting.** We can now employ an operator splitting method on (37), separating the stiff source part, which can be treated by an implicit Euler method, and the transport part, which we can solve by an explicit method such as upwinding:

1. **Stiff source part:**

$$\begin{aligned} \partial_t r &= 0, \\ \partial_t J &= \frac{1}{\varepsilon^2} r\lambda_3(f[y^-] - f[y^+]) + \left(1 - \frac{1}{\varepsilon^2}\right)\gamma\partial_x r \\ &\quad - \frac{1}{\varepsilon^2} J\left(2\lambda_1 + 4\,\varepsilon\,\lambda_2 f\left(K^N * r\right) + \varepsilon\,\lambda_3(f[y^+] + f[y^-])\right). \end{aligned} \tag{38}$$

2. **Transport part:**

$$\begin{aligned} \partial_t r + \gamma\,\partial_x J &= 0, \\ \partial_t J + \gamma\,\partial_x r &= 0. \end{aligned} \tag{39}$$

It can easily be verified that, in the limit $\varepsilon \to 0$, we recover indeed the macroscopic model (11) for $u = 2r$.

4.3. **Alternated Upwind Discretisation.** In the following, we are interested in the numerical implementation of model (1) with the turning rates (2) depending on a non-linear turning function $f$ without a non-directed density-dependent turning term (i.e. $\lambda_2 = 0$). As shown in Section 2.1, in this case, the parabolic limit yields the drift-diffusion equation (11)

$$\partial_t u = D_0\partial_{xx}u - B_0\partial_x\left(u(f^-[u] - f^+[u])\right),$$

with $D_0 = \gamma^2/(2\lambda_1)$ and $B_0 = \lambda_3\gamma/(2\lambda_1)$. Note the shortcut notation $f^{\pm}[u] = f(y_D^{\pm}[u])$. We propose an alternated upwind discretisation with the even part $r$ evaluated at full grid points $x_i = i\,\Delta x$, and the odd part $J$ evaluated at half grid points $x_{i+\frac{1}{2}} = (i+\frac{1}{2})\,\Delta x$. First, we discretise the stiff source part (38) using an implicit Euler discretisation and respecting the direction of the drift. We obtain an explicit expression for $J^*$,

$$J^*_{i+\frac{1}{2}} = \frac{\varepsilon^2 J^n_{i+\frac{1}{2}} + \gamma\frac{\Delta t}{\Delta x}\left(\varepsilon^2-1\right)\left(r^n_{i+1}-r^n_i\right)}{\varepsilon^2 + 2\lambda_1\Delta t + \varepsilon\lambda_3\Delta t\left(f^+[r^n]+f^-[r^n]\right)_{i+\frac{1}{2}}} + \frac{\lambda_3\Delta t\left(\left(f^-[r^n]-f^+[r^n]\right)^+_{i+\frac{1}{2}}\, r^n_i + \left(f^-[r^n]-f^+[r^n]\right)^-_{i+\frac{1}{2}}\, r^n_{i+1}\right)}{\varepsilon^2 + 2\lambda_1\Delta t + \varepsilon\lambda_3\Delta t\left(f^+[r^n]+f^-[r^n]\right)_{i+\frac{1}{2}}},$$

with $r^* = r^n$. Here, $r^n$ and $J^n$ are the numerical solutions of $r$ and $J$ at time $t_n = n\Delta t$. We use the "$*$"-notation for half steps in time. Since $J$ is evaluated at half grid point, the discretisation of the transport part (39) can be chosen independently of the sign of the drift,

$$\frac{1}{\Delta t}\left(r^{n+1}_i - r^*_i\right) + \frac{1}{\Delta x}\left(J^*_{i+\frac{1}{2}} - J^*_{i-\frac{1}{2}}\right) = 0,$$
$$\frac{1}{\Delta t}\left(J^{n+1}_{i+\frac{1}{2}} - J^*_{i+\frac{1}{2}}\right) + \frac{1}{\Delta x}\left(r^*_{i+1} - r^*_i\right) = 0.$$

Taking the limit $\varepsilon \to 0$ in the expression for $J^*_{i+\frac{1}{2}}$ and substituting into the first equation of the transport part, we obtain the following discretisation of the one-dimensional macroscopic model (11):

$$\frac{u^{n+1}_i - u^n_i}{\Delta t} = \frac{D_0}{(\Delta x)^2}\left(\partial^{(c)}_{xx}u^n\right)_i$$
$$- \frac{B_0}{\Delta x}\left(u^n_i\left(f^-[r^n]-f^+[r^n]\right)^+_{i+\frac{1}{2}} - u^n_{i-1}\left(f^-[r^n]-f^+[r^n]\right)^+_{i-\frac{1}{2}}\right)$$
$$- \frac{B_0}{\Delta x}\left(u^n_{i+1}\left(f^-[r^n]-f^+[r^n]\right)^-_{i+\frac{1}{2}} - u^n_i\left(f^-[r^n]-f^+[r^n]\right)^-_{i-\frac{1}{2}}\right).$$

Here, $\partial^{(c)}_{xx}u^n$ denotes the standard central difference discretisations. This illustrates how the choice of discretisation for (38) directly induces a discretisation of model (11). We will now use this scheme to investigate how some of the patterns observed in model (1)-(2) change as $\varepsilon \to 0$.

**Remark 8.** The stability restriction for the proposed AP scheme is less clear. We can expect that the time steps size $\Delta t$ needs to be sufficiently small, with an upper stability bound depending on the space step size $\Delta x$, the diffusion coefficient $D_0$, and the social interaction kernels via the terms $K^N * u$ and $f^{\pm}[u]$.

4.4. **Simulation results.** In Section 2.2 we have seen that for model M4, the two Hopf bifurcations that occurred for the $k_4$ and $k_5$ modes have disappeared as $\varepsilon \to 0$. In this Section, we start with a rotating wave pattern (i.e., travelling pulses) that arises at $\varepsilon = 1$ through a Hopf bifurcation (i.e., for the same parameter values as in Figure 4: $q_a = 1.545$, $q_r = 2.779$, $\lambda_1 = 0.2$, $\lambda_2 = 0$, $\lambda_3 = 0.9$, $\gamma = 0.1$, $A = 2$). Then, we investigate numerically what happens with this pattern as $\varepsilon \to 0$. The initial conditions for the simulations are random perturbations – of maximum amplitude 0.2 – of the spatially homogeneous steady state $u^* = A/2 = 1$. We start with $\varepsilon = 1$,

and run the numerical simulations up to $t = 1000$. Then we decrease $\varepsilon$, and choose the new initial condition to be the final solution obtained with the previous $\varepsilon$ value.

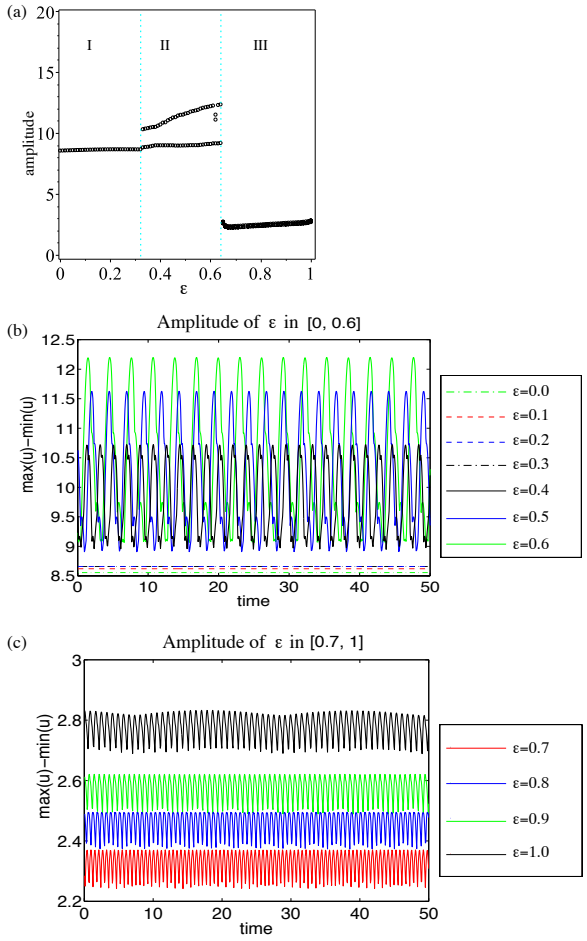


FIGURE 6. The amplitude and density profile of the patterns obtained for $q_a = 1.545$, $q_r = 2.779$, $q_{al} = 0$, $\lambda_1 = 0.2$, $\lambda_2 = 0$, $\lambda_3 = 0.9$ with model M4, as $\varepsilon$ is decreased from 1.0 to 0.0. (a) Bifurcation diagram for the amplitude of the patterns as a function of $\varepsilon$. For $\varepsilon \leq 0.32$ (region I), the amplitude is constant. For $\varepsilon \in (0.32, 0.64)$ (region II) the amplitude oscillates between two different values. For $\varepsilon \geq 0.64$ (region III) there are some very small oscillations in the amplitude, however due to the scale of the plot these oscillations are almost unobservable. (b) Amplitude of the patterns for $\varepsilon \in [0, 0.6]$ and for $t \in (0, 50)$. We show here $\max_{x\in[0,L]} u(x,t) - \min_{x\in[0,L]} u(x,t)$, with $u = u^+ + u^-$. (c) Amplitude of the patterns for $\varepsilon \in [0.7, 1.0]$ and for $t \in (0, 50)$.

Figure 6(a) shows the amplitude of the patterns obtained when $\varepsilon \in [0, 1]$, for the particular parameter values mentioned before. Since some of these amplitudes show

time-oscillations between different values, we graph their maximum and minimum values for each $\varepsilon$. As we decrease $\varepsilon$ from 1.0 towards 0.64 (region III), the amplitude undergoes some very small temporal oscillations (see also panel (c)), corresponding to the rotating wave patterns (with a small time-modulation) shown in Figure 7(c). For $\varepsilon \in (0.32, 0.64)$ (region II), the amplitude oscillates between two large values. This corresponds to the "inside-group" zigzagging behaviour shown in Figure 7(b) near $x = 6$, where the group as a whole does not move in space but individuals inside the group move between the left and right edges of the group. We also note a period-doubling bifurcation at $\varepsilon = 0.61$ (region II, Figure 6(a); see the two dots that appear between the main branches), which leads to a slight decrease in the amplitude. Finally, as $\varepsilon$ is decreased below 0.32 (region I), the movement inside the group is lost and the pattern is described by stationary pulses with fixed amplitude (see Figure 6(a) and Figure 7(a)). Figures 6(b),(c) show the time-variation of the amplitudes of the spatial and spatiotemporal patterns obtained for $\varepsilon \in [0, 1]$. Figures 7(a')-(c') show the density profiles of the patterns observed in regions I-III.

Because the macro-scale models ($\varepsilon = 0$) seem to exhibit stationary pulses (as shown in Figure 7(a)), we now start with these stationary pulses (for $\varepsilon = 1$) and investigate whether they change in any way as $\varepsilon \to 0$. We focus here on model M2 (see Figure 3). Figure 8 shows the amplitude of the stationary pulses obtained with model M2 in a particular parameter region ($q_a = 2.2$, $q_r = 0.93$, $q_{al} = 0$; see also Figure 4), as we decrease the scaling parameter $\varepsilon$. We observe that in this case, the scaling does not affect the patterns or their amplitudes.

**Remark 9.** Note that the rotating wave pattern shown in Figure 7(c) for $\varepsilon = 1$ is obtained near a Hopf/steady-state bifurcation (with $k_5$ the Hopf wavenumber), and hence the 5 rotating peaks that form this pattern. However, as $\varepsilon \to 0$, the wavenumber $k_3$ seems to become unstable (hence the 3 peaks for the patterns shown in Figure 7(a),(b)), even if the dispersion relation shown in Figure 4(b) suggests that $k_3$ should be stable.

5. **Summary and Discussion.** In this study, we investigated the connections between a class of 1D and 2D non-local kinetic models and their limit macroscopic models for self-organised biological aggregations. The non-locality of these models was the result of the assumptions that individuals can interact with neighbours positioned further away, but still within their perception range. To simplify the kinetic models that incorporate microscopic-level interactions (such as individuals' speed and turning rates), we focused on two types of scalings, namely a parabolic and a grazing collision limit, which lead to parabolic models described in terms of average speed and average turning behaviour. We showed that while for the kinetic models the non-local interactions influence the turning rates (i.e., individuals turn to approach their neighbours, to move away from them or to align with them), for the limit parabolic models the non-local interactions influence the dispersion and the drift of the aggregations. In particular, we showed that the assumption that individuals can turn randomly following the non-directional perception of neighbours around them leads, in the macroscopic scaling, to density-dependent diffusion. Moreover, this diffusion decreased with the increase in the population density. Biologically, this means that larger animal groups are less likely to spread out. This phenomenon has been observed for various species. For example, studies have

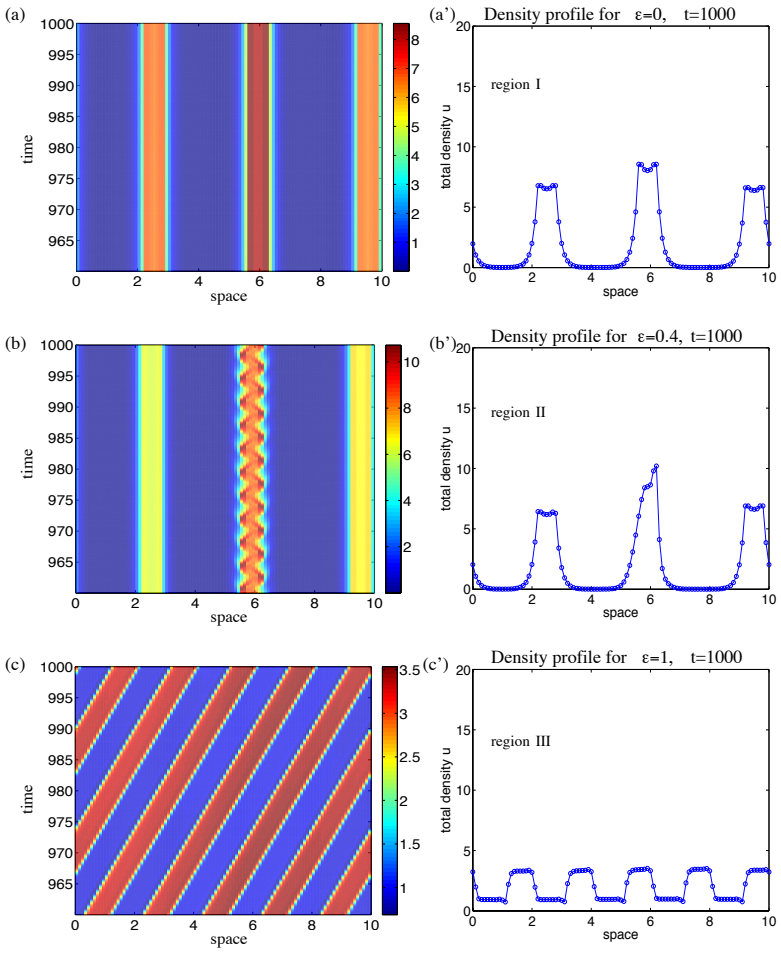


FIGURE 7. The spatial and spatio-temporal patterns obtained with model M4, for $q_a = 1.545$, $q_r = 2.779$, $q_{al} = 0$, $\lambda_1 = 0.2$, $\lambda_2 = 0$, $\lambda_3 = 0.9$, as $\varepsilon$ is decreased from 1.0 to 0.0, using model M4. (a) Stationary pulse patterns observed in region I: $\varepsilon \leq 0.32$; (b) "Inside-group" zigzag patterns observed in region II: $\varepsilon \in (0.32, 0.64)$; (c) Rotating wave (traveling pulse) patterns observed in region III: $\varepsilon \geq 0.64$. Panels (a')-(c') show the density profiles corresponding to patterns in panels (a)-(c), at time $t = 1000$.

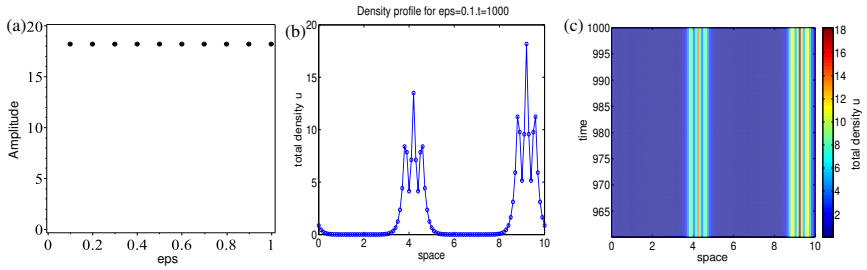


FIGURE 8. The amplitude and density of the patterns obtained for model M2 with $q_a = 2.2$, $q_r = 0.93$, $q_{al} = 0$, $\lambda_1 = 0.2$, $\lambda_2 = 0$, $\lambda_3 = 0.9$, as $\varepsilon$ is decreased from 1.0 to 0.0. (a) Bifurcation diagram for the amplitude of the patterns as a function of $\varepsilon$. (b) Density profile for the stationary patterns. (c) Time-space plot of the density.

shown that aggregations of locusts [13] or ants [4] can persist only if the number of individuals is above a certain threshold.

The introduction in (2) of the term $y_N$ describing random non-directional turning (which generalised the turning rates in [33]) was required by the comparison of the parabolic limit models in 1D and 2D. In particular, the 2D parabolic limit lead to the natural appearance of this term, which is absent from the 1D parabolic model. Therefore, to obtain similar parabolic models in 1D and 2D, we had to explicitly add $y_N$ in equation (2). This suggests that even if the 2D model (18) can be reduced to a special case of the 1D model (1) (as shown in [36]) there are more subtle differences between these nonlocal 1D and 2D models. These differences can impact the types of patterns displayed by the 2D models – an aspect that we will address in a future study.

Next, we investigated how two types of patterns (i.e., travelling and stationary aggregations) displayed by the 1D kinetic models, were preserved in the limit to macroscopic parabolic models. To this end, we first investigated the local stability of spatially homogeneous patterns characterised by individuals spread evenly over the domain, and showed that local Hopf bifurcations are lost in the parabolic limit. These Hopf bifurcations give rise to travelling aggregations (i.e., rotating waves). We then tested this observation numerically, with the help of asymptotic preserving methods. We started with a rotating wave pattern obtained near a Hopf/Steady-state bifurcation for $\varepsilon = 1$ (1D kinetic model; see Figure 7(c)), and studied numerically how does this pattern change when $\varepsilon \to 0$ (1D parabolic model; see Figure 7(a)). By graphing in Figure 6(a) the amplitude of the resulting patterns as the scaling parameter $\varepsilon$ is decreased from $\varepsilon = 1$ to $\varepsilon = 0$, we showed that there were two major transitions. The first transition occurred around $\varepsilon = 0.64$, when the travelling (rotating) groups stopped moving. We note, however, that while the group as a whole was stationary, the individuals inside the group were still moving between the left- and right-edges of the group, leading to an "inside-group" zigzagging behaviour. The second transition occurred around $\varepsilon = 0.32$, when the individuals inside the groups stopped moving, leading to stationary pulses.

We emphasise here that this study is one of the first in the literature to investigate numerically the transitions between different aggregation patterns, as a scaling parameter $\varepsilon$ is varied from values corresponding to mesoscale dynamics ($\varepsilon = 1$) to

values corresponding to macroscale dynamics ($\varepsilon = 0$). Understanding these transitions is important when investigating biological phenomena that occur on multiple scales, since it allows us to make decisions regarding the models that are most suitable to reproduce the observed dynamics.

In this study we investigated the preservation of patterns via the 1D parabolic limit, but similar investigations could be performed for the grazing collision limit. Moreover, as shown previously [32], model (1) can display many more types of complex spatio-temporal patterns then the two types of patterns investigated here. We focused on travelling and stationary aggregations since our aim here was not to investigate how all possible patterns are preserved by all these different scaling approaches. Rather, it was to show that by taking these asymptotic limits, some patterns could be lost. Therefore, even if the macroscopic models are simpler to investigate, they might not exhibit the same patterns as the kinetic models. Our analysis aimed at highlighting the usefulness of asymptotic preserving numerical methods to understand the bifurcation of the solutions as one investigates the transition from mesoscopic-level to macroscopic-level aggregation dynamics.

**Acknowledgments.** JAC, RE and FH acknowledge support from the Engineering and Physical Sciences Research Council (UK) grants number EP/K008404/1, EP/K033689/1 and EP/H023348/1 respectively. JAC acknowledges support from projects MTM2011-27739-C04-02 and the Royal Society through a Wolfson Research Merit Award.

## References

[1] E. D. Angelis and B. Lods, On the kinetic theory for active particles: a model for tumor-immune system competition, *Mathematical and Computer Modeling*, **47** (2008), 196–209.
[2] A. Arnold, J. A. Carrillo, I. Gamba and C.-w. Shu, Low and high field scaling limits for the vlasov- and wigner-poisson-fokker-planck systems, *Transp. Theory Stat. Phys.*, **30** (2001), 121–153.
[3] A. Barbaro and P. Degond, Phase transition and diffusion among socially interacting self-propelled agents, *Discrete Cont Dyn Syst B. In Press.*
[4] M. Beekman, D. J. T. Sumpter and F. L. W. Ratnieks, Phase transitions between disordered and ordered foraging in pharaoh's ants, *Proc. Natl. Acad. Sci.*, **98** (2001), 9703–9706.
[5] N. Bellomo, E. D. Angelis and L. Preziosi, Multiscale modeling and mathematical problems related to tumor evolution and medical therapy, *Journal of Theoretical Medicine*, **5** (2003), 111–136.
[6] N. Bellomo, A. Bellouquid, J. Nieto and J.Soler, Multicellular biological growing systems: hyperbolic limits towards macroscopic description, *Mathematical Models and Methods in Applied Sciences*, **17** (2007), 1675–1693.
[7] N. Bellomo, C. Bianca and M. Delitala, Complexity analysis and mathematical tools towards the modelling of living systems, *Physics of Life Reviews*, **6** (2009), 144–175.
[8] M. G. Bertotti and M. Delitala, Conservation laws and asymptotic behavior of a model of social dynamics, *Nonlinear Analysis RWA*, **9** (2008), 183–196.
[9] A. L. Bertozzi, J. A. Carrillo and T. Laurent, Blow-up in multidimensional aggregation equations with mildly singular interaction kernels, *Nonlinearity*, **22** (2009), 683–710.
[10] A. Blanchet, J. Dolbeault and B. Perthame, Two-dimensional Keller-Segel model: optimal critical mass and qualitative properties of the solutions, *Electron. J. Differential Equations*, No. 44, 32 pp. (electronic).
[11] M. Bodnar and J. J. L. Velazquez, Derivation of macroscopic equations for individual cell-based models: a formal approach, *Math. Meth. Appl. Sci.*, **28** (2005), 1757–1779.
[12] R. Breitwisch and G. Whitesides, Directionality of singing and non-singing behaviour of mated and unmated Northern Mockingbirds, *Mimus polyglottos*, *Anim. Behav.*, **35** (1987), 331–339.
[13] J. Buhl, D. J. T. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller and S. J. Simpson, From disorder to order in marching locusts, *Science*, **312** (2006), 1402–1406.

[14] P.-L. Buono and R. Eftimie, Analysis of Hopf/Hopf bifurcations in nonlocal hyperbolic models for self-organised aggregations, *Math. Models Methods Appl. Sci.*, **24** (2014), 327.
[15] P.-L. Buono and R. Eftimie, Codimension-two bifurcations in animal aggregation models with symmetry, *SIAM J. Appl. Dyn. Syst.*
[16] P.-L. Buono and R. Eftimie, Symmetries and pattern formation in hyperbolic versus parabolic models for self-organised aggregations, *J. Math. Biol.*
[17] M. Burger, V. Capasso and D. Morale, On an aggregation model with long and short range interactions, *Nonlinear Analysis: Real World Applications*, **8** (2007), 939–958.
[18] J. A. Carrillo, M. R. D'Orsogna and V. Panferov, Double milling in self-propelled swarms from kinetic theory, *Kinetic and Related Models*, **2** (2009), 363–378.
[19] J. A. Carrillo, M. Fornasier, J. Rosado and G. Toscani, Asymptotic flocking dynamics for the kinetic Cucker-Smale model, *SIAM J. Math. Anal.*, **42** (2010), 218–236.
[20] J. A. Carrillo, M. Fornasier, G. Toscani and F. Vecil, Particle, kinetic, and hydrodynamic models of swarming, *Mathematical Modelling of Collective Behavior in Socio-Economic and Life Sciences*, 297–336.
[21] J. A. Carrillo, T. Goudon, P. Lafitte and F. Vecil, Numerical schemes of diffusion asymptotics and moment closures for kinetic equations, *J. Sci. Comput.*, **36** (2008), 113–149, URL http://dx.doi.org/10.1007/s10915-007-9181-5.
[22] J. A. Carrillo, Y. Huang and S. Martin, Explicit flock solutions for quasi-morse potentials, *to appear in European Journal of Applied Mathematics*.
[23] J. A. Carrillo, A. Klar, S. Martin and S. Tiwari, Self-propelled interacting particle systems with roosting force, *Math. Models Methods Appl. Sci.*, **20** (2010), 1533.
[24] J. A. Carrillo and B. Yan, An asymptotic preserving scheme for the diffusive limit of kinetic systems for chemotaxis, *Multiscale Model. Simul.*, **11** (2013), 336–361, URL http://dx.doi.org/10.1137/110851687.
[25] A. Chertock, A. Kurganov, A. Polizzi and I. Timofeyev, Pedestrian flow models with slowdown interactions, *Math Models Methods Appl. Sci.*, **24** (2014), 249–275.
[26] Y.-L. Chuang, M. R. D'Orsogna, D. Marthaler, A. L. Bertozzi and L. S. Chayes, State transitions and the continuum limit for a 2d interactiong, self-propelled particle system, *Physica D*, **232** (2007), 33–47.
[27] P. Degond, G. Dimarco and T. Mac, Hydrodynamics of the Kuramoto-Vicsek model of rotating self-propelled particles, *Mathematical Models and Methods in Applied Sciences*, **24** (2014), 277–325.
[28] P. Degond and S. Motsch, Macroscopic limit of self-driven particles with orientation interaction, *C.R. Acad. Sci. Paris Ser. I*, **345** (2007), 555–560.
[29] P. Degond and S. Motsch, Large scale dynamics of the persistent turning awlker model of fish behaviour, *J. Stat. Phys.*, **131** (2008), 989–1021.
[30] R. Eftimie, *Modeling group formation and activity patterns in self-organizing communities of organisms*, PhD thesis, University of Alberta, 2008.
[31] R. Eftimie, Hyperbolic and kinetic models for self-organized biological aggregations and movement: a brief review, *J. Math. Biol.*, **65** (2012), 35–75.
[32] R. Eftimie, G. de Vries and M. A. Lewis, Complex spatial group patterns result from different animal communication mechanisms, *Proc. Natl. Acad. Sci.*, **104** (2007), 6974–6979.
[33] 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** (2007), 1537–1566.
[34] R. Eftimie, G. de Vries and M. Lewis, Weakly nonlinear analysis of a hyperbolic model for animal group formation, *J. Math. Biol.*, **59** (2009), 37–74.
[35] R. Eftimie and U. of Alberta (Canada)., *Modeling Group Formation and Activity Patterns in Self-organizing Communities of Organisms*, Canadian theses, University of Alberta (Canada), 2008, URL http://books.google.co.uk/books?id=YKSsnBUzYToC.
[36] R. Fetecau, Collective behavior of biological aggregations in two dimensions: a nonlocal kinetic model, *Math. Model. Method. Appl. Sci.*, **21** (2011), 1539.
[37] E. Geigant, K. Ladizhansky and A. Mogilner, An integrodifferential model for orientational distributions of F-actin in cells, *SIAM J. Appl. Math.*, **59** (1998), 787–809.
[38] E. Geigant, K. Ladizhansky and A. Mogilner, An integrodifferential model for orientational distributions of $F$-actin in cells, *SIAM J. Appl. Math.*, **59** (1999), 787–809 (electronic).
[39] P. Godillon-Lafitte and T. Goudon, A coupled model for radiative transfer: Doppler effects, equilibrium, and nonequilibrium diffusion asymptotics, *Multiscale Model. Simul.*, **4** (2005), 1245–1279, URL http://dx.doi.org/10.1137/040621041.

[40] T. Goudon, On Boltzmann equations and Fokker-Plank asymptotics: influence of grazing collisions, *J. Stat. Phys.*, **89** (1997), 751–776.
[41] S.-Y. Ha and E. Tadmor, From particle to kinetic and hydrodynamic descriptions of flocking, *Kinetic and Related Models*, **1** (2008), 415–435.
[42] C. K. Hemelrijk and H. Kunz, Density distribution and size sorting in fish schools: an individual-based model, *Behav. Ecol.*, **16** (2004), 178–187.
[43] H. Hildenbrandt, C. Carere and C. K. Hemelrijk, Self-organised complex aerial displays of thousands of starlings: a model, *Behavioral Ecology*, **107** (2010), 1349–1359.
[44] T. Hillen and H. G. Othmer, The diffusion limit of transport equations derived from velocity jump process, *SIAM J.Appl.Math.*, **61** (2000), 751–775.
[45] E. E. Holmes, Are diffusion models too simple? A comparison with telegraph models of invasion, *Am. Nat.*, **142** (1993), 779–795.
[46] A. Klar, An asymptotic-induced scheme for nonstationary transport equations in the diffusive limit, *SIAM J. Numer. Anal.*, **35** (1998), 1073–1094 (electronic), URL http://dx.doi.org/10.1137/S0036142996305558.
[47] A. Klar, An asymptotic preserving numerical scheme for kinetic equations in the low Mach number limit, *SIAM J. Numer. Anal.*, **36** (1999), 1507–1527 (electronic), URL http://dx.doi.org/10.1137/S0036142997321765.
[48] R. Larkin and R. Szafoni, Evidence for widely dispersed birds migrating together at night, *Integrative and comparative biology*, **48** (2008), 40–49.
[49] A. Mogilner and L. Edelstein-Keshet, A non-local model for a swarm, *J. Math. Biol.*, **38** (1999), 534–570.
[50] A. Mogilner and L. Edelstein-Keshet, A non-local model for a swarm, *J. Math. Biol.*, **38** (1999), 534–570, URL http://dx.doi.org/10.1007/s002850050158.
[51] A. Mogilner, L. Edelstein-Keshet and G. B. Ermentrout, Selecting a common direction. II. Peak-like solutions representing total alignment of cell clusters, *J. Math. Biol.*, **34** (1996), 811–842.
[52] D. Morale, V. Capasso and K. Oelschläger, An interacting particle system modelling aggregation behavior: from individuals to populations, *J. Math. Biol.*, **50** (2005), 49–66, URL http://dx.doi.org/10.1007/s00285-004-0279-1.
[53] I. Newton, *The migration ecology of birds*, Academic Press, Elsevier, 2008.
[54] H. G. Othmer and T. Hillen, The diffusion limit of transport equations II: chemotaxis equations, *SIAM J.Appl.Math.*, **62** (2002), 1222–1250.
[55] R. D. Passo and P. de Mottoni, Aggregative effects for a reaction-advection equation, *J. Math. Biology*, **20** (1984), 103–112.
[56] B. Pfistner, A one dimensional model for the swarming behaviour of Myxobacteria, in *Biological Motion, Lecture Notes on Biomathematics, 89* (eds. W. Alt and G. Hoffmann), Springer, 1990, 556–563.
[57] J. Saragosti, V. Calvez, N. Bournaveas, A. Buguin and P. S. B. Perthame, Mathematical description of bacterial traveling pulses, *PLOS Computational Biology*, **6** (2010), e1000890.
[58] J. Saragosti, V. Calvez, N. Bournaveas, A. Buguin, P. Silberzan and B. Perthame, Mathematical description of bacterial traveling pulses, *PLoS Comput. Biol.*, **6** (2010), e1000890, 12, URL http://dx.doi.org/10.1371/journal.pcbi.1000890.
[59] C. M. Topaz, A. L. Bertozzi and M. A. Lewis, A nonlocal continuum model for biological aggregation, *Bull. Math. Bio.*, **68** (2006), 1601–1623.
[60] F. Venuti, L. Bruno and N. Bellomo, Crowd dynamics on a moving platform: mathematical modelling and application to lively footbridges, *Mathematical and Computer Modelling*, **45** (2007), 252–269.

Received xxxx 20xx; revised xxxx 20xx.

*E-mail address*: carrillo@imperial.ac.uk
*E-mail address*: reftimie@maths.dundee.ac.uk
*E-mail address*: fkoh2@cam.ac.uk