POLITECNICO DI TORINO
Repository ISTITUZIONALE

Macroscopic descriptions of follower-leader systems

Original
Macroscopic descriptions of follower-leader systems / Bernardi, S., Estrada-Rodriguez, G., Gimperlein, H., Painter, K.J.. - In: KINETIC AND RELATED MODELS. - ISSN 1937-5093. - 14:6(2021), pp. 981-1002. [10.3934/krm.2021035]

Availability:
This version is available at: 11583/2948265 since: 2022-01-03T16:56:56Z

Publisher:
American Institute of Mathematical Sciences

Published
DOI:10.3934/krm.2021035

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

26 November 2026

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

doi:10.3934/xx.xx.xx.xx

pp. **X–XX**

MACROSCOPIC DESCRIPTIONS OF FOLLOWER-LEADER SYSTEMS

Sara Bernardi

Department of Mathematical Sciences "G. L. Lagrange", Politecnico di Torino
Corso Duca degli Abruzzi 24, 10129 Torino, Italy

Gissell Estrada-Rodriguez
Laboratoire Jacques-Louis Lions, Sorbonne-Université, 4, pl. Jussieu, F-75005 Paris, France

Heiko Gimperlein*

Maxwell Institute for Mathematical Sciences and Department of Mathematics, Heriot-–Watt University
Edinburgh, EH14 4AS, United Kingdom
and Institute for Mathematics, University of Paderborn, Warburger Str. 100, 33098 Paderborn, Germany

Kevin J. Painter

Interuniversity Department of Regional and Urban Studies and Planning, Politecnico di Torino
Torino, 10125, Italy

(Communicated by the associate editor name)

Abstract. The fundamental derivation of macroscopic model equations to describe swarms based on microscopic movement laws and mathematical analyses into their self-organisation capabilities remains a challenge from the perspective of both modelling and analysis. In this paper we clarify relevant continuous macroscopic model equations that describe follower-leader interactions for a swarm where these two populations are fixed. We study the behaviour of the swarm over long and short time scales to shed light on the number of leaders needed to initiate swarm movement, according to the homogeneous or inhomogeneous nature of the interaction (alignment) kernel. The results indicate the crucial role played by the interaction kernel to model transient behaviour.

1. **Introduction.** Collective movements describe the tendency of a group of individuals to coordinate their motion in a manner that generates net flow of the entire population. Examples range from cells to animals, from migrating cell clusters during development and cancer invasion [20, 23] to shifting bird flocks and fish shoals; the latter extend to the kilometre-spanning shoals formed from hundreds of millions of herrings [22]. A point of recent interest concerns the potential division of a population into "leaders" and "followers" and, consequently, how leaders influence swarm dynamics. Clearcut leadership could result from experience, age or prior

2010 *Mathematics Subject Classification.* Primary: 92D25; Secondary: 92D50; 35Q92; 82C22.
Key words and phrases. Follower-leader systems, interacting particle systems, velocity-jump model, macroscopic limit, swarming, bees.

G. E. R. was supported by The Maxwell Institute Graduate School in Analysis and its Applications, a Centre for Doctoral Training funded by the UK Engineering and Physical Sciences Research Council (grant EP/L016508/01), the Scottish Funding Council, Heriot-Watt University and the University of Edinburgh.

* Corresponding author: Heiko Gimperlein.

knowledge: the presence of older birds improves path efficiency in migrating cranes [25]; post-menopausal orcas adopt leading positions during pod foraging [7]; only the "househunting" scouts know the final destination of a new honeybee nest [31]. Subtler leaders arise within superficially identical populations, such as the presence of faster or "braver" individuals in fish shoals and bird flocks [30, 26]. Leader/follower statuses also occur in a host of cellular systems, ranging from collective movements of aggregated amoebae to sprouting blood capillaries during development, physiology and disease [23].

Most theoretical descriptions of collective migration have employed agent/particle-based approaches, e.g. [9, 10, 11, 12, 35], see also [4]. Despite the plethora of models, they typically share a so-called set of "first principles of swarming" [8]: specifically, particle trajectories governed by a combination of repulsion (preventing collisions), attraction/cohesion (preventing dispersal) and alignment of direction/velocity according to neighbour positions, each operating over specified interaction ranges. Models based on these features reproduce a wide variety of collective migration phenomenologies, e.g. see [4]. In the context of follower-leader systems, a key finding has been that swarms can be efficiently guided by a small number of anonymous (i.e. not clearly distinct from the crowd) informed individuals [9]. These individuals influence their neighbours, which in turn influence further followers and knowledge is relayed through the swarm. Surprisingly, as the swarm population increases a diminishing fraction of leaders is needed to achieve the same level of guidance efficiency [9]. Follower-leader alignment strategies have also been incorporated within other swarming studies. For example, in [21] a "transient leadership" model was considered to imitate bird flocks while "hierarchical leadership" was studied in [32] within the framework of a Cucker-Smale model.

Beyond agent-based models, a plethora of continuous models have been proposed, for example see [6, 16, 17, 24, 28, 34]. In common with their individual-based counterparts, movement is governed by a set or subset of attracting/repelling/aligning interactions, typically generating integro-differential equations of parabolic (e.g. [24, 28]) or hyperbolic (e.g. [6, 16]) form. While these models gain analytical tractability, their connection to individual-level behaviour is, inevitably, blurred.

The principal objective of the current paper is to clarify a relevant form of partial differential equation system to describe follower-leader interactions within swarming populations. Specifically, we focus on swarms that are clearly divided into two distinct subpopulations: a leader population with knowledge of the destination, and an unaware follower population. A paradigm for this form of follower-leader system is provided by honey bee swarms. Prior to swarming, a small proportion of scouts search the environment for a new nest, eventually arriving at a consensus for the nest location [31]. Successful translocation of the swarm subsequently relies on this small population of knowledgeable bees leading the uninformed colony to this new location. Scout bees appear to transmit their information by performing a sequence of nest-directed streaks through the upper swarm [2, 31, 19], a behaviour that is believed to accentuate their visibility to followers. Inevitably, such behaviour quickly brings scouts to the leading edge, so a shift in behaviour would be necessary to prevent losing contact with the swarm. One proposal, suggested in [31], is slowly travelling backwards, perhaps along the bottom or sides to minimise their influence, before streaking again. This suggests that leaders switch between active and passive states, e.g. as proposed in [5], where in the former they attempt to maximise their influence on the swarm orientation and in the latter it is minimised.

Taking inspiration from this and other follower-leader systems, we begin with an underlying particle description for movement that is based on a relatively minimal set of assumptions. Specifically. we propose velocity alignment interactions that could lead to coherent and guided swarm movement. We subsequently examine possible hyperbolic and parabolic scaling limits. We discuss the plausibility of these macroscopic models, their relevance in the context of follower-leader systems such as the bee swarms described above, and the new modelling challenges raised. We further present some preliminary simulations, indicating the extent to which the model can describe leader-follower behaviour.

2. **Swarm description.** We assume a swarm attempting to reach some target destination and divided into two main subpopulations: followers (f) and leaders. Leaders are assumed to have knowledge of the target while the followers are completely uninformed. The leader population is further subdivided into those engaging as active leaders (a) and those behaving passively (p). Active leaders are defined as that fraction of the leader population who are moving in the direction of the target, and can potentially engage in behaviour designed to communicate directional information regarding the target. Passive leaders are assumed to move away from the target, e.g. towards the back of the swarm where they can reengage in active leader behaviour. We remark that the total number of leaders is considered to be small, hence we can describe the leaders as a discrete system (it is noted that for bee populations around 3%-5% of the total swarm is believed to be a leader [31], while for fish a minority of informed individuals is known to lead a shoal to food [30]). Transitions are only allowed to occur between passive and active leaders, i.e. there is no transition between follower and leader type. Transitions will be motivated by the objective of maintaining contact with the swarm, i.e. an active leader that has moved to the leading of edge of the swarm changes behaviour to avoid outruning the swarm, while a passive leader moving backwards similarly switches back. Notably, though, we do not explictly incorporate an "attractive" or "cohesive" response into individual behaviours, i.e. where individuals are attracted towards higher swarm density regions.

An obviously important quantity in the model is the position of the swarm with respect to the destination or target. Given a generic member located at $\mathbf{x}_k$, its distance to the target at $\mathbf{x}_{\text{target}}$ is denoted by $I_{\text{target}}(\mathbf{x}_k(t)) = |\mathbf{x}_k(t) - \mathbf{x}_{\text{target}}|$. We fix a corresponding field $\mathbf{b} \simeq \nabla I_{\text{target}}$ within the domain outside the target, such that $-\mathbf{b}$ defines the direction for the leaders to the target. Specifically, we choose $\mathbf{b} = \nabla I_{\text{target}}$ to be divergence free at the target and tangent to the boundary of the domain. As $|\mathbf{x}| \to \infty$, we impose $\mathbf{b}(\mathbf{x}) \to \nabla I_{\text{target}}$.

Following the approach in [1], we define $\bar{\sigma}_i(\mathbf{x}, t, \theta, \tau)$, for $i \in \{f, p, a\}$, as the *microscopic densities* of followers, passive leaders and active leaders at position $\mathbf{x}$, time t, moving in direction θ for some time τ. Integrating with respect to the microscopic quantity τ, we denote $\sigma_i(\mathbf{x}, t, \theta) = \int_0^t \bar{\sigma}_i(\mathbf{x}, t, \theta, \tau) d\tau$ as the density of each subpopulation at position $\mathbf{x}$, time t and moving in direction θ and we call it a *mesoscopic density*. Subsequently integrating over θ generates the *macroscopic densities*

$$\rho_i(\mathbf{x}, t) = \int_0^t \int_S \bar{\sigma}_i(\cdot, \theta, \tau) d\theta d\tau \,, \tag{1}$$

where $S = \{\theta \in \mathbb{R}^n : \ |\theta| = 1\}$ is the unit sphere in $\mathbb{R}^n$ and the measure $d\theta$ is normalized so that the surface area $|S| = 1$. Dropped subscripts are used to

denote total population densities, i.e. $\bar{\sigma} = \bar{\sigma}_f + \bar{\sigma}_p + \bar{\sigma}_a$, $\sigma = \sigma_f + \sigma_p + \sigma_a$ and $\rho = \rho_f + \rho_p + \rho_a$.

Each subpopulation follows a random walk process, described according to a set of rules minimally chosen to describe a generic swarm behaviour but restricting excessive complexity. Motivation behind some of these choices can be found by paradigm examples, such as bee swarms, although we are not aiming to describe a specific system.

Follower behaviour is described by the following set of assumptions.

1F. Trajectories comprise of straight line motions interrupted by (effectively) instantaneous reorientations, where the new direction of motion is randomly chosen. This movement is called a velocity-jump process [27]. Individuals stop (i.e. reorient) with a rate given by a fixed parameter β.

2F. At each reorientation, with probability $\zeta \in (0,1)$ it selects a new direction of motion η, taken to be symmetrically distributed with respect to the previous one according to

$$k(\theta;\eta) = \tilde{k}(|\eta - \theta|) \ .$$

Because $\tilde{k}$ is a probability distribution, it is normalized to $\int_S \tilde{k}(|\theta - e_1|)d\theta = 1$ where $e_1 = (1, 0, ..., 0)$. The turn angle operator $T : L^2(S) \to L^2(S)$ is defined as

$$T\phi(\eta) = \int_S k(\theta;\eta)\phi(\theta)d\theta \ . \tag{2}$$

Trivially, $\tilde{k}$ could be a uniform distribution; more generally, a bias according to the previous orientation would incorporate an element of persistence of orientation, so we consider $\tilde{k}(|\eta - \theta|)$.

3F. With probability $1 - \zeta$ the follower instead aligns with the orientation of the local population, which is defined by

$$\Lambda(\mathbf{x},\theta,t) = \frac{\mathcal{J}(\mathbf{x},t)}{|\mathcal{J}(\mathbf{x},t)|} \ , \ \mathcal{J}(\mathbf{x},t) = \int_{\mathbb{R}^n}\int_S K(|\mathbf{y}-\mathbf{x}|)\sigma(\mathbf{y},t,\theta)\theta d\mathbf{y}d\theta \ . \tag{3}$$

Here K is an interaction kernel. If the flux $\mathcal{J}(\mathbf{x},t) = 0$, we assume that $\Lambda(\mathbf{x},\theta,t)$ takes the value θ [14]. A generalisation of the above would be to choose something of the form:

$$\Lambda(\mathbf{x},\theta,t) = \frac{\mathcal{J}(\mathbf{x},t)}{|\mathcal{J}(\mathbf{x},t)|} \ , \ \mathcal{J}(\mathbf{x},t) = \int_{\mathbb{R}^n}\int_S K(|\mathbf{y}-\mathbf{x}|)\left(\lambda_f\sigma_f + \lambda_p\sigma_p + \lambda_a\sigma_a\right)d\mathbf{y}d\theta \ , \tag{4}$$

where $\lambda_{f,p,a} \geq 0$ would reflect the capacity of followers to differentiate between the various subpopulations during its orientation. A choice $\lambda_a \gg \lambda_{f,p}$ would assume active leaders are considerably more conspicuous than the other populations and dominate the orientation behaviour, for example through engaging in specific movement behaviour or issuing vocal commands; a choice $\lambda_p = 0$, on the other hand, would assume passive leaders make themselves "invisible" to the followers. We note that the limitations of the above homogeneous choice will be discussed in Section 5. In particular, an inhomogeneous variant, given by $\Lambda(\mathbf{x},t) = \nu\mathcal{J}(\mathbf{x},t)$, where ν is a relaxation frequency that depends on the norm of $\mathcal{J}(\mathbf{x},t)$, will subsequently be considered in Section 6.

4F. Followers move with a fixed speed c_f.

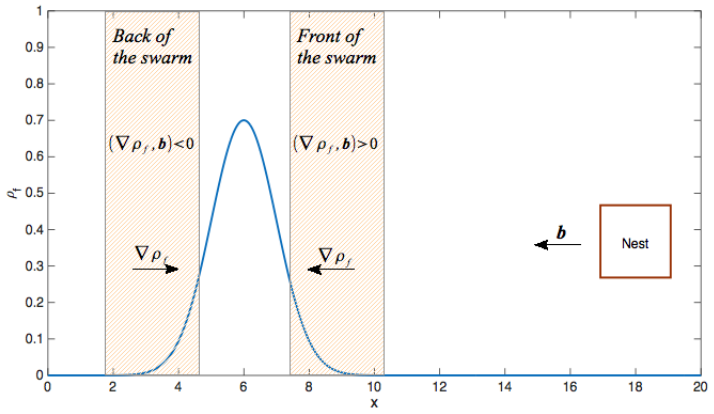


FIGURE 1. Illustration of switching between streakers and passive leaders.

For the leaders dynamics, we simplify the bias random walk process by assuming the turning distribution to be highly concentrated into choosing a specific movement: (i) a flight towards the target for the active leaders, (ii) a flight in the opposite direction for passive leaders. The limiting case of small randomness follows the discrete model proposed in [5], and is predicated on the assumption that leaders have enough knowledge of the target to produce highly directed movements

The dynamics of passive leaders are then specified by the following rules:

1P. No alignment according to other individuals.
2P. Passive leaders move in the direction $\mathbf{c}$, in the opposite direction to the target.
3P. Passive leaders move with a speed c_p.
4P. Transitions between active and passive leader take place at the front and rear edges of the swarm, as described below.

Finally, active leaders move according to:

1A. Given their knowledge of the target, they move as ballistic particles in direction $-\mathbf{b}$ (towards the front of the swarm as represented in Figure 1), with a fixed maximum speed given by c_a. Thus, no randomness is assumed.
2A. Transitions between active and passive leader take place at the front and rear edges of the swarm, as described below.
3A. Active leaders move with a speed c_a.

We note that movements into orientations outside $-\mathbf{b}$ are not permitted for active leaders within this model. Microscopic densities can then be inferred as singular distributions, where the local macroscopic density is concentrated into orientation $-\mathbf{b}$: i.e., $\sigma_a(\mathbf{x}, t, \theta) = \rho_a(\mathbf{x}, t)\delta(\theta + \mathbf{b})$ where δ is the Dirac delta function.

Suppose a Gaussian-type curve for the swarm profile. Physically, the transition rate from active to passive leader, R_{ap}, should be localised to the front edge of a swarm: in a region where $(\nabla\rho_f, \mathbf{b}) > 0$, as illustrated in Figure 1. As follower densities decrease, i.e. $\rho_f \to 0$, R_{ap} should be bounded away from zero: $R_{ap} \geq r_0 > 0$ as $\rho_f \to 0$, where r_0 is the minimal conversion rate outside the swarm. The rules for transition rate from passive to active leaders, R_{pa}, follow a similar description. Transitions should be concentrated to the rear edge of the swarm, i.e. where $(\nabla\rho_f, \mathbf{b}) < 0$. Again, as follower densities decrease we require $R_{pa} \geq r_0 > 0$.

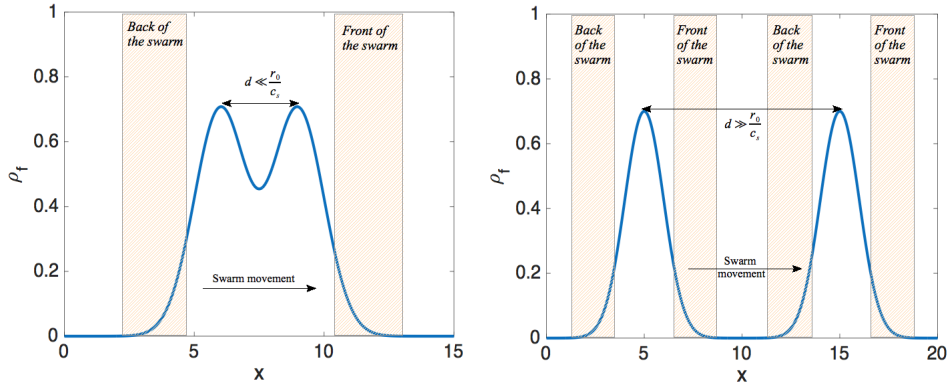


FIGURE 2. Illustration of different swarm shapes.

For more complicated swarm profiles, e.g. dual peaked as in Figure 2, let c_s/r_0 define a length scale on which the density of active leaders (and hence also that of passive leaders) varies. If the spacing between the two peaks is sufficiently larger than c_s/r_0 then active leaders consider the blobs as separate swarms and there are separate fronts and backs for each swarm. For peaks spaced below this distance, the majority of active leaders fly towards the leading edge before converting to passive leaders and it can be considered a single swarm with a single leader population.

We note that while a change from passive to active leader involves transition into a population with fixed orientation, an active to passive change demands transition into a population whose orientation is distributed over the unit sphere. We define $R_{ap} = \int_S \bar{R}_{ap}(\theta)d\theta$ and similarly for R_{pa}.

2.1. **Microscopic description.** Following the swarm description given in Section 2, the population densities satisfy systems of integro-differential equations as follows. For the active leaders we consider

$$\begin{cases} (\partial_\tau + \partial_t - c_a \mathbf{b}\cdot\nabla)\,\bar{\sigma}_a(\cdot,\tau) = -R_{ap}\bar{\sigma}_a(\cdot,\tau)\ , \\ \bar{\sigma}_a(\cdot,\tau=0) = R_{pa}\int_0^t \int_S \bar{\sigma}_p(\cdot,\theta,\tau)d\theta d\tau\ , \end{cases} \tag{5}$$

where $(\cdot)$ denotes the space and time dependence $(\mathbf{x}, t)$. The left hand side of the first equation describes the movement of the individuals towards the nest, for time τ, while the right hand side gives the density of active leaders that change to passive. The density of active leaders that start a new run, at $\tau = 0$, is given by passive leaders switching to active. This is described by the second equation.

Note that the vector field $(\partial_t, \partial_\tau, c_a\mathbf{b}\cdot\nabla)$ in (5) is transversal to the hypersurface $\{\tau = 0\}$, so that the second equation provides the necessary initial conditions.

The passive leaders obey the following system

$$\begin{cases} (\partial_\tau + \partial_t + c_p \mathbf{c}\cdot\nabla)\,\bar{\sigma}_p(\cdot,\tau) = -R_{pa}\bar{\sigma}_p(\cdot,\tau)\ , \\ \bar{\sigma}_p(\cdot,\tau=0) = R_{ap}\int_0^t \int_S \bar{\sigma}_a(\cdot,\theta,\tau)d\theta d\tau\ . \end{cases} \tag{6.a}$$

Including some randomness in the movement of the passive leaders we can write the system (6.a) as follows

$$\begin{cases} (\partial_\tau + \partial_t + c_p\theta\cdot\nabla)\,\bar\sigma_p(\cdot,\theta,\tau) = -R_{pa}\bar\sigma_p(\cdot,\tau) - \beta\bar\sigma_p(\cdot,\theta,\tau)\ , \\ \bar\sigma_p(\cdot,\theta,\tau=0) = \bar R_{ap}(\theta)\int_0^t\int_S \bar\sigma_a(\cdot,\eta,\tau)d\eta d\tau + B(\theta)\int_0^t\int_S \beta\bar\sigma_p(\cdot,\eta,\tau)d\eta d\tau\ . \end{cases} \tag{6.b}$$

The passive leaders in this case also move according to a velocity jump process as described in 1F, with the same stopping rate β. The new direction of motion η is given by the turn angle operator $B(\eta)$. To be more specific, we define $B(\eta) = pT(\eta) + (1-p)T_{\rho_f}$ where $p\in(0,1)$ and we fix $T_{\rho_f} = (\mathbf{b}\cdot\nabla\rho_f)\nabla\rho_f$. The first term in $B(\eta)$ describes the random movement where a new direction of motion is taken arbitrarily, in this case it only depends on the final direction. The second term describes a vector field pointing in the direction of the back of the swarm, where the switch between passive and active leaders occurs. The analysis of equation (6.b) is postponed to Section 3.6.

For the followers we consider the approach in [18] and then we have

$$\begin{cases} (\partial_\tau + \partial_t + c_f\theta\cdot\nabla)\,\bar\sigma_f(\cdot,\theta,\tau) = -\beta\bar\sigma_f(\cdot,\theta,\tau)\ , \\ \bar\sigma_f(\cdot,\eta,\tau=0) = \int_S Q(\eta,\theta)\int_0^t \beta\bar\sigma_f(\cdot,\theta,\tau)d\tau d\theta\ , \end{cases} \tag{7}$$

where

$$Q(\eta,\theta) = \zeta\tilde k(|\eta-\theta|) + (1-\zeta)\Phi(\Lambda\cdot\eta)\ . \tag{8}$$

$\Phi(\Lambda\cdot\eta)$ is the distribution of the new aligned direction and satisfies $\int_S\Phi(\Lambda\cdot\eta)d\eta = 1$. Integrating with respect to τ the systems (5) and (6.a) we obtain, for $\rho_a(\cdot) = \int_0^t\bar\sigma_a(\cdot,\tau)d\tau$ and $\rho_p(\cdot) = \int_0^t\bar\sigma_p(\cdot,\tau)d\tau$, respectively,

$$\partial_t\rho_a - c_a\mathbf{b}\cdot\nabla\rho_a = -R_{ap}\rho_a + R_{pa}\rho_p\ , \tag{9}$$

$$\partial_t\rho_p + c_p\mathbf{c}\cdot\nabla\rho_p = -R_{pa}\rho_p + R_{ap}\rho_a\ . \tag{10}$$

Similarly, integrating system (7) with respect to τ and using the definition of T given in (2) we obtain

$$\partial_t\sigma_f + c_f\theta\cdot\nabla\sigma_f = -\beta\sigma_f + \zeta\beta T\sigma_f + (1-\zeta)\beta\Phi(\Lambda\cdot\theta)\rho_f\ . \tag{11}$$

3. Macroscopic PDE description.

As shown in the previous section, the densities $\rho_a(\mathbf{x},t)$, $\rho_p(\mathbf{x},t)$ and $\sigma_f(\mathbf{x},t,\theta)$ satisfy the following system of kinetic equations,

$$\partial_t\rho_a - c_a\mathbf{b}\cdot\nabla\rho_a = -R_{ap}\rho_a + R_{pa}\rho_p\ , \tag{12}$$

$$\partial_t\rho_p + c_p\mathbf{c}\cdot\nabla\rho_p = -R_{pa}\rho_p + R_{ap}\rho_a\ , \tag{13}$$

$$\partial_t\sigma_f + c_f\theta\cdot\nabla\sigma_f = -\beta\sigma_f + \zeta\beta T\sigma_f + (1-\zeta)\beta\Phi(\Lambda\cdot\theta)\rho_f\ . \tag{14}$$

We control that the population of the followers (14) is conserved:

$$\partial_t\int_S\sigma_f d\theta + c_f\int_S\theta\cdot\nabla\sigma_f d\theta = -\beta\int_S\sigma_f d\theta + \zeta\beta\int_S T\sigma_f d\theta + (1-\zeta)\beta\rho_f = 0\ , \tag{15}$$

since $\int_S\Phi(\Lambda\cdot\theta)d\theta = 1$ and $\int_S T d\theta = 1$ by definition.

In the following subsections we study the macroscopic behaviour of the above systems by considering different scaling limits. In Subsection 3.2 we study a hyperbolic limit of the system (5)-(6.a)-(7), where the leaders are considered to be discrete. Then, in Subsection 3.3, a hyperbolic limit of system (5)-(6.b)-(7), where a random component of the movement is included for the case of passive leaders.

Later, in Subsection 3.4 we consider a diffusion limit of the system (5)-(6.a)-(7) and in Subsection 3.5 a diffusion limit for the system (5)-(6.a)-(7) assuming the population of leaders is "slow" compared to the followers by scaling c_a, c_p. Finally, in Subsection 3.6 we develop a diffusion limit of the system (5)-(6.a)-(7) for slow leaders and including noise in the movement of passive leaders.

3.1. **Properties and asymptotic expansion of operators.** In the following sections we consider

$$T^\varepsilon = T^0 + \varepsilon T^1 + \mathcal{O}(\varepsilon^2)\ , \quad \Phi^\varepsilon = \Phi^0 + \varepsilon\Phi^1 + \mathcal{O}(\varepsilon^2) \quad \text{and} \quad B^\varepsilon = B^0 + \varepsilon B^1 + \mathcal{O}(\varepsilon^2)\ .$$

Moreover, we have that

$$\int_S T^0 1\ d\theta + \varepsilon \int_S T^1 1\ d\theta = 1\ ,$$

so that

$$\int_S T^0 1\ d\theta = 1,\ \int_S T^1 1\ d\theta = 0\ ,$$

and similarly

$$\int_S \Phi^0 d\theta = 1,\ \int_S \Phi^1 d\theta = 0\ ,\ \int_S B^0 1\ d\theta = 1,\ \int_S B^1 1\ d\theta = 0\ . \tag{16}$$

In Section 3.2 we further show that

$$\int_S \theta\Phi^0(\Lambda\cdot\theta)d\theta = z\Lambda^0\ ,\ \int_S \theta B^0 d\theta = \int_S \theta B^1 d\theta = 0\ .$$

3.2. **Hyperbolic limit.** In this section we investigate the macroscopic dynamics of the swarm over short time scales. Consider the following scaling $(\mathbf{x}, t) \mapsto (\mathbf{x}/\varepsilon, t/\varepsilon)$ where $\varepsilon = \bar{\tau}/\mathcal{T} \ll 1$ and $\mathcal{T}$ is the macroscopic time. The transition rates $R^\varepsilon_{ap}, R^\varepsilon_{pa}$ are multiplied by a factor of ε meaning that the switching rates are large enough to be observed at the macroscopic spatial scale. Then we write

$$\varepsilon\partial_t\rho_a - \varepsilon c_a \mathbf{b}\cdot\nabla\rho_a = -\varepsilon R^\varepsilon_{ap}\rho_a + \varepsilon R^\varepsilon_{pa}\rho_p\ , \tag{17}$$

$$\varepsilon\partial_t\rho_p + \varepsilon c_p \mathbf{c}\cdot\nabla\rho_p = \varepsilon R^\varepsilon_{ap}\rho_a - \varepsilon R^\varepsilon_{pa}\rho_p\ , \tag{18}$$

$$\varepsilon\partial_t\sigma_f + \varepsilon c_f\theta\cdot\nabla\sigma_f = -\beta\sigma_f + \zeta\beta T^\varepsilon\sigma_f + (1-\zeta)\beta\Phi^\varepsilon(\Lambda\cdot\theta)\rho_f\ . \tag{19}$$

For the case of the follower population and using the expansions (76) for σ_f, T^ε, Φ^ε and ρ^ε_f from Appendix A we can rewrite (19) as

$$\begin{aligned}\varepsilon\partial_t(\sigma^0_f + \varepsilon\sigma^1_f) + \varepsilon c_f\theta\cdot\nabla(\sigma^0_f + \varepsilon\sigma^1_f) = &-\beta(\sigma^0_f + \varepsilon\sigma^1_f) + \zeta\beta(T^0 + \varepsilon T^1)(\sigma^0_f + \varepsilon\sigma^1_f)\\ &+ (1-\zeta)\beta(\Phi^0 + \varepsilon\Phi^1)(\rho^0_f + \varepsilon\rho^1_f)\ .\end{aligned} \tag{20}$$

For ε^0 we obtain

$$\sigma^0_f = (\zeta + (1-\zeta)\Phi^0(\Lambda\cdot\theta))\rho^0_f\ . \tag{21}$$

Here we considered that $T^0\sigma^0_f = \rho^0_f$ (see Appendix C for properties of the operator T).

Substituting (21) into (19) and integrating with respect to S gives

$$\partial_t\rho^0_f \int_S (\zeta + (1-\zeta)\Phi^0(\Lambda\cdot\theta))d\theta + c_f\int_S \theta\cdot\nabla(\zeta + (1-\zeta)\Phi^0(\Lambda\cdot\theta))\rho^0_f d\theta = 0\ . \tag{22}$$

In the left hand side we use again $\int_S \Phi^0(\Lambda\cdot\theta)d\theta = 1$ and, $\int_S \theta\Phi^0(\Lambda\cdot\theta)d\theta = z\Lambda^0$ where z is given by (81). Hence the conservation equation for the followers reads

$$\partial_t\rho_f^0 + c_f z(1-\zeta)\nabla\cdot(\rho_f^0\Lambda^0) = 0\ . \tag{23}$$

Next we need to find the mean direction $\rho_f^0\Lambda^0$. Following the same analysis as in [15, 18] we substitute the expansion for σ_f^0 given in (21) into (19) and multiply by $\theta\cdot v$, where $v\in\mathbb{R}^n$ is orthogonal to Λ^0. Integrating over S then gives

$$\left(\partial_t\int_S \theta\Psi(\theta)\rho_f^0 d\theta + c_f\int_S \theta\cdot\nabla(\Psi(\theta)\rho_f^0)\theta d\theta\right)\cdot v = \mathcal{O}(\varepsilon)\ . \tag{24}$$

Here $\Psi(\theta) = \zeta + (1-\zeta)\Phi^0(\Lambda\cdot\theta)$, $\mathcal{O}(\varepsilon)$ includes $\zeta\beta\int_S\theta T^1\sigma_f^0 d\theta$, $(1-\zeta)\beta\int_S\theta\Phi^1\rho_f^0 d\theta$ and $-\beta\int_S\sigma_f^1 d\theta$, and letting $\varepsilon\to 0$ in the right hand side we obtain

$$\left(z(1-\zeta)\partial_t(\rho_f^0\Lambda^0) + c_f\int_S\theta\cdot\nabla(\rho_f^0\Psi(\theta))\theta d\theta\right)\cdot v = 0\ .$$

Using the fact that $v\perp\Lambda^0$, we can reformulate the above expression in terms of the orthogonal projection $P_\perp = \mathbb{1} - \Lambda^0\otimes\Lambda^0$ onto $\Lambda_\perp^0$,

$$P_\perp\left(z(1-\zeta)\partial_t(\rho_f^0\Lambda^0) + c_f\nabla\cdot\rho_f^0\int_S(\theta\otimes\theta)\Psi(\theta)d\theta\right) = 0\ . \tag{25}$$

For the first term of the above expression we can write

$$z(1-\zeta)P_\perp(\rho_f^0\partial_t\Lambda^0 + \Lambda^0\partial_t\rho_f^0) = z(1-\zeta)\rho_f^0\partial_t\Lambda^0\ , \tag{26}$$

since $\langle\partial_t\Lambda^0,\Lambda^0\rangle = \frac{1}{2}\partial_t|\Lambda^0|^2 = 0$, i.e., $\Lambda^0\perp\partial_t\Lambda^0$. For the second term we must compute the integral $\int_S(\theta\otimes\theta)\Psi(\theta)d\theta$, where we use $\theta = \cos(s)\Lambda^0 + \sin(s)\Lambda_\perp^0$ in polar coordinates for $n=2$ and spherical coordinates for $n=3$ as in [15]. Finally, we obtain

$$\rho_f^0(z(1-\zeta)\partial_t\Lambda^0 + C_1\Lambda^0\cdot\nabla\Lambda^0) + C_2P_\perp\nabla\rho_f^0 = 0\ , \tag{27}$$

where we have used $\Lambda_\perp^0\otimes\Lambda_\perp^0 = \mathbb{1} - \Lambda^0\otimes\Lambda^0$. Here $C_1 = c_f(1-\zeta)a_3$ and $C_2 = c_f(1-\zeta)\mathbb{1}a_1 + c_f\mathbb{1}\pi\zeta$ for, $a_3 = a_0 - a_1$ and

$$a_0 = \begin{cases}\int_0^{2\pi}\Phi^0(\cos(s))\cos(s)^2 ds\ , & \text{if } n=2\ ,\\ 2\pi\int_0^\pi\Phi^0(\cos(s))\cos(s)^2\sin(s)ds\ , & \text{if } n=3\ ,\end{cases} \tag{28}$$

$$a_1 = \begin{cases}\int_0^{2\pi}\Phi^0(\cos(s))\sin(s)^2 ds\ , & \text{if } n=2\ ,\\ \pi\int_0^\pi\Phi^0(\cos(s))\sin(s)^3\sin(s)ds\ , & \text{if } n=3\ .\end{cases} \tag{29}$$

The final system of equations is given as

$$\partial_t\rho_a^0 - c_a\mathbf{b}\cdot\nabla\rho_a^0 + R_{ap}^0\rho_a^0 - R_{pa}^0\rho_p^0 = 0\ , \tag{30}$$

$$\partial_t\rho_p^0 + c_p\mathbf{c}\cdot\nabla\rho_p^0 - R_{ap}^0\rho_a^0 + R_{pa}^0\rho_p^0 = 0\ , \tag{31}$$

$$\partial_t\rho_f^0 + c_f z(1-\zeta)\nabla\cdot(\rho_f^0\Lambda^0) = 0\ , \tag{32}$$

$$\rho_f^0(z(1-\zeta)\partial_t\Lambda^0 + C_1\Lambda^0\cdot\nabla\Lambda^0) + C_2P_\perp\nabla\rho_f^0 = 0\ . \tag{33}$$

Unlike in the parabolic limit (50)-(53) in Subsection 3.4, we now obtain transport equations for the leader populations, which no longer adjust instantaneously to the swarm of followers.

3.3. **Hyperbolic limit for passive leaders with noise.** Defining the macroscopic density $\rho_p(\mathbf{x},t)=\int_0^t\int_S \bar\sigma_p(\cdot,\eta,\tau)d\eta d\tau$, we study now the hyperbolic limit of (6.b) where, after integrating with respect to τ and using the boundary condition $\bar\sigma_p(\cdot,\eta,\tau=0)$ we write

$$\varepsilon\partial_t\sigma_p+\varepsilon c_p\theta\cdot\nabla\sigma_p=\varepsilon\bar R^\varepsilon_{ap}(\theta)\rho_a-\varepsilon R^\varepsilon_{pa}\sigma_p-\beta\sigma_p+\beta B^\varepsilon\rho_p\ . \tag{34}$$

Following the same procedure as before we consider in this case the expansions (82) and (83) for σ^ε_p, ρ^ε_p, $\bar R^\varepsilon_{ap}$ and R^ε_{ap}, and from (34) we obtain

$$\varepsilon^0:\qquad \sigma^0_p=B^0\rho^0_p\ , \tag{35}$$

$$\begin{aligned}\varepsilon^1:\quad -\beta\sigma^1_p&=\partial\sigma^0_p+c_p\theta\cdot\nabla\sigma^0_p-\bar R^0_{ap}\rho^0_a+R^0_{pa}\sigma^0_p-\beta B^0\rho^1_p-\beta B^1\rho^0_p\\&=B^0\partial_t\rho^0_p+c_pB^0\theta\cdot\nabla\rho^0_p-\bar R^0_{ap}\rho^0_a+R^0_{pa}B^0\rho^0_p\\&\qquad-\beta B^0\rho^1_p-\beta B^1\rho^0_p\ .\end{aligned} \tag{36}$$

The next step is to compute the term $\partial_t\rho^0_p$. For that, we integrate (34) with respect to θ and we have

$$\varepsilon\partial_t(\rho^0_p+\varepsilon\rho^1_p)+\varepsilon c_p\nabla\cdot\int_S\theta(\sigma^0_p+\varepsilon\sigma^1_p)d\theta=\varepsilon(R^0_{ap}+\varepsilon R^1_{ap})(\rho^0_a+\varepsilon\rho^1_a)-\varepsilon(R^0_{pa}+\varepsilon R^1_{pa})(\rho^0_p+\varepsilon\rho^1_p)\ . \tag{37}$$

Here we have used (16). With (37) and since $\int_S\theta B^0d\theta=0$, we obtain

$$\partial_t\rho^0_p=R^0_{ap}\rho^0_a-R^0_{pa}\rho^0_p\ .$$

Substituting the above expression in (36) we finally obtain

$$\sigma^1_p=\frac{-1}{\beta}\Big(B^0R^0_{ap}\rho^0_a+c_pB^0\theta\cdot\nabla\rho^0_p-\bar R^0_{ap}\rho_a-\beta B^0\rho^1_p-\beta B^1\rho^0_p\Big)\ , \tag{38}$$

and hence, $\sigma_p=B^0\rho^0_p+\varepsilon\sigma^1_p+\mathcal{O}(\varepsilon^2)$. Going back to the conservation equation (37) we can write now

$$\partial_t\rho^0_p-R^0_{ap}\rho^0_a+R^0_{pa}\rho^0_p=0\ . \tag{39}$$

Computing the term $-\varepsilon c_p\nabla\cdot\int_S\theta\sigma^1_pd\theta$ using (38) we obtain that

$$-\varepsilon c_p\nabla\cdot\int_S\theta\sigma^1_pd\theta=\varepsilon(\Delta(D_p\rho^0_p)-\nabla\cdot(\mathcal{R}^0_{ap}\rho^0_a))\ , \tag{40}$$

where $D_p=\frac{c_p}{\beta}\int_S\theta\theta^TB^0d\theta$ and $\mathcal{R}^0_{ap}=\frac{c_p}{\beta}\int_S\theta\bar R^0_{ap}d\theta$, which shows that the diffusion term only appears for higher orders of ε. The hyperbolic system of equations reads as

$$\partial_t\rho^0_a-c_a\mathbf{b}\cdot\nabla\rho^0_a+R^0_{ap}\rho^0_a-R^0_{pa}\rho^0_p=0\ , \tag{41}$$

$$\partial_t\rho^0_p-R^0_{ap}\rho^0_a+R^0_{pa}\rho^0_p=0\ , \tag{42}$$

$$\partial_t\rho^0_f+c_fz(1-\zeta)\nabla\cdot(\rho^0_f\Lambda^0)=0\ , \tag{43}$$

$$\rho^0_f(z(1-\zeta)\partial_t\Lambda^0+C_1\Lambda\cdot\nabla\Lambda^0)+C_2P_\perp\nabla\rho^0_f=0\ . \tag{44}$$

3.4. **Diffusion limit.** Next we turn to determining macroscopic equations that describe the interactions between the three different populations over long time regimes. We introduce a parabolic scaling $(\mathbf{x}, t) \mapsto (\mathbf{x}/\varepsilon,\ t/\varepsilon^2)$, where ε is defined in the sense of Section 3.2.

Equations (12), (13) and (14) then become

$$\varepsilon^2 \partial_t \rho_a - \varepsilon c_a \mathbf{b} \cdot \nabla \rho_a = -\varepsilon R^\varepsilon_{ap} \rho_a + \varepsilon R^\varepsilon_{pa} \rho_p \ , \tag{45}$$

$$\varepsilon^2 \partial_t \rho_p + \varepsilon c_p \mathbf{c} \cdot \nabla \rho_p = \varepsilon R^\varepsilon_{ap} \rho_a - \varepsilon R^\varepsilon_{pa} \rho_p \ , \tag{46}$$

$$\varepsilon^2 \partial_t \sigma_f + \varepsilon c_f \theta \cdot \nabla \sigma_f = -\beta \sigma_f + \zeta \beta T^\varepsilon \sigma_f + \varepsilon (1-\zeta) \beta \Phi^\varepsilon(\Lambda \cdot \theta) \rho_f \ . \tag{47}$$

Note that we have also scaled the alignment kernel $\Phi^\varepsilon(\Lambda \cdot \theta)$.

The conservation equation (15) for the follower population is

$$\partial_t \rho_f + n c_f \nabla \cdot w_f = 0 \ , \tag{48}$$

where $w_f = \frac{1}{n} \int_S \theta \sigma_f d\theta$. The next step is to compute the mean direction w_f. We follow the steps in Appendix A and, using asymptotic expansions for the terms σ_f, ρ_f, T^ε and Φ^ε we can re-write the conservation equation (48) as

$$\varepsilon^2 \partial_t \rho^0_f + \varepsilon^2 c_f \nabla \cdot \int_S \theta \sigma^1_f d\theta = 0 \ .$$

Finally, we obtain

$$\partial_t \rho^0_f - c_f \Delta (D_f \rho^0_f) + \nabla \cdot (z \Lambda^0_W \rho^0_f) = 0 \ , \tag{49}$$

where $z\Lambda^0_W = \int_S \theta \Phi^0(\Lambda^0_W \cdot \theta) d\theta$, $D_f = \frac{\zeta c_f}{\beta(1-\zeta)} \int_S \theta\theta^T d\theta$, z is given by (81) and W is the total mean direction of the whole population.

The system describing the macroscopic densities of followers and leaders, in the diffusion limit, reads as follows

$$-c_a \mathbf{b} \cdot \nabla \rho^0_a + R^0_{ap} \rho^0_a - R^0_{pa} \rho^0_p = 0 \ , \tag{50}$$

$$c_p \mathbf{c} \cdot \nabla \rho^0_p - R^0_{ap} \rho^0_a + R^0_{pa} \rho^0_p = 0 \ , \tag{51}$$

$$\partial_t \rho^0_f + c_f \nabla \cdot w_f = 0 \tag{52}$$

$$w_f = -\nabla D_f \rho^0_f + \rho^0_f z \Lambda^0_W \ . \tag{53}$$

Note that the evolution of the swarm is determined by the movement of the followers. There are no time derivatives in the equations for the densities of the leaders, which adjust instantaneously to the movement of the followers. Technically, the time derivatives in the system (45), (46), (47) scale as ε^2, so that the leader's movement is determined at order ε by convection and the transition between their active and passive states.

3.5. **Diffusion limit for slow leaders.** We consider the situation where c_a, c_p, R^ε_{ap} and R^ε_{pa} are all of order ε, in particular the leaders move slowly compared to the followers.

Equations (12), (13) and (14) then become

$$\varepsilon^2 \partial_t \rho_a - \varepsilon^2 c_a \mathbf{b} \cdot \nabla \rho_a = -\varepsilon^2 R^\varepsilon_{ap} \rho_a + \varepsilon^2 R^\varepsilon_{pa} \rho_p \ , \tag{54}$$

$$\varepsilon^2 \partial_t \rho_p + \varepsilon^2 n c_p \mathbf{c} \cdot \nabla \rho_p = \varepsilon^2 R^\varepsilon_{ap} \rho_a - \varepsilon^2 R^\varepsilon_{pa} \rho_p \ , \tag{55}$$

$$\varepsilon^2 \partial_t \sigma_f + \varepsilon c_f \theta \cdot \nabla \sigma_f = -\beta \sigma_f + \zeta \beta T^\varepsilon \sigma_f + (1-\zeta) \beta \Phi^\varepsilon(\Lambda \cdot \theta) \rho_f \ . \tag{56}$$

The system describing the macroscopic densities of followers and leaders, in the diffusion limit, reads as follows

$$\partial_t \rho_a^0 - c_a \mathbf{b} \cdot \nabla \rho_a^0 + R_{ap}^0 \rho_a^0 - R_{pa}^0 \rho_p^0 = 0 \ , \tag{57}$$

$$\partial_t \rho_p^0 + c_p \mathbf{c} \cdot \nabla \rho_p^0 - R_{ap}^0 \rho_a^0 + R_{pa}^0 \rho_p^0 = 0 \ , \tag{58}$$

$$\partial_t \rho_f^0 + c_f \nabla \cdot w_f = 0 \tag{59}$$

$$w_f = -\nabla D_f \rho_f^0 + \rho_f^0 z \Lambda_W^0 \ . \tag{60}$$

3.6. **Random movement for passive leaders.** We start from the mesoscopic equation (34) obtained from (6.b). Introducing the diffusive scaling as in Subsection 3.4 we get

$$\varepsilon^2 \partial_t \rho_a - \varepsilon^2 c_a \mathbf{b} \cdot \nabla \rho_a = -\varepsilon^2 R_{ap}^\varepsilon \rho_a + \varepsilon^2 R_{pa}^\varepsilon \rho_p \ , \tag{61}$$

$$\varepsilon^2 \partial_t \sigma_p + \varepsilon c_p \theta \cdot \nabla \sigma_p = \varepsilon^2 \bar{R}_{ap}^\varepsilon(\theta) \rho_a - \varepsilon^2 R_{pa}^\varepsilon \sigma_p - \beta \sigma_p + \beta B^\varepsilon \rho_p \ , \tag{62}$$

$$\varepsilon^2 \partial_t \sigma_f + \varepsilon c_f \theta \cdot \nabla \sigma_f = -\beta \sigma_f + \zeta \beta T^\varepsilon \sigma_f + \varepsilon (1-\zeta) \beta \Phi^\varepsilon (\Lambda \cdot \theta) \rho_f \ . \tag{63}$$

Note that the speed c_a, as well as the switching rates R_{ap}^ε, R_{pa}^ε and $\bar{R}_{ap}^\varepsilon(\theta)$ are scaled as ε^2. The conservation equation for the passive leaders (62), is then

$$\varepsilon^2 \partial_t \rho_p + \varepsilon^2 c_p \nabla \cdot \int_S \theta \sigma_p d\theta = \varepsilon^2 R_{ap}^\varepsilon \rho_a - \varepsilon^2 R_{pa}^\varepsilon \rho_p \ , \tag{64}$$

where

$$\rho_p = \int_S \sigma_p d\theta \quad \text{and} \quad w_p = \frac{1}{n} \int_S \theta \sigma_p d\theta \ . \tag{65}$$

Following the steps described in Appendix B we find an asymptotic expansion for σ_p, using (84) and (85), as follows

$$\sigma_p = B^0 \rho_p^0 + \varepsilon \sigma_p^1 + \mathcal{O}(\varepsilon^2) \ ,$$

where σ_p^1 is given by (85). Computing the mean direction term $\nabla \cdot \int_S \theta \sigma_p^1 d\theta$ as in (88) we finally obtain, for the passive leaders with random movement,

$$\partial_t \rho_p^0 - c_p \Delta (D_p + \mathbb{T}) \rho_p^0 = R_{ap}^0 \rho_a^0 - R_{pa}^0 \rho_p^0 \ , \tag{66}$$

where

$$D_p = \frac{p c_p}{\beta} \int_S \theta \theta^T T^0(\theta) d\theta \ , \quad \text{and} \quad \mathbb{T} = \frac{(1-p) c_p}{\beta} \int_S \theta \theta^T T_{\rho_f}^0 d\theta \ .$$

Hence the final system is given by

$$\partial_t \rho_a^0 - c_a \mathbf{b} \cdot \nabla \rho_a^0 + R_{ap}^0 \rho_a^0 - R_{pa}^0 \rho_p^0 = 0 \ , \tag{67}$$

$$\partial_t \rho_p^0 - c_p \Delta (D_p + \mathbb{T}) \rho_p^0 - R_{ap}^0 \rho_a^0 + R_{pa}^0 \rho_p^0 = 0 \ , \tag{68}$$

$$\partial_t \rho_f^0 + c_f \nabla \cdot w_f = 0 \tag{69}$$

$$w_f = -\nabla D_f \rho_f^0 + \rho_f^0 z \Lambda_W^0 \ . \tag{70}$$

4. **Numerical solution.** As an example of the macroscopic dynamics we illustrate the time evolution of the swarm dynamics for the system (67) - (70) in a one dimensional domain $x \in (0, L)$, with Neumann conditions imposed at the boundary. The system is solved numerically using a simple finite difference scheme on a sufficiently fine equidistant mesh. More precisely, for the leaders we couple a Lax-Wendroff discretisation for the transport equation (67) to a Crank-Nicolson time stepping for the diffusion equations (68), (70) in a way that conserves the total number of leaders.

Initially we co-localise leaders and followers in $(0, L) = (0, 15)$, setting initial distributions proportional to a Gaussian of width 0.6 and centre $x = 3$, with a total density of 3 for active and passive leaders, respectively 100 for the followers. This reflects certain real world follower-leader systems, where leaders form a fairly small fraction (e.g. [31]). The time evolution of the swarm is illustrated for model parameters $c_a\mathbf{b} = 6$, $c_p(D_p+\mathbb{T}) = c_f D_f = 1$ and Λ^0_W as in (4) which is determined by the leader population alone (formally, $\lambda \to \infty$). The conversion rates between active and passive leaders are taken to be

$$R^0_{ap} = R^0 \left(\exp(-\tfrac{\rho^0_f}{\Theta \max_{x\in(0,L)} \rho^0_f}) - \exp(-1/\Theta)\right)$$

when $\partial_x \rho^0_f > 0$ and $R^0_{ap} = 0$ otherwise, respectively,

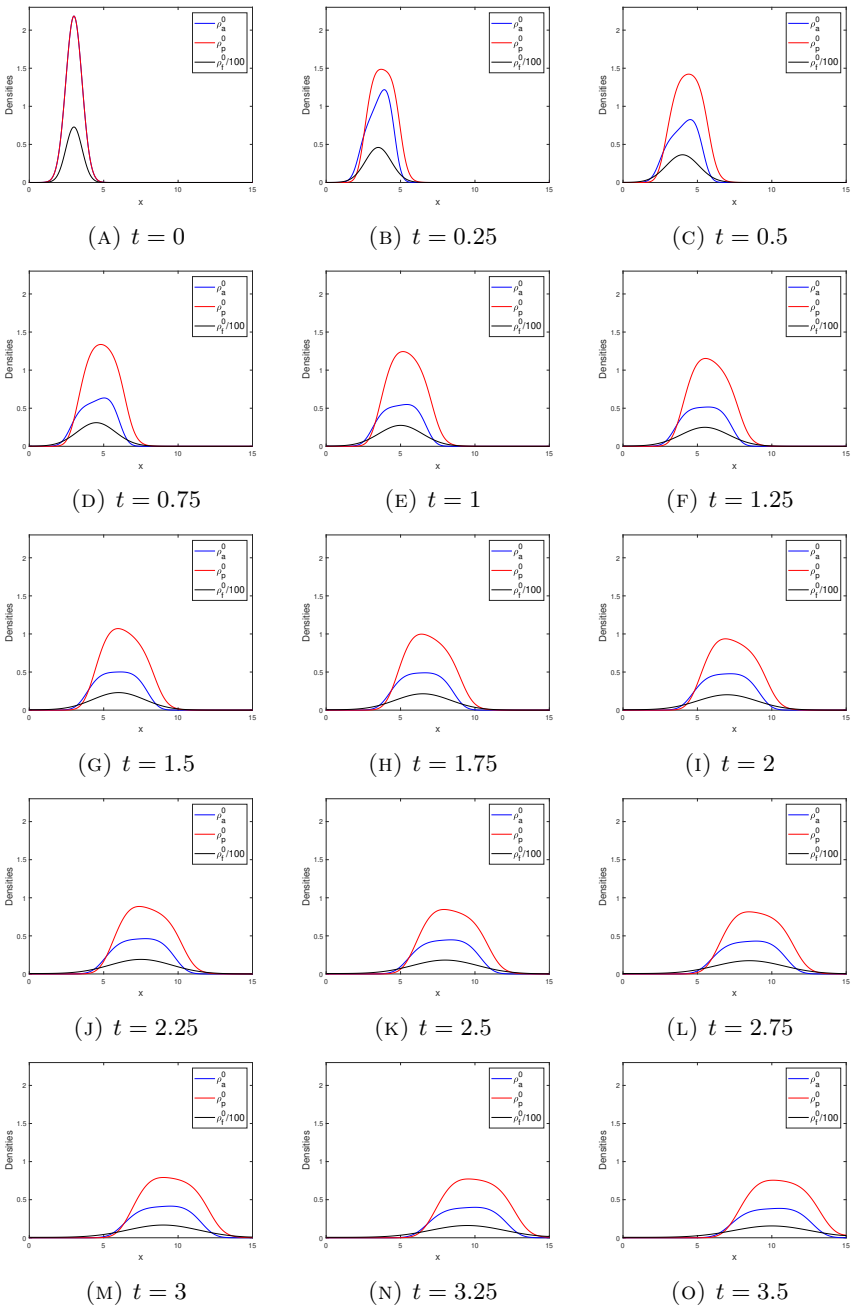
$$R^0_{pa} = R^0 \left(\exp(-\tfrac{\rho^0_f}{\Theta \max_{x\in(0,L)} \rho^0_f}) - \exp(-1/\Theta)\right)$$

when $\partial_x \rho^0_f < 0$ and $R^0_{pa} = 0$ otherwise. For illustration we use a cut-off fraction of $\Theta = 0.2$ and strength $R^0 = 75$.

Figure 3 shows the resulting evolution of the densities of followers (black), active (blue) and passive (red) leaders. For purposes of illustration the density ρ^0_f of followers is divided by a factor 100. The follower and leader populations initially move together as a localised swarm. In the current model the follower population becomes less localised with time, because the diffusion in the current model (70) is not compensated by localising mechanisms. The leader populations widen correspondingly. Note that the fraction of active leaders quickly decreases to a small, but stable fraction, as passive leaders are slow to diffuse to the back of the swarm, where they become active again. This is confirmed by the corresponding percentages of active, respectively passive leaders shown as a function of time in Figure 4. Because of the steep gradient in the conversion rates R^0_{pa} and R^0_{ap}, active and passive leaders remain in the centre of the swarm, with passive leaders generated at the front.

5. **Discussion of macroscopic equations.** The macroscopic equations in the hyperbolic limit, (30) to (33), capture the behaviours depicted in Figures 1 and 2 on hyperbolic time scales. Active leaders follow a transport equation (30), while passive leaders are transported in the opposite direction to that of the active leaders with its respective velocity. At the front of the swarm, the lower bound $R^0_{ap} \geq r_0$ forces ρ_a to decay exponentially fast to zero over a characteristic length scale $\frac{c_a}{r_0}$, as active leaders convert into passive leaders. As passive leaders reach areas of large R^0_{pa}, corresponding to the rear of the swarm, they convert into active and head once again to the swarm front. Here the characteristic length scale is $\frac{c_p}{r_0}$. Note that the movement of active and passive leaders depends on followers only through the transition rates R^0_{ap} and R^0_{pa}: this is logical enough, given that leaders have knowledge of the target site and should not be swayed by the uninformed. A more detailed model may also incorporate some influence of the followers on the leader direction (e.g. due to avoiding collisions), however we have neglected that here for simplicity.

The equations describing dynamics of the followers, (32) and (33), correspond to classical equations for swarming particles [18]. The directional movement of the swarm depends on leaders only through Λ^W, and the swarm impacts on the movement of the leaders only through the switching rates R^0_{ap} and R^0_{pa} as noted



(A) $t = 0$ (B) $t = 0.25$ (C) $t = 0.5$

(D) $t = 0.75$ (E) $t = 1$ (F) $t = 1.25$

(G) $t = 1.5$ (H) $t = 1.75$ (I) $t = 2$

(J) $t = 2.25$ (K) $t = 2.5$ (L) $t = 2.75$

(M) $t = 3$ (N) $t = 3.25$ (O) $t = 3.5$

FIGURE 3. Evolution of follower and leader populations in model example.

above. This illustrates a limitation of the homogeneous choice for the alignment kernel (3): since Λ^W is a unit vector, only a single leader is required to direct the swarm to its target, a result which clearly stretches credulity for very large swarms. To avoid such pathological cases it is possible to consider inhomogeneous

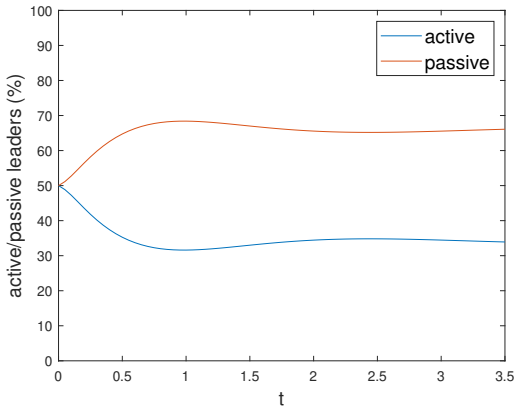


FIGURE 4. Percentage of active and passive leaders as a function of time.

alignment kernels, see Section 6, where the size of the orientation vector Λ^W is taken to increase both with the number of leaders moving along $\mathbf{b}$ and with the strength of the interactions.

A fundamental question concerns whether two follower populations, as in Figure 2, behave as one joint swarm or two separate swarms. Building on the discussion introduced in Section 2, for the model here the distinction lies with the separation of the two populations: for separations much larger than $\frac{c_a}{r_0}$, leaders remain confined to their separate follower swarms; for separations smaller than $\frac{c_a}{r_0}$, the leader populations bridge the gap that separates the peaks to join the swarms together; active to passive leader conversion occurs predominantly within a single zone at the swarm front.

Equations (41) to (44), where a heading noise is incorporated for passive leaders, support the general description above with a few noteworthy modifications. The passive leaders now undergo an additional undirected diffusion through the swarm, and hence their dynamics follow an advection-diffusion equation (42) rather than the transport equation above. The terms of order ε account for secondary effects in the collective movement.

Describing the swarm on parabolic time scales, however, proves more intricate. The macroscopic equations indicate some inherent problems in describing both the advective movement of active leaders and the diffusive movement of the followers. Unlike the hyperbolic limits, the density of followers in (53) is now observed to diffuse over parabolic time scales, implying an eventual loss of swarm cohesion. Further, equations (50) and (51) may not conserve the total number of leaders $\int(\rho_a^0 + \rho_p^0)\, dx$.

Nevertheless, the general shape of the leader population exhibits the form expected for a swarm. To see this, first note that the distribution of active leaders on long (parabolic) time scales is constant along the vector field $\mathbf{b}$ whenever R_{ap}^0 and R_{pa}^0 vanish, i.e. within the interior of the swarm. Further, even within regions of nonzero R_{ap}^0 and R_{pa}^0 we note that

$$-c_a \mathbf{b} \cdot \nabla \rho_a^0 + c_p \mathbf{c} \cdot \nabla \rho_p^0 = 0 \ .$$

For example, in one dimension this implies that $\rho_p^0 = \frac{c_a}{c_p}\rho_a^0$, independent of the transition rates R_{ap}^0 and R_{pa}^0. Therefore

$$c_a \partial_x \rho_a^0 = (R_{ap}^0 - \frac{c_a}{c_p} R_{pa}^0)\rho_a^0 \ .$$

The exponential decay of the density in front of the swarm and behind the swarm easily follows.

When leader populations are assumed to move slowly we derive the parabolic model (57) to (60). Unlike the first parabolic scaling attempt, this system conserves the total number of leaders and the leader populations and the dynamics follow the transport equations known for hyperbolic time scales. Swarms successfully guided through slow leaders have a logical basis, for example consider some adults attempting to guide a group of small children through a park: the latter tend to move in a rather chaotic and random way, so the adults are forced to slow their movements to keep the group intact. The assumption of slow leaders does not hold, however, for certain other swarming systems, an example being bee swarms where active leaders are believed to engage in fast streaking behaviour. Fast directed movements prove problematic to handle in the parabolic limit, which can be attributed to the disparate time scales involved: fast target-directed streaks would result in an active leader rapidly covering a swarm's dimension, hence resulting (on a parabolic time scale) on almost instantaneous transitions from passive to active and back to passive leader. The slow leader model, can perhaps be regarded as a characterisation of such systems in which the circular movement of the leaders (moving to the front of the swarm and back) played out over a number of cycles is regarded as an effective translational movement of the leader population, now slow compared to the random movement of the followers. A careful analysis might, by separating fast and slow scales of the movement, be reduced to this situation. Such an analysis is beyond the scope of the current work.

Similar conclusions are obtained for the model in Section 3.6, corresponding to slow streakers and passive leaders which follow a random walk. Here, the transport equation for the streakers is combined with a diffusion equation for the passive leaders. The total number of leaders is conserved.

The modelling difficulties in the parabolic limit may be circumvented by *prescribing* the macroscopic evolution of the leader population, so that all properties macroscopically expected for the leader population are imposed in a phenomenological model. It replaces Equations (12), (13) by a deterministic set of equations for ρ_a and ρ_p. In this case, the scaling limit is only performed for the follower population using the arguments in Section 3.4. We again obtain (70) for the followers, coupled to the prescribed dynamics of the leaders.

6. **Inhomogeneous alignment kernel.** While (3) and (4) are common model equations to relate the mean direction and density, they do not adequately capture key aspects of the follower-leader interaction. For example, because the mean direction Λ does not depend on the size, but only the direction of $\mathcal{J}$, a single leader at time 0 can determine the direction of the swarm.

In this section we therefore also consider an alignment kernel given by

$$\Lambda^* = \nu \mathcal{J}(\mathbf{x}, t) \ , \tag{71}$$

where we have removed the normalization and $\mathcal{J}(\mathbf{x}, t)$ is given as in (3). Here we limit our discussion to the case $n = 2$. Parameter ν is the relaxation frequency,

previously assumed constant but in this approach taken to depend on the norm of $\mathcal{J}(\mathbf{x},t)$. Here we study the diffusion and hyperbolic limits of the system for follower-leader interactions, as previously done in Section 3. In particular, we focus on the equation for followers (47) and (19) under an alignment given by (71).

The distribution of aligned directions $\Phi(\Lambda^*\cdot\theta)$ in (8) is replaced by

$$\bar{\Phi}(\Lambda^*\cdot\theta)=\frac{\Phi(\Lambda^*\cdot\theta)}{\int_0^{2\pi}\Phi(|\Lambda^*|\cos\theta)d\theta} \tag{72}$$

such that $\int_S\bar{\Phi}(\Lambda^*\cdot\theta)d\theta=1$.

In the diffusion limit, we replace $\Phi^0(\Lambda\cdot\theta)$ by $\bar{\Phi}^0(\Lambda^*\cdot\theta)$ in equation (80). Noting that

$$\int_S\theta\bar{\Phi}^0(\Lambda^*_W\cdot\theta)d\theta=\bar{z}\Lambda^*_W\ ,\quad\text{where}\quad\bar{z}=|\Lambda^*|\int_0^{2\pi}\bar{\Phi}^0(|\Lambda^*|\cos(s))\cos(s)ds\ , \tag{73}$$

for $\theta=\cos(s)\Lambda^*+\sin(s)\Lambda^{*\perp}$ we again obtain (49), where in this case we replace $z\Lambda^0_W$ by $\bar{z}\Lambda^*_W$.

For the hyperbolic limit, let us first define the right hand side of (20) as

$$L(\sigma_f)=-\beta\sigma_f+\beta\zeta T\sigma_f+(1-\zeta)\beta\bar{\Phi}^\varepsilon(\Lambda^*\cdot\theta)\rho_f\ .$$

We know that as $\varepsilon\to0$ the solution $\sigma^0_f=\bar{\Psi}(\theta)\rho_f$, where $\bar{\Psi}(\theta)=\zeta+(1-\zeta)\bar{\Phi}^0(\Lambda^*\cdot\theta)$. The new operator $\bar{\Psi}(\theta)$ needs to be a Generalized Collisional Invariant of the operator $L(\sigma_f)$, as in the following sense [13, 14, 15].

Definition 6.1. A function $\bar{\Psi}(\theta)$ is a Generalized Collisional Invariant of Q if it satisfies

$$\int_S L(\sigma_f)\bar{\Psi}(\theta)d\theta=0\ ,$$

for any σ_f. Equivalently, σ_f satisfies $P_\perp\Big(\int_S\sigma_f(\mathbf{x},t,\theta)\theta d\theta\Big)=0$, where $P_\perp=\mathrm{Id}-\Lambda^*\otimes\Lambda^*$ is an orthogonal projection to Λ^*.

Note that for the case $\zeta=0$, i.e. when only alignment is considered, from Definition 6.1 we conclude that $\int_S L(\sigma_f)\bar{\Phi}^0(\Lambda^*\cdot\theta)d\theta=0$, where $\bar{\Phi}^0$ can be taken as the von Mises-Fisher distribution. Then, the analysis in [13] follows. The system (32)-(33) can be written now as

$$\partial_t\rho_f+c_f\bar{z}(1-\zeta)\nabla\cdot(\rho_f\Lambda^*)=0\ , \tag{74}$$

$$\rho_f(\bar{z}(1-\zeta)\partial_t\Lambda^*+\bar{C}_1\Lambda^*\cdot\nabla\Lambda^*)+\bar{C}_2P_\perp\nabla\rho_f=0\ , \tag{75}$$

where $\bar{z}$ is given by (73) and, similar to previous derivation, $\bar{C}_1=c_f(1-\zeta)\bar{a}_3$, $\bar{C}_2=c_f(1-\zeta)\mathbb{1}\bar{a}_1+c_f\mathbb{1}\pi\zeta$ for, $\bar{a}_3=\bar{a}_0-\bar{a}_1$ with

$$\bar{a}_0=|\Lambda^*|^2\int_0^{2\pi}\bar{\Phi}^0(|\Lambda^*|\cos(s))\cos^2(s)ds\ ,\ \ \bar{a}_1=|\Lambda^*|^2\int_0^{2\pi}\bar{\Phi}^0(|\Lambda^*|\cos(s))\sin^2(s)ds\ .$$

7. **Discussion.** The capacity of individuals to coordinate their movement to generate collective movement is a phenomenon that has attracted significant interest, in both cellular and animal systems. Much of the progress in this area has been facilitated through modelling studies, particularly via the employment of agent-based (or particle) descriptions that consider the movement response of each single member according to its neighbours. Yet the largest swarms can extend over kilometres and contain millions (e.g. herrings, [22]) or even hundreds of billions (e.g. desert locusts, [29, 33]) of members. At such numbers and scales, continuous modelling

approaches become necessary for their efficiency and increased tractability. Consequently, there is a clear interest in clarifying the relevant form of continuous models for swarming systems.

Much recent interest has focussed on follower-leader systems, where the population is decomposed into a leader population which somehow guides a population of followers, e.g. honey bee swarms where only scouts know the swarm's destination. Inspired by this and similar examples we have formulated a minimal microscopic description for a population of followers and leaders, deriving the ensuing macroscopic models under distinct scaling limits. To test the extent to which velocity alignment by itself can propel a guided and coherent swarm, follower orientation is limited to the interaction choice (4): alignment according to the velocity direction. Under both hyperbolic and parabolic scaling, macroscopic models feature drift-terms with advection in the active leader direction. Thus, alignment to velocity alone yields translocation of the swarm towards the target.

Under the hyperbolic scaling a pure-drift equation is generated, implying the potential for travelling-pulses with movement of a cohesive colony towards the target site. The parabolic limit, on the other hand, generates a drift-diffusion equation. While drift is in the direction of the target site, the additional dispersion leads to swarm spreading with time. Early stretching of the colony is to be expected, as the initially tight cluster morphs into a migrating swarm. Once established, though, swarms generally retain a relatively stable speed and shape: continued dispersal would be far from optimal, perhaps resulting in population loss or increasing the risk of predation. This, however, must be viewed in light of our intentionally simple modelling approach, where we have specifically tested the practicalities of an "alignment-only" mechanism. Agent-based approaches for modelling swarms are typically augmented by additional attractive/cohesion behaviours, where individuals are also pulled in the direction of those in their neighbourhood; in non-local continuous models, a similar effect is gained through a nonlocal attraction term that biases movement direction, e.g. [6]. As stressed above, we have reasonably excluded such considerations from the present model for simplicity, however including nonlocal attractions could clearly counteract swarm dispersal. For example, following [3], attraction (or repulsion) between individuals can be considered by

$$\mathcal{A} = \Big(\frac{d_c - \mathrm{dist}(\mathbf{x},\mathbf{y})}{\mathrm{dist}(\mathbf{x},\mathbf{y})}\Big) e^{-(d_c - \mathrm{dist}(\mathbf{x},\mathbf{y}))^2} \vec{e}(\mathbf{x},\mathbf{y}) \ .$$

Here $\mathrm{dist}(\mathbf{x},\mathbf{y})$ denotes the distance between two individuals $\mathbf{x}$ and $\mathbf{y}$, along the direction of the unit vector $\vec{e}(\mathbf{x},\mathbf{y})$. If $\mathrm{dist}(\mathbf{x},\mathbf{y}) < d_c$, where d_c is a critical distance, then the action of $\mathcal{A}$ is repulsive and if $\mathrm{dist}(\mathbf{x},\mathbf{y}) > d_c$ we have attraction.

Recently, there has been considerable interest in the composition and structuring of swarming populations. For example, in the case of bird flocks and fish shoals, the existence of faster and/or braver individuals can lead to hierarchical swarm arrangements [26, 30] and the question is raised as to how much swarm movement is dominated by the choice of a few. For simplicity, the movements of our leaders has been set here somewhat naively: during streaks, they operate as ballistic particles adopting fast movements towards the target. While this may be a reasonable approximation for bee swarms, where leaders have specific *a priori* knowledge, more general "leaders" may be more subtle and get influenced by neighbour movements. Consequently, a logical extension would be to also adopt a velocity jump model for the leader population, where one of their movement contributions stems

from an interaction function similar to (4), distinctly weighted. In other instances, leader/follower statuses may be transient, for example resulting from spatial position within the swarm, and it may be necessary to include switching terms between follower and leader populations.

The study of population dynamics, where a few discrete agents act as leaders, provides an excellent scenario to derive specific control laws and interactions that drive self-organisation, leading the swarm optimally to a desired outcome. This will allow an analytic understanding of follower-leader interactions, with applications not only to understanding the collective dynamics of biological populations but also outside biology, for example to hierarchical swarms of robots.

Appendix A. **Macroscopic diffusion limit for followers.** To compute the mean direction of the followers, w_f in Section 3.4, we start by substituting the following expansions

$$\begin{aligned} &\sigma_f = \sigma_f^0 + \varepsilon\sigma_f^1 + \mathcal{O}(\varepsilon^2)\ , \quad \rho_f = \rho_f^0 + \varepsilon\rho_f^1 + \mathcal{O}(\varepsilon^2)\ , \\ &T^\varepsilon = T^0 + \varepsilon T^1 + \mathcal{O}(\varepsilon^2)\ , \quad \Phi^\varepsilon(\Lambda\cdot\theta) = \Phi^0 + \varepsilon\Phi^1 + \mathcal{O}(\varepsilon^2)\ , \end{aligned} \tag{76}$$

into (47) and we obtain

$$\begin{aligned} \varepsilon^2\partial_t(\sigma_f^0+\varepsilon\sigma_f^1) + \varepsilon c_f\theta\cdot\nabla(\sigma_f^0+\varepsilon\sigma_f^1) = &-\beta(\sigma_f^0+\varepsilon\sigma_f^1) + \zeta\beta(T^0+\varepsilon T^1)(\sigma_f^0+\varepsilon\sigma_f^1) \\ &+\varepsilon(1-\zeta)\beta(\Phi^0+\varepsilon\Phi^1)(\rho_f^0+\varepsilon\rho_f^1)\ . \end{aligned} \tag{77}$$

Grouping in terms of powers of ε and assuming that $T^0\sigma_f^0 = \rho_f^0$ we have

$$\varepsilon^0:\quad \sigma_f^0 = \zeta\rho_f^0\ , \tag{78}$$

$$\varepsilon^1:\quad \sigma_f^1 = \frac{1}{1-\zeta}\Big(-\frac{c_f\zeta}{\beta}\theta\cdot\nabla\rho_f^0 + \zeta^2T^1\rho_f^0 + (1-\zeta)\Phi^0\rho_f^0\Big)\ . \tag{79}$$

The Chapman-Enskog expansion in this case is given by $\sigma_f = \zeta\rho_f^0 + \varepsilon\sigma_f^1 + \mathcal{O}(\varepsilon^2)$. The conservation equation (48) is given by

$$\varepsilon^2\partial_t\rho_f^0 + \varepsilon^2 c_f\nabla\cdot\int_S\theta\sigma_f^1 = 0\ .$$

Finally, we compute the mean direction w_f as follows

$$\nabla\cdot\int_S\theta\sigma_f^1 d\theta = -\Delta(D_f\rho_f^0) + \nabla\cdot\rho_f^0\int_S\theta\Phi^0(\Lambda_W\cdot\theta)d\theta\ , \tag{80}$$

where $D_f = \frac{\zeta c_f}{\beta(1-\zeta)}\int_S\theta\theta^T d\theta$. In the derivation of (80) we used $\int_S\theta T^1 d\theta = 0$ since we assume that T is symmetric.

Also, W is the total mean direction of the whole population and

$$\Lambda_W = \frac{\mathcal{J}_W}{|\mathcal{J}_W|}\quad\text{for}\quad \mathcal{J}_W = n\varepsilon\int_{\mathbf{y}}K^\varepsilon\Big(\frac{|\mathbf{y}-\mathbf{x}|}{\varepsilon}\Big)W(\mathbf{y},t)d\mathbf{y}\ .$$

If we consider $\mathcal{J}$ as in (4) then the mean direction, Λ, will depend only on the mean direction of the active leaders, w_a, as follows

$$\Lambda_a = \frac{\mathcal{J}_a}{|\mathcal{J}_a|}\ \text{where}\ \mathcal{J}_a = n\varepsilon\lambda\int_{\mathbf{y}}K^\varepsilon\Big(\frac{|\mathbf{y}-\mathbf{x}|}{\varepsilon}\Big)w_a d\mathbf{y}\ .$$

The integral over S in (80) is given by $\int_S\theta\Phi^0(\Lambda_W^0\cdot\theta)d\theta = z\Lambda_W^0$ where z can be computed using polar coordinates $\theta = (\cos(s),\sin(s))$ for $n = 2$, or spherical

coordinates $\theta = (\cos\phi\sin(s), \sin\phi\sin(s), \cos(s))$ for $n = 3$. z is given by [15]

$$z = \begin{cases} \int_0^{2\pi} \Phi^0(\cos(s))\cos(s)ds, & \text{if } n = 2, \\ 2\pi \int_0^{\pi} \Phi^0(\cos(s))\cos(s)\sin(s)ds, & \text{if } n = 3. \end{cases} \tag{81}$$

Appendix B. **Diffusion limit for passive leaders with noise.** Let us suppose that the total population of passive leaders, σ_p, can be written in terms of the following expansion:

$$\sigma_p = \sigma_p^0 + \varepsilon\sigma_p^1 + \mathcal{O}(\varepsilon^2) \ . \tag{82}$$

Similarly, we expand $\rho_p = \rho_p^0 + \varepsilon\rho_p^1 + \mathcal{O}(\varepsilon^2)$, $\rho_a = \rho_a^0 + \varepsilon\rho_a^1 + \mathcal{O}(\varepsilon^2)$ and

$$\begin{aligned} B^\varepsilon(\theta) &= B^0(\theta) + \varepsilon B^1(\theta) + \mathcal{O}(\varepsilon^2) \ , \\ \bar{R}^\varepsilon_{ap}(\theta) &= \bar{R}^0_{ap}(\theta) + \varepsilon\bar{R}^1_{ap}(\theta) + \mathcal{O}(\varepsilon^2) \ , \qquad R^\varepsilon_{ap} = R^0_{ap} + \varepsilon R^1_{ap} + \mathcal{O}(\varepsilon^2) \ , \end{aligned} \tag{83}$$

where $B^0(\theta) = pT^0(\theta) + (1-p)T^0_{\rho_f}$ and $B^1(\theta) = pT^1(\theta) + (1-p)T^1_{\rho_f}$.

Substituting (82) and (83) into (46) and grouping the appropriate powers of ε we get

$$\varepsilon^0 : \quad \sigma_p^0 = B^0\rho_p^0 \ , \tag{84}$$

$$\varepsilon^1 : \quad \sigma_p^1 = \frac{1}{\beta}(-c_p\theta\cdot\nabla\sigma_p^0 + \beta B^1\rho_p^0 + \beta B^0\rho_p^1) \ . \tag{85}$$

We obtain a Chapman-Enskog expansion from (82) given by

$$\sigma_p = B^0\rho_p^0 + \varepsilon\sigma_p^1 + \mathcal{O}(\varepsilon^2) \ . \tag{86}$$

Substituting the above expression into the mean direction of the passive leaders, w_p in (65), we obtain

$$w_p = \frac{1}{n}\int_S \theta B^0(\theta)d\theta\rho_p^0 + \frac{\varepsilon}{n}\int_S \theta\sigma_p^1 d\theta = \frac{\varepsilon}{n}\int_S \theta\sigma_p^1 d\theta \ .$$

Since we assume B^0 is symmetric, then $\int_S \theta B^0 d\theta = 0$. The conservation equation for the passive leaders (62), is then

$$\varepsilon^2\partial_t\rho_p^0 + \varepsilon^2 c_p\nabla\cdot\int_S \theta\sigma_p^1 d\theta = \varepsilon^2 R^0_{ap}\rho_a^0 - \varepsilon^2 R^0_{pa}\rho_p^0 \ . \tag{87}$$

The term $\nabla\cdot\int_S \theta\sigma_p^1 d\theta$ can be explicitly computed using (84) and (85) as follows:

$$\begin{aligned} \nabla\cdot\int_S \theta\sigma_p^1 d\theta &= \frac{-c_p}{\beta}\int_S (\nabla\cdot\theta)(\theta\cdot\nabla)B^0\rho_p^0 d\theta \\ &\quad + \frac{1}{\beta}\nabla\cdot\int_S \theta\Big(\beta B^1\rho_p^0 + \beta B^0\rho_p^1\Big)d\theta \ , \\ &= -\Delta(D_p + \mathbb{T})\rho_p^0 \ , \end{aligned} \tag{88}$$

where we have used the fact that $B(\theta)$ is symmetric and therefore $\int_S \theta B^0 d\theta = \int_S \theta B^1 d\theta = 0$ and

$$D_p = \frac{pc_p}{\beta}\int_S \theta\theta^T T^0(\theta)d\theta \ , \quad \text{and} \quad \mathbb{T} = \frac{(1-p)c_p}{\beta}\int_S \theta\theta^T T^0_{\rho_f} d\theta \ .$$

Appendix C. **Turn angle operator.** This section recalls some basic spectral properties of the turn angle operator T defined in (2).

Lemma C.1. *Assume that $\tilde{k}$ is continuous. Then T is a symmetric compact operator. In particular, there exists an orthonormal basis of $L^2(S)$ consisting of eigenfunctions of T.*
With $\theta = (\theta_0, \theta_1, ..., \theta_{n-1}) \in S$, we have

$$\begin{aligned} &\phi_0(\theta) = 1 \qquad \text{is an eigenfunction to the eigenvalue} \quad \nu_0 = 1, \\ &\phi_1^j(\theta) = n\theta_j \quad \text{are eigenfunctions to the eigenvalue} \quad \nu_1 = \int_S \tilde{k}(\cdot, |\eta - e_1|)\eta_1 d\eta < 1. \end{aligned} \tag{89}$$

Any function $\sigma \in L^2(\mathbb{R}^n \times \mathbb{R}^+ \times S)$ admits a unique decomposition

$$\sigma(\mathbf{x}, t, \theta) = u + n\theta \cdot w + \hat{z}, \tag{90}$$

where $\hat{z}$ is orthogonal to all linear polynomials in θ. Explicitly,

$$u(\mathbf{x}, t) = \int_S \sigma(\mathbf{x}, t, \theta)\phi_0(\theta)d\theta, \;\; w^j(\mathbf{x}, t) = \int_S \sigma(\mathbf{x}, t, \theta)\phi_1^j(\theta)d\theta,$$

and $w = (w^1, \dots, w^n)$.

Acknowledgments. This research was partially supported by the Italian Ministry of Education, University and Research (MIUR) through the "Dipartimenti di Eccellenza" Programme (2018-2022) – Dipartimento di Scienze Matematiche "G. L. Lagrange", Politecnico di Torino (CUP: E11G18000350001). SB is member of GNFM (Gruppo Nazionale per la Fisica Matematica) of INdAM (Istituto Nazionale di Alta Matematica), Italy. KJP acknowledges departmental funding through the "MIUR–Dipartimento di Eccellenza" programme.

REFERENCES

[1] W. Alt. Biased random walk models for chemotaxis and related diffusion approximations. *Journal of Mathematical Biology*, 9(2):147–177, 1980.
[2] M. Beekman, R. L. Fathke, and T. D. Seeley. How does an informed minority of scouts guide a honeybee swarm as it flies to its new home? *Animal Behaviour*, 71(1):161–171, 2006.
[3] N. Bellomo. *Modeling complex living systems: a kinetic theory and stochastic game approach*. Springer Science & Business Media, 2008.
[4] A. M. Berdahl, A. B. Kao, A. Flack, P. A. Westley, E. A. Codling, I. D. Couzin, A. I. Dell, and D. Biro. Collective animal navigation and migratory culture: from theoretical models to empirical evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1746):20170009, 2018.
[5] S. Bernardi, A. Colombi, and M. Scianna. A particle model analysing the behavioural rules underlying the collective flight of a bee swarm towards the new nest. *Journal of Biological Dynamics*, 12(1):632–662, 2018.
[6] S. Bernardi, R. Eftimie, and K. J. Painter. Leadership through influence: what mechanisms allow leaders to steer a swarm? *Bulletin of Mathematical Biology*, 83(6):1–33, 2021.
[7] L. J. Brent, D. W. Franks, E. A. Foster, K. C. Balcomb, M. A. Cant, and D. P. Croft. Ecological knowledge, leadership, and the evolution of menopause in killer whales. *Current Biology*, 25(6):746–750, 2015.
[8] J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil. Particle, kinetic, and hydrodynamic models of swarming. In *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, pages 297–336. Springer, 2010.
[9] I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin. Effective leadership and decision-making in animal groups on the move. *Nature*, 433(7025):513, 2005.
[10] I. D. Couzin, J. Krause, R. James, G. D. Ruxton, and N. R. Franks. Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, 218(1):1–11, 2002.

[11] F. Cucker and S. Smale. Emergent behavior in flocks. *IEEE Transactions on Automatic Control*, 52(5):852–862, 2007.
[12] F. Cucker and S. Smale. On the mathematics of emergence. *Japanese Journal of Mathematics*, 2(1):197–227, 2007.
[13] P. Degond, A. Frouvelle, and J.-G. Liu. Macroscopic limits and phase transition in a system of self-propelled particles. *Journal of nonlinear science*, 23(3):427–456, 2013.
[14] P. Degond and S. Motsch. Continuum limit of self-driven particles with orientation interaction. *Mathematical Models and Methods in Applied Sciences*, 18(supp01):1193–1215, 2008.
[15] G. Dimarco and S. Motsch. Self-alignment driven by jump processes: Macroscopic limit and numerical investigation. *Mathematical Models and Methods in Applied Sciences*, 26(07):1385–1410, 2016.
[16] R. Eftimie. Hyperbolic and kinetic models for self-organized biological aggregations and movement: a brief review. *Journal of mathematical biology*, 65(1):35–75, 2012.
[17] R. Eftimie, G. De Vries, M. Lewis, and F. Lutscher. Modeling group formation and activity patterns in self-organizing collectives of individuals. *Bulletin of Mathematical Biology*, 69(5):1537, 2007.
[18] G. Estrada-Rodriguez and H. Gimperlein. Interacting particles with levy strategies: limits of transport equations for swarm robotic systems. *arXiv preprint arxiv:1807.10124*, 2019.
[19] U. Greggers, C. Schoening, J. Degen, and R. Menzel. Scouts behave as streakers in honeybee swarms. *Naturwissenschaften*, 100(8):805–809, 2013.
[20] A. Haeger, K. Wolf, M. M. Zegers, and P. Friedl. Collective cell migration: guidance principles and hierarchies. *Trends in cell biology*, 25(9):556–566, 2015.
[21] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *Departmental Papers (ESE)*, page 29, 2003.
[22] N. C. Makris, P. Ratilal, S. Jagannathan, Z. Gong, M. Andrews, I. Bertsatos, O. R. Godø, R. W. Nero, and J. M. Jech. Critical population density triggers rapid formation of vast oceanic fish shoals. *Science*, 323(5922):1734–1737, 2009.
[23] R. Mayor and S. Etienne-Manneville. The front and rear of collective cell migration. *Nature reviews Molecular cell biology*, 17(2):97, 2016.
[24] A. Mogilner and L. Edelstein-Keshet. A non-local model for a swarm. *Journal of Mathematical Biology*, 38(6):534–570, 1999.
[25] T. Mueller, R. B. O'Hara, S. J. Converse, R. P. Urbanek, and W. F. Fagan. Social learning of migratory performance. *Science*, 341(6149):999–1002, 2013.
[26] M. Nagy, Z. Akos, D. Biro, and T. Vicsek. Hierarchical group dynamics in pigeon flocks. *Nature*, 464(7290):890, 2010.
[27] H. G. Othmer, S. R. Dunbar, and W. Alt. Models of dispersal in biological systems. *Journal of mathematical biology*, 26(3):263–298, 1988.
[28] K. Painter, J. Bloomfield, J. Sherratt, and A. Gerisch. A nonlocal model for contact attraction and repulsion in heterogeneous cell populations. *Bulletin of mathematical biology*, 77(6):1132–1165, 2015.
[29] R. Rainey. Radar observations of locust swarms. *Science*, 157(3784):98–99, 1967.
[30] S. G. Reebs. Can a minority of informed leaders determine the foraging movements of a fish shoal? *Animal behaviour*, 59(2):403–409, 2000.
[31] T. D. Seeley. *Honeybee democracy*. Princeton University Press, 2010.
[32] J. Shen. Cucker–smale flocking under hierarchical leadership. *SIAM Journal on Applied Mathematics*, 68(3):694–719, 2007.
[33] R. Skaf, G. Popov, and J. Roffey. The desert locust: an international challenge. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 328(1251):525–538, 1990.
[34] C. M. Topaz, A. L. Bertozzi, and M. A. Lewis. A nonlocal continuum model for biological aggregation. *Bulletin of mathematical biology*, 68(7):1601, 2006.
[35] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel type of phase transition in a system of self-driven particles. *Physical review letters*, 75(6):1226, 1995.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: sara.bernardi@polito.it
E-mail address: estradarodriguez@ljll.math.upmc.fr
E-mail address: h.gimperlein@hw.ac.uk
E-mail address: kevin.painter@polito.it