PAPER • OPEN ACCESS

Collective response to local perturbations: how to evade threats without losing coherence

View the article online for updates and enhancements.

Physical Biology

PAPER

Collective response to local perturbations: how to evade threats without losing coherence

Emanuele Loffredo[1],[*] , Davide Venturelli[2],[*] and Irene Giardina[3,4,5]

[1] Laboratory of Physics of the École Normale Supérieure, CNRS UMR 8023 and PSL Research, 24 rue Lhomond, 75231 Paris Cedex 05, France
[2] SISSA—International School for Advanced Studies and INFN, via Bonomea 265, 34136 Trieste, Italy
[3] Dipartimento di Fisica, Università Sapienza, 00185 Rome, Italy
[4] Istituto Sistemi Complessi, Consiglio Nazionale delle Ricerche, UOS Sapienza, 00185 Rome, Italy
[5] INFN, Unità di Roma 1, 00185 Rome, Italy
[*] Authors to whom any correspondence should be addressed.

E-mail: emanuele.loffredo@phys.ens.fr and davide.venturelli@sissa.it

Abstract

Living groups move in complex environments and are constantly subject to external stimuli, predatory attacks and disturbances. An efficient response to such perturbations is vital to maintain the group's coherence and cohesion. Perturbations are often local, i.e. they are initially perceived only by few individuals in the group, but can elicit a global response. This is the case of starling flocks, that can turn very quickly to evade predators. In this paper, we investigate the conditions under which a global change of direction can occur upon local perturbations. Using minimal models of self-propelled particles, we show that a collective directional response occurs on timescales that grow with the system size and it is, therefore, a finite-size effect. The larger the group is, the longer it will take to turn. We also show that global coherent turns can only take place if i) the mechanism for information propagation is efficient enough to transmit the local reaction undamped through the whole group; and if ii) motility is not too strong, to avoid that the perturbed individual leaves the group before the turn is complete. No compliance with such conditions results in the group's fragmentation or in a non-efficient response.

1. Introduction

Collective behavior in living aggregations often has a strong anti-predatory function. An efficient group response is crucial to maintain global coherence and cohesion in spite of attacks and disturbances [1–8]. To this end, many groups exhibit collective directional changes in very short times, often triggered by local perturbation events. A beautiful example is the one of bird flocks [9], where collective turns start from a single individual, and the directional information is passed through the group so quickly that the whole flock performs the turn retaining its structure and without attenuation. From the perspective of statistical physics, this behavior is somewhat unusual. Indeed, when considering physical systems with long-range directional order—either at equilibrium or active—perturbing locally the system does not in general change its ordered state, which is another way of saying that the order is stable (i.e. ergodicity is broken in the thermodynamic limit). For instance, if we consider a ferromagnet and apply a magnetic field on a single site of the lattice, the magnetization will remain unaltered. What marks the difference with respect to animal groups is, of course, their size. Living groups might be large, comprising hundreds or thousands of individuals, but they are very far from the order Avogadro numbers that characterize physical systems. Perturbations which would be irrelevant in the thermodynamic limit might in fact change the state of a finite group on observational timescales, if its size is small enough. Response to local perturbations is thus, intrinsically, a finite-size effect. The kind of response exhibited by the system, and the effective timescale for collective adaptation, also depend on the mechanism of information propagation. In

flocks, a linear dispersion law has been observed [9], suggesting that second order dynamics for the birds flight directions should be at play [10, 11]. Other animal groups display instead less efficient information transfer. For instance, experiments on fish schools [3, 4, 7] show that groups exhibit evasion manoeuvres arising from induced alarming stimuli. However—contrary to bird flocks—the speed of information propagation decelerates over time, and thus only part of the group follows the initiator of the evasion event, resulting in a distribution of behavioral cascades. In this case, a dissipative dynamics rules individual directional motion, and the local connectivity of the interaction network might dramatically affect the extension of the collective response [4].

These examples indicate that several factors might contribute to the response behavior of finite groups, and it is not always clear how to disentangle one factor from the other. In this work we perform a systematic study to explore the interplay between size, dynamical rules, motility and boundary conditions in determining the response of the system to local directional perturbations. The framework of our study is the one of self-propelled models of collective motion, where the minimal number of parameters allows to exhaustively explore the model's space.

This paper is organized as follows. In section 2 we introduce the model discussed in our theoretical and numerical analysis: the inertial spin model (ISM [10]). This model is a generalization of the Vicsek model (VM) [12] comprising both inertial and dissipative terms in the dynamical equation for the velocities. In the underdamped limit it reproduces the linear dispersion law observed in flocks of birds (for which it was originally proposed), while in the overdamped limit it reduces to the VM. It therefore represents an ideal playground to investigate different dynamical regimes potentially relevant for different behaviors. In the following sections we consider progressively all the factors that might affect the response behavior. In section 3 we start by looking at the ISM with a fixed interaction network. In this case it is possible to derive analytically the equations describing the evolution of the system, when a local perturbation is applied to a single individual. This allows to clearly understand what is the role of the system's size, and how a collective turn manifests itself in the different dynamical regimes of the model. In section 4 we perform numerical simulations that quantitatively confirm our analytical results. Next, in section 5 we consider the off-lattice, active version of the model, and elucidate how motility affects the fixed-network scenario. Finally, in section 6 we discuss what happens when open boundary conditions (OBCs) are considered, and we show how and when the mechanisms studied in the previous sections can lead to a fragmentation of the group.

2. The modelling framework

Collective behavior in animal groups has been described with a variety of approaches, either considering evolution rules for the individuals in the group [13, 14], or by using coarse-grained equations for mesoscopic fields [15–17]. In this study, we work in the context of self-propelled particle (SPP) models, where the aggregation is modelled in a minimal way as a collection of particles with fixed activity interacting with each other via alignment/imitation rules.

2.1. The VM

The most famous among SPP models of collective motion is the VM ([12]), which we now present in its continuous-time version. Consider a system made of N point-like active particles, each labeled by an index i, which move in a d-dimensional space following the equations of motion

$$\frac{d\mathbf{r}_i}{dt} = \mathbf{v}_i \, , \tag{1}$$

$$\eta \frac{d\mathbf{v}_i}{dt} = -\frac{\partial \mathcal{U}}{\partial \mathbf{v}_i} + \lambda_i \mathbf{v}_i + \boldsymbol{\xi}_i \, . \tag{2}$$

Each particle position $\mathbf{r}_i \in \mathbb{R}^d$ evolves deterministically according to its corresponding velocity vector $\mathbf{v}_i \in \mathbb{R}^{d_v}$ of fixed modulus $|\mathbf{v}_i| = v_0$, while the latter undergoes a Langevin overdamped dynamics with a 'social' alignment force given by

$$\mathcal{U}(\{\mathbf{v}_i\}) = -\frac{J}{2v_0^2} \sum_{ij} n_{ij} \mathbf{v}_i \cdot \mathbf{v}_j, \tag{3}$$

$$-\frac{\partial \mathcal{U}}{\partial \mathbf{v}_i} = \frac{J}{v_0^2} \sum_j n_{ij} \mathbf{v}_j. \tag{4}$$

We recognize in equation (3) the Hamiltonian of the Heisenberg model, with the coefficient $J > 0$ setting the strength of the alignment interactions, and where the connectivity matrix n_{ij} specifies the interacting neighbors. However, in contrast to the standard Heisenberg model, here the connectivity matrix itself evolves in time through the particle positions, i.e. $n_{ij} = n_{ij}(\{\mathbf{r}_i(t)\}_{i=1}^N)$. Among the possible evolution rules, we will adopt in the following the paradigm of *metric* interactions: each particle interacts with those lying within a fixed interaction radius r_c, i.e. $n_{ij}(t) = \Theta(r_c - |\mathbf{r}_i(t) - \mathbf{r}_j(t)|)$.

Featuring in equation (2) are a viscosity coefficient η and a white Gaussian noise $\boldsymbol{\xi}_i$, which are linked by the Einstein-like relation

$$\langle \boldsymbol{\xi}_i(t) \cdot \boldsymbol{\xi}_j(t') \rangle = 2d\eta T \delta_{ij} \delta(t - t') , \tag{5}$$

where the *temperature* T sets the strength of the fluctuations. Finally, the parameter λ_i is a Lagrange multiplier which can be used to enforce the fixed *speed* condition $d|\mathbf{v}_i|^2/dt = 0$ [18].

Due to the time dependence in $n_{ij}(t)$, it is known that the VM admits, even in $d \leqslant 2$, a T-driven transition from an ordered phase (*flock*), where the individual velocity vectors align and add up to a non-zero total velocity

$$\mathbf{V} \equiv \frac{1}{N} \sum_{i=1}^{N} \mathbf{v}_i \,, \qquad (6)$$

to a disordered phase (*swarm*), where the mean group velocity $\mathbf{V}$ is null. A convenient order parameter to describe the transition is the so-called polarization vector

$$\mathbf{\Psi} \equiv \frac{1}{N} \sum_{i=1}^{N} \frac{\mathbf{v}_i}{|\mathbf{v}_i|} = \frac{\mathbf{V}}{v_0} \,, \qquad (7)$$

and its corresponding (scalar) polarization $\Psi = |\mathbf{\Psi}| \in [0,1]$.

2.2. The ISM

The VM successfully describes the large-scale behavior of several living and non-living active systems [13, 16]. However, the VM is not appropriate to explain collective turns in flocks of birds. Experiments indeed show that, in turning flocks, the directional information travels obeying a linear dispersion law, with a propagation speed only depending on the degree of order in the system [9]. The Vicsek dynamics does not reproduce this behavior. An intuition of why this occurs can be grasped by looking at equation (2): its structure is that of an overdamped Langevin equation for the velocity, and this kind of equations usually lead to a diffusive dispersion law [19]. The simplest way to obtain linearly dispersive solutions is to reinstate a second order derivative in equation (2), and this is precisely what the ISM does [10]. As in standard second order equations, it is convenient to write the model as a system of first order equations, by introducing appropriate conjugate variables. The ISM then reads

$$\frac{d\mathbf{r}_i}{dt} = \mathbf{v}_i \,, \qquad (8)$$

$$\frac{d\mathbf{v}_i}{dt} = \frac{1}{\chi} \mathbf{s}_i \times \mathbf{v}_i \,, \qquad (9)$$

$$\frac{d\mathbf{s}_i}{dt} = \frac{\mathbf{v}_i}{v_0} \times \left(-\frac{1}{v_0} \frac{\partial \mathcal{U}}{\partial \mathbf{v}_i} - \frac{\eta}{v_0} \frac{d\mathbf{v}_i}{dt} + \boldsymbol{\xi}_i \right) , \qquad (10)$$

with the noise correlations given in equation (5). In contrast to the VM, in equation (10) the force term does not act directly on the velocity, but rather on its derivative, which is expressed in terms of the new variable $\mathbf{s}_i$. The vectorial products enforce the constraint on the individual speeds. The particles can thus only change their directions: $\mathbf{s}_i$ therefore plays

the role of an internal angular momentum, regulating the rotations of the individual velocity vectors, and this is why it is called a 'spin'. The new parameter χ plays the role of a rotational *inertia*.

2.3. The planar ISM

Let us now specialize the ISM to the planar two-dimensional case, where velocities lie on a plane (i.e. $d_v = 2$). This case is simpler to handle algebraically, and it gives a direct intuition of the physical and biological meaning of the spin and of inertial dynamics. Besides, it is also the relevant one for collective turns in flocks (as we will discuss later on), and for a variety of other biological groups. Generalization to the three-dimensional case is straightforward.

In the planar case, we can write

$$\mathbf{v}_i(t) \equiv v_0 \left(\cos \varphi_i(t), \sin \varphi_i(t) \right) , \qquad (11)$$

where the phase φ_i specifies the angle of the velocity vector with respect to a reference direction. In terms of the phases, equations (9) and (10) assume a much simpler form, namely

$$\dot{\varphi}_i = \frac{s_i}{\chi} \,, \qquad (12)$$

$$\dot{s}_i = -\frac{\partial \mathcal{U}}{\partial \varphi_i} - \frac{\eta}{\chi} s_i + \xi_i \,. \qquad (13)$$

In these equations the spin vector reduces to one single component, whose value s_i identifies the angular velocity of each particle. In the presence of a force, the spin acquires a non-zero value and, as a consequence, the particle turns. It can be shown that the instantaneous radius of curvature of the trajectory is proportional to the inverse value of the spin [10].

Equations (12) and (13) can be rewritten in terms of the phase only, giving

$$\chi \ddot{\varphi}_i + \eta \dot{\varphi}_i = -\frac{\partial \mathcal{U}}{\partial \varphi_i} + \xi_i \,. \qquad (14)$$

In the overdamped limit $\chi/\eta \to 0$ this second order equation reduces to a first-order one, which coincides with the planar Vicsek continuous model derived from equation (2). The ISM is therefore an inertial generalization of the VM, which reduces to the latter in the dissipative limit. The opposite limit in which $\eta \to 0$ renders a deterministic equation with a Hamiltonian reversible structure (see equations (12) and (13)). In this case, since s_i represents the generator of the rotational symmetry of the velocity, its total value $S = (1/N) \sum_i s_i$ is a conserved quantity.

2.4. Collective turns and dynamical regimes

Let us now discuss how the ISM can appropriately describe the dispersion law observed in flocks of

birds [9]. To do so, we recall first what is known from experiments [9, 20]:

- Bird flocks are highly ordered systems, with measured polarization values $\Psi \sim 0.9$;
- Turns are planar, i.e. each trajectory lies approximately on a $2d$ plane;
- Mutual distances do not change during turns, and individuals sweep equal radius paths;
- Turns have a localized spatial origin, and the signal propagates linearly in time through the group with speed $c_s \gg v_0$ and with negligible attenuation.

Given these premises, we can now focus on the planar version of the model, as in equation (14). This expression becomes even simpler if the system is in the deeply ordered phase, as flocks are. In this case, if we choose as a reference direction the one of the global velocity **V**, the individual phases are very small, i.e. $\varphi_i \ll 1$. One can then expand the potential $\mathcal{U}$ in equation (3) up to quadratic order—which is called the *spin-wave* approximation (SWA)—to find

$$\mathcal{U} = -\sum_{ij} n_{ij} \cos(\varphi_j - \varphi_i) \sim \sum_{ij} \Lambda_{ij} \varphi_j \varphi_i \,, \quad (15)$$

where we have introduced the discrete Laplacian Λ with

$$\Lambda_{ij} = -n_{ij} + \delta_{ij} \sum_k n_{ik} \,. \quad (16)$$

Since individuals do not change their mutual positions during turns, we can assume that the interaction network n_{ij} remains approximately the same. If the phases vary slowly from one individual to the next, we can further treat phases as a continuous field $\varphi_i \rightarrow \varphi(\mathbf{x})$. The discrete Laplacian then becomes a continuous one, $\Lambda_{ij} \rightarrow -a^2 \boldsymbol{\nabla}^2$ (a being the mean nearest-neighbor distance), and equation (14) becomes

$$\chi \ddot{\varphi} + \eta \dot{\varphi} - a^2 J n_c \nabla^2 \varphi = \xi \,, \quad (17)$$

with $n_c = \sum_j n_{ij}$ (assumed to be constant on a regular network). While computations can be easily performed also in terms of the discrete variables, the continuous form is more convenient for visualizing the dispersion relation. Indeed, let us now set $\xi = 0$ into equation (17) (or equivalently focus on $\langle \varphi \rangle$), and Fourier transform in both space and time to get

$$\chi \omega^2 - i \eta \omega - a^2 J n_c k^2 = 0 \,. \quad (18)$$

From this relation we can immediately see that in the Vicsek case ($\chi = 0$) the dispersion law becomes purely diffusive, i.e. $\omega \sim i k^2$, leading to strong attenuation and a non-linear relation between space and time. Conversely, when χ is different from zero a more complex dispersion law is obtained,

$$\omega_{\pm} = i\gamma \pm c_s \sqrt{k^2 - k_0^2} \,, \quad (19)$$

where we introduced the propagation speed c_s and the effective dissipation γ as

$$c_s = a \sqrt{J n_c / \chi} \,, \quad (20)$$

$$\gamma = \frac{\eta}{2\chi} \,, \quad (21)$$

and where $k_0 \equiv \gamma / c_s$.

From equation (19) we can clearly evince that the two conditions observed in flocks of birds, i.e. linear dispersion law and no attenuation, are reproduced when $\gamma \ll 1$ and $k_0 \ll k$. In this deeply underdamped regime, the ISM predicts a linear dispersion law with propagation speed given by equation (20). This is a highly non-trivial relation linking together the way directional information propagates through the system, and the degree of order (set by J). Remarkably, this prediction is very nicely verified in experimental data [9], thus supporting the idea that the ISM in the strongly underdamped regime provides a good description for collective turning in flocks.

More generally, according to the ratio of inertia and dissipation, and depending on the size of the system, the ISM interpolates between the dissipative Vicsek limit, and the efficient limit of flocks. For $k \ll k_0$ all the modes are overdamped, while for $k \gg k_0$ the system sustains linear propagation with speed c_s and constant attenuation with damping time $\tau = \gamma^{-1}$. Since k_0 is the inverse of a length scale, it sets a limit on the size L of a flock through which a signal can linearly propagate: indeed, imposing $k_{\min} = L^{-1} \gg k_0$ implies $L \ll c_s / \gamma$. Another equivalent way to understand this constraint comes by thinking in terms of timescales. Indeed, the time needed for a signal of speed c_s to cross the entire flock of size L is

$$\tau_s = \frac{L}{c_s} = \frac{L}{a} \sqrt{\frac{\chi}{J n_c}} \,, \quad (22)$$

but the same signal gets attenuated over a time scale

$$\tau = \gamma^{-1} \,. \quad (23)$$

We can then identify two dynamical propagation regimes:

- *Underdamped regime (inertial propagation)*: $\tau_s < \tau$. In this case directional changes travel linearly through the whole system before attenuation can deteriorate the signal. Flocks of birds observed in experiments [9] belong to the deep extreme of this regime, where attenuation is negligible for all the individuals. In terms of the parameters of the model, the condition defining this regime is $L^2 \eta^2 / (4 a^2 J n_c \chi) < 1$.
- *Overdamped regime (dissipative propagation)*: $\tau_s > \tau$. Propagation of information is inefficient because the signal gets dissipated before reaching the other end of the group. In this case, the signal travels

unaffected up to certain scales $k^{-1} < L$, but it is damped on larger scales. This might result in a strong deformation of the group, or even in its fragmentation. It is likely that fish schools investigated in [4] belong to this regime.

The dispersion law and the way information is propagated through the system determine the dynamical behavior of the scalar polarization and of the correlation functions [18]. In this work, we are interested in understanding how collective turns might arise in finite groups, and we shall find that the dispersion law affects in a crucial way the occurrence of these events. To do so, we will study the response of finite systems to a local perturbation under different boundary and motility conditions, and we will explore the dynamical regimes numerically by tuning the various parameters of the model.

3. Perturbation events

In the previous section we introduced the ISM and we discussed the predictions of the model, under specific approximations, concerning the dispersion law in the system. In this section, we want to generalize our discussion to the study of perturbation events. When a group changes its global direction of motion, this often happens because some external perturbation or disturbance acts upon one or more individuals in the group. This is not the only possibility (see e.g. [21]), but it is certainly one of the most interesting. A reasonable choice is to model such a perturbation as a field $\mathbf{h}_i(t)$ linearly coupled to the individual velocities, in which case

$$\mathcal{U} \rightarrow \mathcal{U}_h = \mathcal{U} - \frac{1}{v_0} \sum_i \mathbf{h}_i \cdot \mathbf{v}_i . \qquad (24)$$

We note that, in principle, in the ISM it would also be possible to couple a field to the spins $\mathbf{s}_i$. This would however change the angular velocity of the perturbed individuals, while it would not bias them directionally, which is here our main purpose (see however [22]). In the planar case, the equations of motion then read

$$\dot{\boldsymbol{r}} = \boldsymbol{v}_i , \qquad (25)$$

$$\dot{\varphi}_i = \frac{s_i}{\chi}, \qquad (26)$$

$$\dot{s}_i = -\frac{\partial \mathcal{U}_h}{\partial \varphi_i} - \frac{\eta}{\chi} s_i + \xi_i . \qquad (27)$$

These are the equations that we will consider in the numerical simulations performed in the remaining of this work. To make analytical progress, however, we still need to perform some simplifying approximations.

3.1. High order and slow network
First of all, we consider a system in its polarized phase. Indeed, our aim is to investigate how a state of collective motion can change under external perturbations. We then perform the following approximations:

- *Spin-wave approximation*: the system is highly polarized, and the individual velocity vectors $\mathbf{v}_i$ deviate weakly from the orientation of the global polarization $\mathbf{V}$. If the latter is chosen to be initially aligned to the *x*-axis when the external field is applied, then $\varphi_i \ll 1$ in the initial phases of the collective turn.
- *Fixed network*: we assume that the adjacency matrix n_{ij} appearing in the equations of motion is no longer a function of time. This assumption is reasonable when the timescale over which the collective turn develops is much shorter than the typical reshuffling time of the interaction network n_{ij}. This condition of *local equilibrium* [23] is precisely what happens in flocks of starlings. A deviation from this condition is however expected to arise for sufficiently large values of v_0. This analysis will be the subject of section 5.

Under these approximations, the equations of motion for s_i and φ_i take the form

$$\dot{\varphi}_i = \frac{s_i}{\chi}, \qquad (28)$$

$$\dot{s}_i \simeq -J \sum_j \Lambda_{ij} \varphi_j - \varphi_i h_i \cos\alpha_i + h_i \sin\alpha_i - \frac{\eta}{\chi} s_i + \xi_i , \qquad (29)$$

where α_i denotes the direction of $\mathbf{h}_i$, while $h_i = h_i(t)$ is its magnitude. We choose as the relevant observable the average polarization angle

$$\Phi(t) = \frac{1}{N} \sum_i \langle \varphi_i(t) \rangle . \qquad (30)$$

This $\Phi(t)$ coincides, for small angles φ_i, with the angle formed by the polarization measured at time t (after the perturbation) with the initial polarization (i.e. the one maintained in the absence of perturbations). It therefore characterizes the way the system changes in time its collective flight direction.

3.2. Analytical results
The derivation of $\Phi(t)$ within the microscopic approach presented above is carried out in detail in appendix A.1 for a generic adjacency network n_{ij}; an alternative derivation, for a regular lattice and in terms of coarse-grained fields $\varphi(\mathbf{x}, t)$, is presented in appendix A.2, leading to the same result. Here we simply present the final expression, which will be compared to numerical simulations in section 4.

As we observed in section 2.1, collective turns in real bird flocks generally present a well-localized origin, as if in response to some punctual external stimulus, and only then they propagate throughout the group. This indicates that the perturbing field should be chosen as *local*, i.e. $\mathbf{h}_i \propto \delta_{ip}$, where p labels a specific particle chosen inside the flock. Applying a local field $\mathbf{h}_p$ on the particle labeled by p induces on the mean polarization defined in equation (30) a response (see appendix A)

$$\Phi(t) = \frac{1}{N\eta} \int_0^t dt' h_p(t') \sin \alpha_p(t') \left[1 - e^{-2\gamma(t-t')} \right]. \tag{31}$$

This equation has been derived under the assumption that the interaction network n_{ij} is symmetric (but generalizations are possible, see equation (A26)). Moreover, it should be considered valid within linear response for small $\mathbf{h}$—however, its validity extends to higher orders by choosing $\alpha_p \equiv \pi/2$, as we show in appendix A. This is the choice we will adopt in our numerical simulations in the following sections.

In the following, we will choose a step-like perturbation in the form $h_p(t) = A_0 \Theta(t)$ applied at a fixed angle $\alpha_p(t) \equiv \alpha_p$, for which equation (31) reduces to

$$\Phi(t) = \frac{A_0 \sin \alpha_p}{\eta N} \left[t - \frac{1 - e^{-2\gamma t}}{2\gamma} \right]. \tag{32}$$

Equation (32) tells us that, when the size N of the system becomes very large (i.e. in the thermodynamic limit), the system will never change the direction of the global order parameter on observational time scales. This is what happens in a ferromagnet at equilibrium: the order is stable at low temperature, and a local field applied on one site cannot change the total magnetization. However, biological groups have much smaller sizes than a physical condensed matter system. In this case, times $t \sim \mathcal{O}(N)$ can in fact be small enough to be reachable in experiments, and the system can thus change its global direction on observational time scales.

We also note that, according to equation (32), the group changes its direction independently of whether the propagation of information is inertial or dissipative. This last feature is however specific to the fixed-network condition: in this case the perturbed individual does not move, and it is always interacting with the rest of group. Even though information arrives damped as it travels through the system, it is injected continuously at the perturbed site and sooner or later everyone will turn. As we will discuss later on, this is not what happens for finite groups of moving individuals, when—if information does not propagate unaltered and quickly enough—the perturbed individual will leave the group before everyone else can follow (see section 6).

The shape of the turn is described by equation (32): for times $t \ll \gamma^{-1}$ the polarization angle $\Phi(t)$ grows with a quadratic dependence, while at larger times a linear behavior is predicted. However, we remind that equation (32) is valid only for small phases φ_i, and it describes the polarization angle at most up to the time at which the turn is complete (i.e. the angles clearly do not increase indefinitely, but a saturation occurs, which is not captured by our equation). Therefore, it is convenient to introduce the timescale of the collective turn τ_{turn}, defined as the time at which $\Phi(t)$ becomes of $\mathcal{O}(1)$, and which represents the observational time window of interest for our analysis. According to the value of τ_{turn} as compared to the typical timescales of the model, one might actually observe either a linear or a quadratic behavior before the saturation:

- *Linear behavior.* If $\tau_{\text{turn}} \gg \gamma^{-1}$, then we can disregard the exponential decay in equation (32) to find

$$\Phi(t) \simeq \frac{A_0 \sin \alpha_p}{\eta N} t, \tag{33}$$

$$\tau_{\text{turn}} \simeq \frac{\eta N}{A_0 \sin \alpha_p}. \tag{34}$$

This condition is met for $\eta^2 N/(2\chi A_0 \sin \alpha_p) \gg 1$, and it is typically what happens when dissipation is large with respect to inertia.

- *Quadratic behavior.* The opposite regime is found when, on the contrary, $\tau_{\text{turn}} \ll \gamma^{-1}$. In this case, we can expand the exponential in equation (32) for small arguments and we find

$$\Phi(t) \simeq \frac{A_0 \sin \alpha_p}{2\chi N} t^2, \tag{35}$$

$$\tau_{\text{turn}} \simeq \left(\frac{\chi N}{A_0 \sin \alpha_p} \right)^{1/2}. \tag{36}$$

This conditions is met for $\eta^2 N/(2\chi A_0 \sin \alpha_p) \ll 1$, and it is typically what happens for very underdamped systems, in which inertia dominates over dissipation. In this regime, however, when the angle α_p is different from $\pi/2$, the expression for $\Phi(t)$ is slightly more complex (see appendix A.1).

Finally, let us note that the conditions for the turn to appear linear or quadratic in the phase growth $\Phi(t)$ (which are stated above, and which involve the field amplitude A_0) are not exactly the same as the conditions for underdamped/overdamped propagation reported in section 2.4 (which involve instead the interaction strength J). One can thus envision realizations of the system in which propagation is underdamped, but the turn still appears linear. As

we mentioned before, the linear/quadratic behaviors represent the two extreme limits of equation (32), which describes the polarization response in general; we distinguished these two behaviors mainly for practical reasons, since they are useful to check numerically the theoretical predictions, as we will do in the next sections.

3.3. Other relevant timescales

The behavior of the polarization angle described by equation (32) is the immediate consequence of a very general feature of the potential $\mathcal{U}(\{\mathbf{v}_i\})$, namely the rotational invariance of the velocity–velocity interactions. In the low-noise phase, the system polarizes in a well-defined direction, thus breaking the symmetry. As in standard O(n) models, the system remains however subject to soft Goldstone modes, i.e. low-energy excitations, in the subspace perpendicular to the polarization vector. In other terms, due to the presence of noise, the vectorial polarization freely fluctuates like a random walk within this zero-mode subspace (see appendix B). When an external field is applied, it provides a bias to this random walk, which results in the (almost) linear time dependence exhibited by equation (32).

However, the spontaneous fluctuations provide another reference time scale. In fact, even in the absence of a field, a finite system subject to fluctuations will eventually depart from its original direction if we wait for long enough. In appendix B we compute explicitly this *wandering* time τ_{w}, which grows with the system size N and is regulated by the amplitude of the noise T. For times $t > \tau_{\mathrm{w}}$, fluctuations have generally already changed (randomly) the direction of the polarization, and it becomes meaningless to speak of perturbation-response events. A condition to be satisfied in our analysis is therefore $\tau_{\mathrm{turn}} \ll \tau_{\mathrm{w}}$. Since the polarization response depends on the field amplitude, for this to happen we roughly need $A_0 \gg T$ (see equations (34) and (36)), a requirement which is easily achieved for very ordered systems. More accurate estimates of this condition are provided in appendix B.

If the system lives in two dimensions (i.e. $d_{\mathrm{v}} = 2 \ \& \ d = 2$), when the network is kept fixed the model is analogous to the XY model. This is not the case for flocks (which perform planar turns but are really three-dimensional systems), but we will adopt this simplification in the numerical simulations discussed in the next sections. It is known that the XY model does not exhibit long-range order in the thermodynamic limit. However, for finite-sized systems the order parameter (the magnetization, equivalent to the polarization defined here) remains finite in the low-temperature phase, and it slowly decreases for $N \to \infty$. In particular, for sizes comparable to those considered here, the system is fully ordered at low temperature (see appendix C).

4. Numerical results—fixed network

In the previous section we derived analytically, under some suitable assumptions, the response of the system to a local directional perturbation, i.e. the time dependence of the polarization angle $\Phi(t)$ in equation (31). In this section we numerically test the validity of such predictions.

We start by addressing the problem in the fixed-network case, where we can build up and validate a suitable perturbation protocol, which will be carried over to the off-lattice case in section 5. For simplicity, we consider as fixed network a regular two-dimensional lattice.

We implemented a time-discretized version of the ISM equations (25)–(27) by generalizing the numerical integration scheme proposed in [24], as reported in appendix D. We adopted a step-like time dependence for the perturbation field $h_p(t)$ (see equation (31)), in the form

$$h_p(t) = \frac{A_0}{2}\left[1 + \tanh\left(\frac{t - t_0}{\tau_{\mathrm{step}}}\right)\right], \qquad (37)$$

where A_0 controls the amplitude of the field, τ_{step} controls the sharpness of the step, and t_0 is a time offset. As we described in section 3, in our numerical simulations we apply the perturbing field on a single individual labelled by p in order to mimic the spatially localized origin of turns in real flocks. We set for simplicity the angle between the individual and the field to be $\alpha_p = \pi/2$, so that the predictions in equations (33) and (35) further simplify since $\sin \alpha_p = 1$. We adopt for the moment periodic boundary conditions (PBCs), to pinpoint the role of the dynamics. In section 6 we will investigate the effect of OBCs, which turns out to be crucial when the system is off-lattice.

The choice of the time scales t_0 and τ_{step} in equation (37) plays no crucial role, as long as the field change remains sharp: we set $\tau_{\mathrm{step}} = 0.1$ to approximate a step function, and $t_0 = 10$ so that the system starts turning after a short transient of $\mathcal{O}(t_0)$ (but other choices would return similar results provided that τ_{step} is chosen not too large). Conversely, the range of values over which A_0 can be chosen requires some discussion. First, equations (33) and (35) tell us that changing the amplitude A_0 of the perturbing field offers a way to tune the velocity of the turn: we thus avoid choosing too large values of A_0, as they would induce a turn which is too fast to be easily studied. Besides, large values of A_0 might push us away from the linear response regime that we wish to explore. Conversely, a lower bound on A_0 is actually imposed by the physics of the system. Indeed, due to its finite size, the polarization angle $\Phi(t)$ is subject both to the action of the perturbing

field, and to the wandering effects described in the previous section (which are non-negligible even if we are working with a highly polarized state). The value of A_0 should then be chosen sufficiently large so as to avoid that—on our time scales of observation, and for our limited number of trials—the effect of the perturbing field gets completely masked by wandering effects. This translates to requiring that the polarization angle $\Phi(t)$ is of $\mathcal{O}(1)$ for times $t \lesssim \tau_w$ in all the dynamical regimes described above. With our set of parameters, this condition is met for $A_0 \sim \mathcal{O}(1)$, hence we simulate our turns in the range $A_0 \in [5, 30]$.

4.1. Collective turns

Here we describe our first simulations of collective turns. Hereafter, the parameters of the model are chosen so that the system is highly polarized ($\Psi > 0.98$, as in real starling flocks [9]). This is useful for two reasons: first, it allows to obtain polarization curves that are not excessively noisy; second, by setting the flock in its deeply ordered state we are enforcing the conditions for the SWA, under which our predictions were derived (see section 3.2).

Let us introduce the following numerical protocol:

(i) The system of size N is initialized on a square lattice of side $L = \sqrt{N}$ in a disordered configuration. We wait for the system to thermalize and then we measure the direction of the order parameter, which is chosen as the reference value, i.e. $\Phi(t = 0) \equiv 0$;

(ii) We select a value for the field amplitude A_0;

(iii) We apply the external field on a single individual p inside the system, and we keep the field switched on for a time T_{turn};

(iv) We switch off the field and let the system thermalize for a time T_{term};

(v) To obtain more robust observations, we repeat steps (iii) and (iv) for n_{turns} times, and we average over the various realizations of the turning event.

We can first use our numerical simulation to check the linear and the quadratic behavior predicted for the polarization curves $\Phi(t)$, see equations (33) and (35). We keep the parameters J, T and η fixed, and we use the inertia χ as the tuning parameter to explore the different regimes of τ_{turn} vs. γ^{-1}. Let us start with the case $\tau_{\text{turn}} > \gamma^{-1}$, where our analytical derivation predicts a linear behavior of the polarization angle $\Phi(t)$—see equation (33). A typical mean polarization curve obtained in this regime is shown in figure 1(a) for two values of χ, together with a linear fit of the initial part of the curve (see inset). Recall that, due to the SWA, the agreement with the analytical prediction is expected to break down as the polarization angle saturates to its final value $\alpha_p = \pi/2$. To
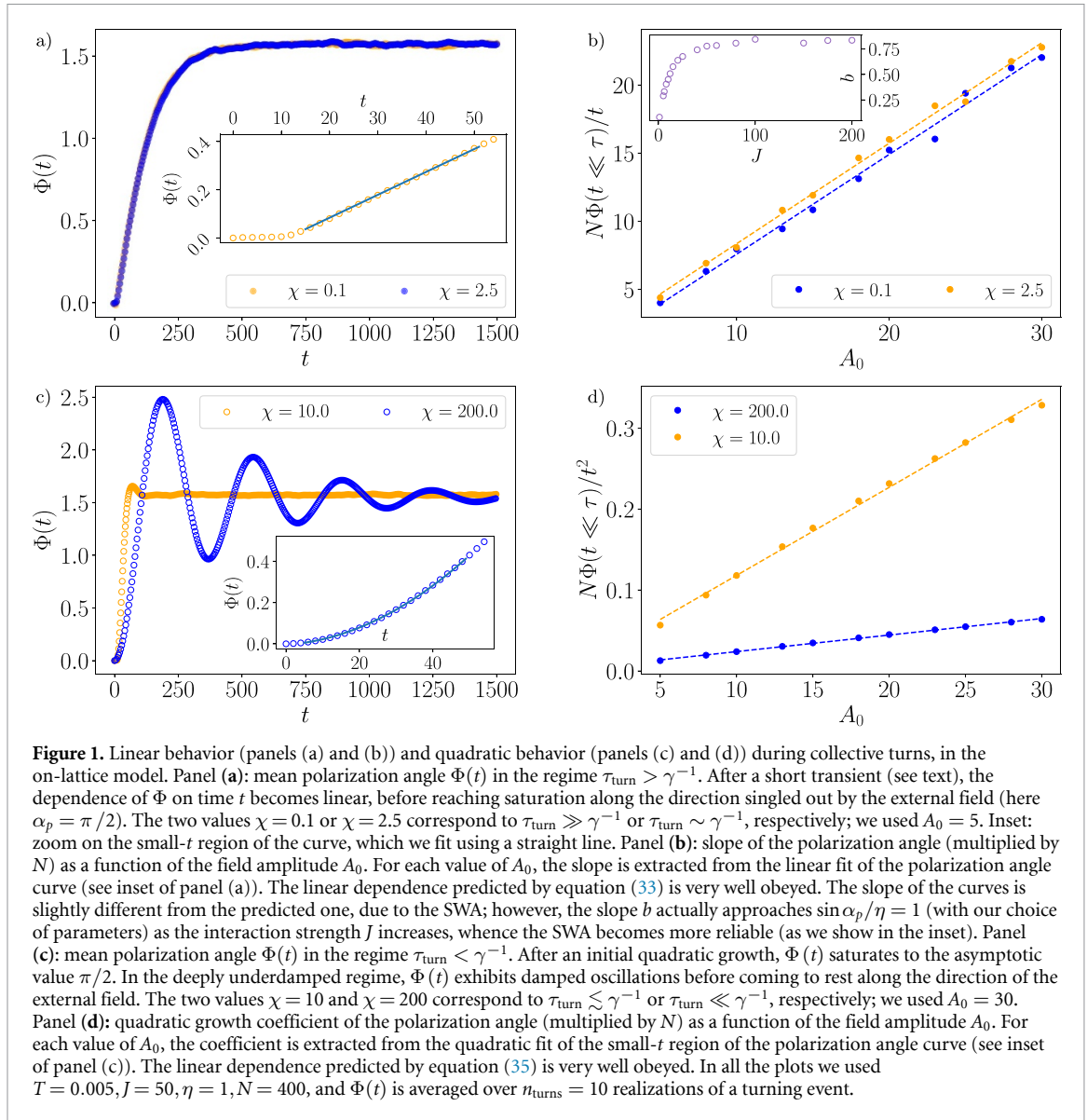
test the agreement of the prefactor in equation (33) with our numerical observations, we have repeated steps (i) to (iv) for several values of A_0 in the range $[5, 30]$, and for each value of A_0 we have fitted the initial part of the mean curve $\Phi(t)$ with a linear function. Figure 1(b) shows our results for the same values of χ as in figure 1(a), which confirm the prediction of a linear dependence of the slope on A_0.

Keeping the values of the parameters J, η and T fixed, while pushing the inertia χ towards higher values, brings the system more into the underdamped region where $\tau_{\text{turn}} < \gamma^{-1}$. Here the predicted time dependence for $\Phi(t)$ is quadratic, as described by equation (35). We note that an initial quadratic behavior is actually expected to occur in any regime, since there is always a time window (however short) in which the condition $\gamma t \ll 1$ is met—see section 3.2. However, as discussed in the previous section, the quadratic behavior holds for the whole duration of the turn only when $\tau_{\text{turn}} \ll \gamma^{-1}$. This occurs if $\eta^2 N/(2\chi A_0 \sin \alpha_p) \ll 1$, which translates to $\chi \gg 5$ for the parameters used in figure 1(c), where we show two polarization curves $\Phi(t)$ for two distinct values of χ. We then repeat the same analysis as for the linear case: for each value of A_0 we fit the initial part of the polarization curve $\Phi(t)$ with a parabola (see inset in figure 1(c)), and check that the estimated prefactor grows linearly with A_0 (see equation (36)). This is precisely what happens, as displayed in figure 1(d).

Finally, a curious feature of the deeply underdamped regime is the occurrence of oscillations in the saturation region of the polarization angle (see the blue curve in figure 1(c)). This is intuitive once we recall that the parameter χ plays the role of a moment of inertia in equation (14). A quantitative estimate (see appendix E) predicts oscillations with frequency $\Omega = \sqrt{A_0/(\chi N)}$ and damping factor γ, in line with numerical observations.

4.2. Propagation law

Our numerical protocol additionally allows us to investigate and check the dispersion law predicted in section 2.4. The behavior of the polarization angle $\Phi(t)$ discussed above describes how the group as a whole rearranges its direction to align with the external field. It does not describe, however, how the perturbation—which is applied locally to a specific individual—is propagated from individual to individual through the system. To explore this issue, we need to consider the phases $\varphi_i(t)$ of the single individuals, and monitor how they change in time. In figures 2(a) and (b) we thus show these individual curves as a function of time, for two different sets of the model's parameters. The perturbed individual (blue curve) is the first to change its direction and to complete the turn. Other individuals display a very similar turning profile, but shifted in time with a certain delay Δt_i, which is larger the farther the

Figure 1. Linear behavior (panels (a) and (b)) and quadratic behavior (panels (c) and (d)) during collective turns, in the on-lattice model. Panel (**a**): mean polarization angle $\Phi(t)$ in the regime $\tau_{\mathrm{turn}} > \gamma^{-1}$. After a short transient (see text), the dependence of Φ on time t becomes linear, before reaching saturation along the direction singled out by the external field (here $\alpha_p = \pi/2$). The two values $\chi = 0.1$ or $\chi = 2.5$ correspond to $\tau_{\mathrm{turn}} \gg \gamma^{-1}$ or $\tau_{\mathrm{turn}} \sim \gamma^{-1}$, respectively; we used $A_0 = 5$. Inset: zoom on the small-t region of the curve, which we fit using a straight line. Panel (**b**): slope of the polarization angle (multiplied by N) as a function of the field amplitude A_0. For each value of A_0, the slope is extracted from the linear fit of the polarization angle curve (see inset of panel (a)). The linear dependence predicted by equation (33) is very well obeyed. The slope of the curves is slightly different from the predicted one, due to the SWA; however, the slope b actually approaches $\sin\alpha_p/\eta = 1$ (with our choice of parameters) as the interaction strength J increases, whence the SWA becomes more reliable (as we show in the inset). Panel (**c**): mean polarization angle $\Phi(t)$ in the regime $\tau_{\mathrm{turn}} < \gamma^{-1}$. After an initial quadratic growth, $\Phi(t)$ saturates to the asymptotic value $\pi/2$. In the deeply underdamped regime, $\Phi(t)$ exhibits damped oscillations before coming to rest along the direction of the external field. The two values $\chi = 10$ and $\chi = 200$ correspond to $\tau_{\mathrm{turn}} \lesssim \gamma^{-1}$ or $\tau_{\mathrm{turn}} \ll \gamma^{-1}$, respectively; we used $A_0 = 30$. Panel (**d**): quadratic growth coefficient of the polarization angle (multiplied by N) as a function of the field amplitude A_0. For each value of A_0, the coefficient is extracted from the quadratic fit of the small-t region of the polarization angle curve (see inset of panel (c)). The linear dependence predicted by equation (35) is very well obeyed. In all the plots we used $T = 0.005, J = 50, \eta = 1, N = 400$, and $\Phi(t)$ is averaged over $n_{\mathrm{turns}} = 10$ realizations of a turning event.
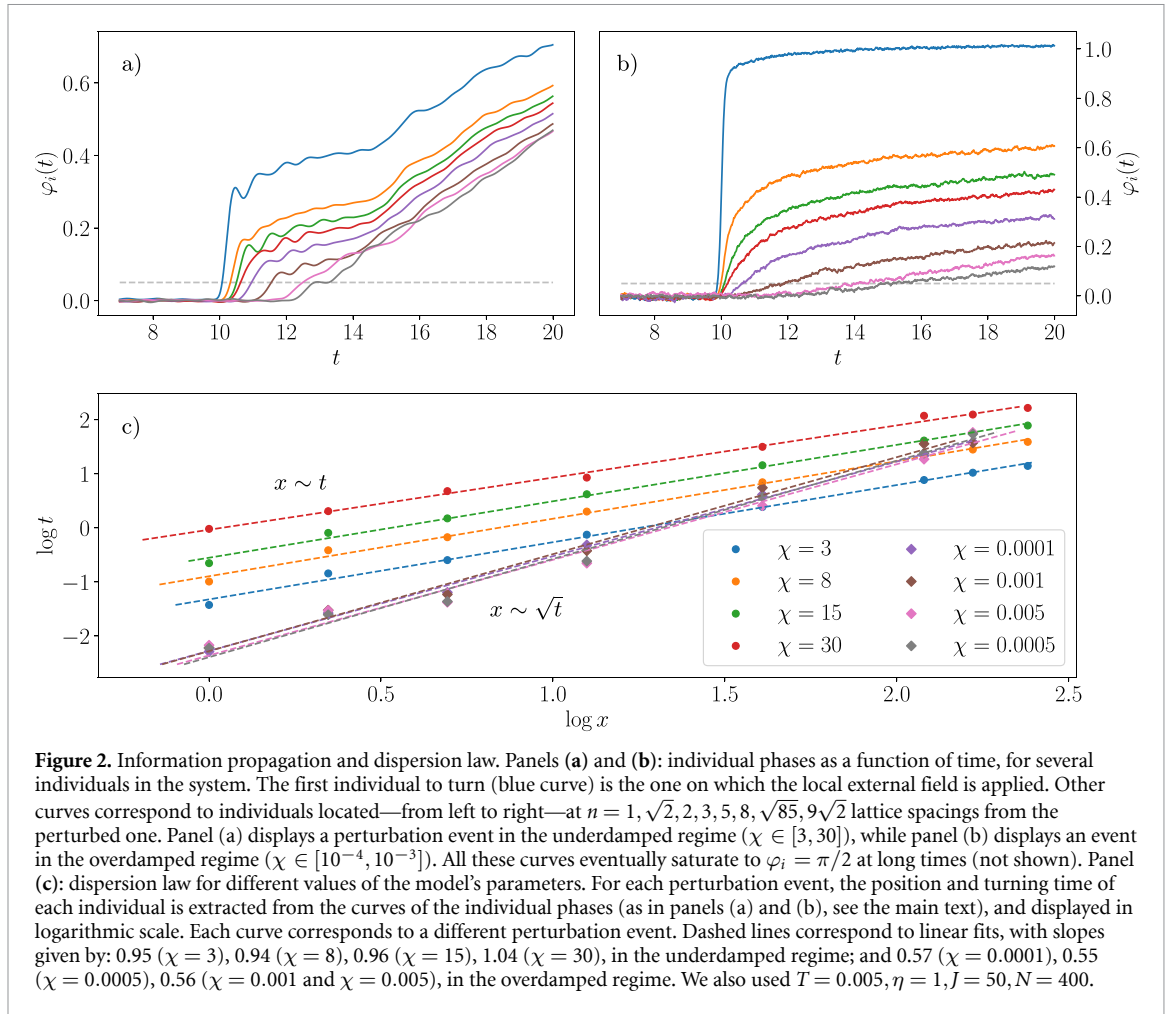
individual is from the location of the perturbation (colored curves from left to right). This behavior indicates that the directional information spreads progressively through the system. Starting from these curves, we can quantify the dispersion law as follows. Let us identify the turning time of a given individual as the time when its direction reaches a threshold angle $\varphi = 0.05$ (dashed horizontal line in figures 2(a) and (b)). The set of individuals which start turning at a given time Δt_i determine the turning front at that time, and their distance x_i from the perturbed individual identifies the location of the front, which travels obeying the dispersion law $x_i(\Delta t_i)$. We plot the latter in figure 2(c) for several simulations performed with different sets of parameters. The figure confirms the picture outlined in section 2.4, where two propagation regimes were identified according to whether the time for information to travel through the entire system, $\tau_s = L/c_s$, is smaller or larger than the time for the attenuation to damp the signal, $\tau =$ γ^{-1}. When $\tau_s \ll \tau$ a linear dispersion law ($x \sim t$) is observed; conversely, when $\tau_s \gg \tau$ we find a dissipative behavior ($x \sim \sqrt{t}$). With the parameters used in figure 2, one has $\tau/\tau_s = 1$ for $\chi = 0.5$.

5. Numerical results off-lattice

In this section we address the question of how a local perturbation affects the global directional order in the full off-lattice model. In particular, we want to test whether and to which extent the analytical predictions derived in section 3.2 are valid when the activity of the model is taken into account. Indeed, the presence of a nonzero activity v_0 introduces a new time scale in the model: its interplay with the previous timescales is expected to foster a rich phenomenology, as it happens already in Vicsek-like models [25]. We will start by studying the unperturbed model, and then move on to analyze the effects of a local perturbation.

Figure 2. Information propagation and dispersion law. Panels (**a**) and (**b**): individual phases as a function of time, for several individuals in the system. The first individual to turn (blue curve) is the one on which the local external field is applied. Other curves correspond to individuals located—from left to right—at $n = 1, \sqrt{2}, 2, 3, 5, 8, \sqrt{85}, 9\sqrt{2}$ lattice spacings from the perturbed one. Panel (a) displays a perturbation event in the underdamped regime ($\chi \in [3, 30]$), while panel (b) displays an event in the overdamped regime ($\chi \in [10^{-4}, 10^{-3}]$). All these curves eventually saturate to $\varphi_i = \pi/2$ at long times (not shown). Panel (**c**): dispersion law for different values of the model's parameters. For each perturbation event, the position and turning time of each individual is extracted from the curves of the individual phases (as in panels (a) and (b), see the main text), and displayed in logarithmic scale. Each curve corresponds to a different perturbation event. Dashed lines correspond to linear fits, with slopes given by: 0.95 ($\chi = 3$), 0.94 ($\chi = 8$), 0.96 ($\chi = 15$), 1.04 ($\chi = 30$), in the underdamped regime; and 0.57 ($\chi = 0.0001$), 0.55 ($\chi = 0.0005$), 0.56 ($\chi = 0.001$ and $\chi = 0.005$), in the overdamped regime. We also used $T = 0.005, \eta = 1, J = 50, N = 400$.

5.1. Unperturbed model

The numerical implementation of the off-lattice model is similar to that of the on-lattice case, but this time the individual positions $\mathbf{r}_i$ evolve ballistically—according to the corresponding velocity vectors $\mathbf{v}_i$—in a squared box with periodic boundary conditions. The interaction network $n_{ij}(t)$ must now be updated regularly: this is obtained by implementing the *cell-list* method (see [26–28] and appendix D). We remind that we chose a metric interaction rule, i.e. n_{ij} is different from zero if individuals i and j have a mutual distance lower than the interaction range r_c.

Studying the phase diagram of the model is made nontrivial by the presence of spatial aggregation effects. Even Vicsek-like models (i.e. first-order models) are known to exhibit strong spatial heterogeneities, for sufficiently low noise, in the case of *additive* interactions [25]. By this we mean an interaction term (in the evolution equation for the individual velocity vector) of the form

$$\sim \sum_{j \in \mathcal{N}_i} \sin(\varphi_j - \varphi_i), \tag{38}$$

which is not normalized by the total number $\mathcal{N}_i$ of particles interacting with the ith (as it happens instead

in non-additive models). As a result, the interaction becomes stronger and stronger as new particles enter the interaction range of the ith, which might cause the formation of clusters. Since the equations of motion of our model feature an interaction term of this form (see section 3.2), we expect a similar mechanism to take place.

Clustering phenomena were indeed observed in our numerical simulations. We considered an initially uniform distribution of individuals; when a density fluctuation brings a few individuals close together, the interaction force increases and eventually dominates over the fluctuations, which are then unable to disrupt the cluster. A detailed analysis of the clustering formation goes beyond the scopes of this work, but one generally finds that these aggregation effects are more prominent at high activity, low temperatures (in the ordered phase of the system), and high values of the inertia (i.e. towards the underdamped regime). On general grounds, we expect the timescale of the clustering process to be mostly influenced by the inertia χ, which slows down the effective motility of the system.

The off-equilibrium phase diagram in figure 3(a) shows the scalar polarization Ψ of the model as a function of the temperature T, for various speeds
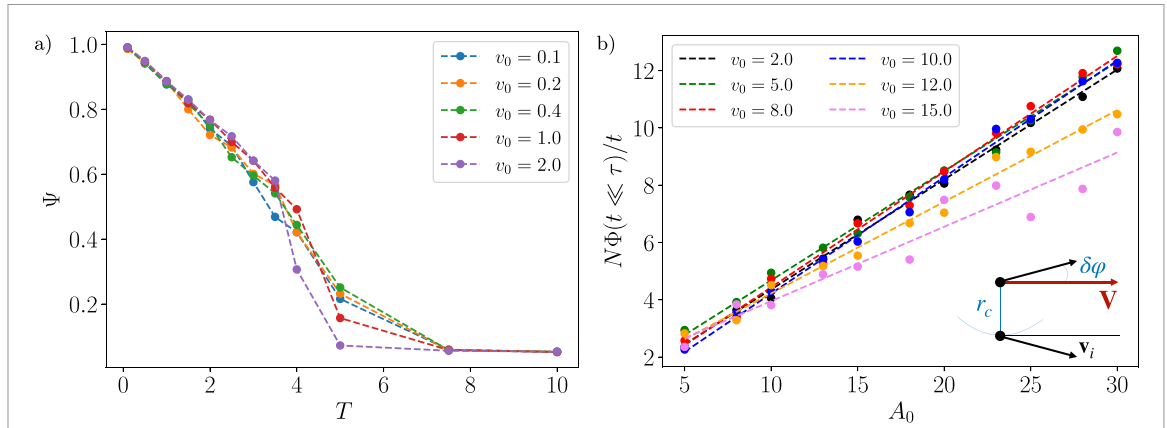
Figure 3. Simulations of the off-lattice model with periodic boundary conditions. Panel (**a**): phase diagram of the model. The scalar polarization Ψ is measured as a function of the temperature T for several values of the activity v_0, in the overdamped regime. The parameters used in the plot are $\chi = 1.25, \eta = 5, J = 0.8, N = 400$. Panel (**b**): effect of the perturbation. As we did in figure 1 for the on-lattice case, we check the linear dependence of the polarization angle $\Phi(t)$ (predicted in the on-lattice case by equation (33)), averaging over $n_{\text{turns}} = 10$ realizations for each value of the activity v_0. The higher the activity, the more the fitted linear prefactor deviates from the prediction of equation (33), until for $v_0 = 15$ we are totally out the validity of the on-lattice derivation, because the network reshuffling is so fast that the fixed-network approximation no longer holds. We used the parameters $\chi = 5, J = 1, T = 0.01, N = 400, \eta = 1$, with scalar polarization $\Psi \sim 0.98$. Inset: sketch showing two interacting individuals moving away from each other. Eventually their distance becomes larger than the interaction radius r_c. Here **V** indicates the common flight direction of the flock, while $\delta\varphi$ is the deviation of a single individual due to thermal fluctuations.

v_0. In constructing the phase diagram, we selected a range of dynamical parameters (η, χ) for which the most severe aggregation effects are absent. As we noted, the density $\rho = N/L^2$ of the system and the interaction radius r_c heavily affect clustering phenomena and, more in general, the dynamics of the model. Hereafter we work with $\rho = 1$ and $r_c = 1.5$, which ensure (within non-clustered, homogeneous systems) an average number of interacting neighbors $n_c \sim \pi r_c^2 \rho \sim 7$. This value reproduces the one measured in real starling flocks, where each bird is found to coordinate with its nearest 7–8 individuals [29].

Note that, while numerical simulations of the 3d ISM have been previously reported in the literature [10, 30, 31], the ones presented in this work are the first numerical simulations of the 2d ISM.

5.2. Perturbed model: role of the dynamics

After analyzing the unperturbed case, we now address the role of the activity in the presence of a local perturbation. We noted in section 4 that a turn *always* takes place in a finite-size, on-lattice system in response to a step-like local perturbation. However, once we bring in a nonzero activity, we are no longer guaranteed that a local perturbation will produce a similar behavior on the entire system. Indeed, two timescales are now expected to compete: that of information propagation (related to the speed c_s, see equation (20)), and the reshuffling time of the network $n_{ij}(t)$, which is the time taken by an individual to change its neighborhood. If the latter process is too fast, then a given individual may leave its interaction neighborhood before the *signal* is able to propagate. We thus generically expect a threshold activity v_0^{lim} such that, for $v_0 < v_0^{\text{lim}}$, the off-lattice system behaves similarly to the on-lattice case, and our analytical

predictions for the time dependence of the polarization angle $\Phi(t)$ still apply.

We can estimate v_0^{lim} by using a simple heuristic argument. Consider two distinct interacting individuals, i.e. at a relative distance smaller than the interaction range r_c. Let $\delta\varphi$ denote the mean angular deviation of their flight direction with respect to that of the flock **V**, as in the inset of figure 3(b). In the fully polarized limit $\Psi = 1$ one has $\delta\varphi = 0$, meaning that all the birds fly straight and never cross; conversely, in the presence of fluctuations $\delta\varphi$, one can derive under the SWA the relation (see appendix B.2)

$$\delta\varphi \equiv \sqrt{\langle\varphi^2\rangle} \sim \sqrt{2(1 - \Psi)}. \tag{39}$$

In the worst-case scenario, the velocity vectors of the two individuals point outwards as in the inset of figure 3(b), so that they drift away from each other and stop interacting after a time t_r defined as

$$2v_0\,\delta\varphi\,t_r \sim r_c. \tag{40}$$

The timescale $t_r \sim r_c/(2v_0\,\delta\varphi)$ thus estimates the *reshuffling* time of the connectivity matrix $n_{ij}(t)$. When an external field acts on a single individual, the information must reach its nearest neighbors before reshuffling happens, in order for a collective turn to take place in the way it occurs on a fixed network. Such information propagates across the interparticle distance a within a time $t_a \sim a/c_s$, where c_s is the information propagation speed introduced in equation (20). We conclude that information propagates similarly to the on-lattice case if $t_a \ll t_r$, so that the network can be seen as quasi-fixed; in turn, this implies

$$v_0 \ll \frac{r_c}{2a}\frac{c_s}{\delta\varphi} = \frac{r_c}{2}\sqrt{\frac{Jn_c}{2\chi(1 - \Psi)}} \equiv v_0^{\text{lim}}. \tag{41}$$

5.3. Numerical results

Using the same protocol that we applied in the on-lattice case (see section 4), we now analyze the effects of a local perturbation. Again, the parameter χ can be tuned in order to explore the dynamical regimes of the model. The choice of the other parameters is instead dictated by two new requirements. First, we should make sure that the system is not in a clustered state. We then check beforehand that the time scale of clustering is much longer than the turning time τ_{turn} of the flock; this requires us to tune the temperature T in order to remain within a well-polarized state, where the system is spatially uniform. Secondly, numerical limitations (such as time discretization) would prevent us from exploring the region with $v_0 > v_0^{\text{lim}}$, if the latter were excessively high. All in all, this leads to the choice of parameters of figure 3(b), which corresponds to $\Psi \sim 0.98$ and $v_0^{\text{lim}} \sim 10$ according to equation (41).

We focus here as an example on the overdamped region $\tau_{\text{turn}} \gg \gamma^{-1}$, but similar results were observed also in the underdamped case. In this regime we know that the polarization angle $\Phi(t)$ should behave linearly (see equation (33)): in figure 3(b) we thus fit its initial part with a straight line, and we plot the so-obtained slope against A_0 for several values of v_0. Most of these lines overlap, which is expected since equation (33) is v_0-independent; conversely, the behavior starts to deviate from a straight line as we approach the region with $v_0 \sim v_0^{\text{lim}}$. This marks the breakdown of the on-lattice approximation.

We finally note that, in the entirety of this section, the activity v_0 of the system does not seem to affect the ability of the local field to trigger a collective reorientation, at least for speeds not much larger than v_0^{lim} (we explored values up to $v_0 = 30$); indeed, the mean polarization angle always saturates to $\Phi(\infty) = \pi/2$, even in the overdamped regime where propagation of information is attenuated by damping. The reason why this happens is crucially related to the presence of periodic boundary conditions: even if the individual feeling the perturbation moves, it never leaves the system, and the perturbation therefore keeps acting on the latter indefinitely. This is not the case, on the contrary, when OBCs are considered. In the next section we will investigate this new setting, where—as we will discuss—the impact of the activity and the regime of information propagation are crucial in determining the occurrence of a turn.

6. Open boundary conditions

In the previous section we investigated the model off-lattice but with periodic boundary conditions. In this case, as we discussed, the local perturbation is able to change the global direction of motion, eve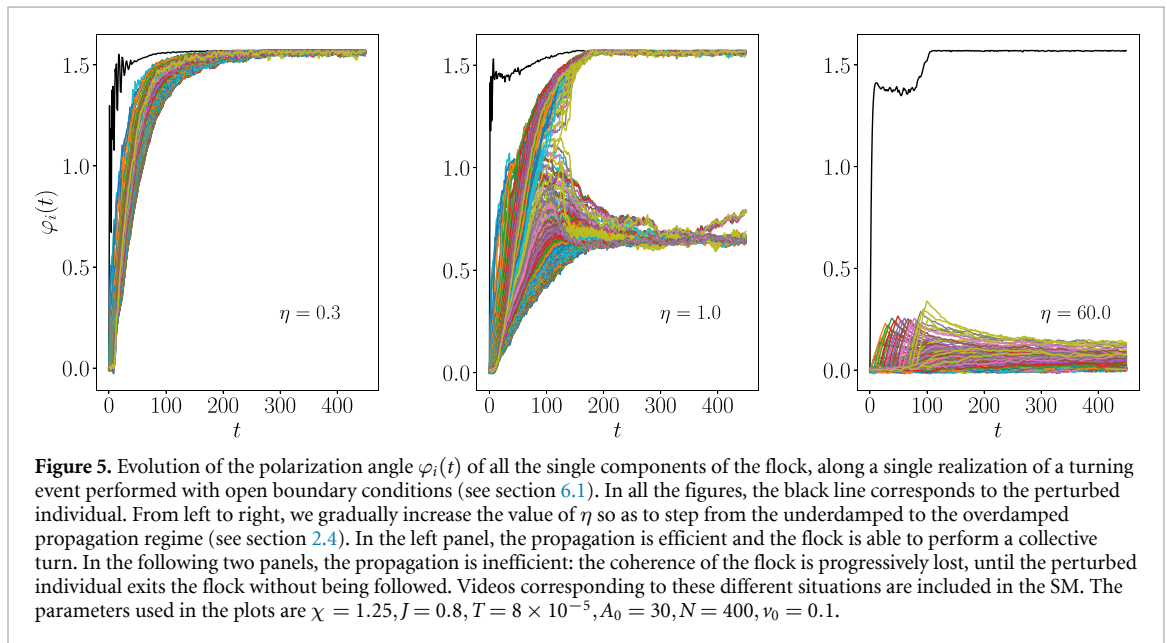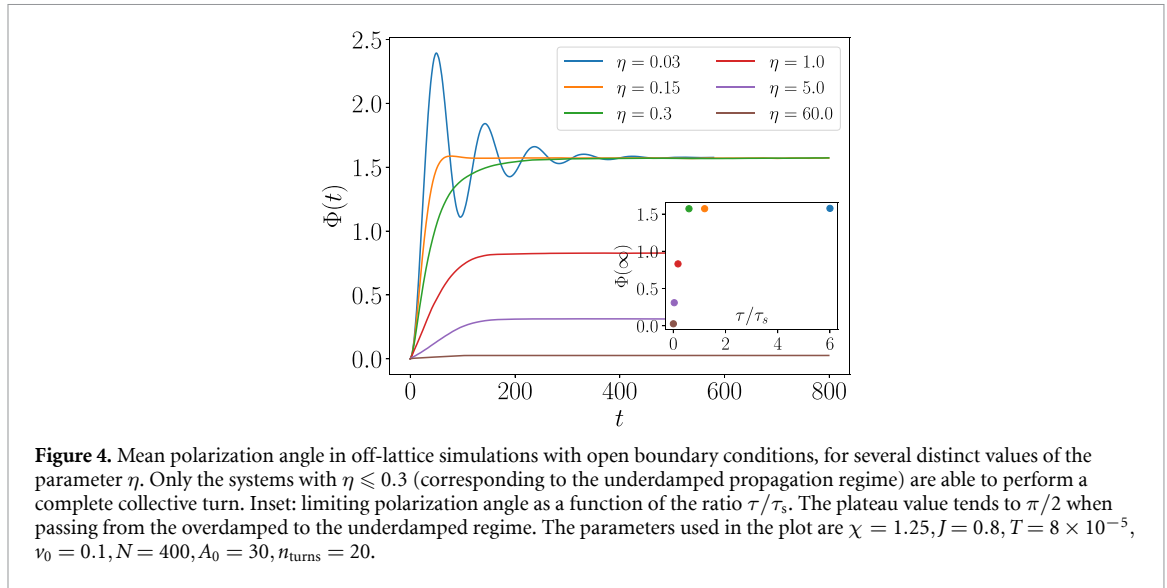n in the presence of strong dissipation. This is apparently at variance with experimental observations, where coherent turning is associated with underdamped inertial propagation. The reason for this mismatch is however related to the artificial nature of a box with periodic boundary conditions. Indeed, PBCs can mask the role of dissipation: whenever the perturbed individual leaves the flock, it enters again from the opposite side of the box, where it can start spreading the signal again. Simulating the system with OBCs, as we do in this section, is expected to eliminate this spurious effect. This new setting mimics more closely the situation of real flocks and other groups, and will unravel the crucial role played by the underdamped propagation regime in allowing the flock to sustain a collective turn.

6.1. Role of the dynamical propagation regimes

To implement a turn with OBCs, we first evolve the system off-lattice for a short time interval, until the scalar polarization reaches the stationary value $\Psi \sim 0.99$, and then we apply the field to an individual placed in the middle of the group. We then proceed, as in the previous sections, to measure the evolution of the mean polarization angle. Since (now) the group has finite boundaries, more stringent conditions must be met in order for a global turn to take place. Indeed, the perturbed individual will quickly follow the field direction and deviate from the group's mean velocity: if all the other individuals do not follow (and turn by the *same* angle) soon enough, a finite difference in orientation with respect to the turn initiator will persist when the latter reaches the boundary. This can result either in the flock remaining compact but turning by a smaller amount, or else in the group's fragmentation. For a full turn to occur we therefore need i) the information about the angular deviation to propagate intact to all individuals (i.e. underdamped inertial propagation), and ii) the turn to be complete before the initiator hits the boundary (i.e. not too large motility v_0).

To verify this picture quantitatively, let us start by considering the first condition. We choose a small value of v_0—for which we know the analytical predictions to hold even off-lattice—and we explore the various dynamical regimes of the model by appropriately varying the other parameters. Since in the off-lattice model large values of χ enhance clustering effects, we keep the inertia fixed and tune instead the dissipation η to span the underdamped/overdamped spectrum.
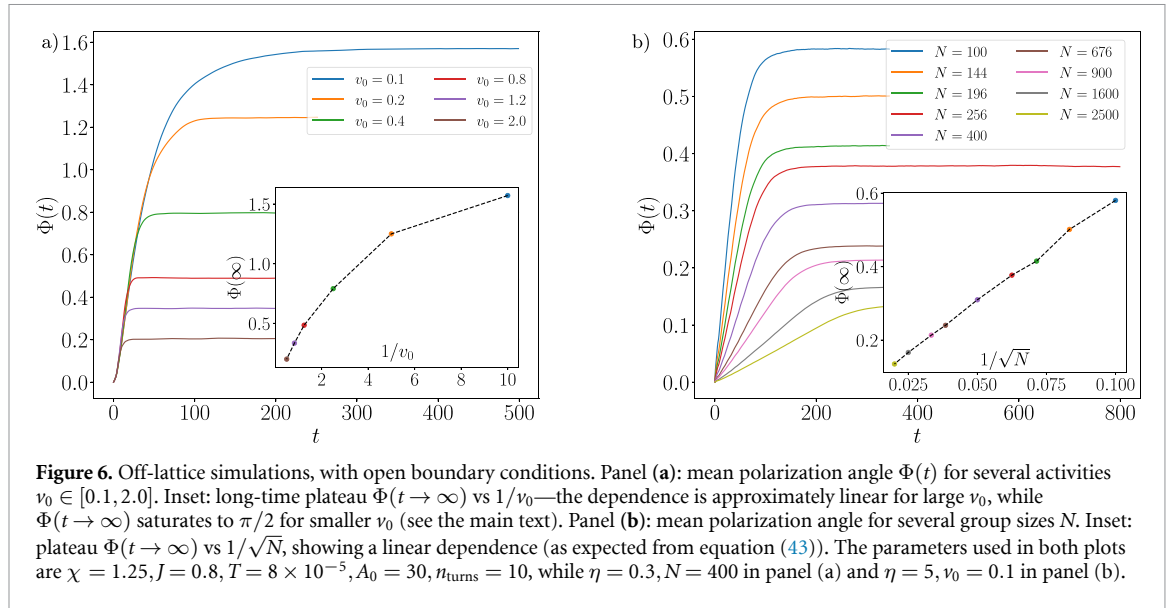
The resulting curves for the mean polarization angle $\Phi(t)$ are reported in figure 4. This figure shows that only for sufficiently low η (i.e. in the underdamped inertial regime) the flock is able to sustain a collective turn and follow the perturbation. At higher values of η (i.e. in the overdamped dissipative regime), the dissipation is stronger and the perturbed individual leaves the flock before the whole group can

Figure 4. Mean polarization angle in off-lattice simulations with open boundary conditions, for several distinct values of the parameter η. Only the systems with $\eta \leqslant 0.3$ (corresponding to the underdamped propagation regime) are able to perform a complete collective turn. Inset: limiting polarization angle as a function of the ratio τ/τ_s. The plateau value tends to $\pi/2$ when passing from the overdamped to the underdamped regime. The parameters used in the plot are $\chi = 1.25, J = 0.8, T = 8 \times 10^{-5}$, $v_0 = 0.1, N = 400, A_0 = 30, n_{\mathrm{turns}} = 20$.

Figure 5. Evolution of the polarization angle $\varphi_i(t)$ of all the single components of the flock, along a single realization of a turning event performed with open boundary conditions (see section 6.1). In all the figures, the black line corresponds to the perturbed individual. From left to right, we gradually increase the value of η so as to step from the underdamped to the overdamped propagation regime (see section 2.4). In the left panel, the propagation is efficient and the flock is able to perform a collective turn. In the following two panels, the propagation is inefficient: the coherence of the flock is progressively lost, until the perturbed individual exits the flock without being followed. Videos corresponding to these different situations are included in the SM. The parameters used in the plots are $\chi = 1.25, J = 0.8, T = 8 \times 10^{-5}, A_0 = 30, N = 400, v_0 = 0.1$.

turn, so that the global flight direction is only partially affected by the perturbation. By definition, the mean polarization angle $\Phi(t)$ saturates to $\alpha_p = \pi/2$ for $t \to \infty$ only if the entire flock remains coherent and aligns with the field, whereas $\Phi(\infty) < \pi/2$ corresponds to a fragmentation of the group. Indeed, if the group splits into two components (one containing the perturbed individual and moving along the field direction, and another one going in another direction), then the resulting $\Phi(\infty)$ will be given by the weighted average over the two flight directions, with the weights represented by the sizes of the two clusters. The fewer individuals follow the perturbed one, the more $\Phi(\infty)$ will differ from $\pi/2$. In the inset of figure 4 we thus plot the limiting polarization angle $\Phi(\infty)$ as a function of τ/τ_s, i.e. the

ratio between the dissipative and inertial timescales. As discussed in section 2.4, this ratio identifies whether the system is in the overdamped ($\tau/\tau_s < 1$) or underdamped ($\tau/\tau_s > 1$) regime of information propagation. This plot therefore shows that only when the system enters the underdamped inertial regime the turn becomes fully efficient and coherent (i.e. $\Phi(\infty) = \pi/2$).

To further illustrate the mechanism described above, we also plot in figure 5 the single polarization angles $\varphi_i(t)$ of all the individuals in the group, for a single realization of the turning event. At small values of η in the underdamped regime, the flock remains coherent and the collective turn can take place. At intermediate values of η, only some individuals follow the perturbed one, and the

Figure 6. Off-lattice simulations, with open boundary conditions. Panel (**a**): mean polarization angle $\Phi(t)$ for several activities $v_0 \in [0.1, 2.0]$. Inset: long-time plateau $\Phi(t \to \infty)$ vs $1/v_0$—the dependence is approximately linear for large v_0, while $\Phi(t \to \infty)$ saturates to $\pi/2$ for smaller v_0 (see the main text). Panel (**b**): mean polarization angle for several group sizes N. Inset: plateau $\Phi(t \to \infty)$ vs $1/\sqrt{N}$, showing a linear dependence (as expected from equation (43)). The parameters used in both plots are $\chi = 1.25, J = 0.8, T = 8 \times 10^{-5}, A_0 = 30, n_{\text{turns}} = 10$, while $\eta = 0.3, N = 400$ in panel (a) and $\eta = 5, v_0 = 0.1$ in panel (b).

group gets disrupted. Finally, at high η, the perturbed individual exits the flock while the flying direction of the rest of the group changes only slightly. Videos showing these turning events in real-time are included in the supplementary material, supporting the scenario we just described.

6.2. Role of the activity and dependence on group size

Let us now discuss what is the role of the activity v_0 in determining the occurrence of a turn. To do so, we consider the system in its underdamped regime, where we know from the previous section that—at small values of v_0—the system is able to fully reorient, following the perturbed individual, in the direction of the applied field. We thus choose in the simulation the same parameters as in figure 4, with $\eta = 0.3$. Then we start increasing the activity of the individuals, to see whether and how much the turning performance is affected.

As discussed before, the ability of the flock to sustain a collective turn is determined not only by the efficiency of information propagation, but also by the time taken by the perturbed individual to exit the group. We can expect that the higher v_0, the harder will be for the flock to turn, since the perturbed individual will leave the system sooner. A simple argument can be used to predict the final direction of flight $\Phi(\infty)$ of the flock depending on its activity. At the beginning of the turn the flock has a large polarization Ψ, so we can imagine the system as if collectively moving straight; when an external step-like field is applied to a bird located in the middle of the flock, the perturbed individual changes its flight direction by $\pi/2$ within a negligible time span. The time taken by the perturbed bird to exit the flock is thus of the order of

$$t_{\text{exit}} \sim L/v_0 \,. \tag{42}$$

In general, we can assume that the system stops turning at $t \sim t_{\text{exit}}$, so that $\Phi(\infty) \sim \Phi(t_{\text{exit}})$ (at least in the absence of oscillations). For the parameters considered in the simulation, the system is in the underdamped propagation regime but has a linear growth (see figure 4). We can adopt the on-lattice prediction in equation (33) to estimate $\Phi(t_{\text{exit}})$, which gives (since $N = L^2$ and $\alpha_p = \pi/2$)

$$\Phi(t = t_{\text{exit}}) \sim \frac{A_0}{N\eta} \frac{L}{v_0} = \frac{A_0}{\eta L v_0} \,. \tag{43}$$

Indeed, our current choice of parameters gives $v_0 \ll c_s \simeq 2.12 \ll v_0^{\text{lim}}$, so that the on-lattice estimate in equation (33) is still reliable according to our discussion in section 5.2. Equation (43) suggests that, when v_0 increases, the final mean polarization angle should decrease. However, this estimate (and in particular the dependence on $1/v_0$) only holds when the exit time t_{exit} falls within the linear growth regime of the polarization angle. This is not the case for sufficiently low activity v_0, because we expect the system to perform a full turn in this limit, and $\Phi(t = t_{\text{exit}})$ must therefore approach $\pi/2$. This picture is confirmed by our numerical simulations. In figure 6(a) we display the mean polarization angle for different values of the activity v_0, and we show that indeed for sufficiently large values of v_0 the group is unable to complete the turn. In the inset, we plot the endpoint $\Phi(t_{\text{exit}})$ as a function of $1/v_0$: as expected, the linear dependence predicted by equation (43) gradually breaks down at small v_0, when $\Phi(t_{\text{exit}})$ approaches $\pi/2$.

It is interesting to compare equation (43) with the response of the polarization $\Phi(t)$ to an impulse-like perturbation of finite duration $\mathcal{T}$. This is derived in appendix F for the *on-lattice* case with a calculation

analogous to the one described in section 3.2, yielding

$$\Phi(t \to \infty) = \frac{A_0 \sin \alpha_p}{N\eta} \mathcal{T}. \qquad (44)$$

The latter clearly agrees with the *off-lattice* prediction in equation (43) upon choosing $\mathcal{T} = t_{\text{exit}}$, i.e. the time during which the perturbed individual remains inside the group (see equation (42)).

Expression (43) also suggests that the final polarization angle, and therefore the efficiency with which the turn is performed, also depends on the size of the group itself. *Ceteris paribus*, large groups tend to be more in the overdamped regime, and they would take more time to perform a turn even on a fixed network. To better pinpoint these effects, we focus on the overdamped regime by using the same parameters as in figure 4 with $\eta = 5$, and then we vary the size $N \in [100, 1600]$. The result is displayed in figure 6(b), where the behavior predicted by equation (43) is very well obeyed, including the linear dependence of the final flight direction on $1/\sqrt{N}$ (see inset).

7. Conclusions

Let us now summarize our results. The main aim of this work was to elucidate the interplay between size, motility and dynamical regimes, in the occurrence of perturbation-response events. More specifically, due to the strong functional relevance of anti-predatory response in finite groups, we focused on the ability of a group to change its global flight direction upon local perturbations. We modelled the perturbation in the form of an external field applied to a single individual of the group (see section 4). Even though the considered scheme is very simplified, it mimics quite reasonably many real instances of perturbations, from attacks of predators, to disturbances and obstacles—which are typically perceived only by a subset of the group, but are eventually transmitted to other individuals thanks to mutual interactions. Experimental data indicate that, in the presence of local triggers, the way the group reacts can be different, ranging from the full coherent turns of starling flocks [9] to the orientational cascades observed in fish schools [4]. Our analysis helps to understand why this might occur.

In the first part of this work (see section 4), we derived the dynamical response to local perturbations of the on-lattice system, and we tested the validity of our analytical expressions in numerical simulations. This allowed us to pinpoint the role of inertia and dissipation, both in the propagation mechanism of information and in the shape of a collective turn. In section 5, we extended our analysis to the off-lattice system in the presence of periodic boundary conditions. However, the PBCs are extremely artificial when thinking of groups. More importantly, they

hide the potentially disastrous effect of a strong dissipation by making a signal (the field), which is naturally finite in a motile finite group, long-standing. To overcome this problem, in section 6 we considered systems with OBCs. In this more realistic setup, we showed that for a coherent complete turn (collective change of direction) to occur, the group must i) live in the underdamped regime, so as to propagate information efficiently without damping, and ii) have a moderate motility, to avoid the 'bullet' effect, when the individual which first perceives the perturbation leaves the group before all the others have the time to fully rearrange their direction—see figure 5.

In our analysis, we considered an external field which is turned on very rapidly, with a step-like profile, and perpendicularly to the original flight direction of the flock. In a way, this represents the worst possible scenario of an abrupt change, forcing the system (if it can) to exhibit a quick response. Of course, less dramatic situations might be possible, where the field changes slowly in time (i.e. the parameter τ_{step} in equation (37) is large), and/or of an angle smaller than $\pi/2$. In this case, we expect less stringent requirements for a collective response with OBCs to occur. Indeed, the initiator changes its direction gradually due to the lesser strength of the field as compared to the social force of neighbors. This also implies that the initiator remains well inside the group for longer, giving more time to information to propagate and to other individuals to catch up. It is possible that in this case the effect of dissipation may therefore be less dramatic [32]. Future analysis of this aspect would certainly be interesting. However, given the variety and extension of possible perturbations occurring in realistic contexts, we believe that our analysis sets some benchmark criterion to understand efficient response behavior.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files). The code used for numerical simulations can be found in [28].

Acknowledgments

Appendix A. Derivation of the response

In this appendix we derive the response on-lattice of the polarization angle $\Phi(t)$ to a local perturbation, chosen in the form of an external field linearly coupled to one of the variables $\mathbf{v}_i$ (see section 3). This will lead to the results anticipated in section 3.2. In the following, we mainly adopt the notation first introduced in [21].

A.1. Microscopic calculation

We start from the coupled Langevin equations (28) and (29), which we recast for convenience in the matrix form

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\Psi} + \hat{L}\boldsymbol{\Psi} = \mathbf{F}(t), \tag{A1}$$

upon defining the $2N$-dimensional vectors

$$\boldsymbol{\Psi} \equiv \begin{pmatrix} \boldsymbol{\varphi} \\ \mathbf{s} \end{pmatrix}, \quad \mathbf{F} \equiv \begin{pmatrix} \mathbf{f}_{(\varphi)} \\ \mathbf{f}_{(s)} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\xi} + \mathbf{h}\sin\boldsymbol{\alpha} \end{pmatrix}. \tag{A2}$$

In particular, the N-dimensional vector $\boldsymbol{\varphi}$ has components $\varphi_i(t)$. We also introduced the block matrix

$$\hat{L} = \begin{pmatrix} \mathbb{O} & -\mathbb{I}/\chi \\ J\tilde{\Lambda} & \eta\mathbb{I}/\chi \end{pmatrix}, \tag{A3}$$

where $\tilde{\Lambda}$ is defined in terms of the discrete Laplacian Λ_{ij} given in equation (16) as

$$\tilde{\Lambda} = \Lambda + \delta\Lambda, \quad (\delta\Lambda)_{ij} = \delta_{ij}\frac{h_i\cos\alpha_i}{J}. \tag{A4}$$

Equation (A1) admits a formal solution as

$$\boldsymbol{\Psi}(t) = e^{-t\hat{L}}\boldsymbol{\Psi}(0) + \int_0^t \mathrm{d}t'\, e^{-(t-t')\hat{L}}\mathbf{F}(t'), \tag{A5}$$

which can be made explicit by diagonalizing $\hat{L}$.

A.1.1. Spectrum of $\hat{L}$

Let us initially set $h_i = 0$ in the expression for $\hat{L}$, which is equivalent to replacing $\tilde{\Lambda}$ with Λ in equation (A3). We will first compute the unperturbed spectrum of $\hat{L}$ and only later add the effect of a small perturbing field h_i. Let us adopt for convenience the *bra-ket* formalism familiar from Quantum Mechanics [33], which allows us to express $\boldsymbol{\varphi} \to |\varphi\rangle$, $\mathbf{s} \to |s\rangle$, and

$$\boldsymbol{\Psi} \to |\Psi\rangle = |\varphi\rangle \otimes |s\rangle, \tag{A6}$$

where the symbol $\otimes$ denotes the tensor product operation. The components of $|\varphi\rangle$ and $|s\rangle$ along the *position* basis

$$\mathcal{B}_{\mathrm{pos}} : \ |i\rangle \otimes |i\rangle, \tag{A7}$$

where $|i\rangle$ is a N-dimensional vector, can be found by taking the scalar products

$$\varphi^i = \langle i|\varphi\rangle, \quad s^i = \langle i|s\rangle. \tag{A8}$$

Conversely, let us denote as

$$\mathcal{B}_a : \ |a\rangle \otimes |a\rangle, \quad \Lambda|a\rangle = \lambda_a|a\rangle \tag{A9}$$

the basis in which Λ is diagonal. The new components φ^a, labelled by the superscript a, are connected to the old ones (φ^i) via

$$\varphi^a = \langle a|\varphi\rangle = \langle a|\sum_i |i\rangle\langle i|\varphi\rangle$$
$$= \sum_i \langle a|i\rangle\varphi^i \equiv \sum_i (U^{-1})^a{}_i\varphi^i, \tag{A10}$$

where in the first line we inserted a decomposition of the identity $\mathbb{I}$, while in the second line we introduced the $N\times N$ matrix U, which enforces the change of basis

$$U : \ \mathcal{B}_a \to \mathcal{B}_{\mathrm{pos}}. \tag{A11}$$

When written in the basis $\mathcal{B}_{\mathrm{pos}}$, the matrix U contains the eigenvectors of Λ as its columns; as a consequence, U diagonalizes Λ as

$$\mathrm{diag}(\lambda_1, \lambda_2, \ldots) = U^{-1}\Lambda U. \tag{A12}$$

In the case of symmetric interactions $n_{ij} = n_{ji}$, the condition $\Lambda_{ij} = \Lambda_{ji}$ makes U unitary ($U^{-1} = U^\dagger$); note that the interaction network does not need to be regular (e.g. a square lattice) in order for this to occur.

Working in the basis $\mathcal{B}_a$, we now look for the eigenvectors of $\hat{L}$:

$$\hat{L}|l\rangle = \omega_l|l\rangle, \quad |l\rangle = |l\rangle_\varphi \otimes |l\rangle_s. \tag{A13}$$

Without loss of generality, we choose both $|l\rangle_\varphi$ and $|l\rangle_s$ to be proportional to $|a\rangle$, i.e.

$$|l\rangle = |a\rangle \otimes c(l)|a\rangle \ \xrightarrow{\mathcal{B}_a} \ \kappa\begin{pmatrix} 1 \\ c(l) \end{pmatrix}, \tag{A14}$$

where κ is some normalization constant. The functional form of $c(l)$ has to be found by solving the eigenvalue equation (A13), which we can rephrase as

$$\begin{pmatrix} \omega_l & 1/\chi \\ -J\lambda_a & \omega_l - \eta/\chi \end{pmatrix}\begin{pmatrix} 1 \\ c(l) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{A15}$$

Solving for $c(l)$ leads to

$$-J\lambda_a - \omega_l^2\chi + \eta\omega_l = 0, \tag{A16}$$

so that, calling as usual $\gamma = \eta/(2\chi)$, we obtain

$$\omega_l = \gamma \pm i\omega_a, \quad \omega_a \equiv \sqrt{J\lambda_a/\chi - \gamma^2}. \tag{A17}$$

Using equation (A15) yields $c(l) = -\chi\omega_l$, which finally gives the eigenvectors

$$|l\rangle \ \xrightarrow{\mathcal{B}_a} \ \kappa\begin{pmatrix} 1 \\ -\chi\omega_l \end{pmatrix}. \tag{A18}$$

For each fixed value of a (there are N such values), equation (A17) shows as expected a twofold degeneracy for $\omega_l \equiv \omega_\pm(\lambda_a)$. We can thus represent

$$\hat{L} \xrightarrow[\text{fixed } a]{\mathcal{B}_a} L \equiv \begin{pmatrix} 0 & -1/\chi \\ J\lambda_a & \eta/\chi \end{pmatrix}, \qquad (A19)$$

each of which is diagonalized by a matrix M such that

$$\text{diag}(\omega_+, \omega_-) = M^{-1}LM. \qquad (A20)$$

The latter is non-unitary (L is not symmetrical even when $\Lambda_{ij} = \Lambda_{ji}$), and reads

$$M = \kappa \begin{pmatrix} 1 & 1 \\ -\chi\omega_+ & -\chi\omega_- \end{pmatrix},$$
$$M^{-1} = \frac{1}{2i\chi\kappa\omega_a} \begin{pmatrix} -\chi\omega_- & -1 \\ \chi\omega_+ & 1 \end{pmatrix}. \qquad (A21)$$

A.1.2. Temporal evolution
Equation (A5) takes a simple form in the basis $\mathcal{B}_a$. Indeed, by expressing

$$|\Psi\rangle = \sum_l |l\rangle\langle l|\Psi\rangle = \sum_l \Psi^l |l\rangle, \quad |F\rangle = \sum_l F^l |l\rangle, \qquad (A22)$$

and by using the completeness relation

$$e^{-t\hat{L}} = \sum_l e^{-t\omega_l} |l\rangle\langle l|, \qquad (A23)$$

one obtains the set of $2N$ decoupled equations

$$\Psi^l(t) = e^{-t\omega_l}\Psi^l(0) + \int_0^t dt' \, e^{-(t-t')\omega_l} F^l(t'). \qquad (A24)$$

In order to translate this result back into the position basis, we need to:

(i) Use the matrices U and U^{-1} on each of the two N-dimensional subspaces of the source term **F** in equation (A2), so as to move to the basis in which the Laplacian is diagonal;
(ii) Apply the matrices M and M^{-1} on each 2-dimensional subspace at fixed a, moving to the basis in which L is diagonal. This gives the components $F^l(t)$, which contain the source terms;
(iii) Use equation (A24) to compute the time evolution of each of the components $\Psi^l(t)$;
(iv) Apply the inverse of the two changes of basis used above (in reverse order), which finally gives $\varphi^i(t)$ and $s^i(t)$.

The above procedure is tedious but straightforward. Here we state the result for the variable $\varphi^i(t)$,

$$\varphi^i(t) = \sum_{a,j=1}^N U_a^i (U^{-1})^a{}_j \int_0^t dt' \, e^{-\gamma(t-t')}$$
$$\times \left\{ f_{(\varphi)}^j(t') \left[\frac{\gamma}{\omega_a}\sin\omega_a(t-t') + \cos\omega_a(t-t') \right] \right.$$
$$\left. + \frac{1}{\chi\omega_a} f_{(s)}^j(t')\sin\omega_a(t-t') \right\}, \qquad (A25)$$

where the second line vanishes in our case since $f_{(\varphi)}^j \equiv 0$. Equation (A25) is still written in terms of the matrix U, whose columns are the normalized eigenvectors of the discrete Laplacian Λ_{ij}: as such, it is valid for any time-independent interaction network n_{ij}, whose eigenvectors could in principle even be obtained numerically. In the following, we will see that we do not need to compute explicitly these eigenvectors in order to derive an expression for $\Phi(t)$ in the case of symmetric interactions.

If we now specialize to the case considered in the main text, where only a field coupled to the phases φ_i is considered, the term $f_{(\varphi)}^j$ is zero (see equations (26) and (27)) and we get

$$\varphi^i(t) = \sum_{a,j=1}^N U_a^i (U^{-1})^a{}_j \int_0^t dt' \, e^{-\gamma(t-t')}$$
$$\times \frac{1}{\chi\omega_a}[h_j(t')\sin\alpha_j + \xi_j(t')]\sin\omega_a(t-t'). \qquad (A26)$$

A.1.3. Corrections to the spectrum of $\hat{L}$
The first-order correction to the unperturbed spectrum when a small external field **h** is switched on can be computed by standard perturbation theory [33],

$$\tilde{\lambda}_a = \lambda_a + \langle a|\delta\Lambda|a\rangle, \quad |\tilde{a}\rangle = |a\rangle + \sum_{b\neq a} \frac{\langle b|\delta\Lambda|a\rangle}{\lambda_a - \lambda_b}|b\rangle, \qquad (A27)$$

where λ_a are the eigenvalues of the discrete Laplacian Λ and $|a\rangle$ are its eigenvectors. This leads to

$$\tilde{\lambda}_a = \lambda_a + \sum_i (U^{-1})^a{}_i \frac{h_i\cos\alpha_i}{J} U_a^i, \qquad (A28)$$

$$\tilde{U}^i{}_a = U_a^i + \sum_{b\neq a}\sum_j \frac{(U^\dagger)^b{}_j(h_j\cos\alpha_j)U_a^j}{J(\lambda_a - \lambda_b)}U^i{}_b, \qquad (A29)$$

which can be used into equation (A25) to replace U_a^i and λ_a, the latter being contained in ω_a via equation (A17). Note that the corrections in equations (A28) and (A29) vanish by choosing $\alpha_i \equiv \pi/2 \; \forall i$.

A.1.4. Average polarization angle
A remarkable property of the discrete Laplacian Λ_{ij}, which can be inferred from its definition in equation (16), is

$$\sum_j \Lambda_{ij} = \sum_j \left(-n_{ij} + \delta_{ij}\sum_k n_{ik} \right) = 0, \qquad (A30)$$

which means that Λ_{ij} always admits a zero mode (i.e. a $\lambda_0 = 0$ eigenvalue) with constant right eigenvector. This implies that the matrix U, which contains the normalized eigenvectors as its columns, must be such that

$$U^i{}_0 = 1/\sqrt{N} \ \ \forall i\,, \tag{A31}$$

or equivalently $(U^\dagger)^0{}_i = 1/\sqrt{N}$, since the transposed of the matrix U contains the same (right) eigenvectors as its *rows*. Note that in general these do *not* coincide with the left eigenvectors of Λ: this is only the case when Λ_{ij} is symmetric, because then the orthogonality condition reads

$$\mathbb{I} = U^{-1}U = U^\dagger U. \tag{A32}$$

In the symmetric case we then have in particular

$$\sum_i U^i_a = \delta_{a0}\sqrt{N}, \tag{A33}$$

and summing over i in equation (A25) in order to obtain $\Phi(t)$, only the term with $a = 0$ survives. Using equation (A2) (or equation (A26)) with $\langle \boldsymbol{\xi} \rangle = 0$, we thus find

$$\Phi(t) = \frac{1}{N}\sum_{j=1}^{N}\int_0^t \mathrm{d}t'\, e^{-\gamma(t-t')}\frac{h_j(t')\sin\alpha_j}{\chi\omega_0}\sin\omega_0(t-t')\,, \tag{A34}$$

where we called $\omega_0 = \omega_{a=0} = i\gamma + \mathcal{O}(h)$ (see equation (A17)). Note that equation (A34) is already of $\mathcal{O}(h)$, so that in general including the corrections to the spectrum of $\hat{L}$ due to the external field does not modify the leading-order contribution in perturbation theory. The validity of this approximation extends to higher orders by choosing $\alpha_j \equiv \pi/2$, because then we have seen that the first order corrections in equations (A28) and (A29) vanish. This is the choice we adopt in our numerical simulations.

By choosing a local field $h_i \propto \delta_{ip} = A_0\Theta(t)$, we finally recover the expression for $\Phi(t)$ reported in the main text, equation (31). When $\alpha_p \neq \pi/2$, however, one has to be careful in some regions of the parameter space. Indeed, the full explicit expression of ω_0 is (including corrections) $\omega_0 = \sqrt{A_0\cos\alpha/(N\chi) - \gamma^2}$. Usually the second term dominates, but for strong underdamping (i.e. $\eta^2 N/(2\chi A_0\cos\alpha) \ll 1$) it does not, and $\Phi(t)$ acquires a more complex structure.

A.2. Coarse-grained calculation

The expression of the response $\varphi^i(t)$ we derived in equation (A25) is general and it applies to any interaction network n_{ij}, even when it is not symmetric. However, we can get to equation (31) of the main text (which refers to the on-lattice case) with much less effort if we work in terms of the coarse-grained fields $\varphi(\mathbf{x}, t)$ and $s(\mathbf{x}, t)$.

The coarse-grained counterpart of our dynamical equations can be easily derived from equations (12) and (13) and read [22]

$$\dot{\varphi} = \frac{\delta \mathcal{U}}{\delta s}\,, \tag{A35}$$

$$\dot{s} = -\frac{\delta \mathcal{U}}{\delta \varphi} - \eta s + \xi\,, \tag{A36}$$

where the continuous version of the potential is

$$\mathcal{U} = \int \frac{\mathrm{d}^d x}{a^d}\left[\frac{Jn_{\mathrm{c}}a^2}{2}|\boldsymbol{\nabla}\varphi|^2 + \frac{1}{2}\cos\alpha\, h(\mathbf{x}, t)\varphi^2 \right. \\ \left. - \sin\alpha\, h(\mathbf{x}, t)\varphi\right]. \tag{A37}$$

Note that the component of the external field **h** parallel to the initial polarization $\mathbf{V}(t=0)$ plays the role of a *mass* term for the field φ. We stress that $\mathcal{H}$ above has been derived under the SWA, hence it has to be considered as a low-temperature expansion of the ISM. In fact, a field theory can still be constructed if one relaxes this assumption, but it will in general be more complicated and possibly contain new, non-Gaussian terms [31]. From equations (A36) and (A35) we get

$$\dot{\varphi} = \frac{s}{\chi}\,,$$
$$\dot{s} = Jn_{\mathrm{c}}a^2\nabla^2\varphi - h\varphi\cos\alpha + h\sin\alpha - \frac{\eta}{\chi}s + \xi\,, \tag{A38}$$

and deriving again the first equation we obtain

$$\chi\ddot{\varphi} + \eta\dot{\varphi} - Jn_{\mathrm{c}}a^2\nabla^2\varphi + h\varphi\cos\alpha = h\sin\alpha + \xi\,. \tag{A39}$$

In analogy with the microscopic derivation, we now discard the term proportional to h on the left-hand side, as we expect it to produce higher-order corrections for small h (note that such term actually vanishes for $\alpha = \pi/2$). This leads us to the linear problem

$$\mathcal{D}_{\mathbf{x}, t}\varphi(\mathbf{x}, t) = f(\mathbf{x}, t) \tag{A40}$$

upon defining

$$\mathcal{D}_{\mathbf{x}, t} \equiv \chi\, \partial_{tt}^2 + \eta\partial_t - Jn_{\mathrm{c}}a^2\nabla_{\mathbf{x}}^2\,, \tag{A41}$$

$$f(\mathbf{x}, t) \equiv \sin\alpha\, h(\mathbf{x}, t) + \xi(\mathbf{x}, t)\,. \tag{A42}$$

The propagator of the differential operator in equation (A41) can be computed by standard methods [18] to give, in the time-momentum domain,

$$G(\mathbf{k}, t) = e^{-\gamma t}\frac{\sin(\omega_{\mathbf{k}}t)}{\chi\omega_{\mathbf{k}}}\,, \tag{A43}$$

where $\omega_{\mathbf{k}} \equiv c_s \sqrt{k^2 - k_0^2}$ is precisely the one appearing on the right-hand side of equation (19). This yields the average response of the system as

$$\langle \varphi(\mathbf{k}, t) \rangle = \int_0^t \mathrm{d}t'\, G(\mathbf{k}, t - t') \langle f(\mathbf{k}, t') \rangle. \quad \text{(A44)}$$

We now turn to the polarization angle in equation (30), whose continuous-space version is

$$\Phi(t) = \frac{1}{\mathcal{V}} \int_{\mathcal{V}} \mathrm{d}^d x \langle \varphi(\mathbf{x}, t) \rangle = \frac{1}{\mathcal{V}} \langle \varphi(\mathbf{k} = \mathbf{0}, t) \rangle. \quad \text{(A45)}$$

Here we called $\mathcal{V}$ the volume occupied by the entire flock (i.e. $\mathcal{V} \sim N$ on a square lattice), and we adopted the continuum prescription

$$\sum_{\mathbf{k}} f(\mathbf{k}) \simeq \frac{\mathcal{V}}{(2\pi)^d} \int \mathrm{d}^d k\, f(\mathbf{k}). \quad \text{(A46)}$$

Using equations (A43) and (A44) then simply gives

$$\begin{aligned} \Phi(t) &= \frac{1}{\mathcal{V}} \int_{-\infty}^t \mathrm{d}t'\, G(\mathbf{k} = \mathbf{0}, t - t') \sin\alpha H(\mathbf{k} = \mathbf{0}, t') \\ &= \frac{\sin\alpha}{\mathcal{V}\eta} \int_{-\infty}^t \mathrm{d}t'\, H(\mathbf{k} = \mathbf{0}, t') \left[1 - e^{-2\gamma(t - t')} \right]. \end{aligned} \quad \text{(A47)}$$

To make contact with equation (31), it is sufficient to choose a local field and give it a stepfunction-like time dependence,

$$H(\mathbf{x}, t) = A_0 \delta(\mathbf{x} - \mathbf{x}_0) \Theta(t) \;\to\; H(\mathbf{k}, t) = A_0 e^{i\mathbf{k}\cdot\mathbf{x}_0} \Theta(t). \quad \text{(A48)}$$

Let us note, however, the difference with respect to the microscopic approach of appendix A.1: in that case, we have studied the effect of a perturbation applied on one of the microscopic variables $\mathbf{v}_i$. Here, conversely, the coarse-graining procedure has washed out the identity of the single microscopic degrees of freedom $\mathbf{v}_i$, and the perturbation in equation (A48) is thus applied locally at position $\mathbf{x}_0$. Moving off-lattice by introducing a large activity v_0 (see section 5), the former approach is expected to fail because of the time dependence in the interaction matrix $n_{ij}(t)$, while the latter approach fails both because the propagator $G(\mathbf{k}, t)$ in equation (A43) was computed on-lattice, and because it erroneously assumes that the perturbation remains fixed in space at $\mathbf{x} = \mathbf{x}_0$.

Appendix B. Spin-wave decomposition and wandering of the order parameter

In this appendix we formally develop the spin-wave decomposition, and then we use it to predict the *wandering time* τ_w of the order parameter $\mathbf{V}$. As explained in section 3.2, the latter is defined as the persistence time of the total polarization $\mathbf{V}$ as it explores its broken-symmetry manifold under the effect of thermal fluctuations. The resulting estimate of τ_w is based on the on-lattice approximation, and it knows nothing about the activity v_0: one can thus interpret τ_w as the wandering time of the total magnetization in the XY model on a fixed lattice, and subject to an underdamped Langevin dynamics.

B.1. Spin-wave decomposition

When the system is in its ordered phase, it is useful to decompose each velocity vector into its components parallel and perpendicular to the global polarization $\mathbf{V}$,

$$\mathbf{v}_i = v_i^L \hat{\boldsymbol{n}} + \boldsymbol{\pi}_i, \quad \text{(B1)}$$

where $\boldsymbol{\pi}_i$ is a $(d_v - 1)$-dimensional vector, and by construction one has

$$\sum_i \boldsymbol{\pi}_i = 0. \quad \text{(B2)}$$

In the planar model, π_i is simply a scalar,

$$\pi_i = v_0 \sin\varphi_i. \quad \text{(B3)}$$

Now let us call $\hat{\boldsymbol{n}}$ the direction assumed by the polarization $\mathbf{V}$ at time $t = 0$, but allow $\mathbf{V}(t)$ to change in time. We can decompose

$$\mathbf{V} = \frac{1}{N} \sum_i \mathbf{v}_i = (\mathbf{V} \cdot \hat{\boldsymbol{n}}) \hat{\boldsymbol{n}} + \delta \boldsymbol{V}^\perp, \quad \text{(B4)}$$

where in the planar case

$$(\mathbf{V} \cdot \hat{\boldsymbol{n}})(t) = \frac{1}{N} \sum_i v_i^L(t), \quad \text{(B5)}$$

$$\begin{aligned} \delta \boldsymbol{V}^\perp(t) &= \frac{1}{N} \sum_i \boldsymbol{\pi}_i = \frac{v_0}{N} \sum_i \sin\varphi_i(t) \hat{\boldsymbol{\pi}} \\ &\simeq \frac{v_0 \hat{\boldsymbol{\pi}}}{N} \sum_i \varphi_i(t), \end{aligned} \quad \text{(B6)}$$

the latter being valid within the SWA.

B.2. Scalar polarization and mean angular deviation

The SWA decomposition can be used to derive a relation between the scalar polarization, i.e. $\Psi = |\mathbf{V}|/v_0$, and the average fluctuation $\delta\varphi$ of a single velocity vector $\mathbf{v}_i$ around the global flight direction $\mathbf{V}$. Indeed, using equations (B1) and (B2) we have

$$\begin{aligned} \Psi &= \frac{1}{v_0 N} \left| \sum_i \mathbf{v_i} \right| = \frac{1}{N} \left| \sum_i \cos\varphi_i \right| \\ &\simeq 1 - \frac{1}{2N} \sum_i \varphi_i^2, \end{aligned} \quad \text{(B7)}$$

where in the last step we used the SWA. Defining the mean fluctuation as

$$\delta\varphi \equiv \sqrt{\langle \varphi^2 \rangle} = \frac{1}{N} \sum_i \varphi_i^2, \quad \text{(B8)}$$

the relation stated in equation (39) follows immediately from equation (B7).

B.3. Persistence time τ_w of the order parameter

We define the wandering time τ_w by the condition

$$\langle |\delta\boldsymbol{\Psi}^{\perp}(t \sim \tau_w)|^2\rangle \sim \mathcal{O}(1), \tag{B9}$$

where the brackets denote the average over the noise, and $\mathbf{V} = v_0\boldsymbol{\Psi}$. Using equation (B5), we find under the SWA

$$\langle |\delta\boldsymbol{\Psi}^{\perp}(t)|^2\rangle = \frac{1}{v_0^2 N^2}\sum_{ik}\langle \boldsymbol{\pi}_i \cdot \boldsymbol{\pi}_k\rangle$$
$$\simeq \frac{1}{N^2}\sum_{ik}\langle \varphi^i(t)\varphi^k(t)\rangle. \tag{B10}$$

The evolution of $\varphi^i(t)$ is given by equation (A25) upon setting the external source field $\mathbf{h} = 0$. This gives, for a *symmetric* interaction network,

$$\varphi^i(t) = \sum_{a,j=1}^{N} U_a^i (U^\dagger)^a{}_j \int_0^t \mathrm{d}t' \frac{e^{-\gamma(t-t')}}{\chi\omega_a}\xi^j(t')$$
$$\times \sin\omega_a(t-t'), \tag{B11}$$

where it is useful to isolate the contribution of the $a = 0$ mode

$$\varphi^i(t) = \frac{1}{N}\sum_{j=1}^{N}\int_0^t \mathrm{d}t' \frac{\xi^j(t')}{\eta}\left[1 - e^{-2\gamma(t-t')}\right]$$
$$+ (a \neq 0). \tag{B12}$$

Indeed, thanks again to the property in equation (A33), the modes $a \neq 0$ give no contribution; using the noise variance in equation (5) we thus obtain

$$\langle |\delta\boldsymbol{\Psi}^{\perp}(t)|^2\rangle = \frac{2T}{N\eta}\left[t - \frac{1-e^{-2\gamma t}}{\gamma} + \frac{1-e^{-4\gamma t}}{4\gamma}\right]. \tag{B13}$$

Specializing this expression for the two regimes $t \gg \gamma^{-1}$ and $t \ll \gamma^{-1}$ and using the condition in equation (B9), we finally get

$$\tau_w = \begin{cases} \frac{\eta N}{2T}, & \text{for } t \gg \gamma^{-1}, \\ \left(\frac{3}{16}\frac{N\eta}{T\gamma^2}\right)^{1/3}, & \text{for } t \ll \gamma^{-1}. \end{cases} \tag{B14}$$

We stress again that equation (B14) can be generically interpreted as the persistence time of the average polarization in the XY model subject to an underdamped Langevin dynamics. Its origin, as discussed above, is related to the presence of the zero modes of the Laplacian, i.e. to the original rotational invariance of the potential function $\mathcal{U}$ of the velocities (see equation (3)). This is a common feature of $O(n)$ models, and it is independent of the kind of dynamics

adopted to let the system evolve. Indeed, the wandering of the order parameter and its persistence time have been computed in a variety of works, see e.g. [34, 35] where a microcanonical dynamics for the XY model was considered, or the more recent [21, 36] where the ISM and the VM were analyzed.

Appendix C. Symmetry breaking in 2d spin systems

In this appendix we address the issue of the presence of a spontaneous magnetization in the XY model (to which our system effectively reduces when the activity v_0 is set to zero). We start by recalling a heuristic argument which is standard in statistical mechanics [37]; a more rigorous proof due to Mermin and Wagner can be found in [38]. Let us call d the dimension of the physical space and let n be the dimension of the order parameter (in our case $d = n = 2$).

Suppose that a nonzero spontaneous magnetization exists near $T = 0$, and let us inspect the stability of the corresponding ordered state against small thermal fluctuations. At low temperature, we can assume all the spins to be almost aligned in one direction, so that we can work within the continuum limit and the SWA. We thus consider a Hamiltonian of the form

$$\mathcal{H} = \frac{\mathcal{J}}{2}\int \mathrm{d}^2x \sum_{\alpha=1}^{n-1}[\nabla\varphi_\alpha(\mathbf{x})]^2, \tag{C1}$$

corresponding to a Gaussian and massless field theory whose Fourier-space propagator reads [37]

$$\widetilde{G}_0(\mathbf{k}) = T/(\mathcal{J}\mathbf{k}^2). \tag{C2}$$

The fluctuation Δ of the order parameter can be estimated as

$$\Delta \equiv \sum_{\alpha=1}^{n-1}\langle \varphi_\alpha^2(\mathbf{x})\rangle = (n-1)\langle \varphi_1^2(\mathbf{x})\rangle$$
$$= (n-1)G_0(\mathbf{r} = 0) = \frac{(n-1)T}{\mathcal{J}}\int_{\pi/L}^{\pi/a}\frac{\mathrm{d}^d k}{(2\pi)^d \mathbf{k}^2}, \tag{C3}$$

where the integration cutoffs are given by the system size L and the lattice spacing a. For $d > 2$ the integral in equation (C3) is infrared-convergent, so that $\Delta \to 0$ as $T \to 0$, which is consistent with our initial assumption that the ordered state is stable against fluctuations. However, in $d = 2$ one has

$$\Delta \sim \frac{(n-1)T}{\mathcal{J}}\ln\frac{L}{a}, \tag{C4}$$

showing that $\Delta \to \infty$ when $L \to \infty$: thus, the long-wavelength fluctuations destabilize the long-range order in the thermodynamic limit. In fact, it is well-known that in $d = 2$ a phase transition is observed at $T = T_{BKT}$ [39]: for $T < T_{BKT}$ the correlation functions

exhibit a scale-free decay, but the average magnetization remains zero (this is known as *quasi long-range order*).

This classical argument explains the absence of long-range order in $d \leqslant 2$ continuous systems in the thermodynamic limit. Conversely, finite-size systems do exhibit a continuous transition at a critical temperature T_c, so that below T_c we observe a nonzero spontaneous magnetization [34, 35, 40–42]. Of course the value of the magnetization (slowly) decays as an inverse power of the system size N, so that there is no contradiction with the BKT theory: a low-temperature, spin-wave analysis renders a total magnetization [40, 43]

$$M(N,T) = \left(\frac{1}{2N}\right)^{T/(8\pi\mathcal{J})}, \qquad (C5)$$

hence for instance $M \sim N^{-1/16}$ at the BKT transition. In fact, one finds $M \sim \mathcal{O}(1)$ for a system with $N \sim \mathcal{O}(10^2 - 10^3)$ like the ones we analyze in this work: this is why in section 4 we can still study the global order parameter **V** even within the on-lattice approximation. On the other hand, in the off-lattice case the coarse-graining procedure which led to equation (C1) breaks down, so that the argument above does not apply—indeed, spontaneous symmetry breaking and long-range order are well-known to take place in $d = 2$ active systems [44].

Finally, we can however ask how the computations performed in this work relate to the stability of order and the Mermin–Wagner result [38]. In our numerical analysis, as stated above, we considered sizes and temperatures for which the magnetization/polarization remains of $\mathcal{O}(1)$ according to equation (C5). What would happen, however, if we considered larger sizes? If we prepare the system in an ordered state and then apply a field, the polarization angle will follow exactly the same behavior as described in the main text, provided that the time scale of the turn is sufficiently fast. Note that actually equation (32) does not depend on the physical dimension d of the space where the system lives. However, there are other quantities that crucially depend on the space dimension: the scalar polarization Ψ is one of them. Once the turn is complete and the instantaneous mean polarization is along the direction of the field, we can compute the scalar polarization within the SWA from equation (B7), assuming as reference direction the one of the field and using the expression for $\varphi_i(t)$ derived in appendix A. It is easy to see that the computation gives back asymptotically the equilibrium estimate, meaning that—once the field is set back to zero—the polarization in $d = 2$ is destroyed by fluctuations, while in larger dimensions it remains finite.

Appendix D. Numerical integration scheme

We start by introducing $\sigma \equiv \sqrt{2T\eta}$ and

$$f_i(\mathbf{x}(t),t) \equiv J \sum_j n_{ij}(t) \sin(x_j(t) - x_i(t)), \qquad (D1)$$

so that the laws of motion in equations (25)–(27) can be expressed as

$$d\varphi_i(t) = \chi^{-1} s_i(t) dt, \qquad (D2)$$

$$ds_i(t) = [f_i(\boldsymbol{\varphi}(t),t) - 2\gamma s_i(t) - h_i(t)\sin\varphi_i(t)] dt \\ + \sigma dW_i(t), \qquad (D3)$$

where $W_i(t)$ is a Wiener process satisfying $\langle \mathbf{W}(t) \rangle = \mathbf{0}$ and

$$\langle W_i(t) W_j(t') \rangle = \delta_{ij} \min(t,t'). \qquad (D4)$$

We aim at discretizing these stochastic differential equations in order to obtain a numerical integration scheme which is at least of the second order in the integration timestep Δt. To this end, we generalize the procedure described in [24] to the case in which $f_i(\mathbf{x}(t),t)$ can depend explicitly on time t. A lengthy but straightforward calculation [22] leads to

$$\varphi_i(t+\Delta t) = \varphi_i(t) + s_i(t)\Delta t/\chi + A_i(t) + \mathcal{O}(\Delta t^{5/2}), \qquad (D5)$$

$$s_i(t+\Delta t) = s_i(t) + \Delta t\left[\overline{f_i} - 2\gamma s_i(t) - \overline{h_i \sin\varphi_i}\right] \\ + \sigma\sqrt{\Delta t}\xi_i - 2\gamma A_i(t) + \mathcal{O}(\Delta t^{5/2}), \qquad (D6)$$

where the bar over a variable stands for its midpoint value:

$$\overline{f_i} = \frac{1}{2}\left[f_i(\boldsymbol{\varphi}(t+\Delta t), t+\Delta t) + f_i(\boldsymbol{\varphi}(t), t)\right]. \qquad (D7)$$

We also introduced the auxiliary variable

$$A_i(t) \equiv [f_i(\boldsymbol{\varphi}(t),t) - 2\gamma s_i(t) - h_i(t)\sin\varphi_i(t)]\Delta t^2/2 \\ + \left(\xi_i + \zeta_i/\sqrt{3}\right)\sigma\Delta t^{3/2}/2, \qquad (D8)$$

where ξ_i and ζ_i are white uncorrelated Gaussian variables with zero mean and unit variance.

There remains to specify how the positions $\mathbf{r}_i(t)$ evolve according to their corresponding velocity $\mathbf{v}_i(t)$. A simple choice is the Euler–Cromer update rule [26, 27],

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t+\Delta t)\Delta t, \qquad (D9)$$

where each component of $\mathbf{r}_i$ has to be recast in $[0, L]$ in the case of periodic boundary conditions. The connectivity matrix $n_{ij}(t)$ contained in the force term

f(**x**(*t*), *t*) in equation (D1) must also be updated regularly, so as to take into account the relative motion of the individual positions. This is obtained by the *cell-list* method [26, 27], which consists of dividing the lattice into cells, and assigning to each particle a label indicating the cell it occupies at a given time *t*; at the same time, we associate to each cell the list of the current occupants. This expedient speeds up the computation of the interacting force contributions, which become of $\mathcal{O}(N)$ (rather than $\mathcal{O}(N^2)$) since we only have to cycle over particles belonging to the same or neighboring cells, and only then check if their distance lies below the metric interaction radius (see section 2.1).

The complete code used for numerical simulations (written in C) is available open source here [28].

Appendix E. Oscillations of the polarization angle in the underdamped regime

We have noted in section 4.1 how, in the deeply underdamped regime, the mean polarization angle $\Phi(t)$ may occasionally overshoot the final angle α imposed by the external perturbation, and perform underdamped oscillations around the latter. An intuitive understanding of this phenomenon can be grasped by adopting again the coarse-grained description of appendix A.2. We start from equation (A39) and set $\alpha = 0$, which corresponds to a SWA around the *final* direction reached by the flock at the end of the turn. We additionally apply the mean-field approximation $\varphi(\mathbf{x}, t) = \varphi_0(t)$, whose validity can be checked *a posteriori*, thus obtaining

$$\chi \ddot{\varphi}_0 + \eta \dot{\varphi}_0 + h\varphi_0 = \xi(t). \quad (E1)$$

We now choose the external field as $h(\mathbf{x}, t) = A_0 \delta(\mathbf{x})\Theta(t)$. Using the definition of $\Phi(t)$ in equation (A45) and taking the average over the noise, we obtain an evolution equation for the total polarization in the form

$$\ddot{\Phi} + \frac{\eta}{\chi}\dot{\Phi} + \frac{A_0}{\chi N}\Phi = 0. \quad (E2)$$

This way we recover the equation of motion of a simple harmonic oscillator with damping coefficient $\gamma = \eta/(2\chi)$ and frequency $\Omega = \sqrt{A_0/(\chi N)}$.

Note that using the mean-field framework means assuming all the individuals to turn coherently, as if the system behaved as a rigid body. Hence, we are neglecting the site-to-site signal propagation which occurs within the flock at finite speed c_s—see section 4.2. However, the latter can generally be ignored when studying the global underdamped oscillations, since their associated timescale Ω^{-1} is much larger than the typical time taken by the information to travel from one individual to another, i.e. $\Omega^{-1} \gg a/c_s$, where a is the lattice spacing. We check this explicitly in figure 7, where we plot $\varphi_i(t)$
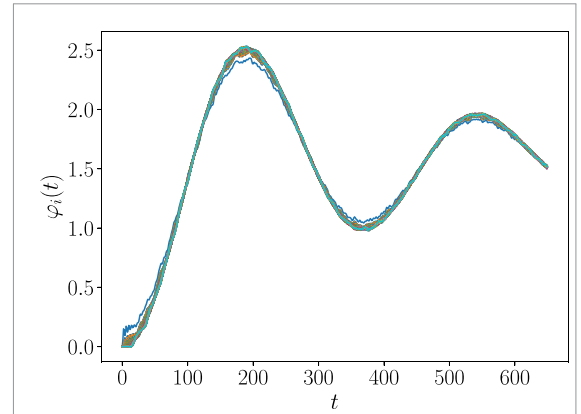
Figure 7. Short-time evolution of $\varphi_i(t)$ for various individuals, in the deeply underdamped regime (on-lattice case). The individual phases φ_i evolve coherently together, hence justifying the use of the mean-field approximation in studying global damped oscillations, as we did in appendix E. We used $\chi = 200, A_0 = 30, J = 50, T = 0.005, \eta = 1, N = 400$.

for various individuals, and observe that the corresponding curves almost coincide.

Appendix F. Impulse-like perturbations

In this appendix we derive the response of the system to a finite-duration perturbation, which provides useful insights for the OBCs case analyzed in section 6.

Let us then specialize the response in equation (31) to the case in which the local external field $\mathbf{h}_p(t)$ has an impulse-like time dependence of the form

$$h_p(t) = A_0 \Theta(t)\Theta(\mathcal{T} - t), \quad (F1)$$

where $\mathcal{T}$ is the impulse duration. This leads to

$$\Phi(t) = \frac{A_0 \sin \alpha_p}{N\eta}\left[\mathcal{T} + \frac{e^{-2\gamma t}}{2\gamma}\left(1 - e^{2\gamma\mathcal{T}}\right)\right], \quad (F2)$$

valid as long as $\Phi(t) \ll \alpha_p$. In the limit $N \to \infty$, the long-*t* value of the response in equation (F2) vanishes, showing (as expected) that a *local* and *impulse-like* perturbation has no influence on the system in the thermodynamic limit.

Curiously, equation (F2) also implies that the endpoint $\Phi(t \to \infty)$ is the one given in equation (44), which does *not* depend on the rotational inertia χ. This may appear counter-intuitive: how can a finite external perturbation win a possibly very large rotational inertia? To answer this point, it is sufficient to note that a system with large inertia χ will keep rotating even after we stop applying the external *torque*. Indeed, consider the analogous problem of a point particle $x(t)$ of *mass* χ subject to viscous friction,

$$\chi \ddot{x} = -\eta\dot{x}. \quad (F3)$$

By the impulse-momentum theorem, applying an impulse $F \cdot \mathcal{T}$ on the particle initially at rest changes

its velocity from $\dot{x} = 0$ to $\dot{x} = F\mathcal{T}/\chi$. It is then a simple exercise to check that the endpoint $x(t \to \infty)$ of the particle trajectory is actually independent of its mass χ.

ORCID iDs

Emanuele Loffredo ⓘ https://orcid.org/0009-0004-4882-8250

Davide Venturelli ⓘ https://orcid.org/0000-0001-6952-9568

References

[1] Lima S L 1995 Back to the basics of anti-predatory vigilance: the group-size effect *Anim. Behav.* **49** 11

[2] Krause J, Krause P, Ruxton G and Ruxton G 2002 *Living in Groups* (Oxford: Oxford Series in Ecology and Evolution (OUP))

[3] Katz Y, Tunstrøm K, Ioannou C C, Huepe C and Couzin I D 2011 Inferring the structure and dynamics of interactions in schooling fish *Proc. Natl Acad. Sci.* **108** 18720

[4] Rosenthal S B, Twomey C R, Hartnett A T, Wu H S and Couzin I D 2015 Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion *Proc. Natl Acad. Sci.* **112** 4690

[5] Handegard N O, Boswell K M, Ioannou C C, Leblanc S P, Tjøstheim D B and Couzin I D 2012 The dynamics of coordinated group hunting and collective information transfer among schooling prey *Curr. Biol.* **22** 1213

[6] Herbert-Read J E, Buhl J, Hu F, Ward A J W and Sumpter D J T 2015 Initiation and spread of escape waves within animal groups *R. Soc. Open Sci.* **2** 140355

[7] Calovi D S, Lopez U, Schuhmacher P, Chaté H, Sire C and Theraulaz G 2015 Collective response to perturbationsin a data-driven fish school model *J. R. Soc. Interface* **12** 20141362

[8] Klamser P P and Romanczuk P 2021 Collective predator evasion: putting the criticality hypothesis to the test *PLOS Comput. Biol.* **17** e1008832

[9] Attanasi A *et al* 2014 Information transfer and behavioural inertia in starling flocks *Nat. Phys.* **10** 691

[10] Cavagna A *et al* 2015 Flocking and turning: a new model for self-organized collective motion *J. Stat. Phys.* **158** 601–27

[11] Cavagna A, Giardina I, Grigera T S, Jelic A, Levine D, Ramaswamy S and Viale M 2015 Silent flocks: constraints on signal propagation across biological groups *Phys. Rev. Lett.* **114** 218101

[12] Vicsek T, Czirók A, Ben-Jacob E, Cohen I and Shochet O 1995 Novel type of phase transition in a system of self-driven particles *Phys. Rev. Lett.* **75** 1226

[13] Vicsek T and Zafeiris A 2012 Collective motion *Phys. Rep.* **517** 71

[14] Couzin I D and Krause J 2003 Self-organization and collective behavior in vertebrates *Advances in the Study of Behavior* vol 32 p 1

[15] Toner J, Tu Y and Ramaswamy S 2005 Hydrodynamics and phases of flocks *Ann. Phys.* **318** 170

[16] Marchetti M C, Joanny J F, Ramaswamy S, Liverpool T B, Prost J, Rao M and Simha R A 2013 Hydrodynamics of soft active matter *Rev. Mod. Phys.* **85** 1143

[17] Ramaswamy S 2010 The mechanics and statistics of active matter *Annu. Rev. Conden. Matter Phys.* **1** 323

[18] Cavagna A, Giardina I and Grigera T S 2018 The physics of flocking: correlation as a compass from experiments to theory *Phys. Rep.* **728** 1

[19] Zwanzig R 2001 *Nonequilibrium Statistical Mechanics* (Oxford: Oxford University Press)

[20] Attanasi A, Cavagna A, Castello L D, Giardina I, Jelic A, Melillo S, Parisi L, Pohl O, Shen E and Viale M 2015 Emergence of collective changes in travel direction of starling flocks from individual birds' fluctuations *J. R. Soc. Interface* **12** 20150319

[21] Cavagna A, Giardina I, Jelic A, Melillo S, Parisi L, Silvestri E and Viale M 2017 Nonsymmetric interactions trigger collective swings in globally ordered systems *Phys. Rev. Lett.* **118** 138003

[22] Venturelli D 2019 Dynamical response to local perturbations of an active matter system with polar order *Master's Thesis* Sapienza, Universitá di Roma (https://doi.org/10.5281/zenodo.7615248)

[23] Mora T, Walczak A M, Castello L D, Ginelli F, Melillo S, Parisi L, Viale M, Cavagna A and Giardina I 2016 Local equilibrium in bird flocks *Nat. Phys.* **12** 1153

[24] Vanden-Eijnden E and Ciccotti G 2006 Second-order integrators for Langevin equations with holonomic constraints *Chem. Phys. Lett.* **429** 310

[25] Chepizhko O, Saintillan D and Peruani F 2021 Revisiting the emergence of order in active matter *Soft Matter* **17** 3113

[26] Frenkel D and Smit B 2001 *Understanding Molecular Simulation* 2nd edn (New york: Academic, Inc.)

[27] Tuckerman M E 2010 *Statistical Mechanics* (Oxford: Oxford University Press)

[28] Venturelli D and Loffredo E 2022 Source code for the 2D inertial spin model (available at: https://github.com/sonarventu/ISM)

[29] Cavagna A, Del Castello L, Dey S, Giardina I, Melillo S, Parisi L and Viale M 2015c Short-range interactions versus long-range correlations in bird flocks *Phys. Rev.* E **92** 012705

[30] Ferretti F, Chardès V, Mora T, Walczak A M and Giardina I 2020 Building generalLangevin models from discrete datasets *Phys. Rev.* X **10** 031018

[31] Cavagna A, Di Carlo L, Giardina I, Grigera T S, Melillo S, Parisi L, Pisegna G and Scandolo M 2021 Natural swarms in **3.99** dimensions (https://doi.org/10.48550/arXiv.2107.04432)

[32] Holubec V 2023 private communication

[33] Sakurai J and Napolitano J 2017 *Modern Quantum Mechanics* (Cambridge: Cambridge University Press)

[34] Leoncini X, Verga A D and Ruffo S 1998 Hamiltonian dynamics and the phase transition of the XY model *Phys. Rev.* E **57** 6377

[35] Lepri S and Ruffo S 2001 Finite-size effects on the Hamiltonian dynamics of the XY-model *Europhys. Lett.* **55** 512

[36] Brambati M, Fava G and Ginelli F 2022 Signatures of directed and spontaneous flocking *Phys. Rev.* E **106** 024608

[37] Le Bellac M 1991 *Quantum and Statistical Field Theory* (Oxford: Clarendon)

[38] Mermin N D and Wagner H 1966 Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic Heisenberg models *Phys. Rev. Lett.* **17** 1133

[39] Kosterlitz J M and Thouless D J 1973 Ordering, metastability and phase transitions in two-dimensional systems *J. Phys.* C **6** 1181

[40] Bramwell S T and Holdsworth P C W 1993 Magnetization and universal sub-critical behaviour in two-dimensional XY magnets *J. Phys.: Condens. Matter* **5** L53

[41] Roomany H H and Wyld H W 1980 Finite-lattice approach to the O(2) and O(3) models in $1+1$ dimensions and the (2+1)-dimensional Ising model *Phys. Rev.* D **21** 3341

[42] Chung S G 1999 Essential finite-size effect in the two-dimensional XY model *Phys. Rev.* B **60** 11761

[43] Tobochnik J and Chester G V 1979 Monte Carlo study of the planar spin model *Phys. Rev.* B **20** 3761

[44] Toner J and Tu Y 1998 Flocks, herds and schools: a quantitative theory of flocking *Phys. Rev.* E **58** 4828