PHYSICAL REVIEW RESEARCH **6**, 023016 (2024)

Modeling collective behaviors from optic flow and retinal cues

Diego Castro,[1,2] Franck Ruffier,[1,*] and Christophe Eloy[2]
[1]*Aix Marseille Université, CNRS, ISM, Marseille, 13288 France*
[2]*Aix Marseille Université, CNRS, Centrale Med, IRPHE, Marseille, 13013 France*

(Received 10 May 2023; accepted 16 February 2024; published 4 April 2024)

Animal collective behavior is often modeled with self-propelled particles, assuming each individual has "omniscient" knowledge of its neighbors. Yet, neighbors may be hidden from view and we do not know the effect of this information loss. To address this question, we propose a visual model of collective behavior where each particle moves according to bioplausible visual cues, in particular the optic flow. This visual model successfully reproduces three classical collective behaviors: swarming, schooling, and milling. This model offers a potential solution for controlling artificial swarms visually.

DOI: 10.1103/PhysRevResearch.6.023016

I. INTRODUCTION

Collective animal behavior is a widespread phenomenon in nature, ranging from the mesmerizing movements of starling murmurations to the coordinated motion of cattle herds [1–4]. These collective behaviors are commonly modeled with self-propelled particles: individuals with an intrinsic speed orient themselves based on a set of rules, alignment, attraction, and avoidance [5–8]. These "3A" rules have been successful in reproducing different collective behaviors, such as swarming (no orientational order), schooling (high orientational order), and milling (the group coordinately swirls in a circular pattern) [9–12]. Furthermore, numerous studies used these rules to replicate some of these phases on artificial robotic swarms [13–15].

In self-propelled particle (SPP) models, the rules of attraction, alignment, and avoidance are typically applied with the simplifying assumption that each individual possess idealized senses, is "omniscient" and gauges perfectly the position, distance, orientation, and velocity of its neighbors [16,17]. However, this assumption may not hold in practice, especially when some neighboring individuals are hidden from view [18–22].

The causal link between visual cues and collective behavior has been shown through several vision-based biological models [23–26]. As a result, different ways of incorporating vision in SPP models have been suggested. The most widespread approach consists of using vision to filter information [27,28]. One of these models suggested that flocks of starlings adjust their density to reach a state of "marginal opacity" [29]. However, this density adjustment does not seem to be widespread, as certain animals such as fish can form opaque schools [7,30].

In a recent study, Bastien and Romanczuk proposed a model of collective behavior based purely on vision capable of reproducing most of the collective behaviors [31]. Their model simulates each individual's response to a projection of the visual field, rather than relying on omniscient information. However, the portrayed milling phase is uncoordinated, meaning that the particles turn in both directions in the same swirl. And, while they claim to use the simplest possible equations of movement that satisfy fundamental symmetries, their model involves six parameters that are hard to relate to the classical 3A rules. In addition, four terms of their model involve spatial and temporal derivatives that would use important computing resources. Instead, we propose to use bioplausible visual cues that can be measured directly by a visual sensor, i.e., the optic flow and the retinal position.

Optic flow refers to the apparent angular velocity of objects in the visual field due to the relative motion between the observer and its surroundings. Numerous animal species perceive and use optic flow for a variety of tasks. Insects use it to navigate in crowded settings [32], evade ground obstacles, and control their landing [33]. Fish use it for navigation [34], and birds use it during takeoff [35]. Optic flow is ubiquitous in nature. It involves specialized neurons, well identified in invertebrates, which have inspired bioinspired sensors dedicated to optic flow [36–39]. These visual sensors can provide panoramic optic flow sensing [37,38] and allows for direct measurement and on-board panoramic use of this visual cue [40,41].

In this article, we propose a nonstereoscopic vision-based model of collective behavior, inspired by animal vision. This model is intended to be implementable on autonomous robots equipped with visual sensors, i.e., the robots do not need to communicate with a central command or/and to know their neighbors' relative coordinates, their relative positions between each other or to be georeferenced. Our model aims at bridging the gap between traditional SPP models that rely on omniscient information and biomimetic visual approaches.

II. MODEL

We consider a system of N self-propelled particles in two dimensions. Each particle is a circular object with radius a, moving with a constant speed U. The position of the ith

---

[*] franck.ruffier@cnrs.fr

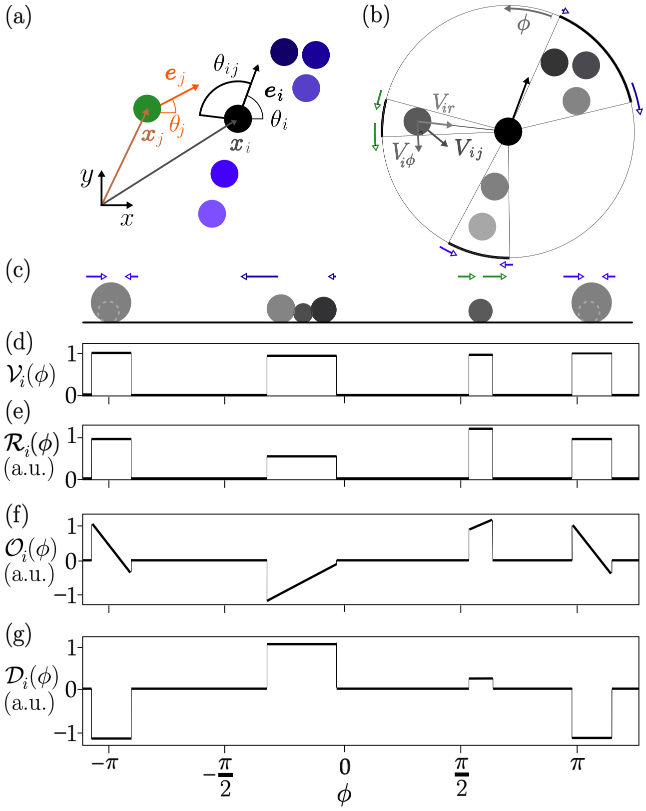

FIG. 1. Notations used and principle of the visual cues. (a) The particle i is located in $\boldsymbol{x}_i$ and θ_i is its heading. θ_{ij} is the retinal position of particle j for the ith particle. (b) The angle ϕ is the line of sight with respect to the particle's heading. The vector $\boldsymbol{V}_{ij} = U(\boldsymbol{e}_j - \boldsymbol{e}_i)$ is the relative velocity of particle j perceived by i. It can be decomposed into radial and azimuthal components, V_{ir} and $V_{i\phi}$. (c) Point of view of the particle i, the arrows represent the optic flow generated by the relative angular velocity of each particle. (d) The visual field $\mathcal{V}_i(\varphi)$ is a binary representation of (c). It is formed of a set of "shades." (e) The function $\mathcal{R}_i(\phi)$ represents the estimated distance associated to shades, assuming they are associated with a single individual of radius a. (f) Optic flow $\mathcal{O}_i(\phi)$ associated with the apparent angular velocity of the shade edges. (g) Optic flow divergence $\mathcal{D}_i(\phi)$.

particle is noted $\boldsymbol{x}_i$ and its direction $\boldsymbol{e}_i = [\cos\theta_i, \sin\theta_i]$, with θ_i its heading [Fig. 1(a)].

We model the interactions between particles using changes in their angular velocity associated with attraction, alignment, and noise. Specifically, the equations of motion can be written as follows:

$$\dot{\boldsymbol{x}}_i = U\boldsymbol{e}_i, \qquad (1a)$$

$$\dot{\theta}_i = k_\odot\,\omega_\odot + k_\parallel\,\omega_\parallel + k_\eta\,\eta, \qquad (1b)$$

where dots denote temporal derivatives and $\eta(t)$ is a standard Wiener process representing rotational noise. The functions $\omega_\odot$ and $\omega_\parallel$ are $O(1)$ functions representing attraction and alignment. The parameters $k_\odot$, $k_\parallel$, k_η control the strength of attraction, alignment, and noise. For simplicity, we did not include an avoidance rule, as it is not a required rule to reproduce collective behaviors [2,12,42].

We begin by introducing an omniscient model that will serve as a reference. This model is inspired by a data-driven fish model [43,44], with the difference being that each particle interacts with all the others. The attraction and alignment terms are given by

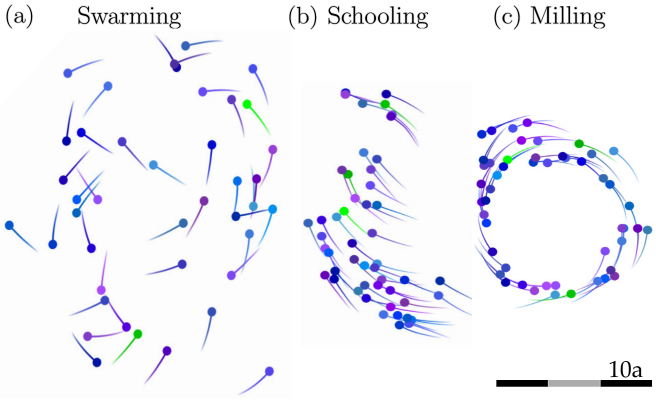

FIG. 2. Illustration of the phases observed for $N = 50$ individuals. The scale bar measures $30a$ in total ($10a$ for each section). The parameters $\epsilon = 1$ and $k_\eta = 0.01$ are fixed. We observe three phases when varying the other two parameters: (a) swarming ($k_\odot = 0.1$, $k_\parallel = 0$); (b) schooling ($k_\odot = 0.06$, $k_\parallel = 0.2$); and (c) milling ($k_\odot = 0.1$, $k_\parallel = 0.04$).

$$\omega_\odot^{\text{omni.}} = \left\langle \sum_{j=1}^{N} \|\boldsymbol{x}_j - \boldsymbol{x}_i\| \sin(\theta_{ij}) b_\epsilon(\theta_{ij}) \right\rangle, \qquad (2a)$$

$$\omega_\parallel^{\text{omni.}} = \left\langle \sum_{j=1}^{N} \frac{\boldsymbol{e}_i \times \boldsymbol{e}_j}{\|\boldsymbol{x}_j - \boldsymbol{x}_i\|^2} b_\epsilon(\theta_{ij}). \right\rangle, \qquad (2b)$$

where $b_\epsilon(\phi) = 1 + \epsilon\cos\phi$ models the blind angle (see also Fig. S1 [45]). When $\epsilon = 0$, b_ϵ is isotropic; when $\epsilon = 1$, the particle cannot see behind itself [45]. The brackets denote a normalization defined as

$$\left\langle \sum_j f(j)\sin\theta_j \right\rangle = \sum_j f(j)\sin\theta_j \Big/ \sum_j |f(j)|. \qquad (3)$$

To model visual perception, we assume that each particle senses a visual field $\mathcal{V}_i(\phi)$, where ϕ represents the angle between the particle's heading and the line of sight [Fig. 1(c)]. The function output is binary, indicating the presence or absence of a shade in the visual field [Figs. 1(b)–1(d)].

Using the information from the visual field $\mathcal{V}_i(\phi)$, we can derive the function $\mathcal{R}_i(\phi) = a/\sin(\Delta\phi/2)$, where $\Delta\phi$ represents the angle of view angle of shades [Fig. 1(e)]. With this definition, when a shade is associated with a single particle, $\mathcal{R}_i(\phi)$ represents its distance from the viewer. Temporal changes in the features of the visual field can be used to calculate the optic flow $\mathcal{O}_i(\phi)$. A simplified optic flow is estimated by assuming that each shade has a pattern that moves and deforms with it. It results that $\mathcal{O}_i(\phi)$ is a linearly interpolated function of the angular velocity between two features of the shade, its rising and falling edges [Fig. 1(f)]. This method to compute the optic flow computation corresponds to a cross correlation of visual field features [46], which are known to occur in animal eyes [47,48]. Similarly, we can compute the optic flow divergence $\mathcal{D}_i(\phi)$ by derivating the optic flow $\mathcal{O}_i(\phi)$

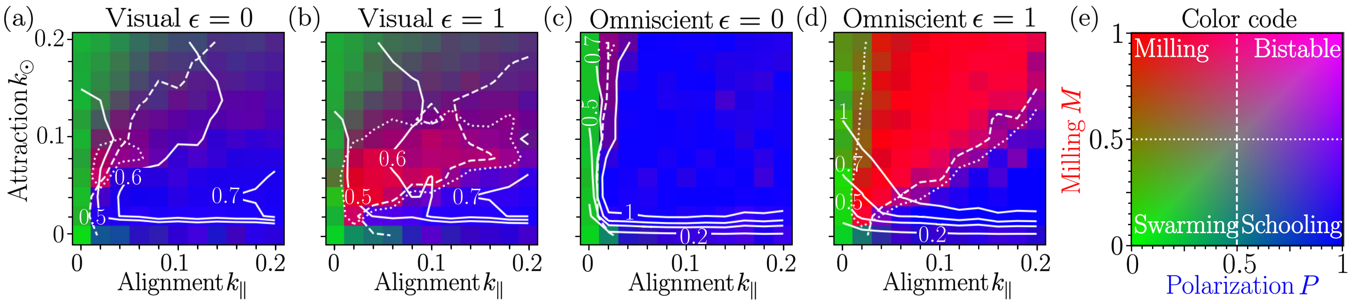

FIG. 3. Phase diagrams for $N = 50$ individuals and noise strength $k_\eta = 0.01$. These phase diagrams compare the visual model (a), (b) and the omniscient model (c), (d) for $\epsilon = 0$ (a), (c) and $\epsilon = 1$ (b), (d). The colors represent different values of P and M as shown in (e). The contours show the values of opacity O (white solid), the line $M = 0.5$ (white dotted), and $P = 0.5$ (white dashed).

[Fig. 1(f)]. Due to the piece-wise linear nature of $\mathcal{O}_i(\phi)$, $\mathcal{D}_i(\phi)$ is a piece-wise constant function.

The optic functions $\mathcal{V}$, $\mathcal{R}$, $\mathcal{O}$, and $\mathcal{D}$ are inspired by animal vision. These functions can easily be computed by a man-made vision system. We will now use these functions for the attraction and alignment terms of a vision model. In this visual model, the attraction and alignment terms are given by

$$\omega_{\odot}^{\text{visu.}} = \left\langle \int_{-\pi}^{\pi} \mathcal{R}_i^2(\phi) b_\epsilon(\phi) \sin\phi \, d\phi \right\rangle, \tag{4a}$$

$$\omega_{\parallel}^{\text{visu.}} = \left\langle \int_{-\pi}^{\pi} \frac{\boldsymbol{e}_i \times \boldsymbol{V}_{ij}}{U \mathcal{R}_i(\phi)} b_\epsilon(\phi) \, d\phi. \right\rangle, \tag{4b}$$

where the brackets denote the normalization given by Eq. (3) with the sum replaced by the integral. The attraction and alignment terms of the visual model are constructed to be similar to those of the omniscient model given in Eqs. (2a) and (2b). The difference in the exponent of $\mathcal{R}$ comes from the additional $\Delta\phi \sim 1/\mathcal{R}$ arising from the integration.

In the alignment term, $\omega_{\parallel}^{\text{visu.}}$, the cross product $\boldsymbol{e}_i \times \boldsymbol{e}_j$ is evaluated from the visual information. This is done by using the optic flow $\mathcal{O}_i(\phi)$ and its divergence $\mathcal{D}_i(\phi)$, which are related to the velocity of particle j with respect to particle i [Fig. 1(b)]. Specifically, the radial component is given by $V_{ir} = -\mathcal{R}_i(\phi)\ \mathcal{D}_i(\phi)$, and the azimuthal component by $V_{i\phi} = \mathcal{R}_i(\phi)\ \mathcal{O}_i(\phi)$. We can use these components to calculate the vector $\boldsymbol{V}_{ij}$ in polar coordinates (r, ϕ) as $\boldsymbol{V}_{ij} = (-\mathcal{D}_i, \mathcal{O}_i)\mathcal{R}_i/U$, while $\boldsymbol{e}_i$ can be expressed as $(\cos\phi, -\sin\phi)$. It results that

$$\frac{\boldsymbol{e}_i \times \boldsymbol{V}_{ij}}{U \mathcal{R}_i(\phi)} = \frac{-\mathcal{D}_i(\phi)\sin\phi + \mathcal{O}_i(\phi)\cos\phi}{U}. \tag{5}$$

The alignment term is thus the sum of two terms: one proportional to the derotated optic flow $\mathcal{O}$ and sensitive to the azimuthal velocity of neighbors, and one proportional to the optic flow divergence $\mathcal{D}$ and sensitive to the radial velocity. When computing $\boldsymbol{e}_i \times \boldsymbol{V}_{ij}$ with Eq. (5), we remove the particle rotation with angular velocity $\dot{\theta}_i$ from the optic flow $\mathcal{O}_i(\phi)$.

III. NUMERICAL SIMULATIONS

The equations of motion presented in Eq. (1), along with the attraction and alignment terms derived in Eqs. (4a) and (4b), provide a model of collective behavior based on realistic visual cues. To make the problem dimensionless, we chose $a = 1$ and $U = 1$. With this approach, four dimensionless parameters remain, the strengths of noise attraction and alignment, k_η, $k_\odot$, $k_\parallel$, and the blind angle parameter ϵ.

To explore the effect of these parameters on collective behaviors, we performed numerical simulations with $N = 50$ particles. Initially, the particles are randomly placed in a square of side aN with random headings [45,49]. The dynamical system described by Eqs. (1a) and (1b) is solved numerically using a discrete implementation of Eqs. (4a) and (4b) (see also Figs. S2-S3 [45]). We examined the effect of the time step δt by conducting simulations with $\delta t = 0.001$, 0.01, and 0.1 (Figs. S17–S22 and Videos S1-S6 [45]). However, no significant differences were observed on the collective behavior, and a time step of $\delta t = 0.1$ was selected for the remaining simulations to ensure computational efficiency.

We first set $k_\eta = 0.01$ and $\epsilon = 1$ (maximum blind angle) and explore the effects of the two remaining parameters $k_\odot$ and $k_\parallel$ in the visual model. Our simulations show three distinct dynamical phases (Fig. 2 for the visual model, Fig. S10 for the omniscient model, and Figs. S23–S28 snapshots for the evolution of the stable phases seen on Videos S7-S9 and S10-S12, [45]). If the alignment is zero, a disordered swarming phase is observed, where individuals form a group without a preferred direction [Fig. 2(a)]. When the alignment strength increases, particles begin to align in the same direction, resulting in the schooling phase [Fig. 2(b)]. If the ratio between the alignment and the attraction strengths is around 0.4, the group exhibits a milling phase [Fig. 2(c)], creating a vortex. These three phases (swarming, schooling, and milling) have regularly been observed in (omniscient) self-propelled-particle models [43,44,50–52].

To quantitatively distinguish between the different dynamical phases, we introduce three global order metrics: polarization P, milling M, and opacity O [43,44]. These metrics are defined as follows:

$$P = \|\overline{\boldsymbol{e}_i}\|, \tag{6a}$$

$$M = \|\overline{\boldsymbol{y}_i \times \boldsymbol{e}_i}\|, \tag{6b}$$

$$O = \overline{\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathcal{V}_i(\phi) d\phi}, \tag{6c}$$

where the overbar represents an average over all individuals and the unit vector $\boldsymbol{y}_i = (\boldsymbol{x}_i - \overline{\boldsymbol{x}_i})/\|\boldsymbol{x}_i - \overline{\boldsymbol{x}_i}\|$ points toward

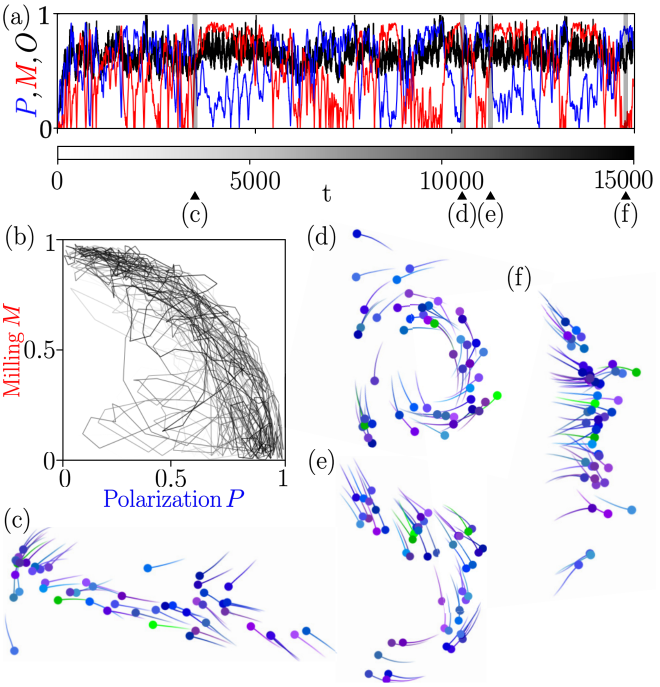

FIG. 4. Example of the bistability observed in the visual model (parameter values: $\epsilon = 1$, $k_{\odot} = 0.1$, $k_{\parallel} = 0.2$, and $k_{\eta} = 0.01$). (a) Time series of the three metrics P, M, and O. (b) Trajectories of the collective behavior in the (M, P) plane. Illustration of the phases: (c) schooling-milling transition; (d) milling; (e) milling-schooling transition; and (f) schooling.

particle i from the center of mass. All three metrics range in the interval [0, 1]. The polarization P measures the alignment: $P = 0$ corresponds to particles pointing in all directions, $P = 1$ corresponds to a perfectly aligned school. The milling M represents the normalized angular momentum: straight-line formation gives $M = 0$ and perfect milling gives $M = 1$. The opacity O measures the “occupancy” of the visual fields: $O = 0$ when there is no object in the visual field, and $O = 1$ when the entire visual field is obscured.

We now compare the visual and omniscient models by setting the value of the noise to $k_{\eta} = 0.01$ and exploring the parameter space $(k_{\odot}, k_{\parallel}) \in [0, 0.2] \times [0, 0.2]$ for two values of the blind angle parameter ($\epsilon = 0$ or 1). For each parameter set, we ran ten simulations over long durations ($\Delta t = 5000$). The mean values of P, M, and O were determined by averaging over the last 1000 time units to ensure that the transient has no influence. The outcomes of these simulations are synthesized in phase diagrams (Fig. 3, and expanded on each metric independently on Figs. S4, S6–S9, S11, S13-S16 [45]).

Somewhat arbitrarily, we chose to identify the collective phases from the values of the polarization P and the milling M parameters: schooling when $P > 0.5$ and $M < 0.5$; milling when $P < 0.5$ and $M > 0.5$; swarming when $P < 0.5$ and $M < 0.5$; and bistable when $P > 0.5$ and $M > 0.5$ (we will come back to this particular phase below).

IV. DISCUSSION

Our vision model qualitatively reproduces the phases observed in the omniscient model (Fig. 3). Specifically, we observe the three phases in the vision model: schooling when $k_{\odot} \lesssim 0.5 k_{\parallel}$; swarming when $k_{\parallel} \lesssim 0.3 k_{\odot}$; and milling or bistability otherwise. When the blind angle parameter increases, it tends to stabilize the milling phase both in the visual model and the omniscient model, as observed in the literature [43].

The difference between the two models increases at high values of the alignment and attraction strengths when the opacity O is maximum. In the visual model, the opacity does not exceed 0.7, regardless of the values of ϵ and k_{η}. In the omniscient model, however, the opacity exceeds 0.8 and even reaches one [Figs. 3(c) and 3(d)]. This is because the radius a does not play any role in the omniscient model and the characteristic length is set by k_{η}^2. In both models, as expected, larger noise strength causes a decrease in average opacity (Figs. S4 and S11) [45].

Now, let us examine the bistable phase. Figure 4 and S29 show this phase in the visual model for $\epsilon = 1$, $k_{\odot} = 0.1$, $k_{\parallel} = 0.2$, and $k_{\eta} = 0.01$. After a transient, the group forms a milling phase until $t \approx 6000$, but eventually it transitions to a schooling phase for $6000 \lesssim t \lesssim 7500$ before returning to a schooling phase again, and so on. These transitions show that the system exhibits a noise-induced intermittency between two stable states: milling and schooling. The schooling phase far from the transition resembles a front of parallel individuals [Fig. 4(f)]. Just before the transition to milling, some individuals move ahead [Fig. 4(c)]. Reciprocally, the milling phase just before the transition to schooling opens up, generating a C shape [Fig. 4(e)]. These transitions are fairly stereotyped as they tend to follow the same path in the (M, P) plane [Fig. 4(b)]. The existence of these intermittent transitions mediated by noise is likely due to a second-order transition between the milling and schooling phases as seen on the metric distribution along the boundaries of both phases [Figs. S6–S9 and S13–S16]. This multistability has been present on similar 3A models [43].

We changed the group size from $N = 5$ to 300 (Figs. S5 and S12 [45]). Although the group size does not seem to impact qualitatively the phase diagram, small groups tend to favor schooling, whereas large groups tend to favor swarming.

V. CONCLUSIONS

In conclusion, we proposed a model based on biologically plausible visual cues. This model successfully reproduces the three classical phases of animal collective behavior: swarming, schooling, and milling. These findings show that visual cues provide enough information to enable collective behavior.

Furthermore, our findings imply potential practical uses for synchronizing groups of artificial drones, which may be governed by analogous visual stimuli. In future studies, we aim to investigate these opportunities more thoroughly and enhance our model for a more accurate depiction of animal collective behavior in a three-dimensional space.

[1] B. Liebchen and D. Levis, Collective behavior of chiral active matter: Pattern formation and enhanced flocking, Phys. Rev. Lett. **119**, 058002 (2017).

[2] J.-B. Caussin, A. Solon, A. Peshkov, H. Chaté, T. Dauxois, J. Tailleur, V. Vitelli, and D. Bartolo, Emergent spatial structures in flocking models: A dynamical system insight, Phys. Rev. Lett. **112**, 148102 (2014).

[3] L. Barberis and F. Peruani, Large-scale patterns in a minimal cognitive flocking model: Incidental leaders, nematic patterns, and aggregates, Phys. Rev. Lett. **117**, 248001 (2016).

[4] J. L. Silverberg, M. Bierbaum, J. P. Sethna, and I. Cohen, Collective motion of humans in mosh and circle pits at heavy metal concerts, Phys. Rev. Lett. **110**, 228701 (2013).

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

[6] C. W. Reynolds, Flocks, herds and schools: a distributed behavioral model, in *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques* (ACM, New York, 1987), pp. 25–34.

[7] F. Ginelli and H. Chaté, Relevance of metric-free interactions in flocking phenomena, Phys. Rev. Lett. **105**, 168103 (2010).

[8] E. V. Albano, Self-organized collective displacements of self-driven individuals, Phys. Rev. Lett. **77**, 2129 (1996).

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

[10] S. Mishra, A. Baskaran, and M. C. Marchetti, Fluctuations and pattern formation in self-propelled particles, Phys. Rev. E **81**, 061916 (2010).

[11] A. P. Solon, J.-B. Caussin, D. Bartolo, H. Chaté, and J. Tailleur, Pattern formation in flocking models: A hydrodynamic description, Phys. Rev. E **92**, 062111 (2015).

[12] T. Vicsek and A. Zafeiris, Collective motion, Phys. Rep. **517**, 71 (2012).

[13] C. Virágh, G. Vásárhelyi, N. Tarcai, T. Szörényi, G. Somorjai, T. Nepusz, and T. Vicsek, Flocking algorithm for autonomous flying robots, Bioinspir. Biomim. **9**, 025012 (2014).

[14] X. Wang and J. Lu, Collective behaviors through social interactions in bird flocks, IEEE Circuits Syst. Mag. **19**, 6 (2019).

[15] Y. Shen and C. Wei, Multi-UAV flocking control with individual properties inspired by bird behavior, Aerosp. Sci. Technol. **130**, 107882 (2022).

[16] N. W. Bode, J. J. Faria, D. W. Franks, J. Krause, and A. J. Wood, How perceived threat increases synchronization in collectively moving animal groups, Proc. R. Soc. **277**, 3065 (2010).

[17] D. Mateo, N. Horsevad, V. Hassani, M. Chamanbaz, and R. Bouffanais, Optimal network topology for responsive collective behavior, Sci. Adv. **5**, eaau0999 (2019).

[18] E. Soria, F. Schiano, and D. Floreano, The influence of limited visual sensing on the Reynolds flocking algorithm, in *2019 Third IEEE International Conference on Robotic Computing (IRC)* (IEEE, Piscataway, NJ, 2019), pp. 138–145.

[19] J. D. Davidson, M. M. Sosna, C. R. Twomey, V. H. Sridhar, S. P. Leblanc, and I. D. Couzin, Collective detection based on visual information in animal groups, J. R. Soc. Interface. **18**, 20210142 (2021).

[20] A. Strandburg-Peshkin, C. R. Twomey, N. W. Bode, A. B. Kao, Y. Katz, C. C. Ioannou, S. B. Rosenthal, C. J. Torney, H. S. Wu, S. A. Levin *et al.*, Visual sensory networks and effective information transfer in animal groups, Curr. Biol. **23**, R709 (2013).

[21] B. H. Lemasson, J. J. Anderson, and R. A. Goodwin, Motion-guided attention promotes adaptive communications during social navigation, Proc. R. Soc. B. **280**, 20122003 (2013).

[22] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini *et al.*, Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study, Proc. Natl. Acad. Sci. USA **105**, 1232 (2008).

[23] B. H. Lemasson, J. J. Anderson, and R. A. Goodwin, Collective motion in animal groups from a neurobiological perspective: The adaptive benefits of dynamic sensory loads and selective attention, J. Theor. Biol. **261**, 501 (2009).

[24] B. Collignon, A. Séguret, and J. Halloy, A stochastic vision-based model inspired by zebrafish collective behaviour in heterogeneous environments, R. Soc. Open Sci. **3**, 150473 (2016).

[25] J. Qi, L. Bai, Y. Xiao, Y. Wei, and W. Wu, The emergence of collective obstacle avoidance based on a visual perception mechanism, Inf. Sci. **582**, 850 (2022).

[26] I. Aoki, Internal dynamics of fish schools in relation to inter-fish distance, Bull. Jap. Soc. Sci. Fish. **50**, 751 (1984).

[27] D. L. Krongauz, A. Ayali, and G. A. Kaminka, Vision-based collective motion: A locust-inspired reductionist model, PLoS Comput. Biol. **20**, e1011796 (2024).

[28] G. Ariel, Y. Ophir, S. Levi, E. Ben-Jacob, and A. Ayali, Individual pause-and-go motion is instrumental to the formation and maintenance of swarms of marching locust nymphs, PLoS ONE **9**, e101636 (2014).

[29] D. J. Pearce, A. M. Miller, G. Rowlands, and M. S. Turner, Role of projection in the control of bird flocks, Proc. Natl. Acad. Sci. USA **111**, 10422 (2014).

[30] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, A. Orlandi, G. Parisi, A. Procaccini, M. Viale *et al.*, Empirical investigation of starling flocks: a benchmark study in collective animal behaviour, Animal Behaviour **76**, 201 (2008).

[31] R. Bastien and P. Romanczuk, A model of collective behavior based purely on vision, Sci. Adv. **6**, eaay0792 (2020).

[32] G. Portelli, F. Ruffier, and N. Franceschini, Honeybees change their height to restore their optic flow, J. Comp. Physiol. A **196**, 307 (2010).

[33] N. Franceschini, F. Ruffier, and J. Serres, A bio-inspired flying robot sheds light on insect piloting abilities, Curr. Biol. **17**, 329 (2007).

[34] K. Bonnen, Motion vision: Fish swimming to see, Curr. Biol. **33**, R30 (2023).

[35] J. R. Serres, T. J. Evans, S. Åkesson, O. Duriez, J. Shamoun-Baranes, F. Ruffier, and A. Hedenström, Optic flow cues help explain altitude control over sea in freely flying gulls, J. R. Soc. Interface **16**, 20190486 (2019).

[36] F. Ruffier and N. Franceschini, Optic flow regulation: the key to aircraft automatic guidance, Rob. Auton. Syst. **50**, 177 (2005).

[37] F. Expert and F. Ruffier, Flying over uneven moving terrain based on optic-flow cues without any need for reference frames or accelerometers, Bioinspir. Biomim. **10**, 026003 (2015).

[38] D. Floreano, R. Pericet-Camara, S. Viollet, F. Ruffier, A. Brückner, R. Leitel, W. Buss, M. Menouni, F. Expert, R. Juston

et al., Miniature curved artificial compound eyes, Proc. Natl. Acad. Sci. USA **110**, 9267 (2013).

[39] F. Schilling, E. Soria, and D. Floreano, On the scalability of vision-based drone swarms in the presence of occlusions, IEEE Access **10**, 28133 (2022).

[40] L. Bergantin, N. Harbaoui, T. Raharijaona, and F. Ruffier, Oscillations make a self-scaled model for honeybees' visual odometer reliable regardless of flight trajectory, J. R. Soc. Interface **18**, 20210567 (2021).

[41] G. C. De Croon, J. J. Dupeyroux, C. De Wagter, A. Chatterjee, D. A. Olejnik, and F. Ruffier, Accommodating unobservability to control flight attitude with optic flow, Nature (London) **610**, 485 (2022).

[42] A. Morin, J.-B. Caussin, C. Eloy, and D. Bartolo, Collective motion with anticipation: Flocking, spinning, and swarming, Phys. Rev. E **91**, 012134 (2015).

[43] D. S. Calovi, U. Lopez, S. Ngo, C. Sire, H. Chaté, and G. Theraulaz, Swarming, schooling, milling: phase diagram of a data-driven fish school model, New J. Phys. **16**, 015026 (2014).

[44] A. Filella, F. Nadal, C. Sire, E. Kanso, and C. Eloy, Model of collective fish behavior with hydrodynamic interactions, Phys. Rev. Lett. **120**, 198101 (2018).

[45] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.6.023016 for additional mathematical proof, numerical implementation, optic flow algorithm, several in detail figures, and videos.

[46] S. Ullman, Analysis of visual motion by biological and computer systems, Computer **14**, 57 (1981).

[47] N. Franceschini, A. Riehle, and A. Le Nestour, Directionally selective motion detection by insect neurons, in *Facets of Vision* (Springer, Berlin, Heidelberg, 1989), pp. 360–390.

[48] H. Eichner, M. Joesch, B. Schnell, D. F. Reiff, and A. Borst, Internal structure of the fly elementary motion detector, Neuron **70**, 1155 (2011).

[49] D. Castro, F. Ruffier, and C. Eloy, Source Code for the Flocking by Eye Project, Gitlab (Accessed 25-Feb-2024), https://gitlab.com/FlokingByEyeDiego/flokingbyeye.

[50] H. J. Charlesworth and M. S. Turner, Intrinsically motivated collective motion, Proc. Natl. Acad. Sci. USA **116**, 15362 (2019).

[51] X. Lu, C. Zhang, and B. Qin, An improved Vicsek model of swarm based on remote neighbors strategy, Physica A **587**, 126553 (2022).

[52] Y. Shang and R. Bouffanais, Influence of the number of topologically interacting neighbors on swarm dynamics, Sci. Rep. **4**, 4184 (2014).