

# Modeling collective behaviors from optic flow and retinal cues

Diego Castro,<sup>1,2</sup> Franck Ruffier,<sup>1,\*</sup> and Christophe Eloy<sup>2</sup>

<sup>1</sup>*Aix Marseille Université, CNRS, ISM, Marseille, 13288, France*

<sup>2</sup>*Aix Marseille Université, CNRS, Centrale Med, IRPHE, Marseille, 13013, France*

(Dated: May 22, 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 bio-plausible 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.

## 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 bio-plausible 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 bio-inspired 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 non-stereoscopic 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  $i$ th particle is noted  $\mathbf{x}_i$  and its direction  $\mathbf{e}_i = [\cos \theta_i, \sin \theta_i]$ , with  $\theta_i$  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

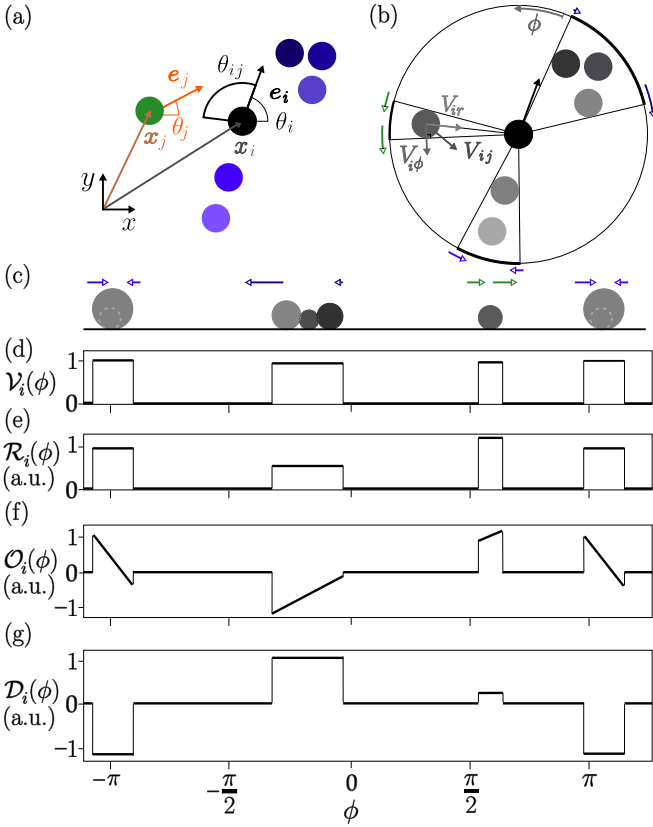


FIG. 1. Notations used and principle of the visual cues. (a) The particle  $i$  is located in  $\mathbf{x}_i$  and  $\theta_i$  is its heading.  $\theta_{ij}$  is the retinal position of particle  $j$  for the  $i$ -th particle. (b) The angle  $\phi$  is the line of sight with respect to the particle's heading. The vector  $\mathbf{V}_{ij} = U(\mathbf{e}_j - \mathbf{e}_i)$  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  $\mathcal{V}_i(\phi)$  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  $\mathcal{D}_i(\phi)$ .

motion can be written as follows

$$\dot{\mathbf{x}}_i = U\mathbf{e}_i, \quad (1a)$$

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

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  $k_{\odot}$ ,  $k_{\parallel}$ ,  $k_{\eta}$  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 that each particle interacts with all the others. The attraction and alignment terms are given by

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

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

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

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

To model visual perception, we assume that each particle senses a visual field  $\mathcal{V}_i(\phi)$ , where  $\phi$  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 [Fig. 1(b)-(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. 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 to the attraction and alignment terms of a vision model.

In this visual model, the attraction and alignment terms are given by

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

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

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)-(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  $\mathbf{e}_i \times \mathbf{e}_j$  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  $\mathbf{V}_{ij}$  in polar coordinates  $(r, \phi)$  as  $\mathbf{V}_{ij} = (-\mathcal{D}_i, \mathcal{O}_i)\mathcal{R}_i/U$ , while  $\mathbf{e}_i$  can be expressed as  $(\cos\phi, -\sin\phi)$ . It results that

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

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  $\mathbf{e}_i \times \mathbf{V}_{ij}$  with Eq. (5), we remove the particle rotation with angular velocity  $\theta_i$  from the optic flow  $\mathcal{O}_i(\phi)$ .

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,  $k_{\eta}$ ,  $k_{\odot}$ ,  $k_{\parallel}$ , and the blind angle parameter  $\epsilon$ .

### III. NUMERICAL SIMULATIONS

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 (4a) (See also Figs. S2-S3, [45]). We examined the effect of the time step  $\delta t$  by conducting simulations with  $\delta t = 0.001, 0.01, \text{ and } 0.1$  (Figs. S17–S22 and Video 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. 6 for the visual model, Fig. S10 for the omniscient model and

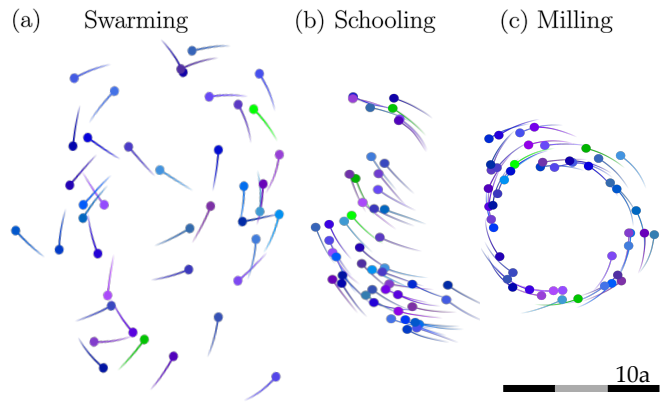


FIG. 2. Illustration of the phases observed for  $N = 50$  individuals. The scale bar measures  $30a$  in total ( $10a$  for each section). The parameters  $\epsilon = 1$  and  $k_{\eta} = 0.01$  are fixed. We observe three phases when varying the other two parameters: (a) swarming ( $k_{\odot} = 0.1,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

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. 6(a)]. When the alignment strength increases, particles begin to align in the same direction, resulting in the schooling phase [Fig. 6(b)]. If the ratio between the alignment and the attraction strengths is around 0.4, the group exhibits a milling phase [Fig. 6(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,

$$P = \|\overline{\mathbf{e}_i}\|, \quad (6a)$$

$$M = \|\overline{\mathbf{y}_i \times \mathbf{e}_i}\|, \quad (6b)$$

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

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

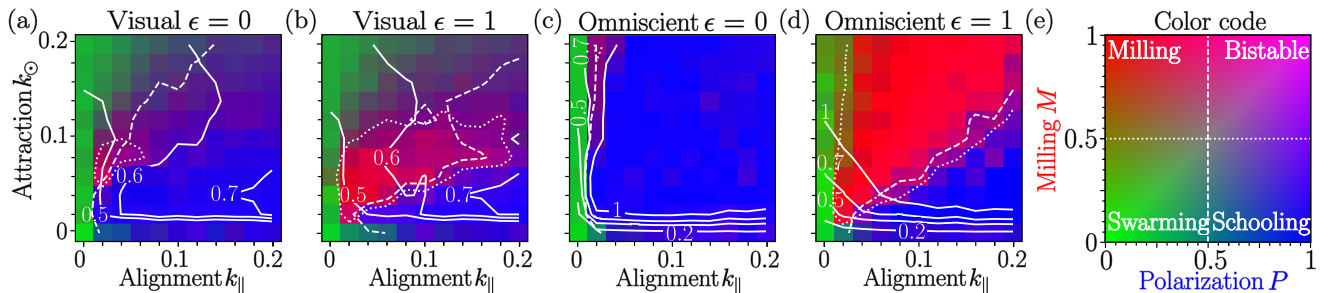


FIG. 3. Phase diagrams for  $N = 50$  individuals and noise strength  $k_\eta = 0.01$ . These phase diagrams compare the visual model (a), (b) and the omniscient model (c), (d) for  $\epsilon = 0$  (a), (c) and  $\epsilon = 1$  (b), (d). The colors represent different values of  $P$  and  $M$  as shown 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).

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 10 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  $k_\odot \lesssim 0.5k_\parallel$ ; swarming when  $k_\parallel \lesssim 0.3k_\odot$ ; 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  $\epsilon$  and  $k_\eta$ . In the omniscient model, however, the opacity exceeds 0.8 and even reaches one [Fig. 3(c) and (d)]. This is because the radius  $a$  does not play any role in the omniscient model and the characteristic length is set by  $k_\eta^2$ . 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. Figures 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  $t \approx 6000$ ,

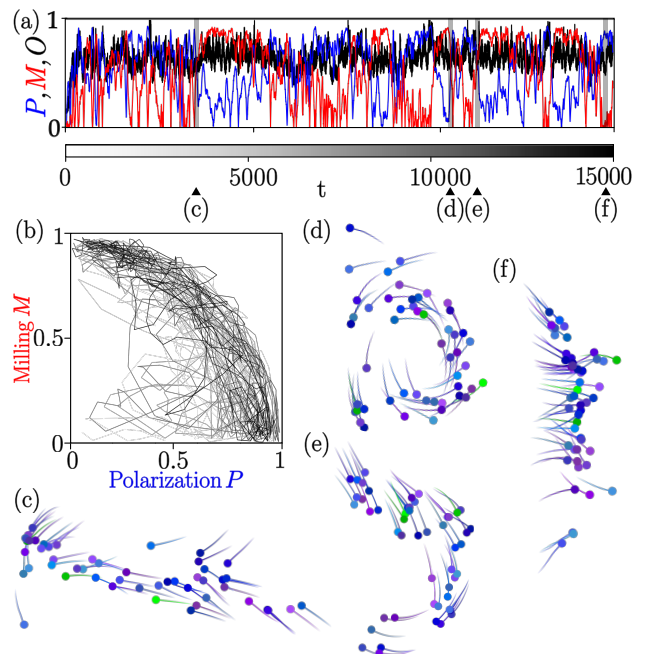


FIG. 4. Example of the bistability observed in the visual model (parameter values:  $\epsilon = 1$ ,  $k_\odot = 0.1, k_\parallel = 0.2$ , and  $k_\eta = 0.01$ ). (a) Time series of the three metrics  $P$ ,  $M$ , and  $O$ . (b) Trajectories of the collective behavior in the  $(M, P)$ -plane. Illustration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

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 multi-stability 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. CONCLUSION

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.

---

\* franck.ruffier@cnr.fr

- [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* (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, *Proceedings of the National Academy of Sciences* **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, *Physics reports* **517**, 71 (2012).
- [13] C. Virág, 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 and Systems Magazine* **19**, 6 (2019).
- [15] Y. Shen and C. Wei, Multi-uav flocking control with individual properties inspired by bird behavior, *Aerospace Science and Technology* **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**, 3065–3070 (2010).
- [17] D. Mateo, N. Horsevad, V. Hassani, M. Chamanbaz, and R. Bouffanais, Optimal network topology for responsive collective behavior, *Science advances* **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, 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, *Journal of the Royal Society 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, *Current Biology* **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, *Proceedings of the national academy of sciences* **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, *Journal of theoretical biology* **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, *Information Sciences* **582**, 850 (2022).
- [26] I. Aoki, Internal dynamics of fish schools in relation to

- inter-fish distance\*1, Bulletin of the Japanese Society of Scientific Fisheries **50**, 751 (1984).
- [27] D. L. Krongauz, A. Ayali, and G. A. Kaminka, Vision-based collective motion: a locust-inspired reductionist model, PLOS Computational Biology **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. **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, Science advances **6**, eaay0792 (2020).
- [32] G. Portelli, F. Ruffier, and N. Franceschini, Honeybees change their height to restore their optic flow, Journal of Comparative Physiology 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, Current Biology **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, Robotics and Autonomous Systems **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. **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.
- [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 journal of Physics **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 the Supplemental Material after the bibliography for Additional mathematical proof, numerical implementation, optic flow algorithm, several detailed 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, 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, <https://gitlab.com/FlokingByEyeDiego/> (Accessed 25-Feb-2024)
- [50] H. J. Charlesworth and M. S. Turner, Intrinsically motivated collective motion, Proc. Natl. Acad. Sci. **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**, 1 (2014).

**Supplemental Material for**  
**“Modeling collective behaviors from optic flow**  
**and retinal cues”**  
**D. Castro, F. Ruffier, and C. Eloy**

**CONTENTS**

S1 Mathematical development	7
S1.1 Visual and Omniscient attraction equivalence proof	7
S1.2 Normalization for non-trigonometric equation	7
S1.3 Equivalence of the visual and omniscient model for the alignment term	7
S1.4 Blind spot	8
S2 Digital implementation	8
S2.1 Omniscient model	8
S2.2 Visual model	8
S2.3 Algorithm to detect corresponding features necessary to compute optic flow	9
S3 Detailed Phase Diagrams (Fig. 3)	10
S3.1 Visual Model	10
S3.2 Omniscient Model	11
S4 Video Snapshots of Time step stability analysis	20
S4.1 Visual model	20
S4.2 Omniscient model	23
S5 Video Snapshots of Phases Reproduction (Fig. 2) ( $k_\eta = 0.01, N = 50$ )	26
S5.1 Visual Model	26
S5.2 Omniscient Model	29
S6 Video Snapshots of Bistable Phase Visual model (Fig. 4)	32
Supplemental References	32

**S1 MATHEMATICAL DEVELOPMENT**

**S1.1 Visual and Omniscient attraction equivalence proof**

We defined the attraction rule for the omniscient and visual models to be equivalent one to the other. But changing the discrete sum of the omniscient model into an integral needed for the visual model, requires to multiply the integrand by  $\mathcal{R}_i(\phi)$ .

Here we present a step by step calculation of this integral to show the equivalence with the discrete sum. We start by the integral of Eq. (4a),

$$\omega_{\odot}^{\text{visu.}} = \left\langle \int_{-\pi}^{\pi} \mathcal{R}_i^2(\phi) b_\epsilon(\phi) \sin \phi d\phi, \right\rangle \quad (\text{S1})$$

where  $b_\epsilon(\phi) = 1 + \epsilon \cos \phi$ .

$\mathcal{V}_i(\phi)$  is a piece-wise constant function that can be expressed as a set unit boxcar functions (a shade). Each, can be described as a rising  $\varphi_{\uparrow k}$  and its corresponding falling edge  $\varphi_{\downarrow k}$ . Furthermore, from these edges, we can compute the shade’s mid-point  $\varphi_k = (\varphi_{\uparrow k} + \varphi_{\downarrow k})/2$ , and the shade’s half-width  $\Delta\varphi_k = (\varphi_{\downarrow k} - \varphi_{\uparrow k})/2$ .  $\mathcal{R}_i(\phi)$  is  $\mathcal{V}_i(\phi)$  weighted by  $\frac{a}{\sin \Delta\varphi_k}$ . This means that  $\mathcal{R}_i(\phi)$  for a binary  $\mathcal{V}_i(\phi)$ , is independent of  $\phi$  and is only be dependent on the  $k$ -shade’s half-width ( $\mathcal{R}_k(\Delta\varphi_k)$ ). The subscript changes from  $i$  to  $k$ , as  $i$  indicate is a value referred to continuous information,  $k$  to discrete information of a single shade. A shade  $k$  id different from an individual  $j$ , as in a shade can be occlusion or aggregation of any  $j$ -particles. Then, we can rewrite Eq. (S1) as:

$$\omega_{\odot}^{\text{visu.}} = \left\langle \sum_k \mathcal{R}_k^2(\Delta\varphi_k) \int_{\varphi_{\uparrow k}}^{\varphi_{\downarrow k}} b_\epsilon(\phi) \sin \phi d\phi, \right\rangle \quad (\text{S2})$$

$$\omega_{\odot}^{\text{visu.}} = \left\langle \sum_k 2\mathcal{R}_k^2(\Delta\varphi_k) \sin(\Delta\varphi_k) \sin(\varphi_k) (1 + \epsilon \cos \varphi_k \cos \Delta\varphi_k) \right\rangle \quad (\text{S3})$$

And here  $\mathcal{R}_k(\Delta\varphi_k) = \frac{a}{\sin \Delta\varphi_k}$ , Resulting in:

$$\omega_{\odot}^{\text{visu.}} = \left\langle \sum_k 2 \frac{a^2 \sin(\Delta\varphi_k)}{\sin^2(\Delta\varphi_k)} \sin(\varphi_k) (1 + \epsilon \cos \varphi_k \cos \Delta\varphi_k) \right\rangle \quad (\text{S4})$$

$$\omega_{\odot}^{\text{visu.}} = \left\langle \sum_k 2a\mathcal{R}_i(\Delta\varphi_k) \sin(\varphi_k) (1 + \epsilon \cos \varphi_k \cos \Delta\varphi_k) \right\rangle \quad (\text{S5})$$

For the case where there is no occlusion nor aggregation  $k = j$  and taking into account the physical constrains of the particles:  $\cos \Delta\varphi_k \approx 1$ ,  $\mathcal{R}_i(\Delta\varphi_k) = \|\mathbf{x}_j - \mathbf{x}_i\|$  and  $\varphi_k = \theta_{ij}$ , then Eq. (S5) becomes

$$\omega_{\odot}^{\text{visu.}} \approx \left\langle \sum_{j \neq i}^N 2a \|\mathbf{x}_j - \mathbf{x}_i\| \sin(\theta_{ij}) (b_\epsilon(\theta_{ij})) \right\rangle \quad (\text{S6})$$

Finally the  $2a$  factor is nullified by the  $\langle \cdot \rangle$  normalization. So it becomes identical to the attraction term in Eq. (2a), proving the equivalence between the two models.

**S1.2 Normalization for non-trigonometric equation**

We presented a normalization to guarantee the  $O(1)$  for all the terms of the models. Here we proof that the the non-trigonometric equation presented that are normalized by this equation, does in fact maintains the same form as shown on Eq. (3). We start by taking the alignment term of the omniscient model,

$$\omega_{\parallel}^{\text{omni.}} = \left\langle \sum_{j \neq i}^N \frac{\mathbf{e}_i \times \mathbf{e}_j}{\|\mathbf{x}_j - \mathbf{x}_i\|^2} b_{\epsilon}(\theta_{ij}) \cdot \right\rangle, \quad (\text{S7})$$

this cross product can be expressed as:

$$\mathbf{e}_i \times \mathbf{e}_j = U \|\mathbf{e}_i\| \|\mathbf{e}_j\| \sin(\theta_i - \theta_j) \quad (\text{S8})$$

Rewriting equation (2b) to better isolate the terms of the normalization described on equation (3)

$$\omega_{\parallel}^{\text{omni.}} = \left\langle \sum_{j=1}^N \frac{U \|\mathbf{e}_i\| \|\mathbf{e}_j\|}{\|\mathbf{x}_j - \mathbf{x}_i\|^2} b_{\epsilon}(\theta_{ij}) \sin(\theta_i - \theta_j) \right\rangle, \quad (\text{S9})$$

where,  $\mathbf{e}_i$  and  $\mathbf{e}_j$  are a unit vectors. Here  $\sin(\theta_j) = \sin(\theta_i - \theta_j)$  and  $f(j) = \frac{U \|\mathbf{e}_i\| \|\mathbf{e}_j\|}{\|\mathbf{x}_j - \mathbf{x}_i\|^2} b_{\epsilon}(\theta_{ij})$ . Which maintains the form of the normalization presented on Eq. (3).

### S1.3 Equivalence of the visual and omniscient model for the alignment term

Here we show the equivalence between visual and omniscient model for the alignment term if we consider that each shade is associated to a single particle.

The visual alignment terms is defined as:

$$\omega_{\parallel}^{\text{visu.}} = \sum_k \left( \int_{\varphi_{\uparrow k}}^{\varphi_{\downarrow k}} (\mathcal{O}_i(\phi) \cos \phi - \mathcal{D}_i(\phi) \sin \phi) b_{\epsilon}(\phi) d\phi \right) \quad (\text{S10})$$

On the previous section we discussed the piece-wise constant properties of  $\mathcal{R}_i(\phi)$ , the same is true for  $\mathcal{D}_i(\phi)$  and  $\mathcal{O}_i(\phi)$ , being independent of  $\phi$  Taking Eq. (S10) and Eq. (4b), we can express the alignment for the visual model as two components, a radial and a azimuthal to be:

$$\omega_{\parallel}^{\text{visu.}} = \left\langle \frac{\omega_{\parallel, \phi} + \omega_{\parallel, r}}{U} \right\rangle \quad (\text{S11})$$

$$\omega_{\parallel, \phi} = \sum_k \mathcal{O}_k(\Delta\varphi_k) \left( \int_{\varphi_{\uparrow k}}^{\varphi_{\downarrow k}} \cos \phi b_{\epsilon}(\phi) d\phi \right) \quad (\text{S12})$$

$$\omega_{\parallel, r} = \sum_k -\mathcal{D}_k(\Delta\varphi_k) \left( \int_{\varphi_{\uparrow k}}^{\varphi_{\downarrow k}} \sin \phi b_{\epsilon}(\phi) d\phi \right) \quad (\text{S13})$$

Eq. (S13) is the exact form as Eq. (S2) with the weight of the integral being  $\mathcal{D}_i(\Delta\varphi_k)$  instead of  $\mathcal{R}_i^2(\Delta\varphi_k)$  resulting in

$$\omega_{\parallel, r} = \sum_k \frac{-2\mathcal{D}_k(\Delta\varphi_k)}{\mathcal{R}_k(\Delta\varphi_k)} \sin(\varphi_k) (1 + \epsilon \cos \varphi_k \cos \Delta\varphi_k) \quad (\text{S14})$$

As for Eq. S12 the developing the integral and assuming  $\Delta\phi \approx \sin \Delta\phi$  we have:

$$\omega_{\parallel, \phi} = \sum_k \frac{\mathcal{O}_k(\Delta\varphi_k)}{\mathcal{R}_k(\Delta\varphi_k)} (2 \cos \varphi_k (1 + \epsilon \cos \Delta\varphi_k \cos \varphi_k) + \epsilon (1 - \cos \Delta\varphi_k)) \quad (\text{S15})$$

Once more, for the case where there is no occlusion nor aggregation  $k = j$  and taking into account the physical constrains of the particles:  $\cos \Delta\varphi_k \approx 1$  and  $\varphi_k = \theta_{ij}$ , then Eqs. (S14-15) becomes

$$\omega_{\parallel, r} \approx \sum_{j \neq i}^N \frac{-2\mathcal{D}_k(\Delta\varphi_k)}{\mathcal{R}_k(\Delta\varphi_k)} \sin(\theta_{ij}) b_{\epsilon}(\theta_{ij}) \quad (\text{S16})$$

$$\omega_{\parallel, \phi} \approx \sum_{j \neq i}^N \frac{2\mathcal{O}_k(\Delta\varphi_k)}{\mathcal{R}_k(\Delta\varphi_k)} \cos(\theta_{ij}) b_{\epsilon}(\theta_{ij}) \quad (\text{S17})$$

This leads to

$$\omega_{\parallel}^{\text{visu.}} \approx \left\langle \sum_{j \neq i}^N 2 \frac{\mathbf{e}_i \times \mathbf{V}_{ij}}{\mathcal{R}_k^2(\Delta\varphi_k)} b_{\epsilon}(\theta_{ij}) \right\rangle \quad (\text{S18})$$

Now  $\mathbf{V}_{ij} = U(\mathbf{e}_j - \mathbf{e}_i)$  then  $\mathbf{e}_i \times \mathbf{e}_j = \frac{\mathbf{e}_i \times \mathbf{V}_{ij}}{U}$ . And  $\mathcal{R}_i(\Delta\varphi_k) = \|\mathbf{x}_j - \mathbf{x}_i\|$ , resulting in

$$\omega_{\parallel}^{\text{visu.}} \approx \left\langle \sum_{j \neq i}^N 2 \frac{\mathbf{e}_i \times \mathbf{e}_j}{\|\mathbf{x}_j - \mathbf{x}_i\|^2} b_{\epsilon}(\theta_{ij}) \right\rangle \quad (\text{S19})$$

The normalization nullifies the factor 2 that is not present in the omniscient model and shows that the two models are equivalent when shades are formed by single particles.



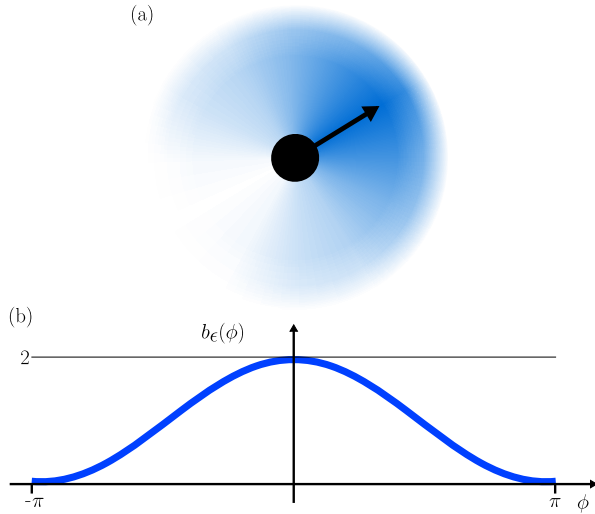


Fig. S 5. Effect of a blind spot. (a) 2D view: When  $\epsilon = 1$ , each particle cannot see at its rear (transparent), but the view is normal in front (blue). (b) Function  $b_\epsilon(\phi) = 1 + \epsilon \cos \phi$  when  $\epsilon = 1$

#### S1.4 Blind spot

Figure Fig. S5 illustrates the function  $b_\epsilon(\varphi)$  used to model a blind spot in the visual field. When  $\epsilon$  is equal to 0 there is no effect and with 1 the effect is maximum.

## S2 DIGITAL IMPLEMENTATION

### S2.1 Omniscient model

As previously stated, the model that we propose has an omniscient implementation. The alignment term is shown on Eq. (S7) can be normalized by Eq. (5), which lead to the implementation of:

$$\omega_{\parallel} = \left\langle \sum_{j \neq i}^N \frac{\mathbf{e}_i \times \mathbf{e}_j}{\|\mathbf{x}_j - \mathbf{x}_i\|^2} b_\epsilon(\theta_{ij}) \right\rangle \quad (\text{S20})$$

For attraction term, the implemented equation is:

$$\omega_{\odot} = \left\langle \sum_{j \neq i}^N \|\mathbf{x}_j - \mathbf{x}_i\| \sin(\theta_{ij}) b_\epsilon(\theta_{ij}) \right\rangle \quad (\text{S21})$$

### S2.2 Visual model

On the main text, we described the continuous form of the equations for attraction (Eq. (4a)) and for alignment (Eq. (4b)) and Sections S1.1 and S1.3 expanded on equivalency between these visual terms when there is no

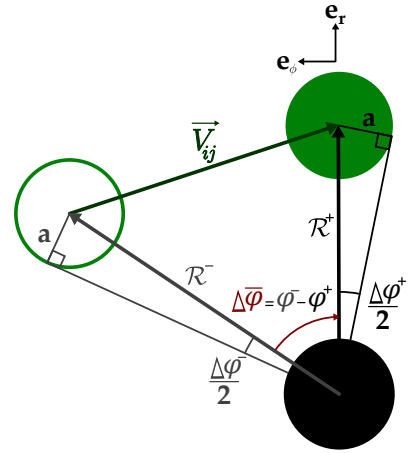


Fig. S 6. Point of view of the  $i$ -th particle (black) of the  $j$ -th particle in  $t-1$  (contour) and  $t$  (solid color)

occlusion or aggregation. Here, we go in more depth on how the visual components are derived and calculated, as well as the formulation of the implementation for equations for the numerical simulations.

Consider the point of view of the  $i$ -th particle in 2 consecutive time steps for the  $j$ -th particle:

Each particle has a radius  $a$ , the over index represent + time  $t$  and - time  $t-dt$ .  $\mathbf{e}_r$  and  $\mathbf{e}_\phi$  are the unit vectors of the polar coordinates referred to  $\theta_i$ , where the former is radial and the later azimuthal.  $\vec{V}_{ij}$  is the relative velocity vector perceived by the particle  $i$  of the net velocity of the particle  $j$ . Then from the geometry of the problem we have

$$\mathcal{R}_i^\pm = \frac{a}{\sin \Delta\varphi^\pm} \quad (\text{S22})$$

Then the  $\mathcal{R}_i$  vectors can be described as follows:

$$\vec{\mathcal{R}}_i^+ = \mathcal{R}_i^+ \mathbf{e}_r \quad (\text{S23})$$

$$\vec{\mathcal{R}}_i^- = \mathcal{R}_i^- [\cos \Delta\bar{\varphi} \mathbf{e}_r - \sin \Delta\bar{\varphi} \mathbf{e}_\phi] \quad (\text{S24})$$

Consequently the distance associated with the movement of  $\vec{V}_{ij}$  is:

$$U \vec{V}_{ij} \delta t = (\mathcal{R}_i^+ - \mathcal{R}_i^- \cos \Delta\bar{\varphi}) \mathbf{e}_r + \mathcal{R}_i^- \sin \Delta\bar{\varphi} \mathbf{e}_\phi \quad (\text{S25})$$

$$\vec{V}_{ij} = \frac{[\mathcal{R}_i^+ - \mathcal{R}_i^- \cos \Delta\bar{\varphi}, \mathcal{R}_i^- \sin \Delta\bar{\varphi}]}{U \delta t} \quad (\text{S26})$$

From Eq. (S26) we can do the decomposition in radial and azimuthal components to calculate  $\mathcal{O}_i$  and  $\mathcal{D}_i$

$$\mathcal{O}_i = \frac{1}{U \delta t} - \frac{\mathcal{R}_i^- \cos \Delta\bar{\varphi}}{\mathcal{R}_i^+ U \delta t} \quad (\text{S27})$$

$$\mathcal{D}_i = \frac{\mathcal{R}_i^- \sin \Delta\bar{\varphi}}{\mathcal{R}_i^+ U \delta t} \quad (\text{S28})$$

Note that on a robotic implementation this method of calculating the optic flow and its divergence ( $\mathcal{O}_i$  and  $\mathcal{D}_i$ ), would be entirely replaced by its direct measurement via a specialized optic flow sensor.

For the computer implementation of the alignment term of the visual model, we use Eq. (S26) to calculate the cross product of the relative vectors over  $\mathcal{R}_i^2(\Delta\varphi_k)$

$$\omega_{\parallel} = \left\langle \sum_k \sin^2(\Delta\varphi_k) \frac{\vec{V}_{ik} \times \mathbf{e}_i}{Ua^2} (1 + \epsilon \cos \varphi_k) \right\rangle \quad (\text{S29})$$

Note that there are two approximations ( $\cos \Delta\varphi_k \approx 1$  and  $\sin \Delta\varphi_k \approx \Delta\varphi_k$ ) on Eq. (S29).

As for attraction, we use at Eq. (S6) written on the terms described above.

$$\omega_{\odot} = \left\langle \sum_k 2a^2 \frac{\sin(\varphi_k)}{\sin(\Delta\varphi_k)} (1 + \epsilon \cos \varphi_k \cos \Delta\varphi_k) \right\rangle \quad (\text{S30})$$

### S2.3 Algorithm to detect corresponding features necessary to compute optic flow

The optic flow is the perception of relative movement and we use a feature matching-inspired algorithm to calculate it in numeric simulation. But this feature match requires that you discriminate and detect the same feature in two consecutive time steps. Here we present the algorithm to match different features given the binary characteristic of  $\mathcal{V}_i(\phi)$ . In the model this will be the assumption to use Eqs. (S27) and (S28). Note that shades can split or join in between time steps. To detect these events, we employ a minimum distance error to correlate two set of features in two consecutive time steps.

In Fig. S7, we illustrate the algorithm used, where yellow boxes are associated to corresponding shades. To correspond to each other, shades should be associated to radial and azimuthal velocities bounded by  $U$ .

This algorithm is also illustrated on the following pseudo-code (see Algorithm ).

The model is implemented on python 3.10 (it will be available for download on [1] and [2]). This includes the implementation of the algorithm described by Fig. S7, all model equations, data management and figure creation. There is a markdown file that only include the model implementation without data logging.

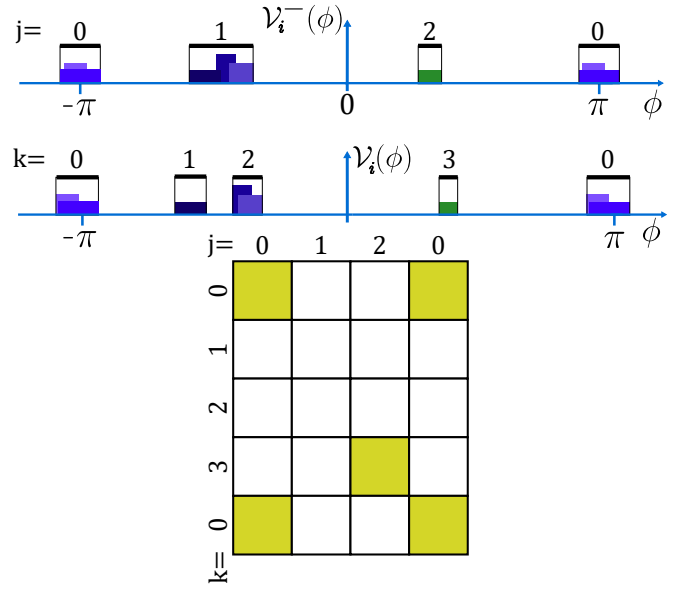


Fig. S 7. Example of two consecutive visual fields. The matrix at the bottom illustrate the algorithm we use to find corresponding shades between two time steps.  $\Delta\bar{\phi}$  is the delta between the mid point of the shades between two consecutive time steps.

```

 $\overline{V}_{i,t-1} \leftarrow [(\varphi_{\uparrow l} + \varphi_{\downarrow l})/2]$  at t-1  $\forall \varphi$  in  $\phi_{i,t-1}$ 
 $\Delta V_{i,t-1} \leftarrow [\varphi_{\downarrow l} - \varphi_{\uparrow l}]$  at t-1  $\forall \varphi$  in  $\phi_{i,t-1}$ 
 $\overline{V}_{i,t} \leftarrow [(\varphi_{\uparrow k} + \varphi_{\downarrow k})/2]$  at t  $\forall \varphi$  in  $\phi_{i,t}$ 
 $\Delta V_{i,t} \leftarrow [\varphi_{\downarrow k} - \varphi_{\uparrow k}]$  at t  $\forall \varphi$  in  $\phi_{i,t}$ 
 $\omega \leftarrow \omega_{k,t-1} \delta t$   $\triangleright$  Previous known rotation
Require:  $l \& k \geq 0$ 
 $VisualDist \leftarrow \|\overline{V}_{i,t-1} - \overline{V}_{i,t}\|$   $\triangleright$  A  $k \times l$  matrix
 $RefDist \leftarrow \min(\sin \Delta V_{i,t}, \sin \Delta V_{i,t-1}) 2U\delta t/a + \omega$ 
 $VisualArea \leftarrow \|\frac{1}{\Delta V_{i,t}} - \frac{1}{\sin \Delta V_{i,t-1}}\|$   $\triangleright$  A  $k \times l$  matrix
 $RefArea \leftarrow 2U\delta t/a$ 
for all values in  $VisualArea, VisualDist, RefDist$  &  $RefArea$  do
  if  $ValueVisualArea \downarrow RefDist$  &  $ValueVisualArea \downarrow RefArea$  then
     $Pair_t \leftarrow [\varphi_{\uparrow k}, \varphi_{\downarrow k}]$ 
     $Pair_{t-1} \leftarrow [\varphi_{\uparrow l}, \varphi_{\downarrow l}]$ 
  end if
end for
Corresponding shades between time steps

```

## S3 DETAILED PHASE DIAGRAMS (FIG. 3)

## S3.1 Visual Model

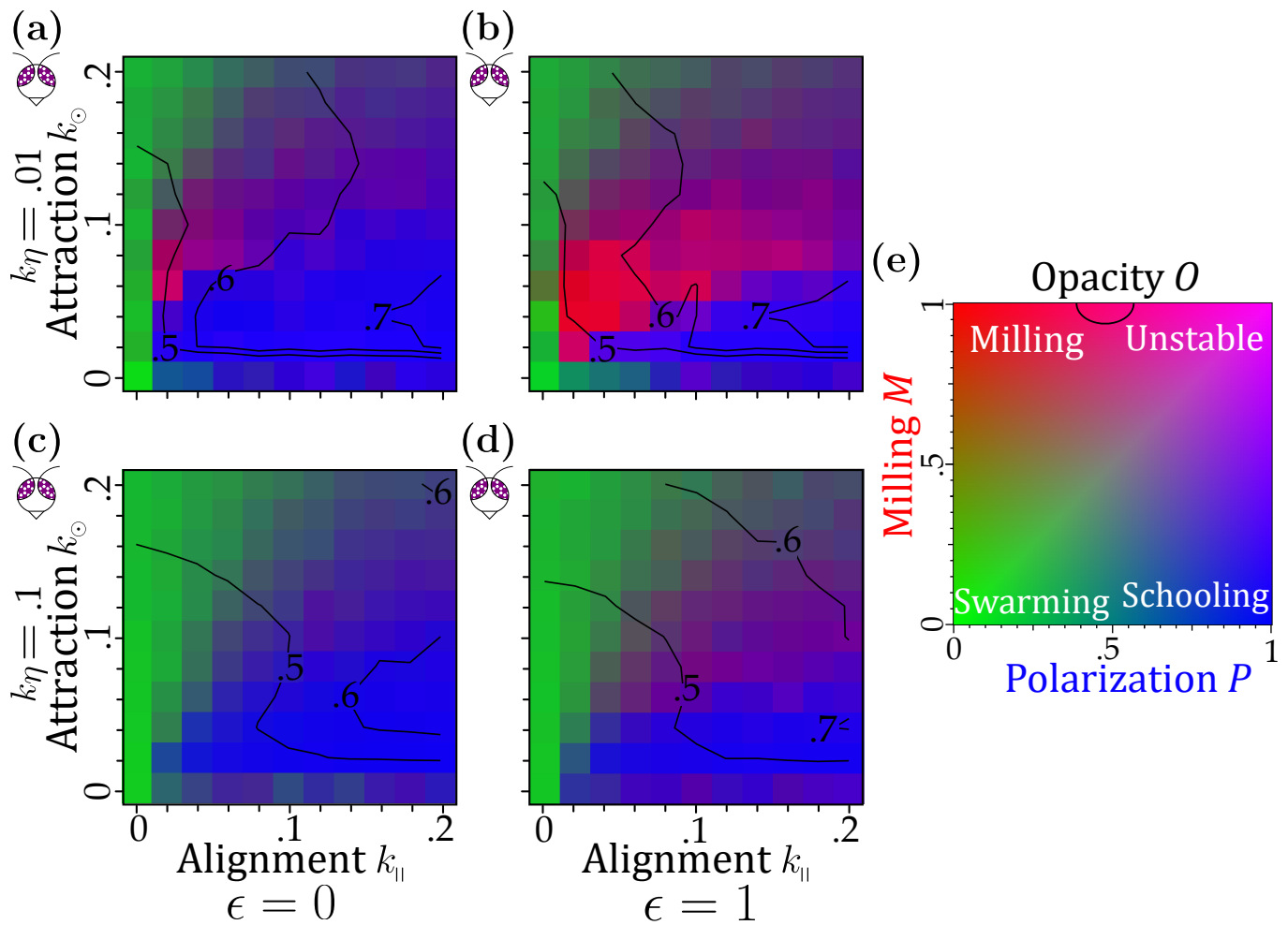


Fig. S 8. Phase diagrams for  $N = 50$  individuals. These phase diagrams are for: (a)-(c) noise strength  $k_\eta = 0.01$ , where there is no blind angle ( $\epsilon = 0$ ) and maximum ( $\epsilon = 1$ ), respectively. (b)-(d) noise strength  $k_\eta = 0.1$ , where there is no blind angle ( $\epsilon = 0$ ) and maximum ( $\epsilon = 1$ ), respectively. In (e), is illustrated how, in the phase space  $(k_\parallel, k_\odot)$ , the colors represents different values of  $P$  and  $M$ , and contours show the opacity  $O$ ; these colors also represents the phases discussed earlier.

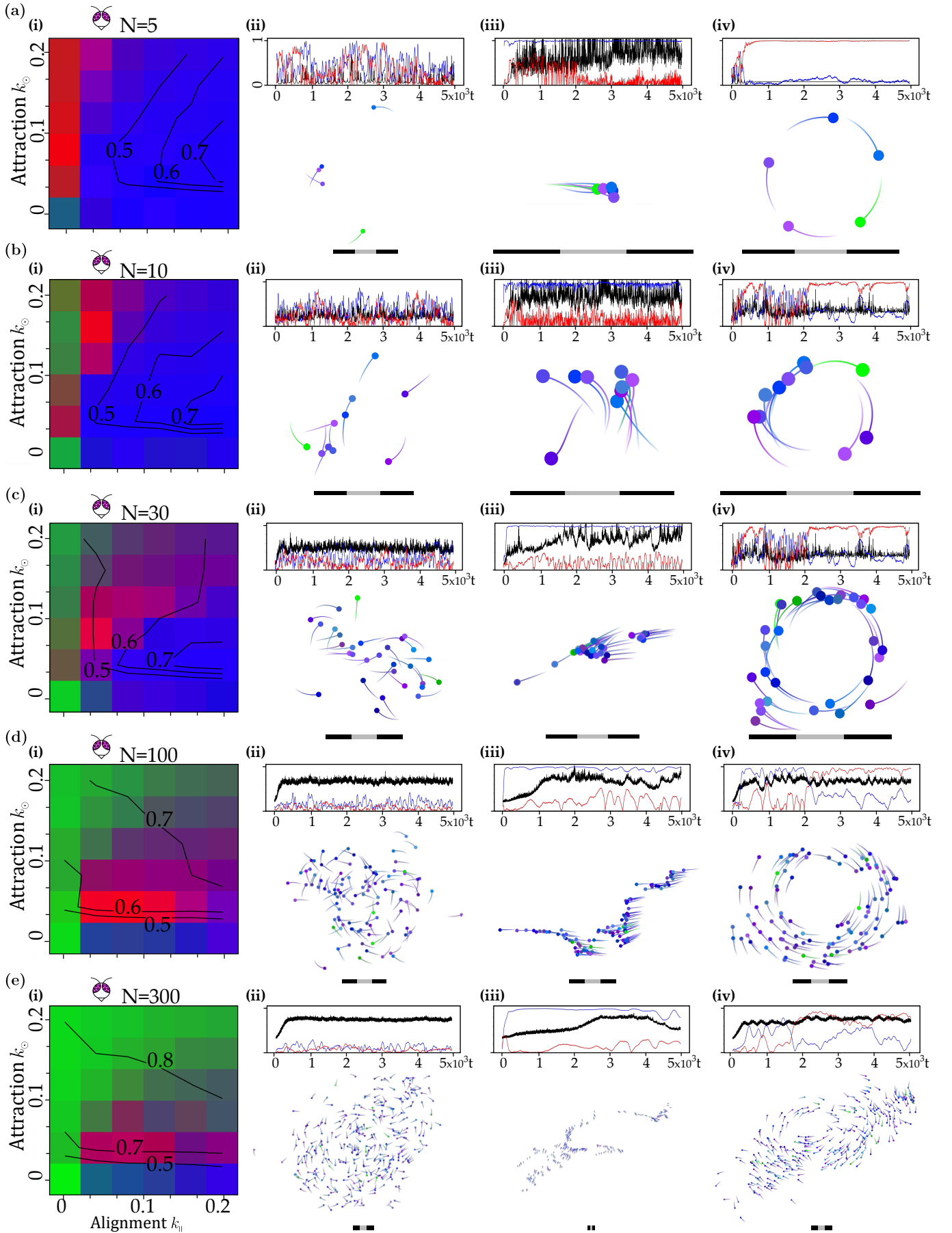


Fig. S 9. Phase Diagram and Phase Reproduction for different Flock sizes for the Visual model. i) Phase Diagram, ii)Swarming, iii)Schooling, iv)Milling. a)  $N=5$ , b)  $N=10$ , c)  $N=30$ , d)  $N=100$ , e)  $N=300$ .

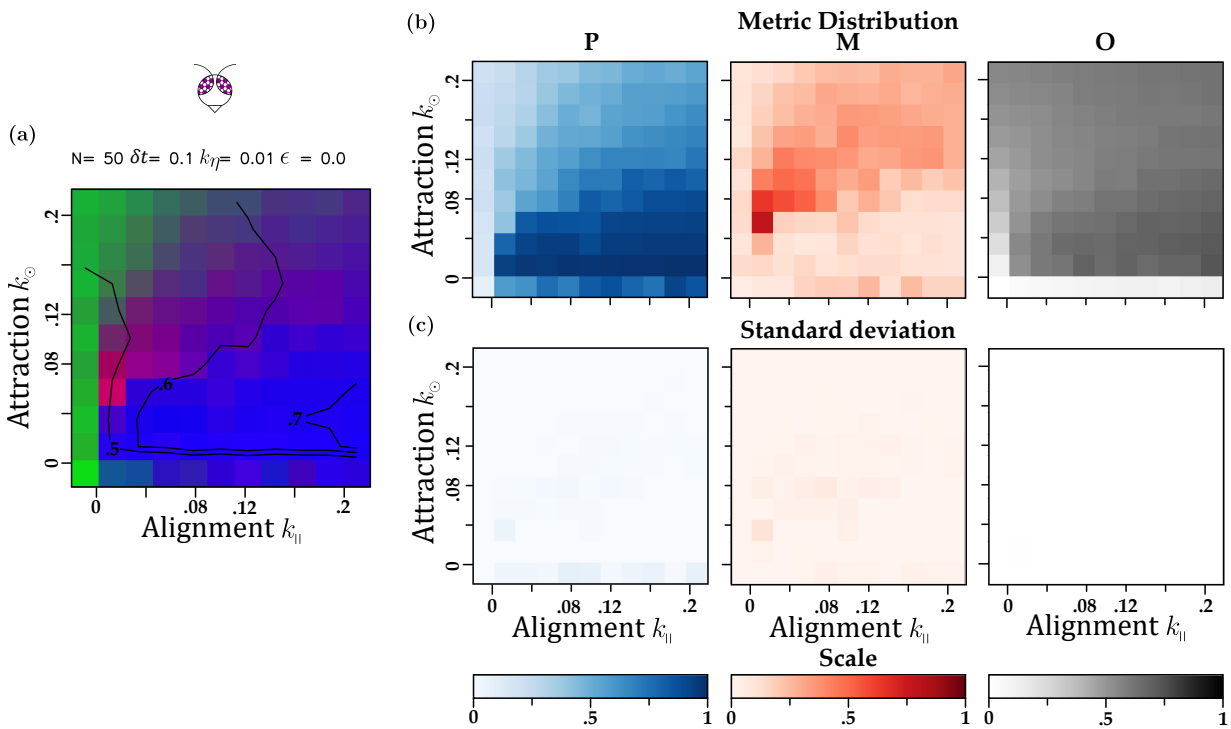


Fig. S 10. Phase diagram for the Visual model of a flock size of 50 individuals,  $k_{\eta} = 0.01$  and no blind angle. a) Integrated metrics, b) Separated metrics and c) Variance between the 10 experiments final metrics.

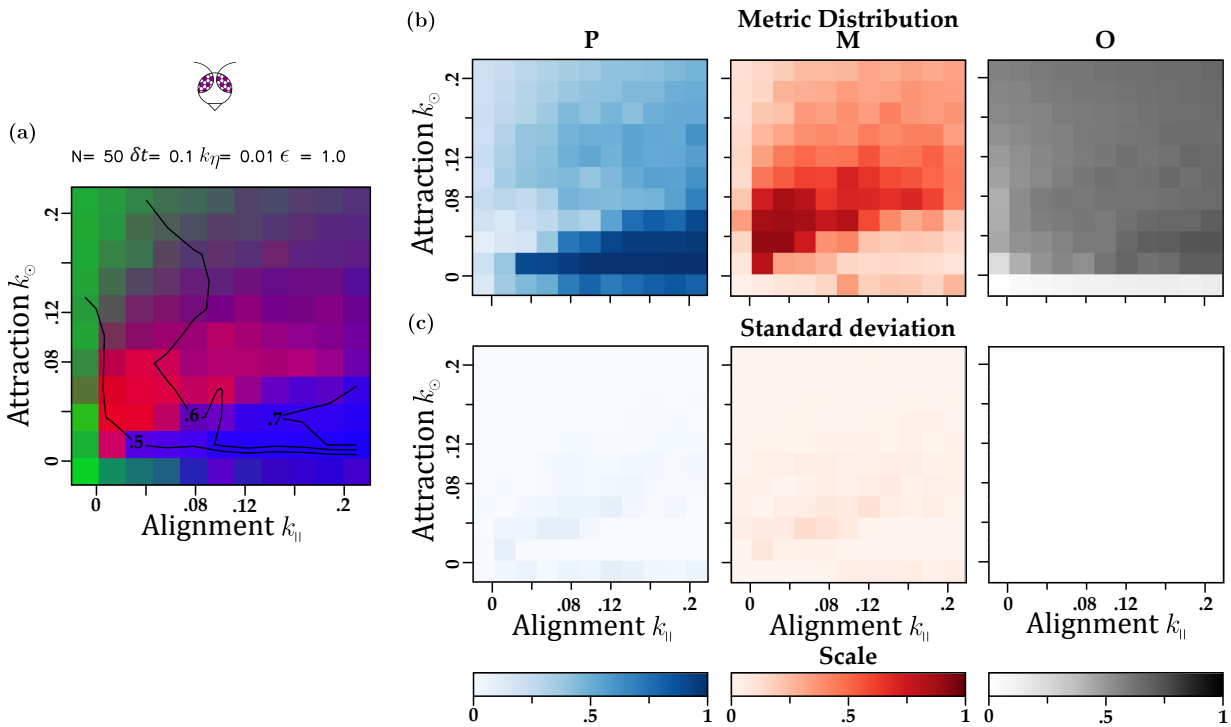


Fig. S 11. Phase diagram for the Visual model of a flock size of 50 individuals,  $k_{\eta} = 0.01$  and maximum blind angle. a) Integrated metrics, b) Separated metrics and c) Variance between the 10 experiments final metrics.

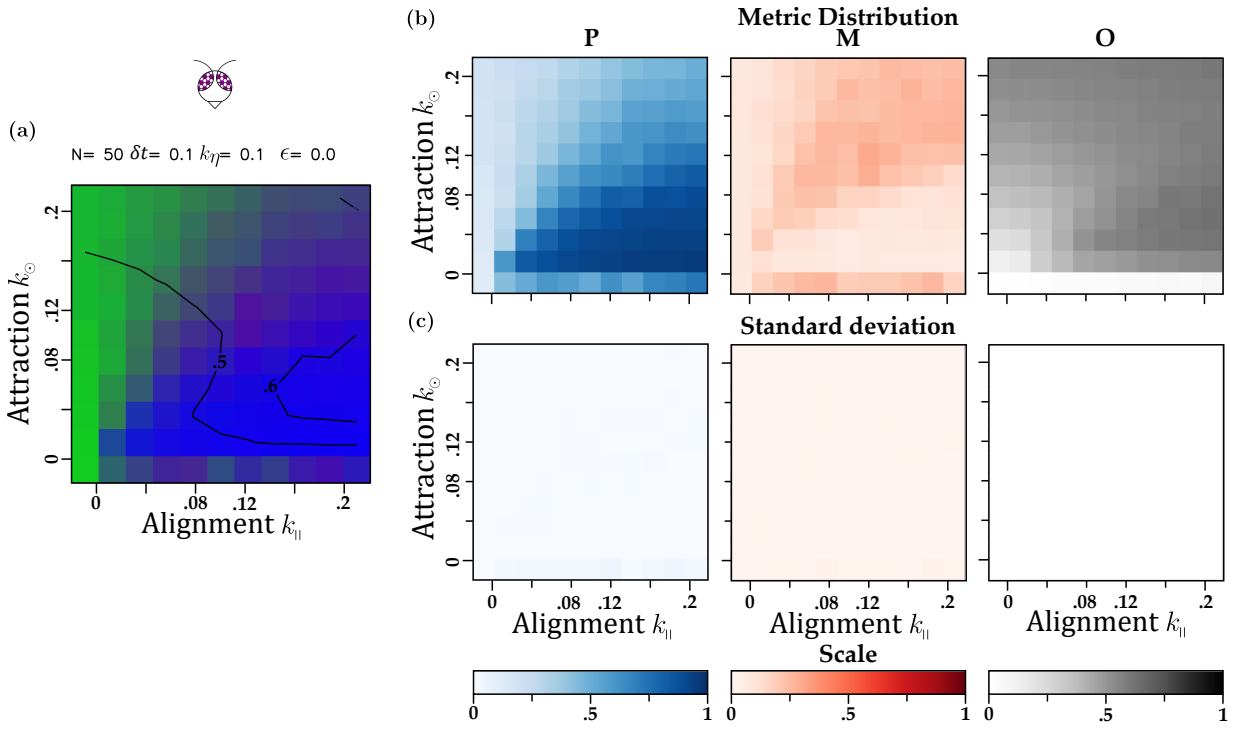


Fig. S 12. Phase diagram for the Visual model of a flock size of 50 individuals,  $k_{\eta} = 0.1$  and no blind angle. a) Integrated metrics, b) Separated metrics and c) Variance between the 10 experiments final metrics.

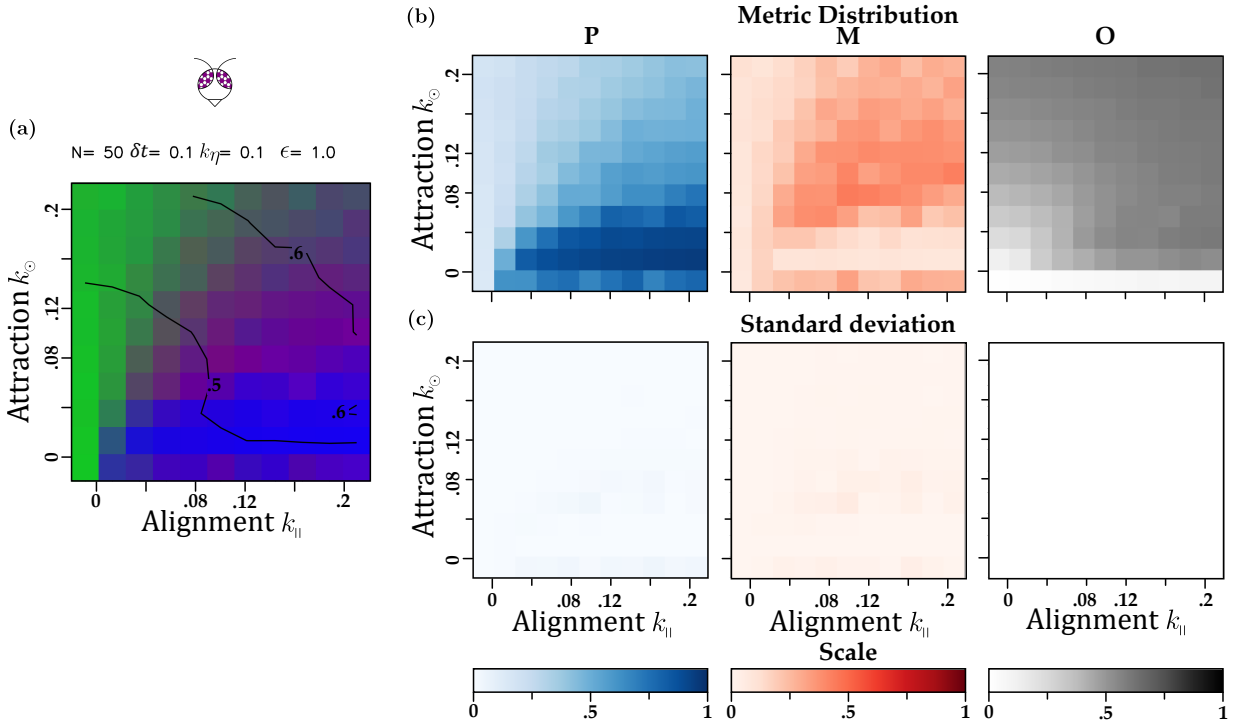


Fig. S 13. Phase diagram for the Visual model of a flock size of 50 individuals,  $k_{\eta} = 0.1$  and maximum blind angle. a) Integrated metrics, b) Separated metrics and c) Variance between the 10 experiments final metrics.

## S3.2 Omniscient Model

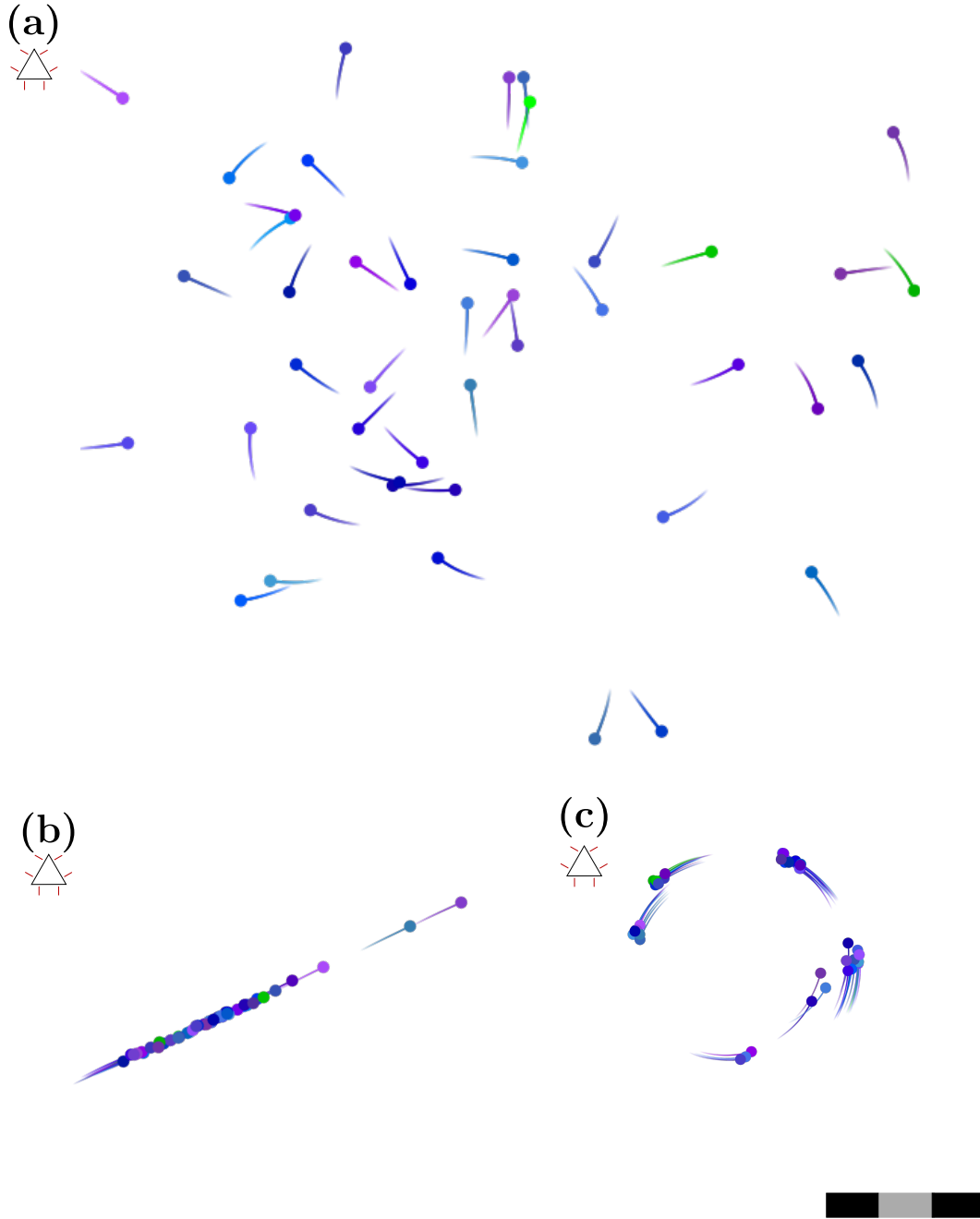


Fig. S 14. Illustration of the phases observed for the omniscient model and a flock size of  $N = 50$  individuals. With  $k_\eta = 0.01$ . The scale bar measures  $30a$  ( $10a$  each section). We observe three phases: (a) Swarming ( $k_\odot = 0.06, k_\parallel = 0, \epsilon = 0$ ); (b) Schooling ( $k_\odot = 0.06, k_\parallel = 0.2, \epsilon = 1$ ); (c) Milling ( $k_\odot = 0.06, k_\parallel = 0.06, \epsilon = 1$ ).

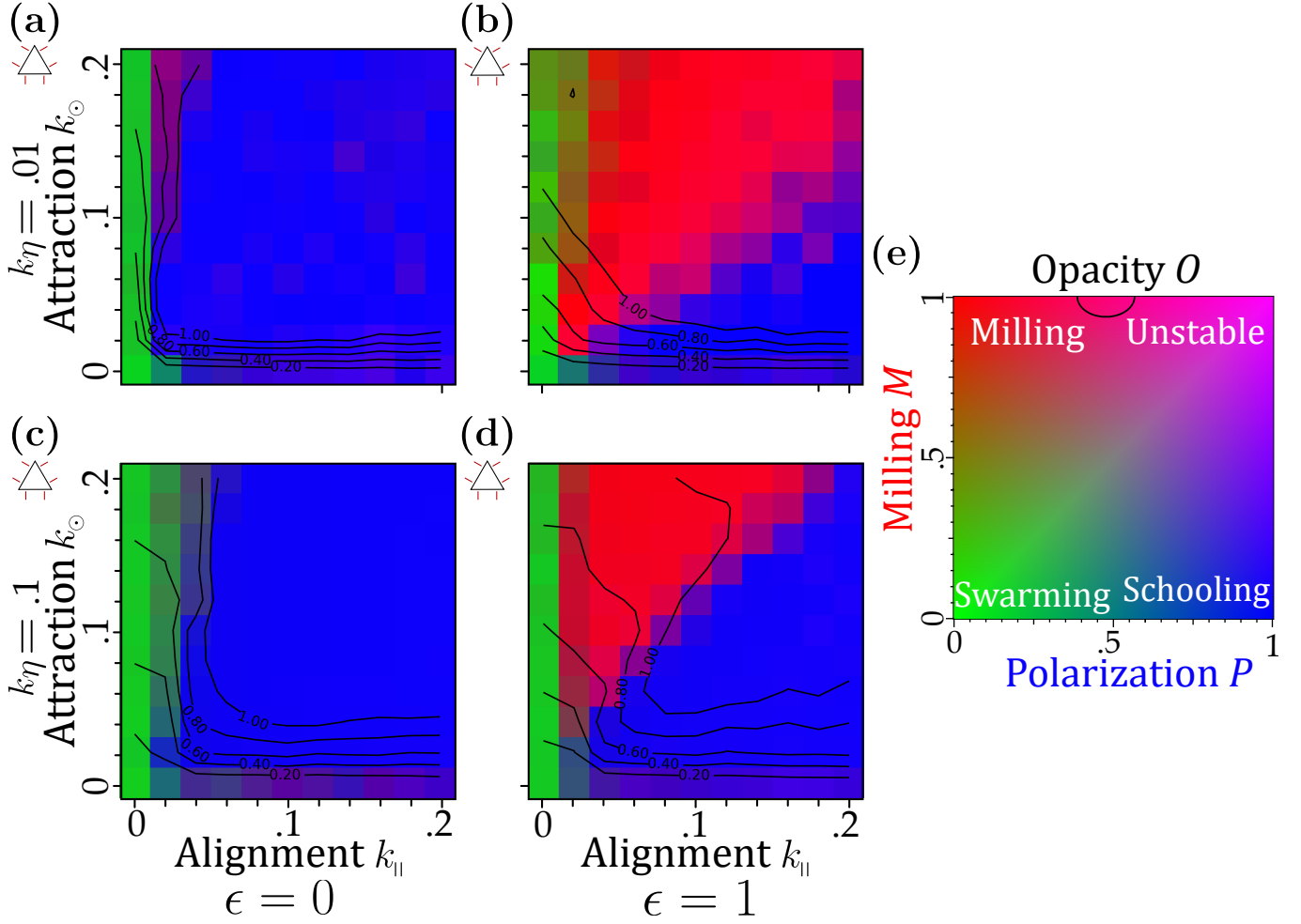


Fig. S 15. Phase diagrams for omniscient model and flock size of  $N = 50$  individuals. These phase diagrams are for: (a)-(c) noise strength  $k_{\eta} = 0.01, \epsilon = 0$  and  $\epsilon = 1$ , respectively. (b)-(d) noise strength  $k_{\eta} = 0.1, \epsilon = 0$  and  $\epsilon = 1$  respectively. In (e), it is illustrated how, in the phase space  $(k_{||}, k_{\odot})$ , the colors represent different values of  $P$  and  $M$ , and contours show the opacity  $O$ ; these colors also represent the phases discussed earlier.



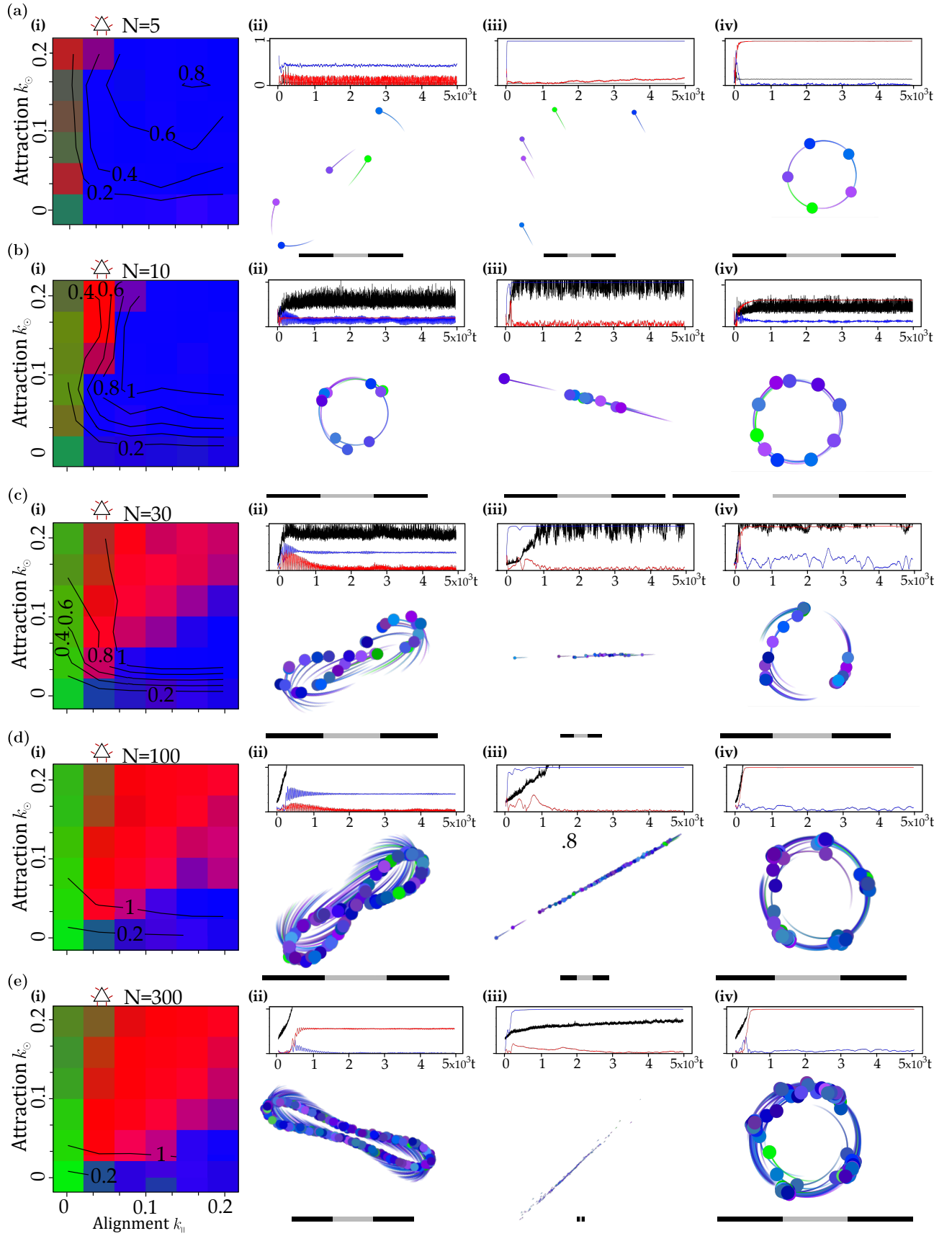


Fig. S 16. Phase Diagram and Phase Reproduction for different Flock sizes for the Visual model. i) Phase Diagram, ii)Swarming, iii)Schooling, iv)Milling. a) $N=5$ , b) $N=10$ , c) $N=30$ , d) $N=100$ , e) $N=300$ .

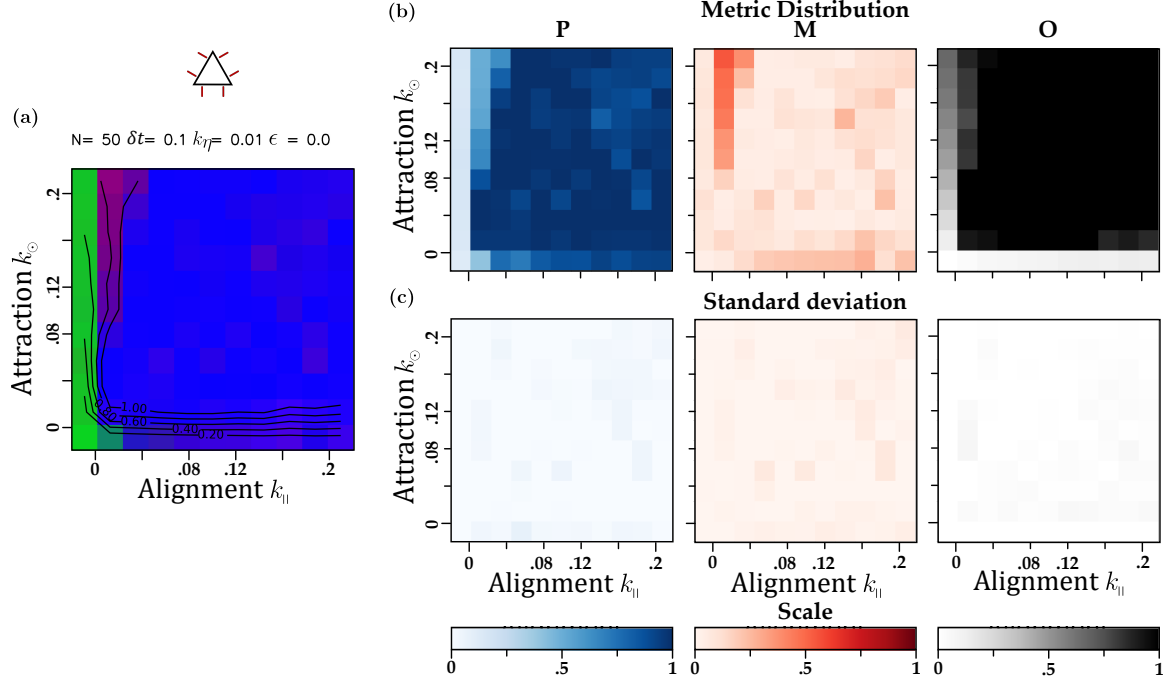


Fig. S 17. Phase diagram for the Omniscient model of a flock size of 50 individuals,  $k_\eta=0.01$  and no blind angle. a) Integrated metrics, b) Separated metrics and c) Variance between the 10 experiments final metrics.

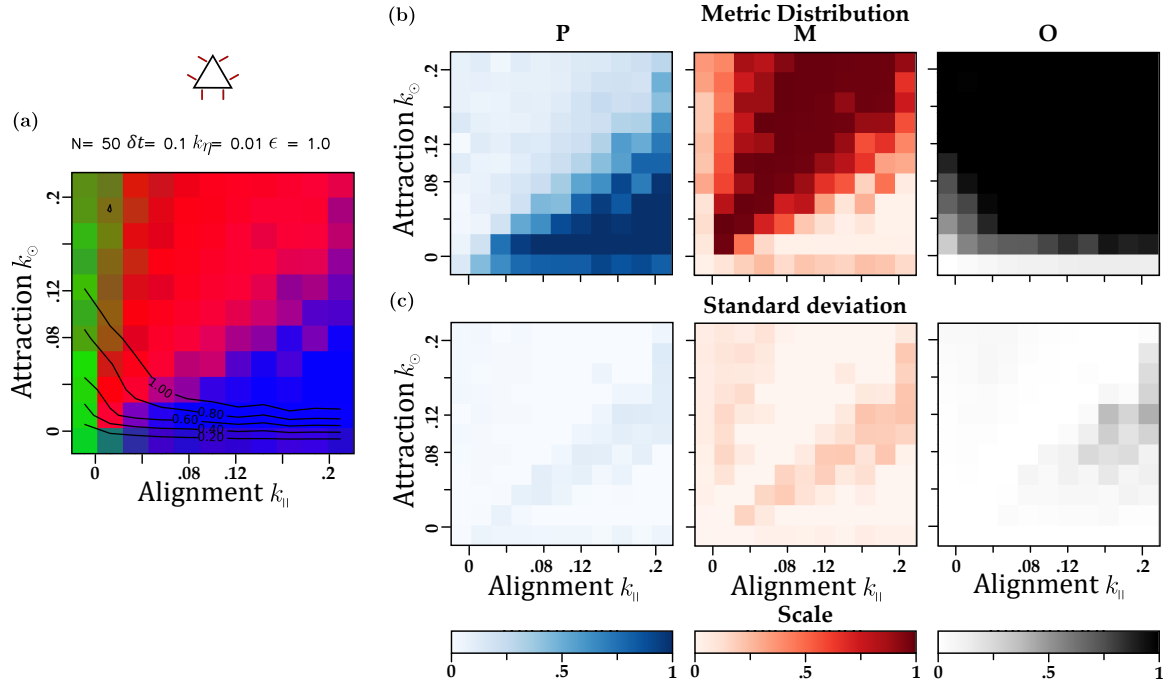


Fig. S 18. Phase diagram for the Omniscient model of a flock size of 50 individuals,  $k_\eta=0.01$  and maximum blind angle. a) Integrated metrics, b) Separated metrics and c) Variance between the 10 experiments final metrics.

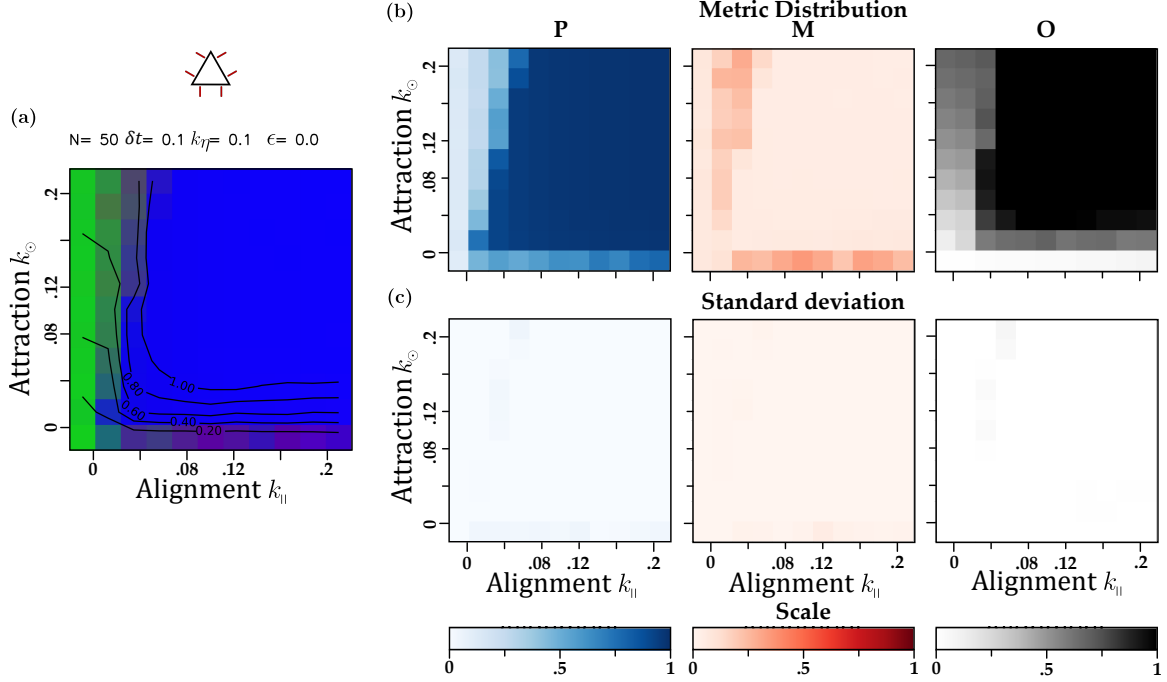


Fig. S 19. Phase diagram for the Omniscient model of a flock size of 50 individuals,  $k_\eta=0.1$  and no blind angle. a) Integrated metrics, b) Separated metrics and c) Variance between the 10 experiments final metrics.

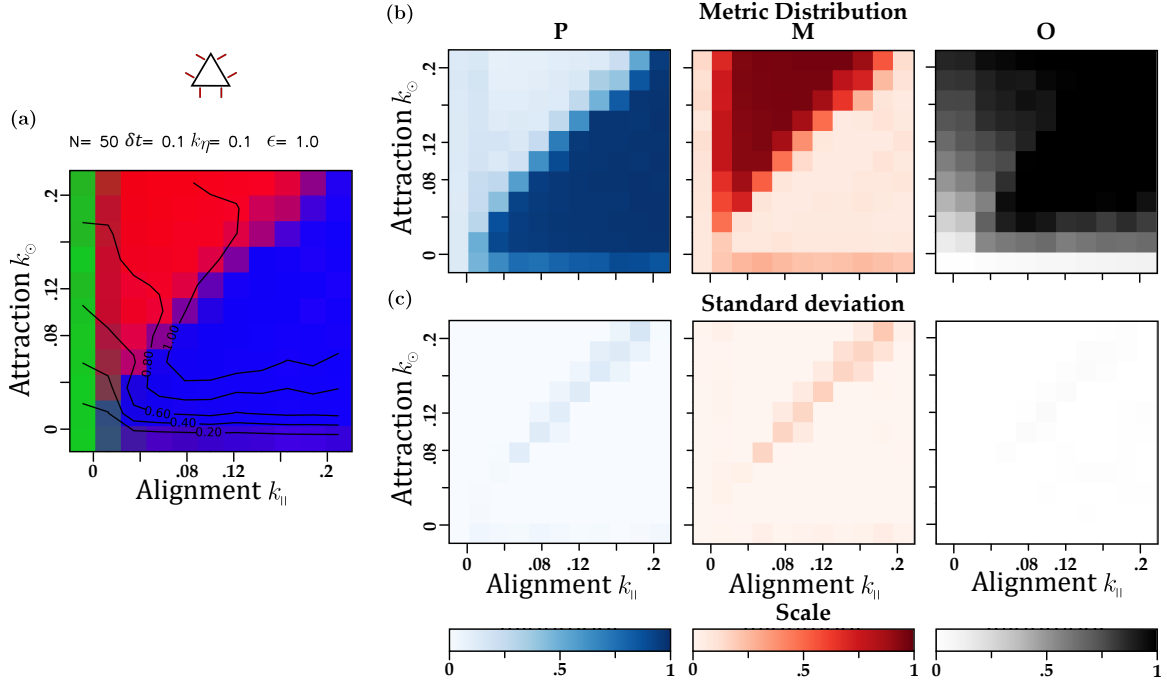


Fig. S 20. Phase diagram for the Omniscient model of a flock size of 50 individuals,  $k_\eta=0.1$  and maximum blind angle. a) Integrated metrics, b) Separated metrics and c) Variance between the 10 experiments final metrics.

## S4 VIDEO SNAPSHOTS OF TIME STEP STABILITY ANALYSIS

## S4.1 Visual model

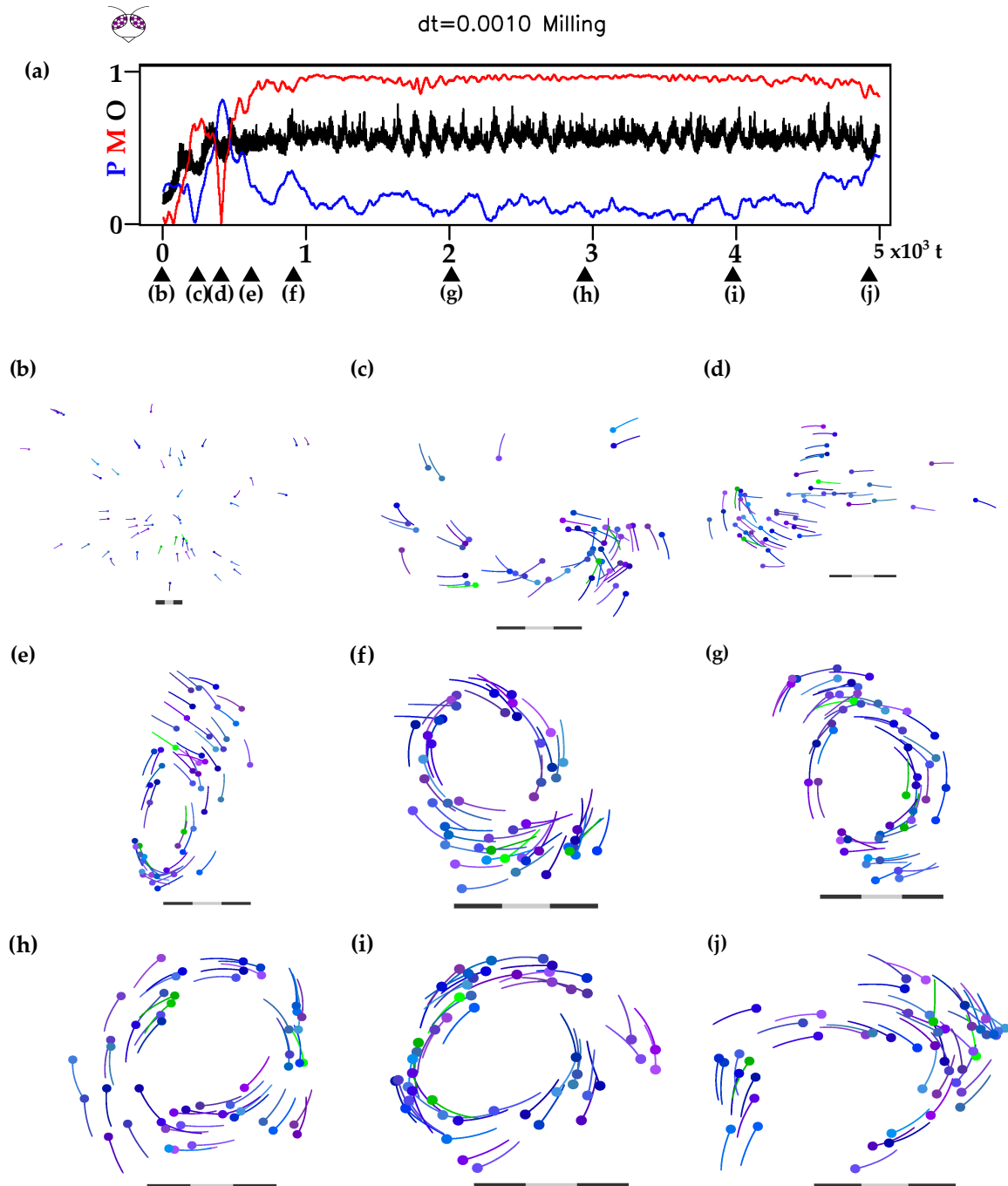


Fig. S 21. Snapshot of the time step stability at  $\delta t=0.001$  for the Visual model. Reproduction of a milling phase for a flock size of 50 individuals,  $k_{\odot}=0.06, k_{\parallel}=0.06$  and  $k_{\eta}=0.01$ . a) Evolution of the flock metrics, b-j different snapshots. , Video S1 (<https://youtu.be/DDjb5NUwKIU>).

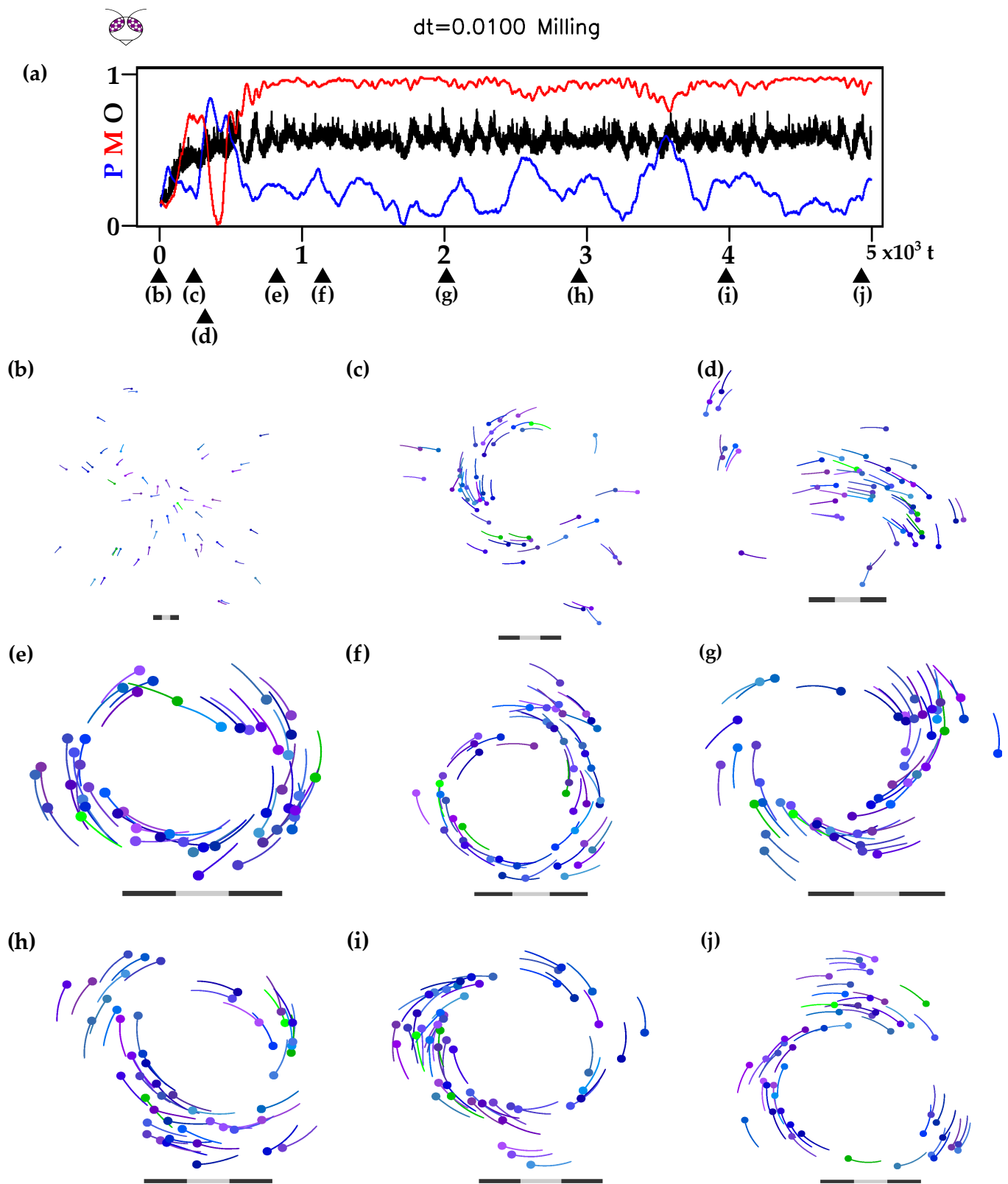


Fig. S 22. Snapshot of the time step stability at  $\delta t=0.01$  for the Visual model. Reproduction of a milling phase for a flock size of 50 individuals,  $k_{\odot}=0.06, k_{\parallel}=0.06$  and  $k_{\eta}=0.01$ . a) Evolution of the flock metrics, b-j different snapshots. , Video S2 (<https://youtube.com/shorts/d0o7oabdQKo>).

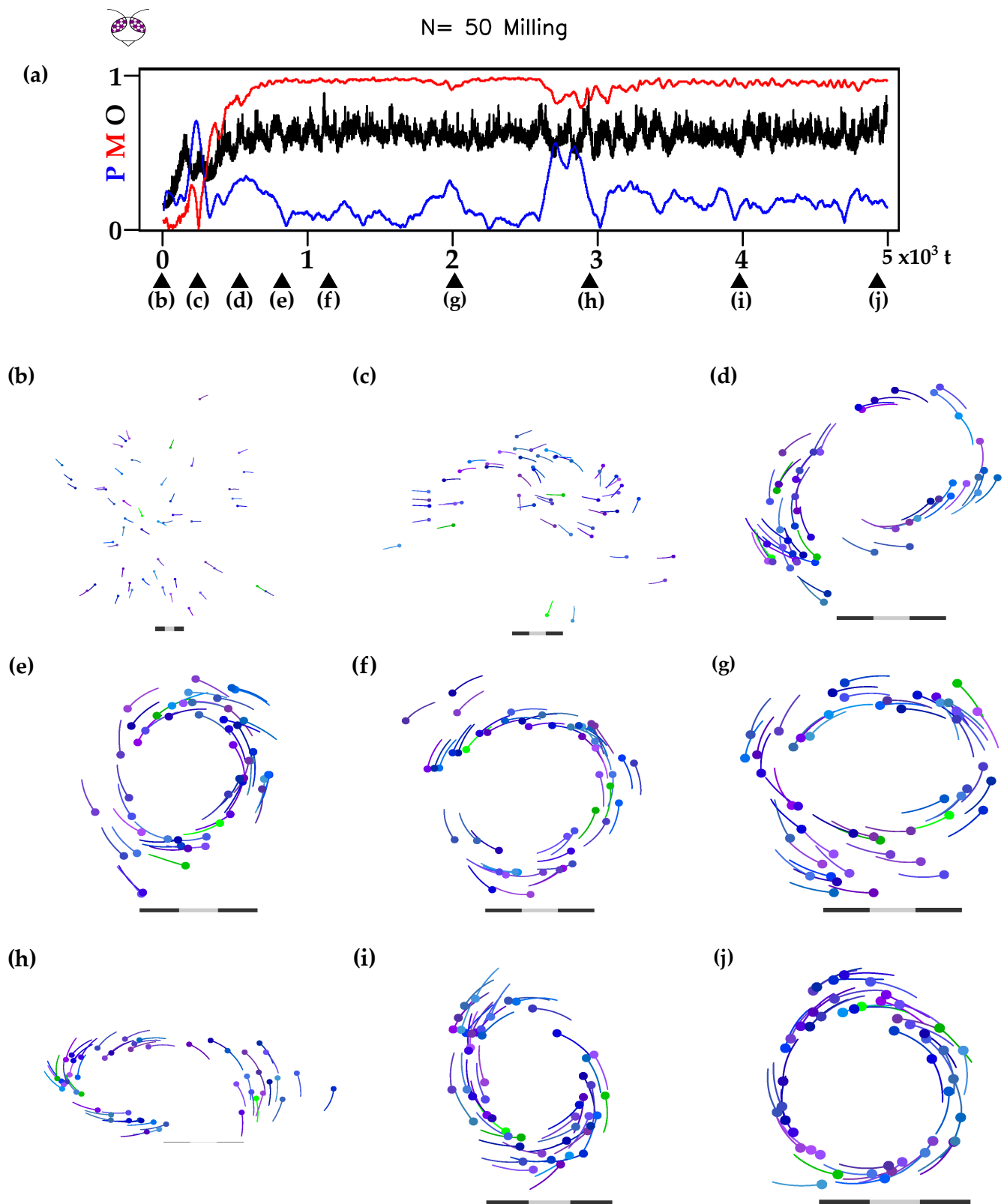


Fig. S 23. Snapshot of the time step stability at  $\delta t=0.1$  for the Visual model. Reproduction of a milling phase for a flock size of 50 individuals,  $k_{\odot} = 0.06, k_{\parallel} = 0.06$  and  $k_{\eta} = 0.01$ . a) Evolution of the flock metrics, b-j different snapshots. , Video S3 (<https://youtu.be/1KMUSwcKyN8>).

## S4.2 Omniscient model

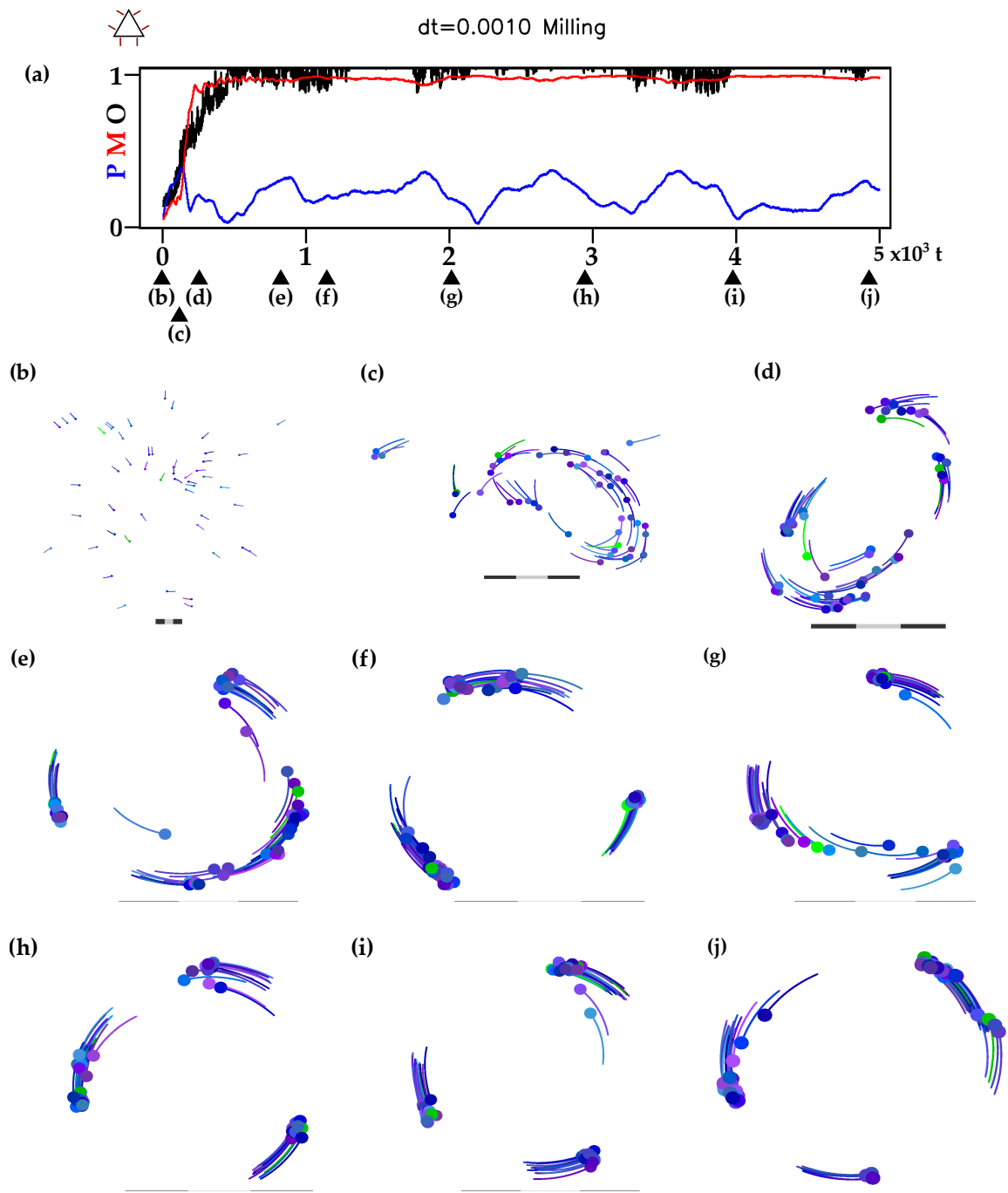


Fig. S 24. Snapshot of the time step stability at  $\delta t=0.001$  for the Omniscient model. Reproduction of a milling phase for a flock size of 50 individuals,  $k_{\odot}=0.06, k_{\parallel}=0.06$  and  $k_{\eta}=0.01$ . a) Evolution of the flock metrics, b-j different snapshots. , Video S4 ([https://youtu.be/r77\\_\\_ozLFvc](https://youtu.be/r77__ozLFvc)).

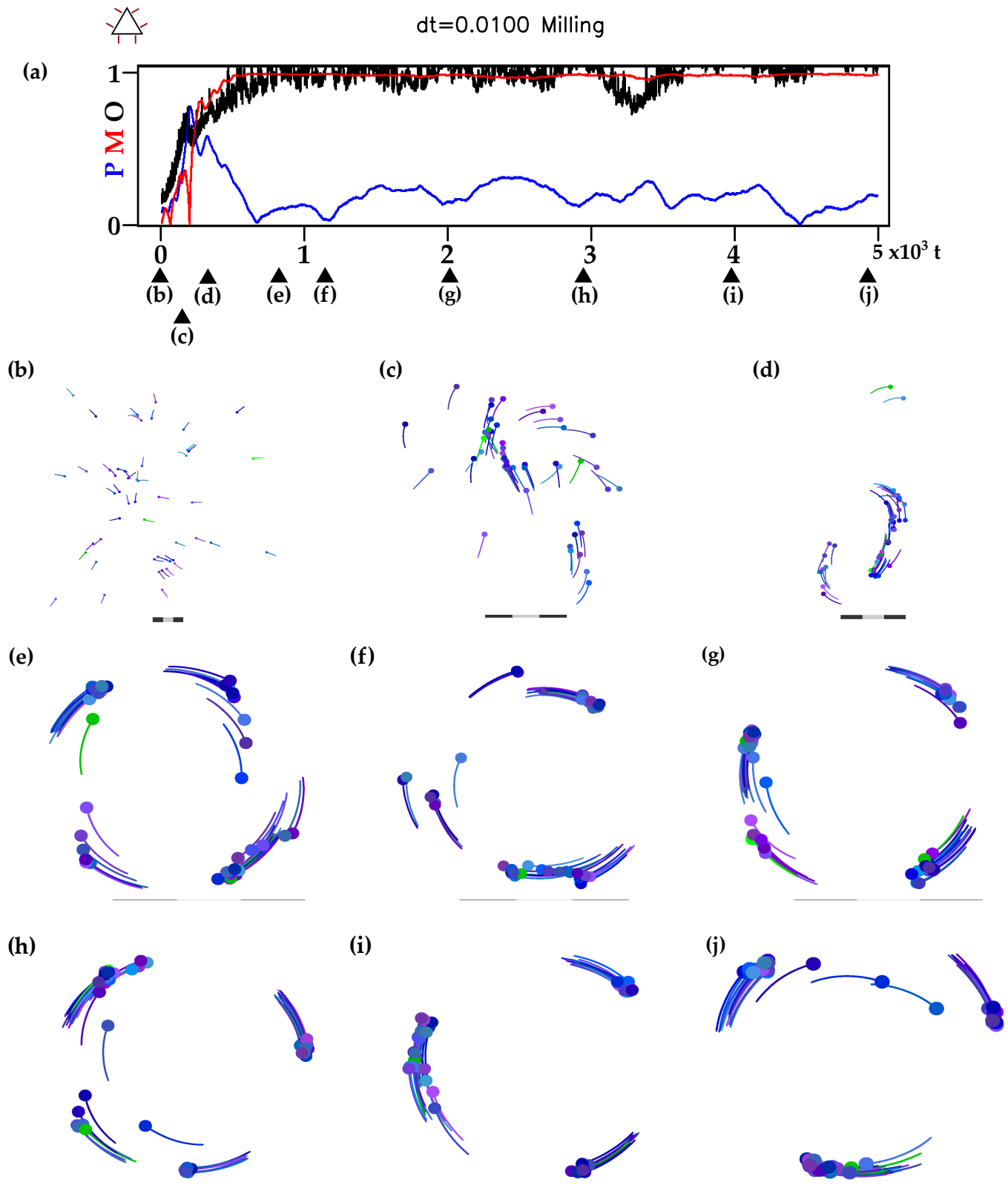


Fig. S 25. Snapshot of the time step stability at  $\delta t=0.01$  for the Omniscient model. Reproduction of a milling phase for a flock size of 50 individuals,  $k_{\odot}=0.06, k_{\parallel}=0.06$  and  $k_{\eta}=0.01$ . a) Evolution of the flock metrics, b-j different snapshots. , Video S5 ([https://youtube.com/shorts/f\\_McHAyXwzU](https://youtube.com/shorts/f_McHAyXwzU)).



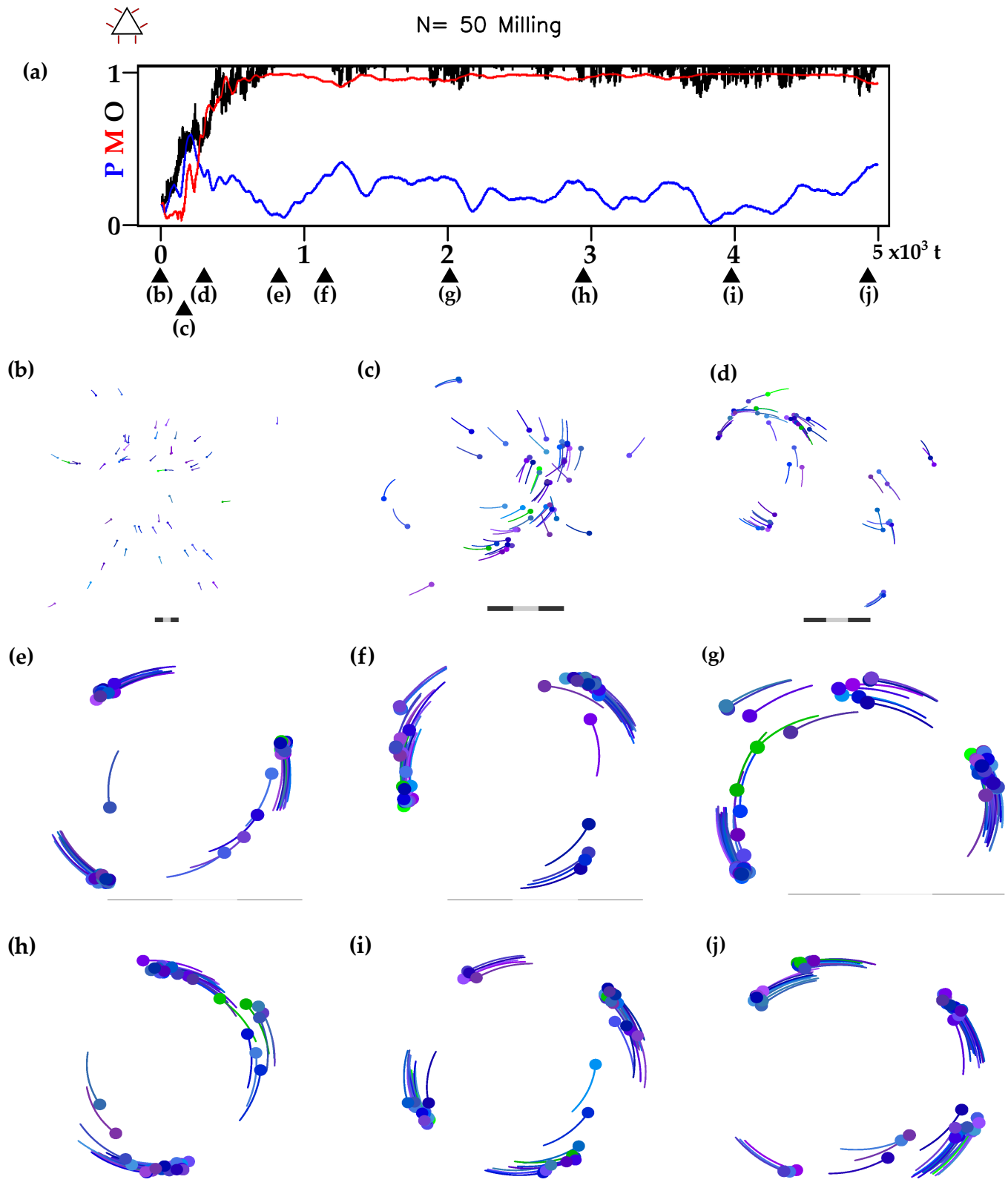


Fig. S 26. Snapshot of the time step stability at  $\delta t=0.1$  for the Omniscient model. Reproduction of a milling phase for a flock size of 50 individuals,  $k_{\odot}=0.06, k_{\parallel}=0.06$  and  $k_{\eta}=0.01$ . a) Evolution of the flock metrics, b-j different snapshots. , Video S6 (<https://youtu.be/NiVgYHt-Fkc>).

S5 VIDEO SNAPSHOTS OF PHASES REPRODUCTION (FIG. 2) ( $k_\eta = 0.01$ ,  $N = 50$ )

S5.1 Visual Model

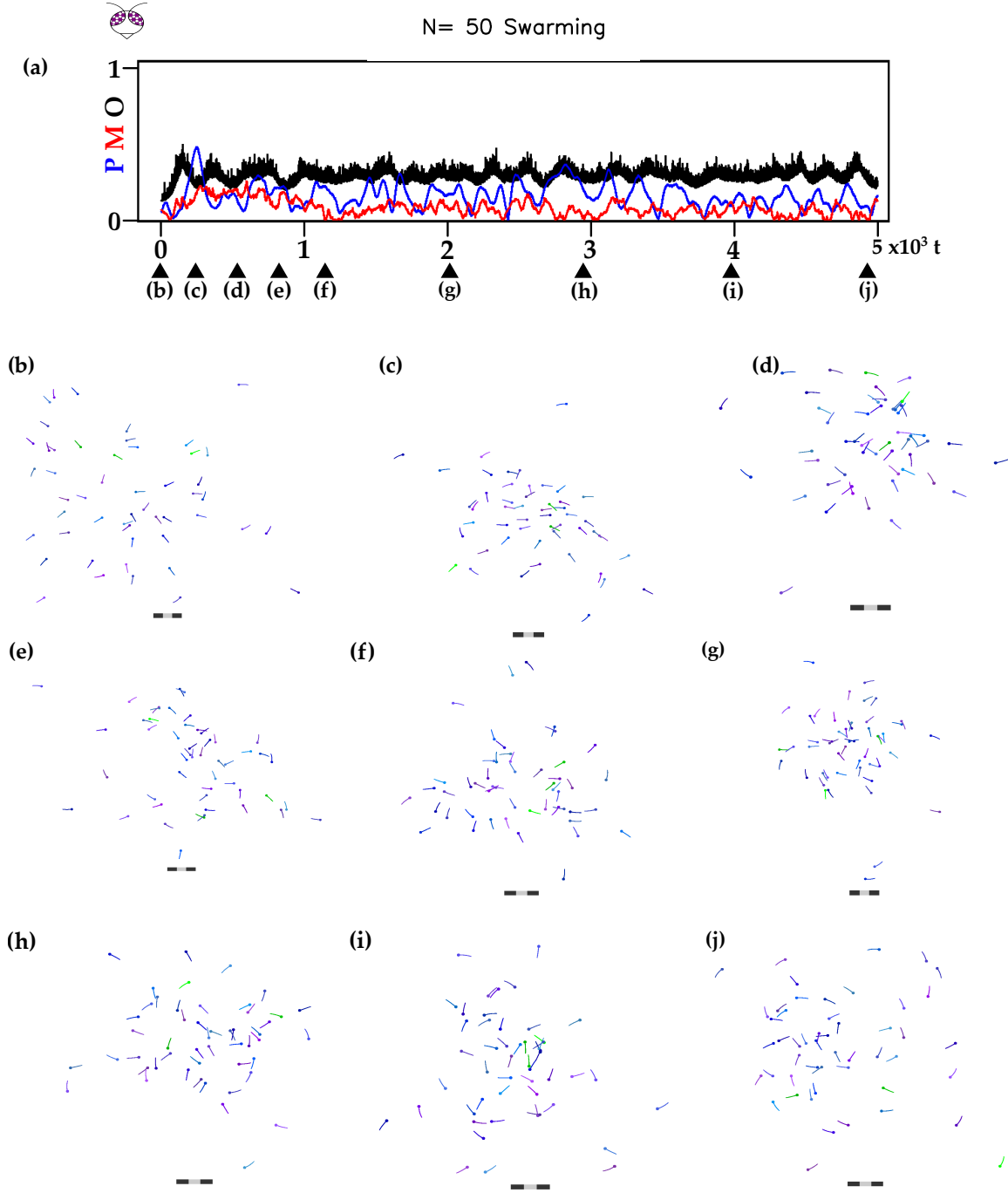


Fig. S 27. Swarming phase reproduction for a flock size of 50 individuals for the Visual model.  $k_\eta = 0.06$   $k_\eta = 0.0$   $k_\eta = 0.01$  and no blind angle , Video 7 (<https://youtu.be/U5yeh-7TOUE>).

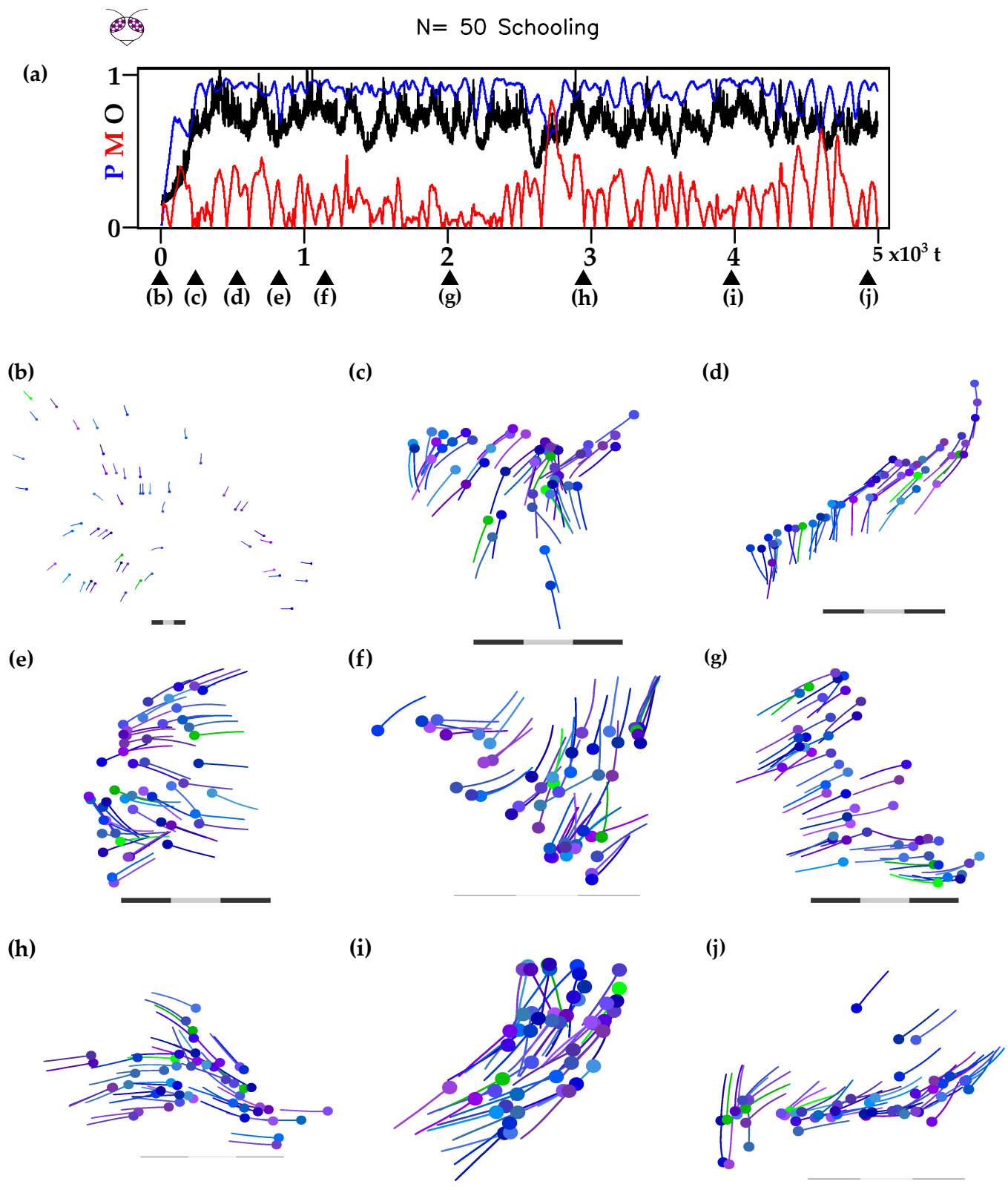


Fig. S 28. Schooling phase reproduction for a flock size of 50 individuals for the Visual model.  $k_\eta = 0.06$   $k_\eta = 0.2$   $k_\eta = 0.01$  and maximum blind angle, Video S8 (<https://youtu.be/2sMePYecmzk>).

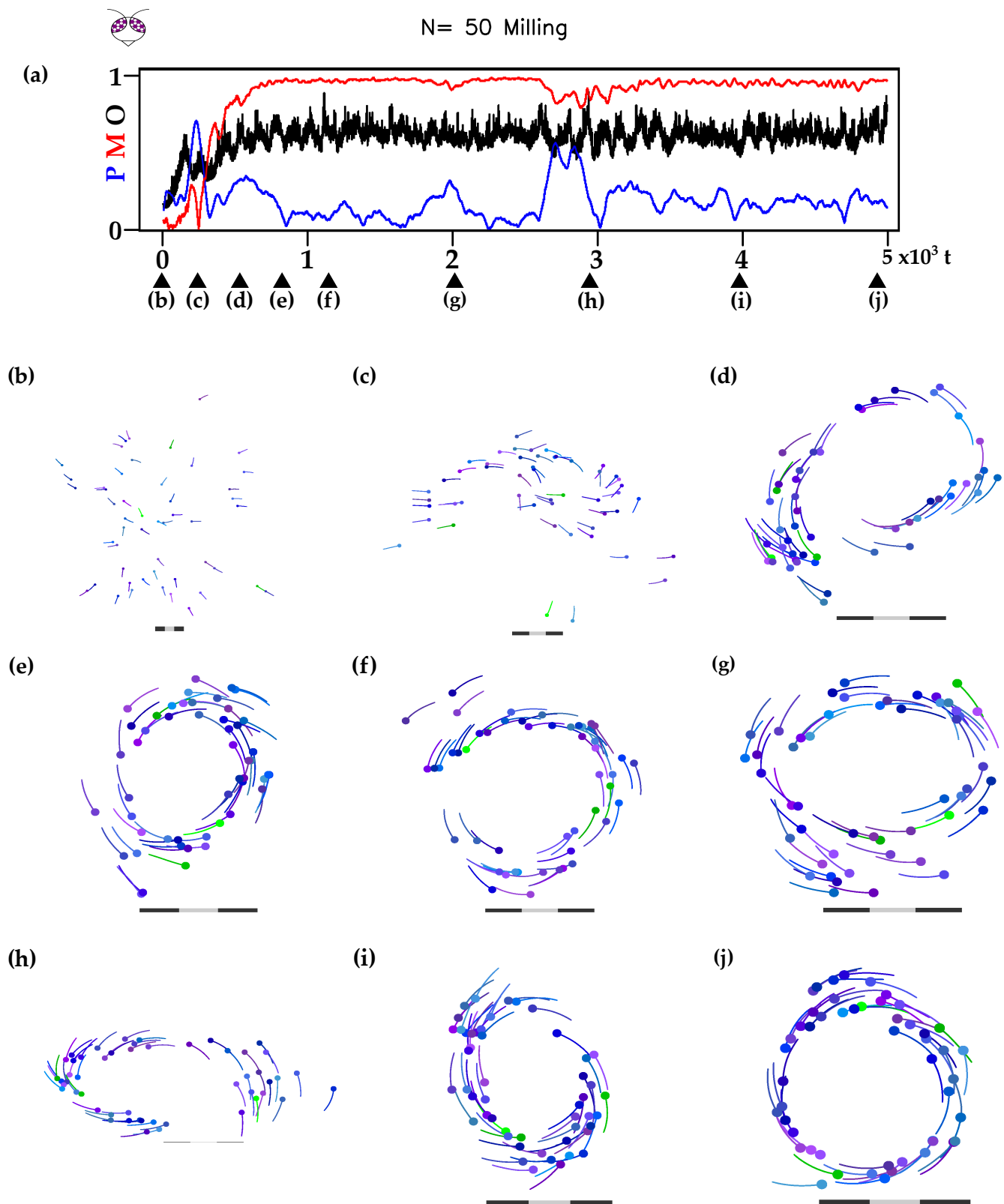


Fig. S 29. Milling phase reproduction for a flock size of 50 individuals for the Visual model.  $k_\eta = 0.06$   $k_\eta = 0.06$   $k_\eta = 0.01$  and maximum blind angle, Video S9 (<https://youtu.be/8dFZNKr1Afw>).

## S5.2 Omniscient Model

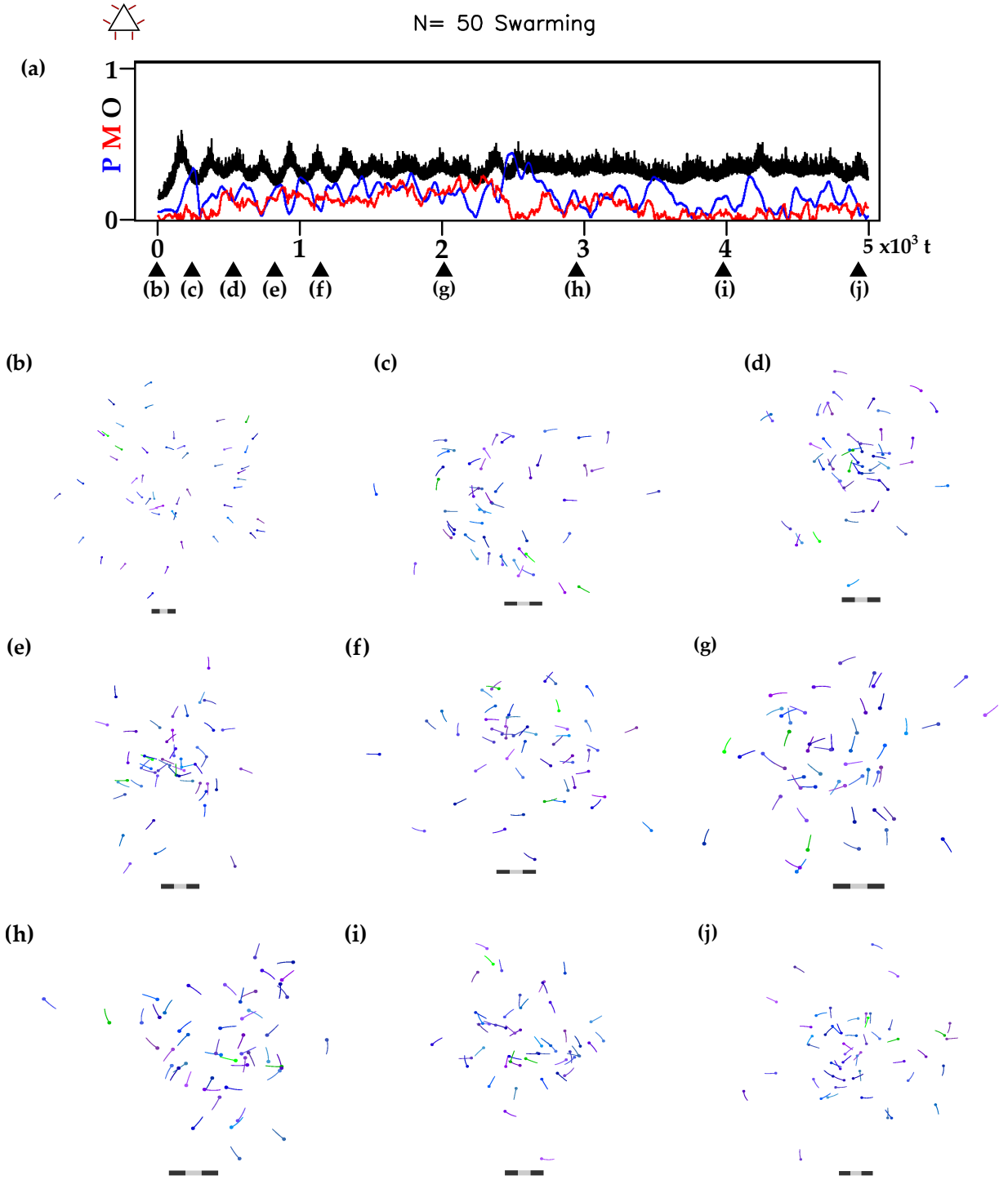


Fig. S 30. Swarming phase reproduction for a flock size of 50 individuals for the Omniscient model.  $k_\eta = 0.06$   $k_\eta = 0.0$   $k_\eta = 0.01$  and no blind angle, Video S10 (<https://youtu.be/NgZ8PCJ102k>).

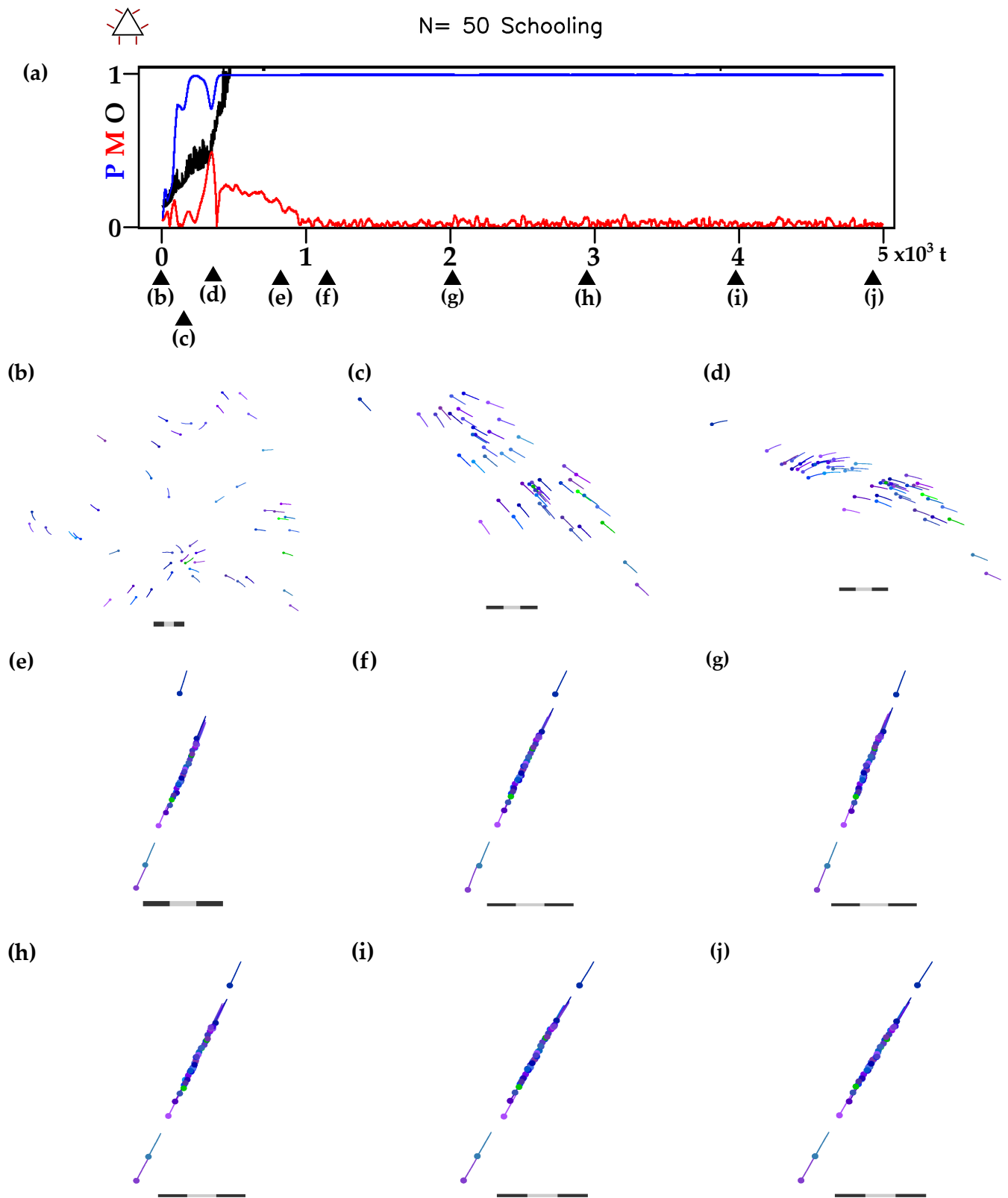


Fig. S 31. Schooling phase reproduction for a flock size of 50 individuals for the Omniscient model.  $k_\eta = 0.06$   $k_\eta = 0.2$   $k_\eta = 0.01$  and maximum blind angle, Video S11 ([https://youtu.be/x73EZqu\\_2a4](https://youtu.be/x73EZqu_2a4)).

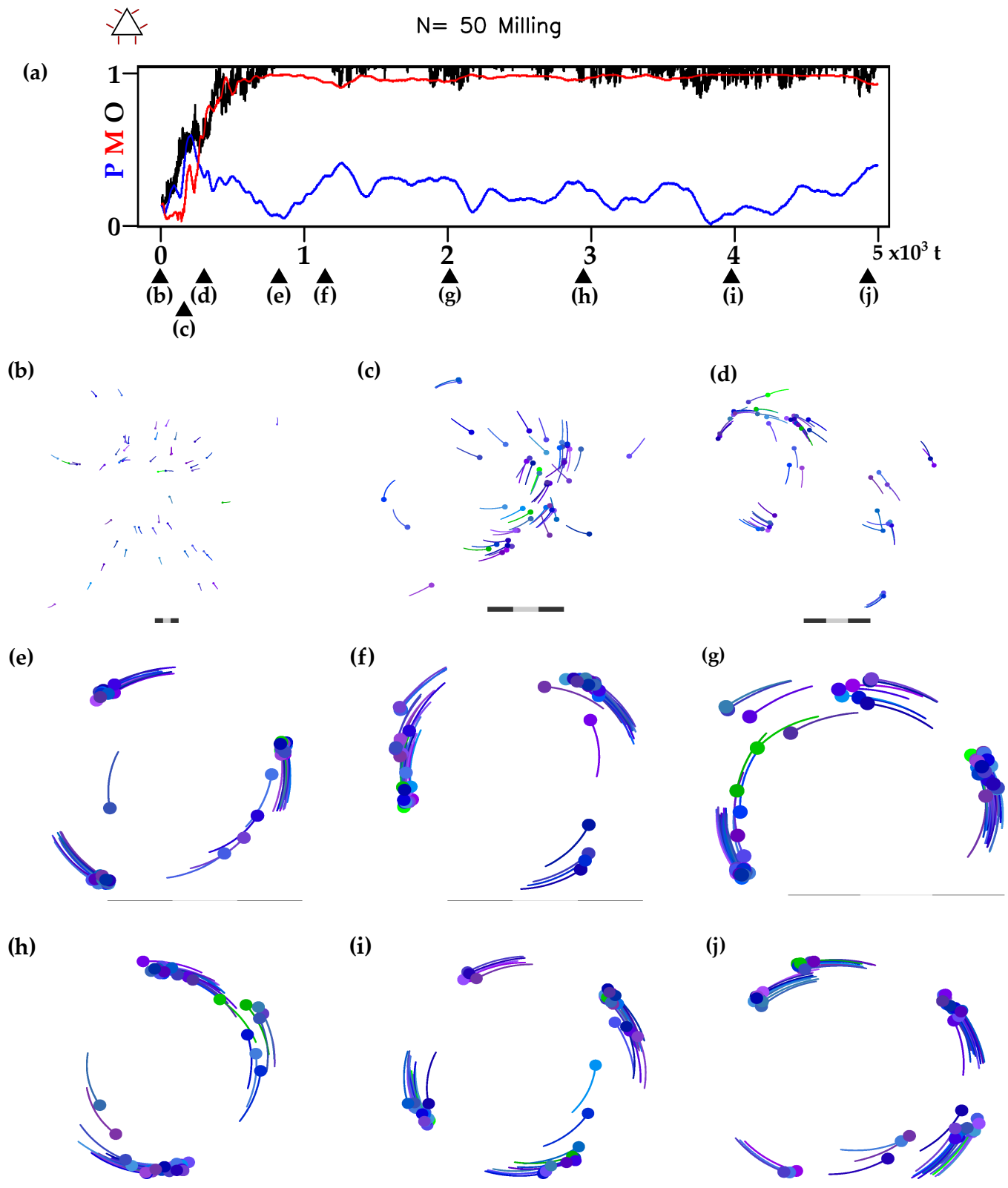


Fig. S 32. Milling phase reproduction for a flock size of 50 individuals for the Omniscient model.  $k_\eta = 0.06$   $k_\eta = 0.06$   $k_\eta = 0.01$  and maximum blind angle , Video S12 ([https://youtu.be/CWpTSvr\\_zBk](https://youtu.be/CWpTSvr_zBk)).

## S6 VIDEO SNAPSHOTS OF BISTABLE PHASE VISUAL MODEL (FIG. 4)

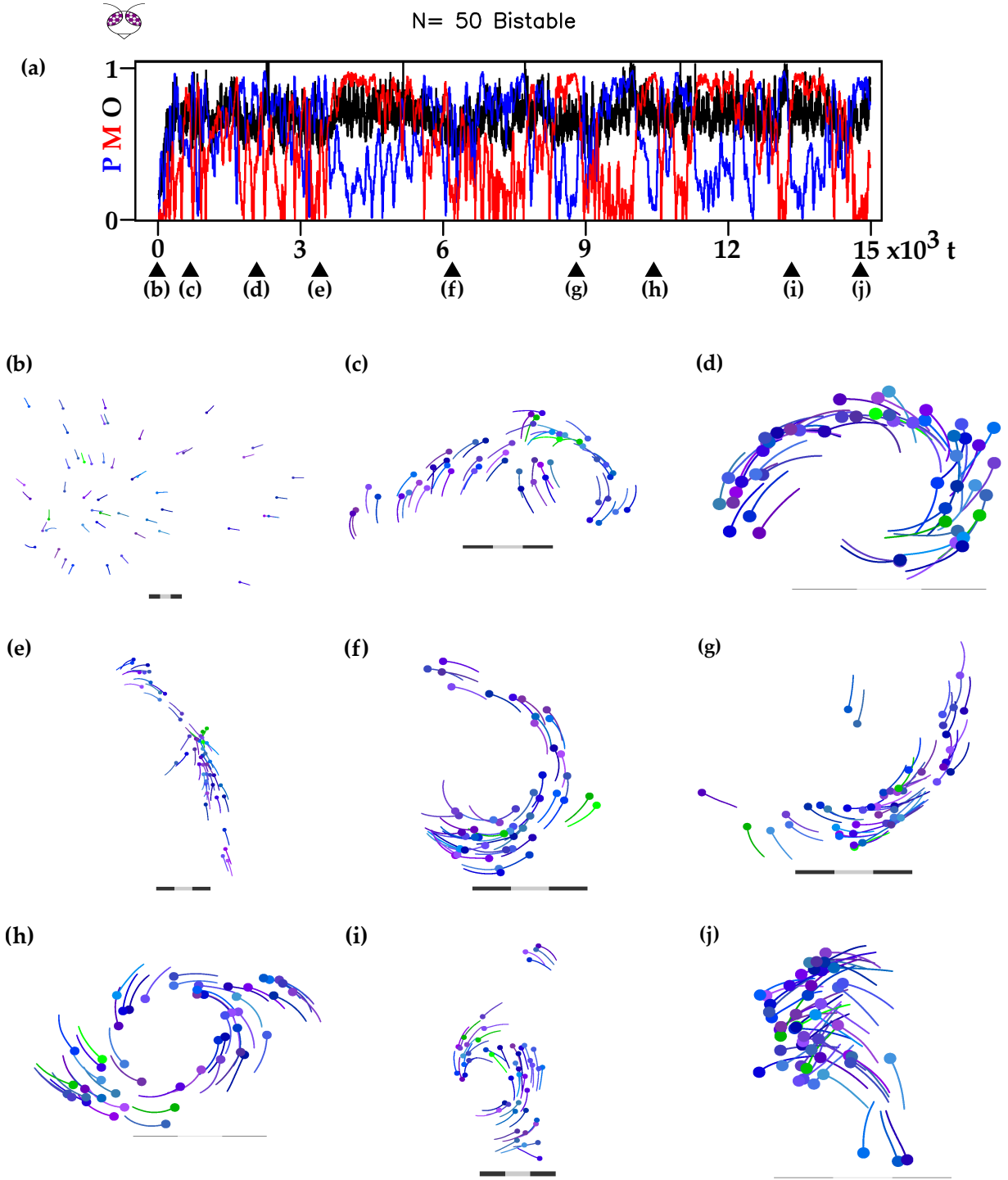


Fig. S 33. Bistable phase reproduction for a flock size of 50 individuals for the Visual model.  $k_\eta = 0.1$   $k_\eta = 0.2$   $k_\eta = 0.01$  and maximum blind angle, Video S13 (<https://youtu.be/JVCp65SVX2Q>).

[1] D. Castro, F. Ruffier, and C. Eloy, Flocking by



eye model, <https://gitlab.com/FlockingByEyeDiego/flockingbyeye> (), [Accessed 15-Feb-2024].

[2] D. Castro, F. Ruffier, and C. Eloy, Repository for model-

ing collective behaviours from optic flow and retinal cues, <https://osf.io/46fqx/> (), [Accessed 15-Feb-2024].