

Research



Cite this article: Lecheval V, Mann RP. 2023 Smart self-propelled particles: a framework to investigate the cognitive bases of movement. *J. R. Soc. Interface* **20**: 20230127. <https://doi.org/10.1098/rsif.2023.0127>

Received: 7 March 2023

Accepted: 5 July 2023

Subject Category:

Life Sciences—Mathematics interface

Subject Areas:

computational biology

Keywords:

movement, cognition, wall avoidance, normative model

Author for correspondence:

Valentin Lecheval

e-mail: v.lecheval@leeds.ac.uk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.6742215>.

Smart self-propelled particles: a framework to investigate the cognitive bases of movement

Valentin Lecheval and Richard P. Mann

School of Mathematics, University of Leeds, Leeds, UK

VL, 0000-0001-6041-2718; RPM, 0000-0003-0701-1274

Decision-making and movement of single animals or group of animals are often treated and investigated as separate processes. However, many decisions are taken while moving in a given space. In other words, both processes are optimized at the same time, and optimal decision-making processes are only understood in the light of movement constraints. To fully understand the rationale of decisions embedded in an environment (and therefore the underlying evolutionary processes), it is instrumental to develop theories of spatial decision-making. Here, we present a framework specifically developed to address this issue by the means of artificial neural networks and genetic algorithms. Specifically, we investigate a simple task in which single agents need to learn to explore their square arena without leaving its boundaries. We show that agents evolve by developing increasingly optimal strategies to solve a spatially embedded learning task while not having an initial arbitrary model of movements. The process allows the agents to learn how to move (i.e. by avoiding the arena walls) in order to make increasingly optimal decisions (improving their exploration of the arena). Ultimately, this framework makes predictions of possibly optimal behavioural strategies for tasks combining learning and movement.

1. Introduction

Animals display a diversity of complex behaviours, both social and in relation with their physical environment. Foraging, courtship, predator avoidance, learning and so on are a few examples of the behaviours in which animals engage—deciding at any particular time which activities are suited to their current social and physical contexts.

Natural selection is expected to shape the decision-making processes of animals such that their behaviours are optimized to accomplish fitness-relevant goals, within the constraints the animals face. In many circumstances, decisions are made through movements; an animal moving, for instance, towards a food source rather than towards a member of its group. Any constraints altering movements, such as locomotor or cognitive abilities or environmental forces (e.g. water or air flow or obstacles) may therefore be entangled in a decision-making process observed in fine details [1,2]. Decision-making processes combined with movement involve a great variety of behaviours across taxa, including social insects' building behaviour [3], the effect of the environment on animal movements [4], predator detection and avoidance [5], learning experiments in a spatial context [6] and collective motion [7,8]. To evaluate the optimality of decision-making processes when movement is involved is not trivial to formulate mathematically, especially given the high number of factors at play. There has been extensive research on Lévy-walk models to investigate what are the optimal movement patterns to optimize foraging [9–12].

This research effort, however, faces several limits to understand animal movements at the scale of the individuals and their behaviours in response to

biotic and abiotic stimuli [13,14]. Here, we investigate a class of models to explore how rules of interaction emerge from primary goals (finding food, avoiding predators or mating), given a set of constraints regarding cognitive abilities (in terms of perception or learning) or locomotor capacities. Our aim with this framework is therefore to find out whether or not certain behavioural interactions can arise from first principles considerations.

This framework is based on self-propelled agents imbued with cognitive abilities by means of an artificial neural network, which, given a set of stimuli regarding an agent's environment, determines how the agent moves [15,16]. During their lives, the fitness of agents is calculated from objective functions representing primary goals, and the parameters of this artificial neural network evolve via a simple evolutionary algorithm favouring agents with high fitness. This framework has been used in movement ecology, as a paradigm to study how animals explore and navigate in their environment on relatively long time and large spatial scales [15,17]. Similar approaches have explored the behavioural rules underlying the emergence of collective behaviours with reinforcement learning [18,19]. In these studies, the goal of the optimization process is to find individual-level behavioural rules responsible for specific collective configurations observed in natural flocking systems. However, in the existing literature of agent-based models in which agents have a cognitive model, the behavioural rules underlying collective patterns are not *emerging* from first principles (foraging, predation or reproduction), and the reasons *why* animals would behave likewise are still unknown. We note an early attempt of addressing these goals in the literature, which lacked details in the quantification of the emerging behaviours and collective patterns [20]. Investigating the rationale behind the proximate causes of animal spatially embedded decisions is crucial to get an integrated understanding of animal movements, including collective behaviours, even on small temporal and spatial scales.

In this article, we introduce this class of models for the study of animal behaviour in general and of the cognitive bases on animal movements in particular. To stress our intention to use this class of models for the study of animal behaviour, we name it *smart self-propelled particles*, an explicit reference to self-propelled particles, agent-based models widely used to investigate collective motion [21–23], although the study of social interactions is beyond the scope of this article. Here, we describe the framework in detail, and show how to use it for the study of spatially embedded animal behaviours. Furthermore, we demonstrate and discuss an approach to develop the models based on empirical data and results. In the present work, we will investigate the optimal movement behaviours of single animals in a bounded space, given different fitness functions. We will also investigate the respective effects of different locomotor constraints in affecting the emerging rules of interaction.

2. Methods

The framework we introduce for the study of animal behaviour consists of a class of agent-based models, with cognitive agents. 'Cognitive agents' refers to the movements of agents being controlled by their brain (an artificial neural network) given a set of stimuli. The parameters of these brains are optimized with a genetic algorithm to achieve a particular task, given an ecological

context as well as locomotor and cognitive abilities. In this class of models, we expect the structure of the brains (artificial neural networks) to be important in determining the outcome of their predictions. The structure concerns the number of neurons, their arrangement (distributed in one or several so-called hidden layers) and their connections. One possibility to design an architecture of the artificial neural network appropriate to the problem under consideration is to optimize the structure of the network with a genetic algorithm, as part of the brain optimization [24]. Another simpler approach is to incrementally increase or decrease the number of neurons and to constrain the structure with empirical approach. The latter is the chosen approach in the current study. It allows to focus on the effect of the optimization of the network parameters in this first investigation but also to closely monitor the effect of structural changes and to potentially infer the cognitive difficulty of a task.

2.1. Case study: the emergence of a wall-avoidance behaviour

In this study, we will focus on one particular situation: a single individual moving in an enclosed set-up (e.g. a tank or a cage). Although this seems to be an artificial context, this situation reproduces the context of animals in captivity, enclosed in an arena or a tank, a situation of wide interest in Ethology, given the number of experimental designs (for instance involving learning or personality research) set in such environments. Moreover, there are ecological contexts in the wild in which the study of movements when facing boundaries are particularly relevant, such as flying insects avoiding spider webs during their flight. By choosing this situation in this study, we pave the way towards more complex contexts in the future.

Here, cognitive agents are navigating alone (i.e. without interactions with others) in a square arena. They have a lifespan and their trial either stops (i.e. they die, in genetic algorithm jargon) when hitting the arena walls or at the end of their lifetime. We explore how different constraints and goals affect the emerging rules of interaction of the agent with the walls and how this relates to empirical results found in fish.

2.2. Perception, decisions and movements of agents

2.2.1. Perception

Agents perceive two cues in their environment: their distance from the closest point of any wall and their angle to that wall (figure 1a). Agents have a simulated brain consisting of an artificial neural network of two input neurons, a hidden layer and one output neuron. The two input neurons correspond to each cue the agents perceive (distance and angle relative to the closest wall), normalized between 0 and 1.

2.2.2. Decisions

Decisions regarding movement are made at each time step according to the cues (distance and angle relative to the closest wall) the agent perceives. These cues are transmitted to the input neurons of the artificial neural network of the agent (a multi-layer perceptron), which predicts one output value $o_1 \in [0, 1]$ given by the output neuron. The activation function used in these artificial neural networks is the sigmoid function $f(z) = 1/(1 + e^{-z})$, where z is the weighted sum of the input connections, added to the bias parameter of the focal neuron (figure 1b). The output value of o_1 controls the direction changes of the agent.

The structure of the artificial neural network (i.e. the number of neurons in hidden layers) controls the complexity of the behaviours displayed by the agents. There is therefore a trade-off between the emerging complexity, the generalization power of the networks and the difficulty to optimize the values of the

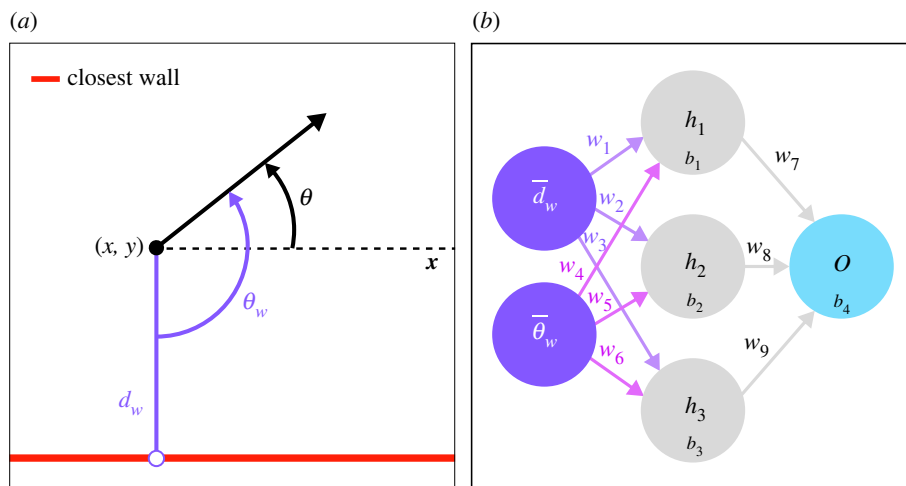


Figure 1. Notations used to depict the location of an agent relative to the closest wall and structure of the artificial neural network. (a) The agent (black dot) is located in (x, y) with a heading θ . Its distance and angle to the closest point of any wall (white dot) are, respectively, d_w and θ_w . (b) Structure of the artificial neural network of reference in our study, with two input neurons receiving the normalized distance and angle relative to the closest point of any wall \bar{d}_w and $\bar{\theta}_w$, one hidden layer with three neurons and one output neuron. The 13 parameters are made of nine weights w_i and four biases b_i .

artificial neural network parameters due to the number of parameters. In this study, we restrict our analysis to networks with one hidden layer. We evaluate, qualitatively, the optimal number of neurons in one hidden layer, between one and six neurons, by ensuring the chosen number of neurons can at least reproduce the rule of interaction with the wall of actual fish.

2.2.3. Movements

Based on the output value $o_1 \in [0, 1]$, a coefficient $c_{\text{steering}} = o_1 - 0.5$ is calculated ($c_{\text{steering}} \in [-0.5, 0.5]$) and translated into a turning angle $\Delta\phi \in [-\pi, \pi]$,

$$\Delta\phi = c_{\text{steering}} \cdot i_{\text{steering}},$$

with i_{steering} the steering increment $i_{\text{steering}} \in [0, 2\pi]$, a constant parameter of the model setting the maximal angle turned by agents at each time step. The heading of the agent at time t is therefore $\theta' = \theta + \Delta\phi$. The position (x, y) of the agent at time t is then updated at each increment of time (set to 1 in an arbitrary time unit),

$$\begin{cases} x' = x + s \cos(\theta') \\ y' = y + s \sin(\theta') \end{cases}$$

with s the speed of the agent, which is a constant parameter in this study.

2.3. Genetic algorithm

2.3.1. Initialization

At each generation, the same number n of individuals is generated. To optimize the exploration of the parameter space across the n agents, each parameter of the artificial neural networks (i.e. weights and biases alike) is initialized to a value in $\mathcal{U}(-50, 50)$ with the Latin hypercube sampling method. We use the Python implementation provided by the Surrogate Modeling Toolbox Python package, with the `maximin` criterion, which maximizes the minimum distance between points and places the points in a randomized location within their interval [25]. Each agent is simulated four times with different initial locations, randomly taken in the arena, with one initial heading $h \in (0; -\pi; -\pi/2; \pi/2)$. A fifth simulation round is executed with the initial location always set to the centre of the arena, and a heading $h = 0$. This additional round is not considered to calculate the fitness of agents but is used to compare the score of agents across generations with a constant initial location, given that the set of initial locations influences the final score in the other four rounds.

2.3.2. Selection

During each round, agents move in a square arena. We have investigated two distinct goals for agents: to survive for the longest time or to maximize the exploration of the arena. With the surviving goal, the score of agents is the sum of time steps during which they managed to stay within the boundaries of the arena. In the exploration goal, agents score one point each time they discover a new cell of the discretized square arena. This mimics a situation where an animal explores an environment in which a resource, such as food, is randomly distributed in space, with a constant probability to be found and collected in any location.

In both goals, the round of an agent ends when the agent goes out of the arena boundaries or if its lifespan exceeds the maximum lifespan of agents (T_{max}). When all agents have finished their four rounds, their scores S are summed over the rounds and transformed into a fitness.

We investigate the effect of locomotor constraints by focusing on the effects of speed and turning. In many species, turning is costly and affects movement strategies [26–28]. We introduce a penalty to turn, proportional to the squared amount of angle turned in order to penalize very sharp turns. The penalty is set to a constant p_t , which is applied at each turn an agent makes such that the total cost to turn across all rounds is $C_t = \sum p_t |\Delta\phi|^2$. The penalized score $S_p = S - C_t$ is then interpolated between 0 and 1, where 0 stands for the minimum score of the generation the current agent belongs to, and 1 for the maximum of the current generation.

The normalized fitness of an agent i is

$$f_i = \frac{S_p}{\sum_{j=1}^n S_p}.$$

The new generation of n agents is generated from the previous generation by favouring the highest fitnesses. This sampling is stochastic (with replacement), with probability of getting picked proportional to the normalized fitness. Each individual picked to get through the next generation gets the parameters (weights and biases) of its artificial neural network copied and mutated. Small and large mutations occur randomly (33% of the parameters affected by small mutations, 1% by large mutations), by applying a Gaussian noise of standard deviation 0.5 and 5, respectively. These mutation levels have been set by trial and error and the standard deviations have been scaled to the order of magnitude of the parameter values which lie in $[-50, 50]$.

Table 1. List of parameters of all simulations. Each row stands for a set of simulation including 60 trials with different initial conditions, each trial having 20 000 agents simulated for 150 generations with 5000 time steps each. HL, hidden layer.

name	max. turning angle	no. of neurons in HL	goal	speed	turning penalty
parameters of reference	6.28	3	explore	50	0
survive objective function	6.28	3	survive	50	0
parameters with a turning penalty	6.28	3	explore	50	0.33
parameters with a slow speed	6.28	3	explore	5	0

2.4. Simulations

We simulate the model to investigate its predictions. Each generation consists of $n = 20\,000$ agents navigating a 600×600 pixels square discretized in a 400-cell grid. Each round can last up to 5000 time steps if the agent does not leave the square boundaries beforehand.

Each simulation runs for 150 generations and is repeated in 60 trials. Parameters explored in our simulations are listed in table 1. Simulations are implemented in Python 3.10 and analyses performed with both Python and R v. 4.1.3. We infer rules of interaction with the wall by looking at the average turning angle as a function of the distance and angle to the closest wall. Average turning angles are all calculated using the circular mean

$$\bar{\alpha} = a \tan 2 \left(\frac{\sum_{j=1}^n \sin \alpha_j}{\sum_{j=1}^n \cos \alpha_j} \right).$$

2.5. Empirical data

We use results from a study published on the rummy-nose tetra *Hemigrammus rhodostomus* [29]. This study provided a method to infer rules of interaction with a wall from trajectories of a single fish. We use these results (i) to validate the structure of our artificial neural network, (ii) as an empirical reference to our own simulated results, and (iii) use their method to analyse simulated trajectories to gain further understanding on the relationship between behaviour, trajectories and analyses. We use a simplified rule of interaction of wall avoidance inspired by [29] describing how fish avoid walls by turning away from them at short distances. This rule of interaction can be depicted by the product of two functions, one standing for the direction of wall avoidance, odd function of θ_w and a second one standing for the decay of the intensity of the wall repulsion as distance to the wall r_w increases, such as the angle turned by fish due to wall avoidance is

$$\Delta\phi = \sin(\theta_w) \exp\left(-\left(\frac{d_w}{l_w}\right)^2\right), \quad (2.1)$$

where l_w is a parameter controlling the range of the wall interaction.

3. Results

Before analysing simulations in which we varied the goals and constraints of cognition and locomotion, we validated the structure of the artificial neural networks used as artificial brains of the agents in our model. In other words, the capabilities of an artificial neural network structure in predicting empirical results was evaluated by varying the number of neurons in the hidden layer. Our goal was to have a minimal number of parameters (i.e. a small artificial neural network) to facilitate convergence to optimal solutions and to get a parsimonious

and tractable model, while having an artificial brain structure allowing actually observed behaviours to emerge.

3.1. Validating the artificial neural network structure

To validate the structure of the artificial neural networks, we checked what would be the least complex artificial neural network that could be trained to predict a known and measured wall-avoidance behaviour of fish swimming in a circular tank [29]. The minimal artificial neural network structure consists of one hidden layer made of one neuron. We trained artificial neural networks with one hidden layer with a number of neurons ranging from one to six neurons. Each artificial neural network is trained to fit the (simplified) rule of interaction with the wall of *H. rhodostomus*. We assessed that a hidden layer made of at least three neurons is sufficient to broadly reproduce the behaviour measured in *H. rhodostomus* (figure 2). Networks with fewer neurons could not completely reproduce the empirically measured behaviour, and networks with more neurons did not quantitatively improve the fit (electronic supplementary material, figure S1). This validation process is effectively a method to evaluate the cognitive difficulty of any given task. By considering iteratively the effect of the number of neurons, the minimal number of neurons required to achieve a particular task in a specific manner can be inferred and used as an effective measure of the cognitive difficulty of a task. We therefore use this structure (one hidden layer made of three neurons) in the subsequent sections.

3.2. Influence of the objective function

Objective functions define how the fitness of agents is calculated. Fitness is used to favour agents with the highest fitness scores in generating the following generation. To test the effect of the objective function on the emergence of rules of interaction, we defined two different ecological contexts of the goal of a single agent in a bounded environment:

- (i) the agent has to survive for the longest period of time (i.e. the largest number of time steps);
- (ii) the agent has to explore its environment.

In both situations, as soon as an agent leaves its bounded environment (i.e. its arena), its life stops and so ends its capacity to improve its fitness.

3.2.1. Survive for the longest period of time

In this first objective function, the fitness score of an agent corresponds to the number of time steps in which the agent stays within the boundaries of its arena. The resulting best rules of interaction with the wall averaged across generations

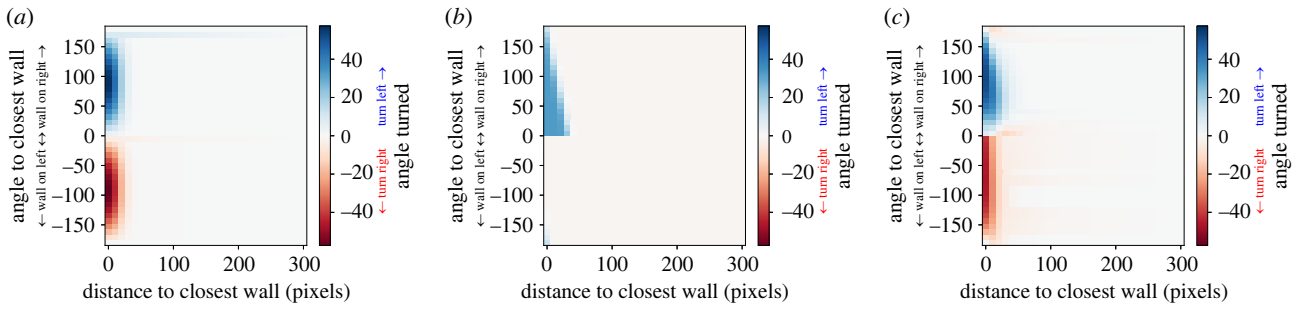


Figure 2. Possibilities of the artificial neural networks used as brains of agents. (a) Simplified empirical rule of interaction of fish with the closest point of a circular wall found in *H. rhodostomus* [29]. The figure shows how much an agent turns right (red) or left (blue) depending on its position (angle and distance) relative to the closest point of the wall. An angle of 0 represents an agent orthogonally facing the wall. (b) Fit of an artificial neural network with one neuron in the hidden layer with the empirical rule of interaction (c). (c) Fit of the artificial neural network of reference (figure 1b) with the empirical rule of interaction (a).

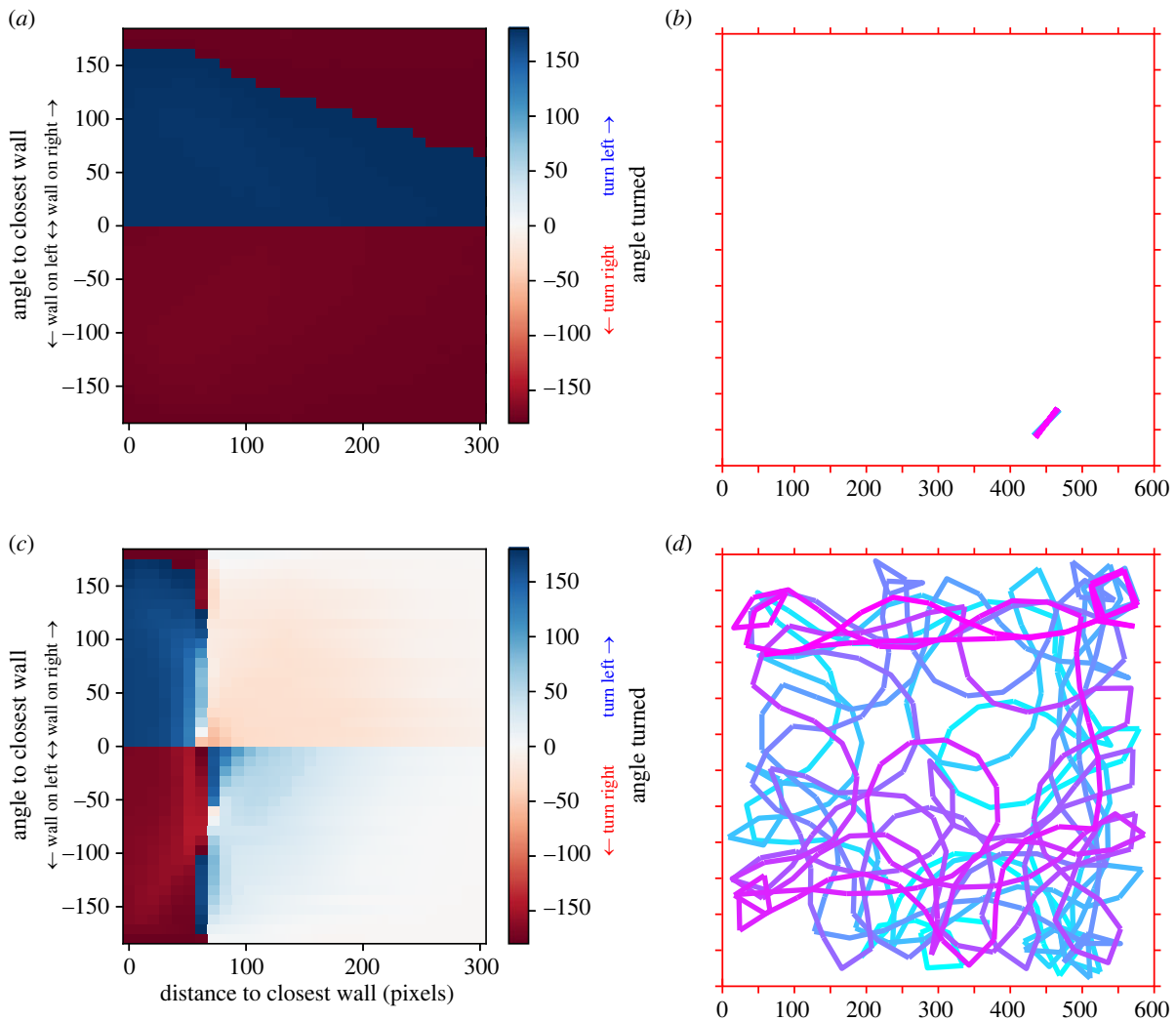


Figure 3. Effect of the objective function on the emergence of the interaction rules. (a) Best rules of interaction with the wall averaged across generations and trials when the objective is to survive for the longest period of time. (b) Example of a trajectory of an agent trained with the objective to survive for the longest period of time. The trajectory effectively consists of overlapped segments, minimizing the risks of hitting the edges of the arena. Colour variation depicts the passage of time. (c) Average of the best rules of interaction with the wall when the objective is to explore the arena. (d) Example of a trajectory of an agent trained when the objective is to explore the arena. Colour variation depicts the passage of time.

and trials (figure 3a) shows almost no effect of the distance to the closest point of the wall. The trajectories of optimal agents in this situation consist of back-and-forth segments of the shortest length (50 pixels) obtained by doing U-turns at the end of each segment bout (figure 3b). The directions of these sharp U-turns made to stay close to a segment-like trajectory are on average performed away from the wall (i.e.

turning left when the closest wall is on the right-hand side and right when the closest wall is on the left-hand side). Trying to move back-and-forth on a segment is a strategy effectively minimizing the explored area for agents moving at a constant speed (i.e. not being able to slow down or to stop), therefore minimizing the risks of hitting a wall. This is effectively analogous to a freezing behaviour for

agents that cannot stop moving. Freezing behaviour is a very common strategy of animals facing predatory threats [30,31]. Although this population-level strategy differs greatly from the empirical rule of interaction observed in fish, we already see the emergence of the anti-symmetry resulting in agents turning away from the closest wall.

3.2.2. Explore the arena

An alternative ecological context is for the agents to increase their fitness when exploring this bounded environment. In this case, the fitness score of an agent corresponds to the number of cells of the discretized arena explored *at least* once (i.e. exploring the same cell several times does not increase the score).

We find that in this context, there is a strong interaction between the angle to the closest point of the wall and the distance to the closest point of the wall (figure 3c). In agreement with the rules of interaction measured in experiments with *H. rhodostomus*, we find that at the population level, when close to the wall, the direction at which an agent turns (i.e. left or right) depends on its angle to the wall, turning left (respectively, right) when the wall is at its right-hand side (respectively, left-hand side). This corresponds to a wall-avoidance behaviour, as measured empirically. We also note a weak interaction of attraction to the wall at distances from the wall corresponding to slightly more than the smallest path length of 50 pixels. The optimal trajectories found in this context result in clear wall-following behaviour (figure 3d), where most of the time is spent near the edges of the arena with a few visits towards the centre of the arena to increase the number of cells of the discretized arena explored. This is an already known result found in fish and bacteria with run-and-tumble motion, in which a wall-avoidance behaviour coupled with discrete reorientations actually results in wall-following behaviour [29,32,33]. We find a similar behaviour when the same agents are set to move in a smaller or a larger arena (electronic supplementary material, figure S2), supporting the robustness of the emergent strategies found by agents. This alternative objective function produces at the population level interaction rules which are very different from the previous one, in which agents simply needed to survive for as long as they could. This highlights how the definition of the ecological context in our framework affects the resulting optimal rules of interaction.

We now turn to finer quantitative details regarding the movements of agents optimizing the exploration criterion (figure 4; electronic supplementary material, figure S3). We find that in most of the cases, agents are found at a distance of more than 50 pixels away from the closest wall (figure 4a). The exception is when agents are facing the closest wall (angle to the wall around 0): in these cases, agents can be found closer to the wall. Three main behaviours regarding angle turned can be identified: very small angle turns, angle turns close to 90° and U-turns (figure 4b). When agents are close to the wall (within 50 pixels from the wall), they favour turning away from the wall, specifically with large angles, around 90° and, mainly, complete U-turns (figure 4c). For agents far away from the wall (at least at 150 pixels from the wall), we observe the three modes of the global distribution (small angles, right angles and U-turns), with the predominance of small turning angles (figure 4d). We note that the distribution of these small angles (less than 60° in magnitude)

is compatible with a normal distribution (figure 4d, inset). This shows the emergence of an apparently stochastic behaviour at the population level—with agents turning as often to the right as to the left when far away from the wall—in a strictly deterministic model of movement. This is in agreement with empirical results found in *H. rhodostomus* [29].

3.3. Individual-level interaction rules versus population-level interaction rules

The average behaviour shown figure 3c is not the one actually used by the agents (figure 5). We can indeed find equally good performances by individuals whose rule of interaction with the wall greatly differ from the average one. The behaviour of these agents is very similar in regions of the input variable space under strong selection—here, when agents are close to a wall, we see a strong wall-avoidance behaviour. However, when the agents are far from a wall and therefore in regions under less selection pressures, we see more variation; for instance, the distance at which individuals start to avoid the wall or the strength or existence of a parameter region in which agents are attracted to the closest wall. This leads to a diversity of behaviours and forms of trajectories even though, on average, agents all successfully avoid the wall by turning left when the wall is on their right and vice versa. This suggests the possibility for diversity in the presence of strong selective pressures, where early experience leads to a diverse yet equally satisfying pool of solutions. It is worth noting that this diversity is difficult to note in empirical data and generally overlooked. Most of the empirical studies investigating rules of interaction in animals report results obtained from behaviours averaged across individuals, trials, time of the day and days of experiment [34–37], preventing the finding of diverse solutions from different individuals or depending on statuses hidden to the experimenter, such as the state of hunger of a specific individual.

3.4. Effect of locomotor constraints

We study the effect of locomotor constraints by investigating the costs associated with turning (figure 6a–c; electronic supplementary material, figure S4). We find that overall, the optimal behaviour stays similar: agents still need to turn left when the wall is at their right-hand side and vice versa (figure 6a). We note that agents start turning earlier (i.e. further away from the wall) when they get to face the wall and turn later on (i.e. at closest distances from the wall) when they do not directly face the wall, unlike the rule of interaction of reference (figure 3c). This is probably the result of agents minimizing the instances when they need to turn a lot to avoid the wall.

We also investigated the effect of speed by running simulations in which agents move 10 times slower than in the conditions of reference—namely five pixels per time unit instead of 50 (figure 6d,e; electronic supplementary material, figure S5). This change in the locomotor capacities relative to the size of the bounded space affects mainly the distance at which agents start to avoid the closest wall (figure 6d). Agents also make much smaller turns, resulting in a distribution of angle turned very different from the distribution obtained with the parameters of reference (electronic supplementary material, figure S5a). This results in trajectories being less tortuous than other rules of interaction found for

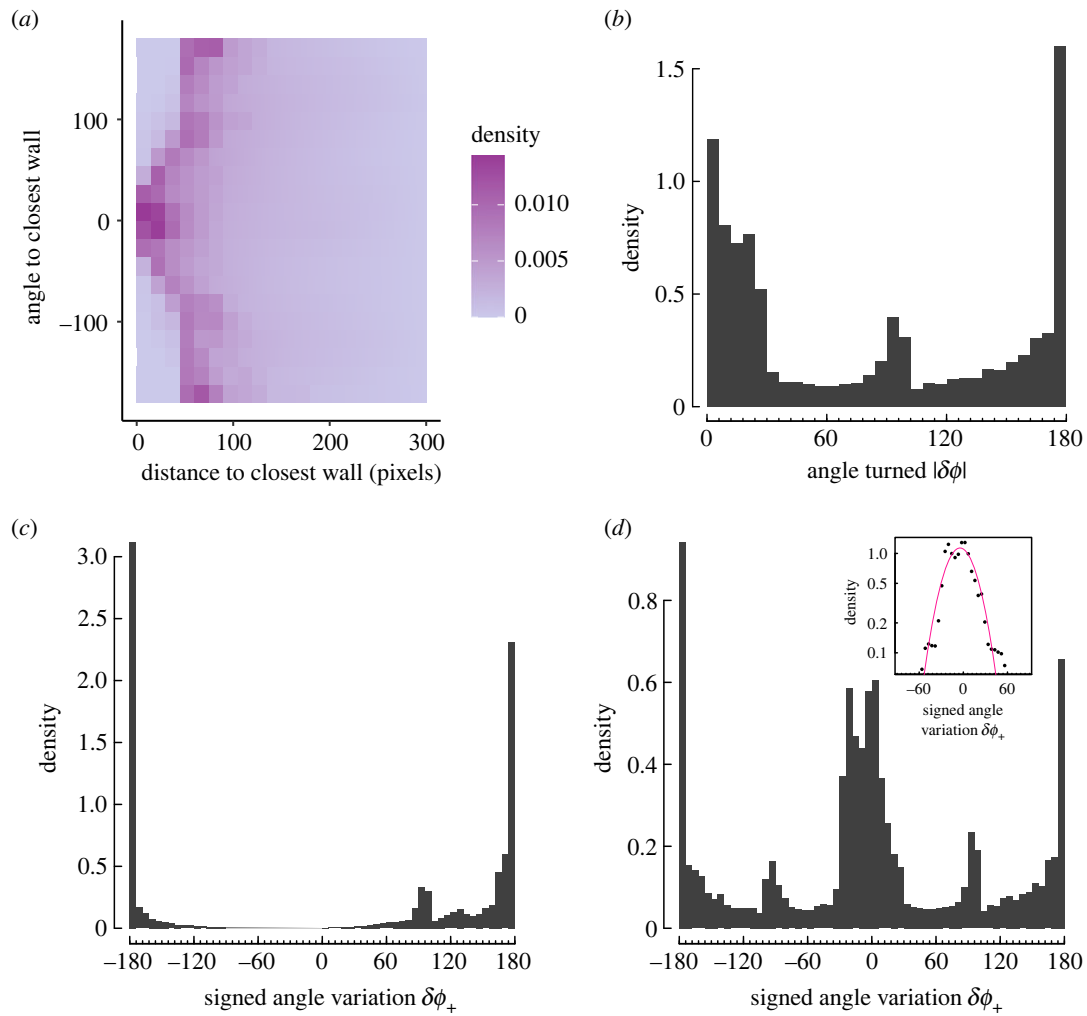


Figure 4. Statistics of exploration and angle changes. (a) Density of presence of the best agents as a function of distance and angle to the closest wall. (b) Distribution of the absolute value of the angle turned. (c) Distribution of the signed angle variation restricted to agents close to the wall (i.e. less than 50 pixels from the wall), $\delta\phi_+ = \delta\phi \times \text{sign}(\theta_w)$, positive for angle turned away from the wall, negative otherwise. (d) Distribution of the signed angle variation restricted to agents far away from the wall (i.e. more than 150 pixels from the wall). Inset: zoom of the distribution between -60 and 60° . Purple line shows a Gaussian fit.

previous constraints (figure 6f). This leads agents to being distributed closer to the wall on average (electronic supplementary material, figure S5b) with two clear angles relative to the wall around 54 and 132° .

4. Discussion

We find that the objective function affects the movement behaviours that emerge, with a great variety of solutions found to the same problem. In particular, we show that a wall-following trajectory emerges from wall avoidance behaviour of agents actually optimizing the exploration of their bounded environment. They do so by successfully learning how to avoid the obstacles of their environment, based on very simple local cues. The emerging rules of interaction with the wall resembles the one found in the fish *H. rhodostomus*, where single fish were studied when swimming freely in small circular tanks [29]. At the population level, we find indeed a clear anti-symmetric wall avoidance behaviour at close distances from the wall, in agreement with empirical data [29]. This is achieved with an extremely simple cognitive model made of an artificial neural network with only 13 parameters (weights and biases), and whose input is also kept

very simple—only information regarding the position of the agent relative to the closest point of any wall of the arena is provided. We note that our cognitive model is not a direct representation of an actual brain. Future work could use biologically motivated brain structures of the network [1,2] to investigate the influence of neural network structures on the behaviours emerging from a smart self-propelled particles model. Locomotor abilities such as the speed or turning abilities also affect the emerging behaviours, changing both the interplay between distance and angle to the wall in determining wall avoidance and the resulting trajectories and arena exploration.

We stress the importance of combining the use of this class of models with empirical data. Such a theoretical approach provides multiple strategies, many of them potentially being biologically unrealistic. Empirical data therefore provide additional context and constraints to guide the development of meaningful models. Moreover, adding constraints reduces the diversity of optimal strategies successfully solving the problem under study and helps exclude suboptimal solutions.

Our results obtained with an objective function requiring individuals to maximize their time spent without touching the wall suggest that freezing behaviours are compatible

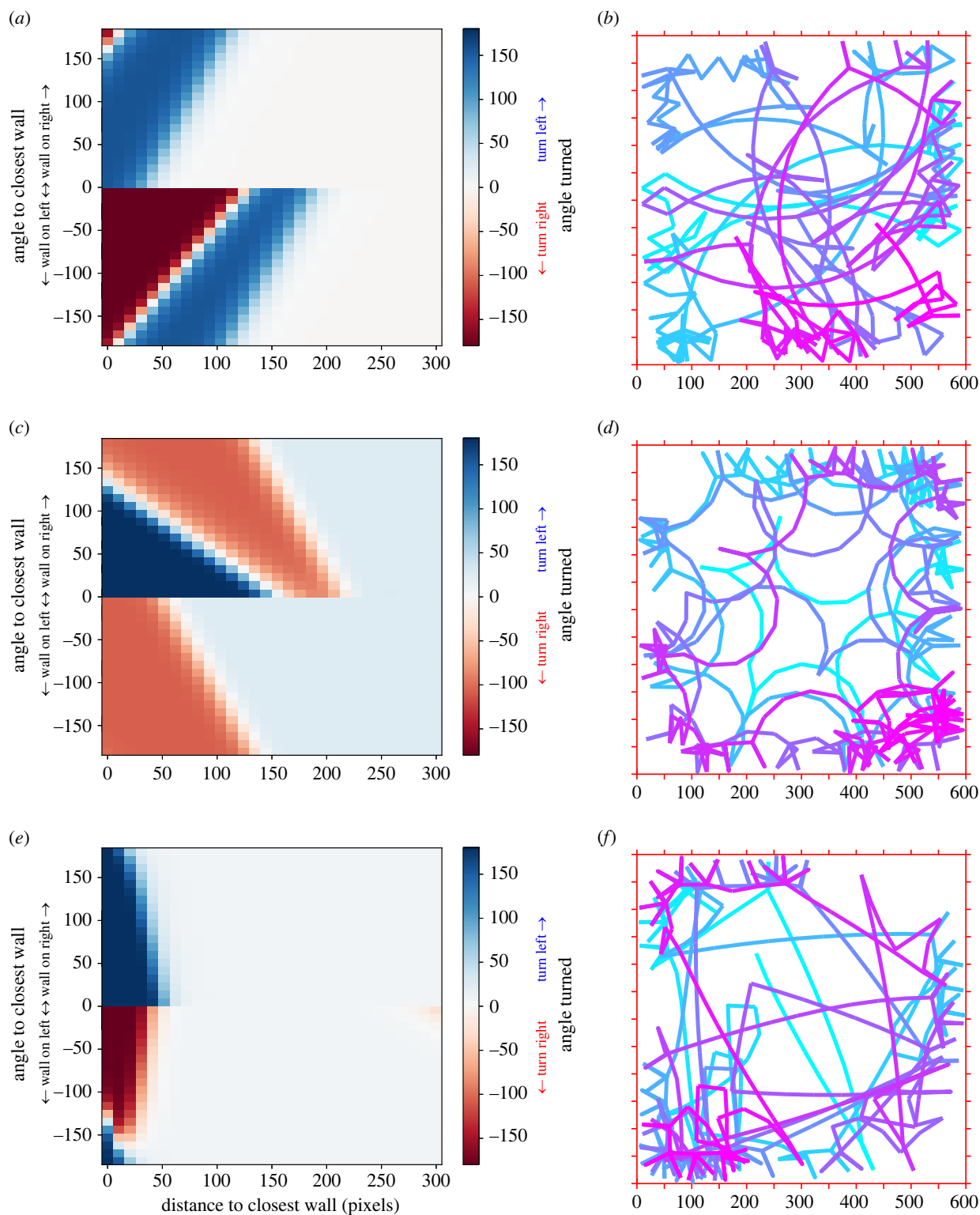


Figure 5. Diversity of equally efficient strategies to solve the exploration problem. (*a,c,e*) Angle turned as a function of the angle and distance to the closest point on the wall of three distinct agents, solving the exploration problem with the maximum exploration score. (*b,d,f*) Trajectories corresponding to, respectively, rules of interaction (*a,c,e*). Colour variation depicts the passage of time.

with individuals trying to favour survival against exploration when facing extreme threats [30,31].

The good agreement between our findings and empirical results in fish occur at the population level, on aggregated and averaged simulated data. However, this fit overlies a complexity largely unexplored in biology. Our simulation results look similar once averaged across multiple agents but the individual behaviours actually encompass an important variety of strategies for exploring the arena while avoiding the wall, although not in high-risk zones, which

are less diverse in the exhibited behaviours. The discrepancy between individual- and population-level behaviours and the heterogeneity of behaviours at the individual level are rarely investigated when looking for rules of interaction, while it could occur that none of the individuals actually behaves as the average behaviour. Identifying whether the average behaviour is exhibited by many individuals may allow to identify behaviours emerging from either high or low selective pressures. The very simplicity of our model and the convergence to many different solutions stress that particular

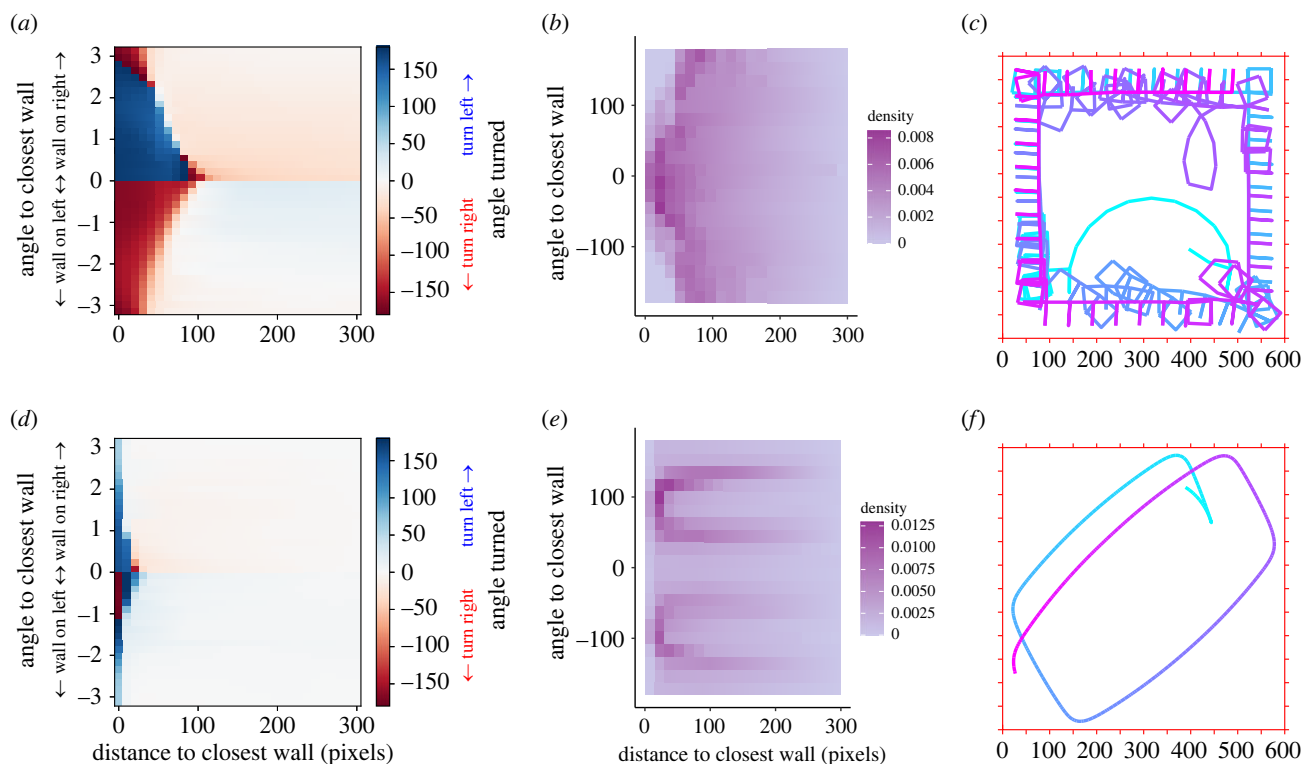


Figure 6. Effect of locomotor constraints on the optimal rule of interaction with the wall. (a–c) Results when introducing a turning penalty every time the agent turns. (d–f) Results when agents are moving 10 times slower than the speed of reference (five pixels per time step in contrast to 50). (a,d) Best rules of interaction with the wall averaged across generations and trials. (b,e) Density of presence of the best agents as a function of distance and angle to the closest wall. (c,f) Example of trajectories of the best agents of each locomotor constraint. Colour depicts the passage of time.

attention should be given to heterogeneity of behaviours at the individual level when analysing empirical data.

In our study, we find apparently stochastic behaviour emerging at the population level: when agents are moving far away from the closest wall, their average behaviour is Gaussian and apparently stochastic at the population level—they turn as often to the left as to the right on average. This apparently stochastic behaviour emerges in a model of movement which is strictly deterministic: when submitted with the same cues, the agents will always move exactly in the same way. Stochastic behaviours are often interpreted as the signature of free will or noise in empirical data [29]. Our result here stresses that, in practice, one needs to be careful when pooling empirical data from multiple individuals together, e.g. to calculate an average behaviour: apparent stochasticity may arise even when agents with various strategies have a rationale and unnoisy behaviour. This underlines the importance of providing evidence that animals exhibit a stochastic behaviour at the individual level rather than at the group or population level.

Our results highlight the value of our approach to test the direct effect of an ecological context in evaluating the optimality of rules of interaction of animals with their social and physical environments. It is an holistic approach providing a class of models to explore how animals solve tasks, fulfil their needs, in the presence of complex trade-offs and constraints brought by their environment and their own perceptory, cognitive, physiological and locomotor machineries. This class of models therefore fills a gap by providing a method to investigate systems combining two distinct aspects of spatially embedded behaviours: decision-making and movements. In other words, this class of models allows integrated research of the cognitive bases of animal movements, specifically by unifying two research bodies often divided. For instance, in the

field of collective motion which studies how groups of animals such as bird flocks or fish schools coordinate and move in a synchronized manner, the literature is split between studies focusing on the decision-making underlying movements in groups and studies looking at the spatial mechanisms of movements. Recent studies, however, demonstrate how the interplay between movement and how individuals represent space underlies a continuous feedback between decision-making and movements [1,2]. Here, we propose to explore optimal decision-making *under the constraints* of spatially embedded tasks—that is taking into account the respective effects of intent, perception, locomotion and obstacles. Although this class of models pre-existed our study [15], it was mainly used to study the optimality of animal movements on large spatial and temporal scales, i.e. on scales on which the resolution of empirical data does not allow the study of cognition—or lacked analysis of emerging behaviours in relation to empirical data [16,20].

We show here how this framework can be used with empirical data on fine temporal and spatial scales, to both constrain the exploration and guide the modelling choices to successfully study the cognitive bases of animal movements. The smart self-propelled particles approach potentially benefits several scales of description. In addition to its use in behavioural ecology to investigate animal movements, for instance investigating the optimal spatial distribution of animals given a set of constraints, the class of models can also be applied in Ethology, for instance to compare experimental results of learning experiments involving movements with theoretical predictions. While most empirical studies are descriptive and do not really explain why animals behave the way they do, we propose here, from the descriptive empirical knowledge of animal behaviours, to reverse engineer the

fitness functions shaping the behaviours of animals. Doing so will help understand why animals adopt the rules of interaction predicted by the smart self-propelled particles framework. On the other hand, should the predictions differ from empirical observations, it will help formulating an improved model by adding or removing constraints or changing the goal agents have to achieve, essentially inferring the fitness function of the animal. This study paves the way towards the development of normative models aimed at understanding better the cognitive bases of decision-making of moving animals with an holistic and integrated approach allowing to evaluate the respective effects of cognitive, locomotor and environmental constraints.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

References

- Sridhar VH, Li L, Gorbonos D, Nagy M, Schell BR, Sorochkin T, Gov NS, Couzin ID. 2021 The geometry of decision-making in individuals and collectives. *Proc. Natl Acad. Sci. USA* **118**, e2102157118. (doi:10.1073/pnas.2102157118)
- Oscar L, Li L, Gorbonos D, Couzin ID, Gov NS. 2023 A simple cognitive model explains movement decisions in zebrafish while following leaders. *Phys. Biol.* **20**, 045002. (doi:10.1088/1478-3975/acd298)
- Khuong A, Gautrais J, Perna A, Kuntz P, Jost C, Theraulaz G. 2016 Stigmergic construction and topochemical information shape ant nest architecture. *Proc. Natl Acad. Sci. USA* **113**, 1303–1308. (doi:10.1073/pnas.1509829113)
- Khuong A, Lecheval V, Fournier R, Blanco S, Weitz S, Beziau JJ, Gautrais J. 2013 How do ants make sense of gravity? A Boltzmann Walker analysis of *Lasius niger* trajectories on various inclines. *PLoS ONE* **8**, e76531. (doi:10.1371/journal.pone.0076531)
- Rodríguez-Morales D, Tapia-McClung H, Robledo-Ospina LE, Rao D. 2021 Colour and motion affect a dune wasp's ability to detect its cryptic spider predators. *Sci. Rep.* **11**, 15442. (doi:10.1038/s41598-021-94926-7)
- Pillot MH, Gautrais J, Gouello J, Michelena P. 2010 Moving together: incidental leaders and naive followers. *Behav. Processes* **83**, 235–241. (doi:10.1016/j.beproc.2009.11.006)
- Ward AJW, Sumpter DJT, Couzin ID, Hart PJB, Krause J. 2008 Quorum decision-making facilitates information transfer in fish shoals. *Proc. Natl Acad. Sci. USA* **105**, 6948–6953. (doi:10.1073/pnas.0710344105)
- Ling H, Mclvor GE, van der Vaart K, Vaughan RT, Thornton A, Ouellette NT. 2019 Costs and benefits of social relationships in the collective motion of bird flocks. *Nat. Ecol. Evol.* **3**, 943–948. (doi:10.1038/s41559-019-0891-5)
- Viswanathan GM, Buldyrev SV, Havlin S, Da Luz M, Raposo E, Stanley HE. 1999 Optimizing the success of random searches. *Nature* **401**, 911–914. (doi:10.1038/44831)
- Viswanathan GM, Da Luz MG, Raposo EP, Stanley HE. 2011 *The physics of foraging: an introduction to random searches and biological encounters*. Cambridge, UK: Cambridge University Press.
- James A, Plank MJ, Edwards AM. 2011 Assessing Lévy walks as models of animal foraging. *J. R. Soc. Interface* **8**, 1233–1247. (doi:10.1098/rsif.2011.0200)
- Zaburdaev V, Denisov S, Klafter J. 2015 Lévy walks. *Rev. Mod. Phys.* **87**, 483–530. (doi:10.1103/RevModPhys.87.483)
- Pyke GH. 2015 Understanding movements of organisms: it's time to abandon the Lévy foraging hypothesis. *Methods Ecol. Evol.* **6**, 1–16. (doi:10.1111/2041-210X.12298)
- Reynolds A. 2015 Liberating Lévy walk research from the shackles of optimal foraging. *Phys. Life Rev.* **14**, 59–83. (doi:10.1016/j.plrev.2015.03.002)
- Mueller T, Fagan WF, Grimm V. 2011 Integrating individual search and navigation behaviors in mechanistic movement models. *Theor. Ecol.* **4**, 341–355. (doi:10.1007/s12080-010-0081-1)
- Bredeche N, Fontbonne N. 2022 Social learning in swarm robotics. *Phil. Trans. R. Soc. B* **377**, 20200309. (doi:10.1098/rstb.2020.0309)
- Nathan R, Getz WM, Revilla E, Holyoak M, Kadmon R, Saltz D, Smouse PE. 2008 A movement ecology paradigm for unifying organismal movement research. *Proc. Natl Acad. Sci. USA* **105**, 19 052–19 059. (doi:10.1073/pnas.0800375105)
- Durve M, Peruani F, Celani A. 2020 Learning to flock through reinforcement. *Phys. Rev. E* **102**, 012601. (doi:10.1103/PhysRevE.102.012601)
- Costa T, Laan A, Heras FJH, de Polavieja GG. 2020 Automated discovery of local rules for desired collective-level behavior through reinforcement learning. *Front. Phys.* **8**, 200. (doi:10.3389/fphys.2020.00200)
- Ward CR, Gobet F, Kendall G. 2001 Evolving collective behavior in an artificial ecology. *Artif. Life* **7**, 191–209. (doi:10.1162/106454601753139005)
- Aoki I. 1982 A simulation study on the schooling mechanism in fish. *Bull. Jpn. Soc. Sci. Fish.* **48**, 1081–1088. (doi:10.2331/suisan.48.1081)
- Vicsek T, Czirók A, Ben-Jacob E, Cohen I, Shochet O. 1995 Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.* **75**, 1226–1229. (doi:10.1103/PhysRevLett.75.1226)
- Lopez U, Gautrais J, Couzin ID, Theraulaz G. 2012 From behavioural analyses to models of collective motion in fish schools. *Interface Focus* **2**, 693–707. (doi:10.1098/rsfs.2012.0033)
- Leung F, Lam H, Ling S, Tam P. 2003 Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *IEEE Trans. Neural Netw.* **14**, 79–88. (doi:10.1109/TNN.2002.804317)
- Bouhleh MA, Hwang JT, Bartoli N, Lafage R, Morlier J, Martins JRRA. 2019 A Python surrogate modeling framework with derivatives. *Adv. Eng. Softw.* **135**, 102662. (doi:10.1016/j.advengsoft.2019.03.005)
- Wilson RP, Griffiths IW, Legg PA, Friswell MI, Bidder OR, Halsey LG, Lambertucci SA, Shepard ELC. 2013 Turn costs change the value of animal search paths. *Ecol. Lett.* **16**, 1145–1150. (doi:10.1111/ele.12149)
- Munden R, Börger L, Wilson RP, Redcliffe J, Brown R, Garel M, Potts JR. 2021 Why did the animal turn? Time-varying step selection analysis for inference between observed turning-points in high frequency data. *Methods Ecol. Evol.* **12**, 921–932. (doi:10.1111/2041-210X.13574)
- Klappstein NJ, Potts JR, Michelot T, Börger L, Pilfold NW, Lewis MA, Derocher AE. 2022 Energy-based step selection analysis: modelling the energetic drivers of animal movement and habitat use. *J. Anim. Ecol.* **91**, 946–957. (doi:10.1111/1365-2656.13687)
- Calovi DS, Litchinko A, Lecheval V, Lopez U, Pérez Escudero A, Chaté H, Sire C, Theraulaz G. 2017 Disentangling and modeling interactions in fish with burst-and-coast swimming reveal distinct alignment and attraction behaviors. *PLoS Comput. Biol.* **14**, 1–28. (doi:10.1371/journal.pcbi.1005933)
- Eilam D. 2005 Die hard: a blend of freezing and fleeing as a dynamic defense—implications for the control of defensive behavior. *Neurosci. Biobehav.*

- Rev.* **29**, 1181–1191. (doi:10.1016/j.neubiorev.2005.03.027)
31. Roelofs K. 2017 Freeze for action: neurobiological mechanisms in animal and human freezing. *Phil. Trans. R. Soc. B* **372**, 20160206. (doi:10.1098/rstb.2016.0206)
 32. Vladescu ID, Marsden EJ, Schwarz-Linek J, Martinez VA, Arlt J, Morozov AN, Marenduzzo D, Cates ME, Poon WCK. 2014 Filling an emulsion drop with motile bacteria. *Phys. Rev. Lett.* **113**, 268101. (doi:10.1103/PhysRevLett.113.268101)
 33. Elgeti J, Gompper G. 2015 Run-and-tumble dynamics of self-propelled particles in confinement. *Europhys. Lett.* **109**, 58003. (doi:10.1209/0295-5075/109/58003)
 34. Herbert-Read JE, Perna A, Mann RP, Schaerf TM, Sumpter DJT, Ward AJW. 2011 Inferring the rules of interaction of shoaling fish. *Proc. Natl Acad. Sci. USA* **108**, 18 726–18 731. (doi:10.1073/pnas.1109355108)
 35. Katz Y, Tunström K, Ioannou CC, Huepe C, Couzin ID. 2011 Inferring the structure and dynamics of interactions in schooling fish. *Proc. Natl Acad. Sci. USA* **108**, 18 720–18 725. (doi:10.1073/pnas.1107583108)
 36. Zienkiewicz AK, Ladu F, Barton DA, Porfiri M, Bernardo MD. 2018 Data-driven modelling of social forces and collective behaviour in zebrafish. *J. Theor. Biol.* **443**, 39–51. (doi:10.1016/j.jtbi.2018.01.011)
 37. Escobedo R, Lecheval V, Papaspyros V, Bonnet F, Mondada F, Sire C, Theraulaz G. 2020 A data-driven method for reconstructing and modelling social interactions in moving animal groups. *Phil. Trans. R. Soc. B* **375**, 20190380. (doi:10.1098/rstb.2019.0380)
 38. Lecheval V, Mann RP. 2023 Data from: Smart self-propelled particles: a framework to investigate the cognitive bases of movement. Dryad Digital Repository. (doi:10.5061/dryad.ftdz08z9)
 39. Lecheval V, Mann RP. 2023 Smart self-propelled particles: a framework to investigate the cognitive bases of movement. Figshare. (doi:10.6084/m9.figshare.c.6742215)