

VU Research Portal

Environment induced emergence of collective behaviour in evolving swarms with limited sensing

Diggelen, F. V.; Luo, J.; Karagüzel, T. A.; Cambier, N.; Ferrante, E.; Eiben, A. E.

2022

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Diggelen, F. V., Luo, J., Karagüzel, T. A., Cambier, N., Ferrante, E., & Eiben, A. E. (2022). *Environment induced emergence of collective behaviour in evolving swarms with limited sensing*.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Environment induced emergence of collective behaviour in evolving swarms with limited sensing

Fuda van Diggelen*

Jie Luo*

Tugay Alperen Karagüzel*

fuda.van.diggelen@vu.nl

Vrije Universiteit Amsterdam

Noord-Holland, The Netherlands

Eliseo Ferrante

Vrije Universiteit Amsterdam

Noord-Holland, The Netherlands

Technology Innovation Institute

Abu Dhabi, United Arab Emirates

e.ferrante@vu.nl

Nicolas Cambier

Vrije Universiteit Amsterdam

Noord-Holland, The Netherlands

n.p.a.cambier@vu.nl

A.E. Eiben

Vrije Universiteit Amsterdam

Noord-Holland, The Netherlands

a.e.eiben@vu.nl

ABSTRACT

Designing controllers for robot swarms is challenging, because human developers have typically no good understanding of the link between the details of a controller that governs individual robots and the swarm behavior that is an indirect result of the interactions between swarm members and the environment. In this paper we investigate whether an evolutionary approach can mitigate this problem. We consider a very challenging task where robots with limited sensing and communication abilities must follow the gradient of an environmental feature and use Differential Evolution to evolve a neural network controller for simulated robots. We conduct a systematic study to measure the flexibility and scalability of the method by varying the size of the arena and number of robots in the swarm. The experiments confirm the feasibility of our approach, the evolved robot controllers induced swarm behavior that solved the task. We found that solutions evolved under the harshest conditions (where the environmental clues were the weakest) were the most flexible and that there is a sweet spot regarding the swarm size. Furthermore, we observed collective motion of the swarm, showcasing truly emergent behavior that was not represented in- and selected for during evolution.

CCS CONCEPTS

• **Computer systems organization** → **Evolutionary robotics**;
• **Computing methodologies** → **Multi-agent systems**; **Cooperation and coordination**; • **Applied computing** → *Computer-aided design*.

*Three authors contributed equally to this research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://permissions.acm.org).

GECCO '22, July 9–13, 2022, Boston, USA

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00

<https://doi.org/10.1145/1122445.1122456>

KEYWORDS

Evolutionary robotics, Embodied AI, Differential evolution, Evolutionary swarm robotics

ACM Reference Format:

Fuda van Diggelen, Jie Luo, Tugay Alperen Karagüzel, Nicolas Cambier, Eliseo Ferrante, and A.E. Eiben. 2022. Environment induced emergence of collective behaviour in evolving swarms with limited sensing. In *Proceedings of The Genetic and Evolutionary Computation Conference 2022 (GECCO '22)*. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/1122445.1122456>

1 INTRODUCTION

Animal groups in nature have evolved collective motion behaviors for certain benefits, like increasing environmental awareness [6, 15] and safety from predation [19]. Implementing such behaviors in robot swarms has several applications, but designing adequate robot controllers is a great challenge. The principal problem is that the controller can only govern the individual robots directly, but the desired behavior is defined at the group level [12]. Human developers have typically no good understanding of the link between the controller details and the induced swarm behavior that is an indirect result of the interactions between individual swarm members and the environment. In the current practice, controllers in swarm robotics require an extensive manual design and fine-tuning process before obtaining desired behavior.

We developed an artificial evolutionary pipeline to automatically design and evaluate a robot controller for the swarm to complete a collective level task. On a conceptual level, we may note that we hereby address a higher level of complexity than traditional EC, even one step higher than evolutionary robotics (ER). As explained in [8], in traditional EC we have a 3-step chain from genotypes to fitness values, genotype → phenotype → fitness, while in ER the chain is 4-fold, genotype → phenotype → behavior → fitness. In the case of robot swarms, we have one more level, that of the group behavior: genotype → phenotype → individual behavior → group behavior → fitness. Obviously, the arrow from individual behavior to group behavior is complex in itself, thus optimizing along the

whole chain from genotype to fitness is very intricate. This makes the endeavour of evolving controllers for swarms far from trivial.

For this paper we consider a very challenging task, where a swarm consisting of robots with very limited sensing and communication abilities must follow the gradient of an environmental feature, e.g., light intensity, temperature, or radiation. To solve this problem, we define genotypes that represent neural networks (phenotypes) in a robot and postulate that all robots of the swarm have the same controller. The fitness of a genotype / phenotype will be defined by the behavior of the swarm. Our goal is to investigate whether an evolutionary algorithm is capable of finding controllers that enable the swarm to follow the gradient.

By constraining a single agent’s sensor capabilities to only sense the local scalar value of the gradient, we know that this task is not solvable for an agent in isolation without any memory to compute the gradient on its own. This problem is especially interesting as we thus investigate if an environment can induce a collective behavior which would normally not evolve, to complete a task that otherwise could not be solved. In contrast to other works, the optimization does not directly solve for emergent behavior but might exhibit it as a successful strategy. To this end, we are seeking answers to the following specific research questions.

Research Question 1: Will there be any emergent collective behavior among the members of the swarm?

Research Question 2: How flexible and scalable are the evolved solutions with respect to changes in the size of the robot arena and the number of robots in the swarm?

To answer our research questions, we set up a system where the controllers of (simulated) swarms of differential drive robots are evolved by Differential Evolution. We measure the flexibility and scalability by running experiments in three arenas and with three different group sizes in a simulator called Isaac Gym [17].

2 RELATED WORK

Automated design for (swarm) robotics has a long history with varying approaches [4, 22], which often require an optimization method that can handle high non-linearity and non-smooth objective functions. EAs proved themselves to be capable of handling such a hard task, which resulted in the development of fields like evolutionary robotics [10]. In this paper, we draw inspiration from this field and create an automated design pipeline suitable for swarm robotics. A strong difference in our approach is the level of control that we specify. Evolutionary robotics is often focused on learning low-level controls on a single robot for basic tasks like locomotion [18, 39]. In contrast, a swarm controller indirectly specifies a higher level of behavior on the group level [32]. This additional level of abstraction creates a situation where we cannot expect a group behaviour to directly correlate with the objective. Nevertheless as mentioned before, the ability to optimize control when there is no clear mapping from variables to objective is what makes evolution attractive to apply on swarms as well.

Indeed, non-automated designs usually struggle to resolve this mapping between agent-based (i.e. microscopic) and group-level (i.e. macroscopic) models, and must therefore resort to trial-and-error. In some cases, it is however possible to establish a quantitative micro-macro link so that mathematical equations are able to indicate the

settings required to achieve an objective or, vice versa, to predict the outcome of an experiment from variables’ values [25]. Such a micro-macro link is possible for binary (leftward or rightward) alignment of multiple agents [11], but it is inapplicable to more advanced alignment-dependent behaviors such as flocking [26]. As a matter of fact, existing approaches are only relevant to a fairly limited set of tasks; specifically collective decision-making between a finite quantity of options, i.e., best-of- n problems [34].

Typically, collective perception can be represented as a best-of- n problem [33], but continuous representations are more susceptible to be found in real environments [3, 23, 36]. In formal definition collective perception is; *where social interactions among individuals lead to collective computation of an environmental property by only allowing scalar measurements made by these individuals* [3]. This approach has been proven viable in biology [2, 23], where schools of fish –incapable of sensing the light gradient– can nevertheless follow the gradient to hide in the dark, by only considering their interactions with other fish.

Collective perception is also studied on artificial systems. In [36], a honey-bee inspired algorithm was used for aggregation of a robot swarm as a result of local interactions between each other and an environmental feature. Unfortunately, the authors provided a discrete environmental feature to the robots plus some additional (limited) communication capabilities, which made the resulting swarm behavior not truly obtained through a collective estimation alone. Another collective perception application can be found in [33] where robots collectively estimate certain environmental features and decide about the most frequent one (i.e. white tiles or black tiles in a tile grid). Yet again collective perception was not strictly necessary as the agents in [33] could communicate through voting and direct information exchange. There are many more papers demonstrating collective behavior of robot swarms where some form of collectively perceived feature(s) is used. Unfortunately, most of them either use communication capabilities or directional information related to the environment and/or other agents [28, 35]. An exception to the rule is presented by [14]. Here, a true collective perception method is proposed to control a flocking robot swarm that can sense a gradient, in an emergent way, and follow it in the environment, using only local and scalar measurements. The only downside to this solution is the manual optimization and design of the governing rules, which is a slow process that requires a lot of knowledge. Additionally, such an approach is not easily generalizable among other tasks and can easily be affected from individual failures or varying environmental conditions.

To cope with setups that are not fully specified at design time, automated design has been applied for collective perception tasks as well. [5] employs an evolutionary optimization framework to optimize controller parameters of robots in a swarm, to learn to discriminate different shelters in the environment and aggregate at the best to exploit. Although successful, [5] presented a limited usage of automated design with collective perception, due to a highly specific task and setup, with very little freedom for artificial evolution (only parameters of pre-designed controller rules were optimized). Another example for automated design on a controller with collective perception is in [29], where an underwater swarm could locate an radio signal source and then move towards it as a cohesive group. Unfortunately, these real-life applications

–although showing successful collective perception– are limited in their automated design as they only optimize pre-designed control parameters (a biased controller with model parameters). A much more expressive optimization was shown by [9], where a swarm of robots, controlled by evolved neural networks (which are model agnostic approximators), learned to locate the food in their environment by emerging a communication scheme. This study showed that swarm behavior can develop with model-agnostic controllers undergoing evolution, even if such a behavior is not directly specified in the fitness function.

More similar to the work presented here, [24] and [38] presented automated design approaches to optimize not only controller parameters but controllers themselves (neural networks in these particular examples). [24] showed that a simulated swarm could learn similar flocking behavior, when directly optimizing for fitness functions that closely relate to well-known flocking rules (e.g. [26]). In our opinion, these results are unsurprising as the flocking is directly optimized for, rather than emerged. Additionally, such an approach can be unsuitable for robots, as it may limit the controller’s robustness when conditions change and with possible failures [10]. In our opinion, a more robust approach would be to favor a resultant swarm behavior of swarm (behavioral fitness) rather than directly rewarding the desired actions of individuals (functional fitness) [10]. [38] showed that flocking-like behavior emerged when evolving a swarm of simulated agents only for exploiting a vital source for survival and reproduction. These results show that with an environment-driven task, swarm-level behavior can emerge if it provides an advantage with respect to solving the task collectively. In this paper, we hope to see the similar collective behavior emerge, by using a very generic controller (i.e. model agnostic) and optimizer in a more strict setting.

To summarize, this study positions itself uniquely in that it hopes to show that, with limited sensing and true collective perception (no explicit communication or direct sensing), emergent collective behavior can be induced by the environment using a model-agnostic controller that undergoes evolution.

3 METHODOLOGY

Full code base can be found in the following git repository:
https://github.com/fudavd/EC_swarm/tree/GECCO_2022

3.1 Controller

As a single behavior is considered during the evolutionary process, we use the same controller on every robot in the swarm. Thus, an individual in our EA refers to a single homogeneous swarm with the controllers of its constituents being an exact copy of the brain encoded by the genome. To be clear, the controller for each member of the swarm will be the same, but their respective input will likely differ due to their different positions and orientation within the swarm. Each controller provides a target velocity to the two wheels of the corresponding differential drive robot. For this we use a fully connected reservoir Neural Network (NN) controller where we only optimize the output layer using DE. The choice of this controller is based on two advantages: 1) NN are known to be model agnostic function approximators which means that there is no prior bias on the possible behaviours. 2) reservoir NN reduce the search space

significantly (only the output layer is optimized) thereby increasing the learning rate.

In more detail, as an input to the reservoir NN we provide relative positional and heading information from 4 directional sensors, plus a local value sensor reading of a scalar field map (thus in total 9 different inputs) every 0.1s. The 4 directional sensors cover a combined 360 degrees view of the robot’s surrounding, with a single sensor only sensing the nearest neighbour within a specific locally defined quadrant: Front-, Back-, Left-, and Right- quadrant (90 degrees each). The information obtained are the distance and relative heading of this neighbour with respect to the controlled robot (2 dimensions of information per sensor, in 4 quadrants). If no neighbour is present within the quadrant’s maximum range (2m) the sensor set its distance measured to 2.0 and relative heading to 0. Lastly, local values of the scalar field are provided based on the robot’s position in the environment. All sensor inputs are pre-processed to a $[-1, 1]$ range, before being fed into the NN. The two NN outputs specify the heading- ($w \in [-1, 1]$) and forward velocity ($v \in [0, 1]$), which are rescaled and translated to correct wheel speeds (up to ± 14 cm/s per wheel).

The reservoir NN architecture consist of an input layer (s_{in}) with 9 neurons, 2 hidden layers (with Softplus functions) of the same size as the input layer, and an output layer (\tanh^{-1} of size 2). All weights are randomly initialized between $[-2, 2]$ with a uniform distribution. We set all biases to 0 and will only optimize the weights of the output layer during evolution (18 weights). The final NN controller can be formalized as follows:

$$NN = \tanh^{-1} (W_{out} \text{ReLU} (W_{h2} \text{ReLU} (W_{h1} s_{in})))$$

with, $W_{h1,h2} \in \mathbb{R}^{9 \times 9}$ and $W_{out} \in \mathbb{R}^{2 \times 9}$

Each individual in the EA population receives the same reservoir ($W_{h1,h2}$) in their NN controller but with varying output layers W_{out} . We encode the reservoir NN controller’s phenotype by a vector of weights with $2 \times 9 = 18$ floating point numbers (min/max value ± 10.0) as its genotype (x), specifying the rows in W_{out} .

$$x = [W_{out1}, W_{out2}] \quad \text{with, } x \in \mathbb{R}^{18}$$

The phenotype of the controller is illustrated in Figure 1.

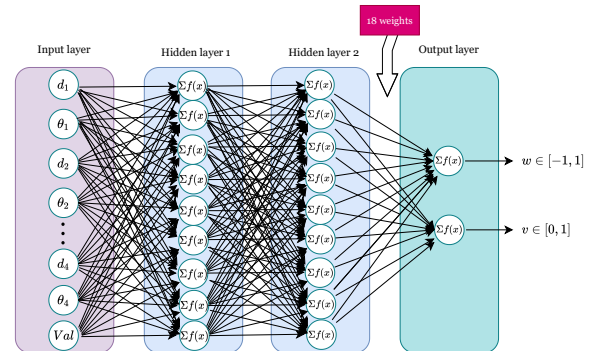


Figure 1: Phenotype of the controller - Neuron Network

3.2 Evolutionary Algorithm

Evolutionary methods for solving NP-hard optimization problems have become a very popular research topic in recent years. Some of the more established methods are genetic algorithm (GA), particle swarm optimization (PSO), and differential evolution (DE), which all have shown to be able to solve hard optimization problems for various domains [1, 7, 37]. [13] states that GA is more well-established because of its much earlier introduction however it has less ability to reach good solution without local search, while the more recent PSO and DE algorithms have started to attract more attention especially for continuous optimization problems. However in PSO, the best particle in the swarm exerts its one-way influence over all the remaining solutions in the population. This often leads to premature clustering around the best particle, especially if the fitness gaps are large. This is definitely an unwanted property, therefore in this paper, we choose DE to search our best solution. DE is a population-based Evolutionary Algorithm (EA) that samples new candidates by perturbing the current population [30]. The three main components in this method are as follows:

Differential mutation operator: a new candidate is generated by randomly picking a triplet from the population, $(x_i, x_j, x_k) \in X$, then x_i is perturbed by adding a scaled difference between x_j and x_k , that is:

$$y = x_i + F(x_j - x_k) \quad (1)$$

where $F \in R_+$ is the scaling factor.

Uniform crossover operator: sample a binary mask $m \in \{0, 1\}^D$ according to the Bernoulli distribution with probability $p = P(\text{md} = 1)$ shared across all D dimensions, and calculate the final candidate according to the following formula:

$$v = m \odot y + (1 - m) \odot x_i \quad (2)$$

The last component is a selection mechanism: the authors of [30] proposed to use the ‘‘survival of the fittest’’ approach, i.e., combine the previous population with the new one and select N candidates with the highest fitness values, i.e., the deterministic $(\mu + \lambda)$ selection.

Studies [16, 20, 31] have shown that the schemes for the trial vector (candidate) generation can have a significant influence on the algorithm’s performance. Here we follow general recommendations which can be found in literature [21] for stable exploration behavior, namely the crossover probability (CR) being fixed to a value of 0.9 and the scaling mutation factor (F) being fixed to a value of 0.5 (see Table 1).

We apply DE to change the weights of the NNs of the robots to improve their controllers for the tasks. The whole process is illustrated in Figure 2. The pseudocode of DE for evolving the controllers is shown in Algorithm 1.

3.3 Performance measures

3.3.1 Fitness function. In a homogeneous swarm, we evaluate fitness on group level. In this research, we are mainly interested in a pure evolution-driven effects on swarm behavior without directly encouraging any form of communication, collaboration or collectiveness in the evolutionary framework. Therefore, the fitness was solely dependent on the swarm’s ability to follow the increasing gradient of the scalar field which is defined in the environment

Algorithm 1 Differential Evolution

```

1: INITIALIZE population (X controllers w/ vectors of weights)
2: while iter < MAXiter do
3:   for i=1:i<=X; i++ do
4:     GENERATE three individuals  $x_i, x_j, x_k$  from X randomly
5:     MUTATION using formula:  $y = x_i + F(x_j - x_k)$ 
6:     Crossover using formula:  $v = m \odot y + (1 - m) \odot x_i$ 
7:     EVALUATE controllers (based on fitness value)
8:     SELECT survivors / UPDATE population
9:   end for
10: end while

```

(shown in Figure 3-a). Additionally, we would like to emphasize that our fitness function does not distinguish between robots that are sensing and following the increasing gradient collectively or as solitaires. The fitness function F is defined as follows:

$$F = \frac{\sum_{t=0}^T f_t}{G_{max} \cdot T} \quad \text{and} \quad f_t = \frac{\sum_{n=1}^N G_n}{N} \quad (3)$$

Where G_n is the scalar value of the grid cell in which agent n (of all agents, N) is located at a time t . Thus the fitness at a specific time (f_t) is calculated as the mean scalar value of all swarm members. Final fitness (F) is calculated by averaging all f_t over total simulation time T . At last, we normalize using the maximum scalar value G_{max} , always equal to 255 for all experiments.

3.3.2 Behavioral measures. Besides our fitness function F , we want to investigate the possible emergence of flocking behavior. For this, we analyse two additional behavioral measures. The first measure: order (Φ), is defined as follows:

$$\Phi = \frac{\sum_{n=1}^N \varphi_n}{N} \quad \text{and} \quad \varphi_n = \frac{\left\| \left(\sum_{p=1}^P \angle e^{j\theta_p} \right) + \angle e^{j\theta_n} \right\|}{P + 1} \quad (4)$$

Here, φ_n defines the order value calculated for agent n . Which is the average current heading direction of agent n (noted as $\angle e^{j\theta_n}$) and all its perceived neighbors P (noted as $\angle e^{j\theta_p}$). The total swarm order Φ is then defined as the average φ_n over all agents in the swarm. Order measure gives a powerful insight about the alignment of agent’s direction of motion. If all agents move towards the same direction, then the order measure approaches to 1; and if they move in different directions, order approaches to 0.

In addition to order, the number of collisions between robots are also checked periodically. This check is not trivial since robots are modelled in a physics simulator, all collisions are realistically modeled and certainly have affect on the movement capabilities of colliding robots. In other words, when agents are collided, they can tip over, stop or drag each other because of the physical interaction between them. We will report on collision by plotting the total number of collision present in the swarm as a function of time.

The final measure to analyze is the trajectories of robots during an experiment. By including and highlighting the start and end points as well, presenting the trajectories of robots on the scalar field provides a clear intuition about the swarm’s motion characteristics. Trajectories can show us if there is any correlated forward motion or rotation among the swarm and the relative difference of the local values of the scalar field between start and end points. In addition, trajectories can reveal any kind of behavioral differences between controllers that are evolved in different conditions and obtaining

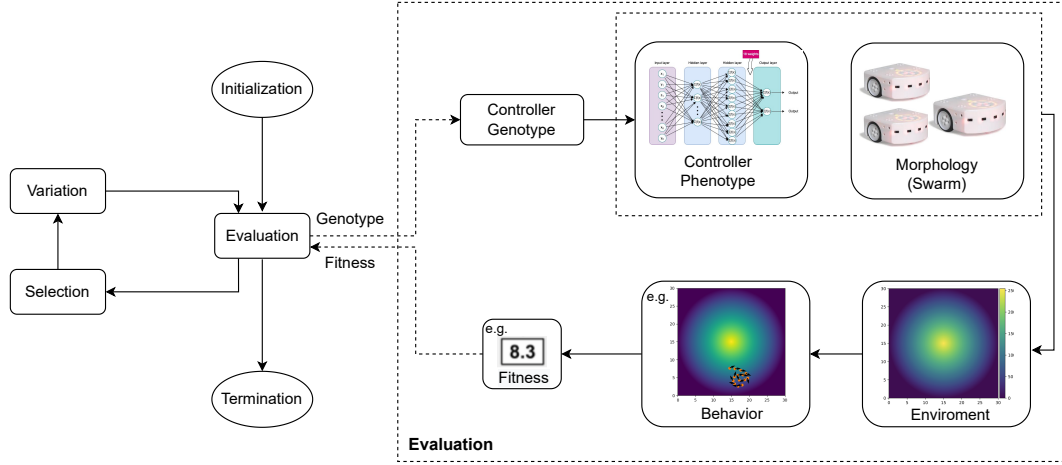


Figure 2: Differential Evolution Framework: This is a DE framework to change the weights of the NNs of the robots to improve their controllers for the tasks. In the Evaluation box, we show examples of a robot morphology, controller, environment, behavior and fitness value.

similar fitness. Finally visual inspection of the trajectories will be made.

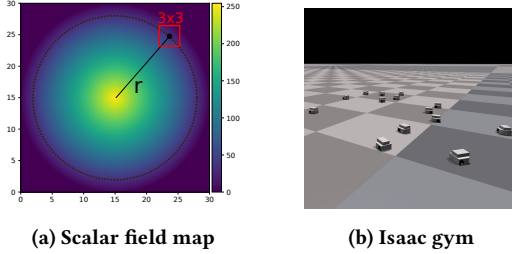


Figure 3: The swarm environment. (a) the scalar field map with its maximum value (255) in the center. Initial swarm position (shown as a black dot) is randomly placed on a circle with distance r away from the center. Individuals are randomly placed within a 3×3 m bounding box (shown in red). (b) a random initial state is shown in simulation.

3.4 Experimental setup

3.4.1 Simulator. At the start of an evaluation, a single swarm is randomly placed at a fixed distance away from the center of the arena (sampled uniformly on a circle with radius r). Each swarm member is then randomly placed inside a 3×3 m box (the red square in Figure 3-a) around this position with a random heading. We simulate the behavior swarm for 10 minutes ($dt = 0.05s$) and sample the controller every 0.1s for new motor inputs. Simulation of the swarm is done using Isaac Gym (see Figure 3-b), which allows us to simulate the whole population of 25 different swarm individuals in parallel. To reduce the effect of lucky runs, we evaluate every individual 2 times, in which the lowest value is taken as the final fitness. Thus, in total we evolve 25 individuals with 2 repeated tests for 100 generation in 30 different runs with 3 different evolutionary experiments (45,000 evaluations of 10 minutes in simulation).

3.4.2 Evolutionary Experiment. In total we perform DE optimization 3 times in different environments (3 arenas), all with the same

task of collective gradient sensing. The individuals in the population are not a single robot (which is more conventional in evolutionary robotics) but rather describe a homogeneous swarm. An initial population of 25 individuals (i.e. 25 different homogeneous swarms) is randomly generated as the first generation. We test each individual twice and use the result of the lowest fitness. In each generation, 3 individuals are randomly picked to generate a new individual by crossover and mutation. The next generation is formed by the top 25 individuals from the previous generation plus the new candidates according to fitness value. The evolutionary process is terminated after 100 generations. All the evolutionary experiments are repeated 30 times independently to get a robust assessment of the performance per data set.

3.4.3 Collective behavior, Flexibility & Scalability. After the three evolutionary experiments we re-test the single best controller over all 30 runs per arena (thus in total three different best controller). The best controller is tested in the same environment to evaluate if collective motion has emerged. We look at three different aspects of collective behavior over time: *alignment* in terms of order (Equation 4, Φ), *performance* as mean scalar value of the swarm (Equation 3, f_i), and *collision* counted at every controller update. These metrics are plotted with respect to time.

Additionally, we investigate two types of swarm properties, namely flexibility and scalability [27]. Firstly, we (cross-)validate the overall best controllers per environment, meaning that the same swarm was re-tested in its own environment and the other two arenas for comparison. A similar performance in the other environments indicates flexibility of the found solution since changing environment sizes indicate a slower or faster change in the local values of the gradient for the same amount of displacement. Since the best controllers for each environment size is evolved to be the best with the corresponding rate of change of local values, different environment size and different rate of change of local values can be regarded as environmental variations. Secondly, to see the scalability of our control solution, we re-tested the overall best controller per arena, in the same arena with different swarm sizes (here we

resized the initialization box accordingly). For both flexibility and scalability experiments we test the best controller 30 times per condition. The experimental parameters we used in the experiments are described in Table 1.

Table 1: Main experiment parameters

DE	Value	Description
Swarm size	14	Number of robots in a swarm
Arena size	10/30/45	Swarm environment
Population size	25	Number of individuals (swarms) per generation
Generations	100	Termination condition for each run
Mutation (F)	0.5	Mutation factor for individuals
Crossover (CR)	0.9	Probability of crossover for individuals
Evaluation time	10	Duration of an evaluation in minutes
Repetitions	30	Number of repetitions per experiment
Flexibility		
Swarm size	14	Number of robots in a swarm
Arena size	10/30/45	Swarm environment
Evaluation time	10	Duration of an evaluation in minutes
Repetitions	30	Number of repetitions per experiment
Scalability		
Swarm size	5/14/50	Number of robots in a swarm
Arena size	10/30/45	Swarm environment
Evaluation time	10	Duration of an evaluation in minutes
Repetitions	30	Number of repetitions per experiment

4 EXPERIMENT RESULTS

All reported results and analysis data/scripts can be found in the following database: <https://doi.org/10.34894/IREUW1>

4.0.1 Efficacy. We measure efficacy by the mean and maximum fitness averaged over the 30 independent runs at every generation.

For each evolutionary experiment with swarm size 14 we run the learning task 30 times at three different arenas. Figure 4 shows that the experiment in arena 45x45 has the highest average fitness at the start, followed by the experiment in arena 30x30. Then the fitness of all the evolutionary experiments increase steadily generation by generation. From around generation 28, the average fitness from arena 10x10 surpasses the other two, so does the max fitness. Similarly, around the same cross-over point, the average fitness from arena 45x45 becomes the lowest. In the end, the mean best solution over 30 runs were found in the 10x10 condition ($F_{10} = 0.46 \pm 0.038$), followed by 30x30 ($F_{30} = 0.39 \pm 0.055$), and 45x45 ($F_{45} = 0.33 \pm 0.040$). Statistical analysis showed that these difference were significant between all conditions ($p \ll 0.05$).

4.0.2 Flexibility. The final swarm controller should be able to operate despite variations in its environment. We re-evaluate the overall best controllers' fitness from each evolutionary optimization (3 arenas with swarm size 14) in 9 additional tests across three different arenas, namely 10x10m, 30x30m and 45x45m, with 30 repetitions.

Figure 5 shows that the best controller from the 10x10 evolutionary run (in color purple) performs best in arena10, however it significantly drops in performance when re-tested in the other two larger arenas ($p < 0.05$). For the 30x30 run (blue boxplots), the best controller performs similarly well in both the smaller arena10 and

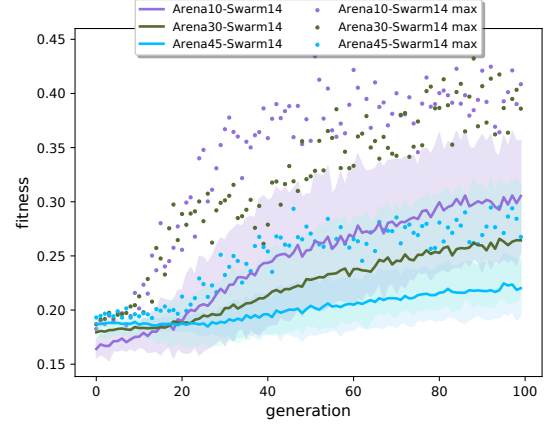


Figure 4: The mean \pm STD (line) and average max (dot) fitness over 100 generations (averaged over 30 runs) for all experimental condition (10x10: purple, 30x30: green, 45x45: blue.

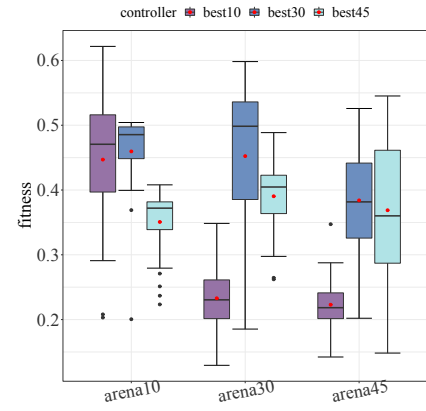


Figure 5: Flexibility boxplot. Cross-validation of best controller in other arenas. Red dots show mean values.

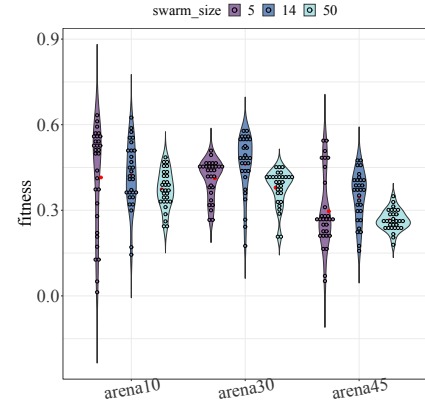


Figure 6: Scalability violin plot. Re-test of best controllers with different swarm sizes (5: purple, 14: blue, 50: cyan) for 30 runs. Each circle represents a single run with red dots as mean value.

its own (arena30), while the fitness is lower in the bigger arena45. The last best controller from the 45x45 run (cyan boxplots) has a relatively steady performance across all three arenas. From this plot,

we may conclude that the best controller from the biggest arena 45x45 is the most flexible one. During evolution, the most flexible controller was evolved in the hardest environment. Apparently controllers evolved in larger arenas, perform relatively (equally) well in smaller arenas, while the reverse is not the case.

4.0.3 Scalability. The swarm should be able to operate with a wide range of group sizes and support large numbers of individuals without impacting performance considerably. That is, the coordination mechanisms and strategies to be developed for swarm robotic systems should ensure cooperation of the swarm with varying swarm sizes. To measure the scalability property, we choose the best controllers from evolutionary experiments (Arena10,30,45 with swarm size 14) and apply them on three different swarm sizes –being 5, 14, 50– within the same environment.

Figure 6 shows that all the best controllers are quite scalable across three arena with close mean fitness. Experiment with swarm size 14 (blue violins) has the highest mean fitness across three arenas. Experiment with swarm size 50 has a relatively lower standard deviation which indicates a more robust performance. In other words, a bigger swarm size creates a bigger sample of the map (because it has a bigger coverage) which makes the estimation of the gradient better and the performance more stable.

4.0.4 Collective behavior. Figure 7 shows the correlation of alignment (Order, Φ) within the swarm members and their performance (Mean scalar value, f_t) over time. We observe that when order peaks, i.e. namely red point 1, 3, 4 and 5 in arena 10x10m, red points 2, 3 and 5 in arena 30x30m and red points 1,3 in arena 45x45m, the mean scalar value increase (i.e. the tangent of the dashed black line is highly positive). Vice versa, when the colored lines reach to lower order value, the mean scalar value decrease or stay the same (i.e. tangent is near zero).

The snapshots on the right are correspond with the same numbered points in the order/mean scalar value plot. We can see that in most cases when a swarm is becoming more aligned (i.e. order is high), the direction of movement is similar to the direction of increasing gradient (i.e. towards the center). In the low alignment case (i.e. low order) most of the members disperse in different directions, causing no net difference in the mean scalar value.

Impressively, here we can see a collective ‘search-like’ behavior emerge that is reminiscent of an optimization algorithm trying to maximize a function. Scalar values are periodically and collectively sensed by the swarm. When the order is low, the swarm samples in different direction for higher values i.e. “exploration”. In contrast, during the high peaks the swarm collectively moves in the direction of the gradient which causes the mean scalar value to increase i.e. “exploiting”. These exploration exploitation states are periodically changed to escape local optima and find a global maximum. Even more exciting, when mean scalar value plateaus (i.e. swarm arriving near maximum) there is a clear switch in behavior and a new pattern emerges. In the 10x10 the swarm approaches the center and overshoots it, directly followed by a fast recovery toward the center. For 30x30 controller, the agents randomly spread around the maximum, slowly creeping inward. Unfortunately, the 45x45 experiment run was too short to see a final pattern emerge.

Figure 8-a shows the number of collisions over 600s timestamps in three arenas. We observe that the smaller arena size, the

higher no. of collisions. Moreover, at the start of each run the swarm is very chaotic and collisions occur in all environments. Fortunately, it can be observed that the swarm resolves these collision itself. In addition, occasional group splits can also be observed. Since neither the ordered swarm behaviour is the main objective nor controllers are manually designed for it, we expect to see this splits time to time. What is interesting is that despite this splits, clusters continue to behave in the same way and eventually merge with the rest under the influence of the commonly sensed environmental feature. Figure 8-(b,c,d) shows the trajectories of the best controller per arena re-tested in the same environments. The black blocks are the starting points of a swarm and the white blocks are the ending points. All of them are following the gradient well. They start with circling around, then moving towards the brighter area once they sense the different gradients. When they reach the brightest zone (yellow), they start circling again. A video of these behaviors can be found at <https://youtu.be/yhKFvpLa9iI>.

5 DISCUSSION

The goal of this paper was to optimize the controllers of robots in a swarm for a task solely related to an environmental property and test whether a collective behavior emerges. Optimizing swarm behavior is often hard and requires vast knowledge on the specific task and the limitations of its members. Handwritten controllers therefore require a long process of prototyping which is sensitive to human biases, and limitation in engineered metrics for ‘good’ collective behavior. In our setup we showed that even with a very simple task description, good performance can be reached with learned coordinated swarm behavior as an emergent property.

From the results, the final best controller experiments showed that evolving in a bigger environment improves the controller in terms of flexibility. We suspect this to be caused by a less pronounced gradient in the bigger arena making the task more difficult. In other words, when evolved in a bigger arena a swarm needs to sample a bigger area to be able to sense the increasing gradient. This is because ratio between the swarm area and the total scalar field area is the smallest in the 45x45m environment. To cope with less salience of local values differences, the 45x45m flocks adopts a line formation strategy covering a big circle with more area (see Figure 7-C1). This is slightly different than the spiral formation seen in Figure 7-B1 employed by swarms evolved in the other arenas, and seem to maintain high fitness values when cross-validated. These behaviors can also be seen in the swarm trajectories (Figure 8). Lastly, we also see different behavior back in a less aligned rotation (low order) and more aligned forward motion (high order) with the best controller evolved in the largest environment.

The alignment within the swarm members (Figure 7) provides some interesting information on the swarm behavior. Here, we see a correlation between increasing order value and increasing mean scalar value seen by the swarm. Vice versa, lower order value correspond to a plateau/decrease in mean scalar value. This indicates that, when a swarm is becoming more aligned, the direction of movement is similar to the direction of increasing gradient.

As a final remark, for our specific optimization we wanted the swarm to sense the increasing gradient (while a single one is not capable of doing it) and move towards the increasing local values.

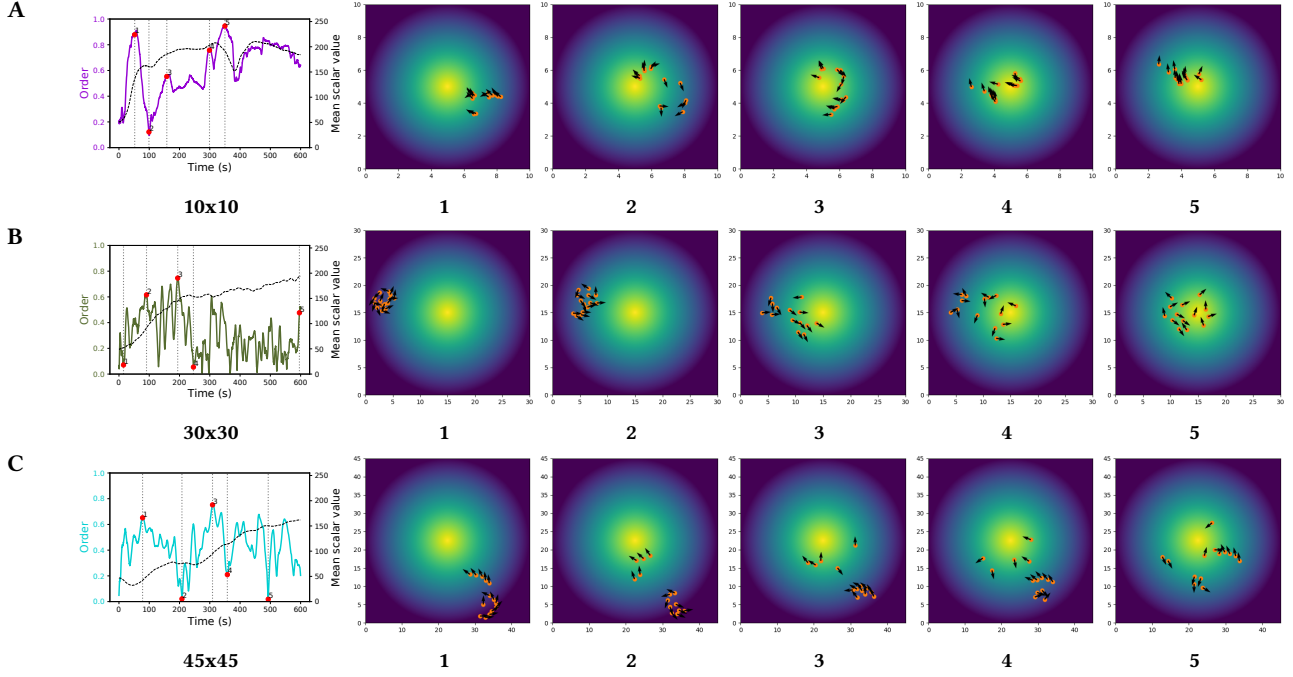


Figure 7: Flocking behavior of the final controller re-tested in the same environment. The three line plots on the left show the correlation between order (alignment) and mean scalar value (performance). The snapshots on the right are corresponding to each line plot. The dashed black lines show the mean scalar value and the colored lines show the order value. Distinct peaks in alignment are highlighted with numbers 1-5. Please note the different scales on each map.

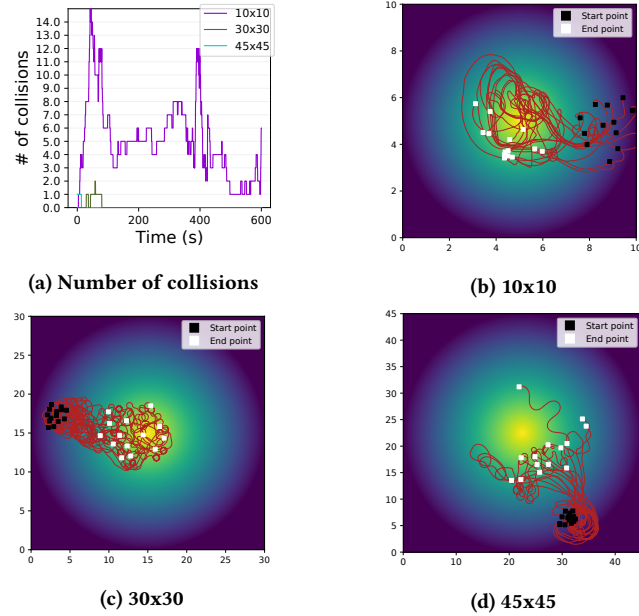


Figure 8: No. of collisions and trajectories of the best controller per arena re-tested in the same environments. Please note the different scales in subplots b, c, d.

It should be noted that such a fitness could be any function that we want. This opens up the possibility to ignore standard engineering practices that enforce certain swarm behavior, optimize flocking rule parameters or design complex controllers –all of which may

only be optimal for a specific task– and continuously update our swarm as an adaptive system. The model-agnostic nature of a NN with an EA enables the emergence of different swarm strategies that we don't have to pre-define (think of different formations, sub-group sizes, search behavior, etc.).

6 CONCLUSIONS AND FUTURE WORK

We have successfully developed an evolutionary system for the automated design of controllers for robot swarms. As explained in the Introduction, this is a highly nontrivial feat because of the long and complex chain between genotypes and fitness values. To test our pipeline we evolved several swarms for the task of gradient following. Here, we found that the environmental property induced the emergence of collective behavior, a result that was not explicitly formalised in the objective function. To our best knowledge, there is only one other study that has achieved this [38], but we should note that it considered simulated point agents with signalling abilities which significantly simplified the task. Ongoing work is devoted to additional validation of our results using real robots.

Interesting directions for future research include investigating more and different gradient landscapes in addition to the circular and unimodal one used in this study. Different gradient landscapes might be circular as well but with a multimodal structure or they can be in a linear form instead of circular. Moreover, instead of having a homogeneous swarm in terms of sensing capabilities, more investigation can be conducted where the sensing abilities of the individual robots differ (some may have more or better sensors than others, while using the same generic controller). This investigation

can even be stretched by extending the evolutionary approach to swarms where the controllers are not necessarily identical as they are in this study.

ACKNOWLEDGMENTS

This work is supported by Technology Innovation Institute, UAE.

REFERENCES

- [1] The Jin Ai and Voratas Kachitvichyanukul. 2009. Particle swarm optimization and two solution representations for solving the capacitated vehicle routing problem. *Computers and Industrial Engineering* 56, 1 (2009), 380–387. <https://doi.org/10.1016/j.cie.2008.06.012>
- [2] Andrew Berdahl, Colin J Torney, Christos C Ioannou, Jolyon J Faria, and Iain D Couzin. 2013. Emergent sensing of complex environments by mobile animal groups. *Science* 339, 6119 (2013), 574–576.
- [3] Andrew M Berdahl, Albert B Kao, Andrea Flack, Peter AH Westley, Edward A Codling, Iain D Couzin, Anthony I Dell, and Dora Biro. 2018. Collective animal navigation and migratory culture: from theoretical models to empirical evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences* 373, 1746 (2018), 20170009. <https://doi.org/10.1098/rstb.2017.0009>
- [4] Mauro Birattari, Antoine Ligot, and Gianpiero Francesca. 2021. AutoMoDe: a modular approach to the automatic off-line design and fine-tuning of control software for robot swarms. In *Automated Design of Machine Learning and Search Algorithms*. Springer, 73–90.
- [5] Alexandre Campo, Simon Garnier, Olivier Dédriche, Mouhcine Zekkri, and Marco Dorigo. 2011. Self-organized discrimination of resources. *PLoS One* 6, 5 (2011), e19888.
- [6] Iain D Couzin, Jens Krause, Nigel R Franks, and Simon A Levin. 2005. Effective leadership and decision-making in animal groups on the move. *Nature* 433, 7025 (2005), 513–516. <https://doi.org/10.1038/nature03236>
- [7] Meng Deng, Hui Jun Xiao, and Feng Mei Yang. 2006. Genetic algorithm for the open vehicle routing problem. *Beijing Huagong Daxue Xuebao (Ziran Kexueban)/Journal of Beijing University of Chemical Technology (Natural Science Edition)* 33, 4 (2006), 84–87.
- [8] A.E. Eiben and J.E. Smith. 2015. *Introduction to Evolutionary Computing* (2nd ed.). Springer-Verlag, London.
- [9] Dario Floreano, Sara Mitri, Stéphane Magnenat, and Laurent Keller. 2007. Evolutionary conditions for the emergence of communication in robots. *Current biology* 17, 6 (2007), 514–519.
- [10] Dario Floreano and Joseba Urzelai. 2000. Evolutionary robots with on-line self-organization and behavioral fitness. *Neural Networks* 13, 4-5 (2000), 431–443.
- [11] Heiko Hamann, Gabriele Valentini, Yara Khaluf, and Marco Dorigo. 2014. Derivation of a micro-macro link for collective decision-making systems. In *International Conference on Parallel Problem Solving from Nature*. Springer, 181–190.
- [12] Ken Hasselmann, Antoine Ligot, Julian Ruddick, and Mauro Birattari. 2021. Empirical assessment and comparison of neuro-evolutionary methods for the automatic off-line design of robot swarms. *Nature communications* 12, 1 (2021), 1–11.
- [13] Voratas Kachitvichyanukul. 2012. Comparison of three evolutionary algorithms: GA, PSO, and DE. *Industrial Engineering and Management Systems* 12 (09 2012), 215–223. <https://doi.org/10.7232/iems.2012.11.3.215>
- [14] Tugay Alperen Karagüzel, Ali Emre Turgut, and Eliseo Ferrante. 2020. Collective Gradient Perception in a Flocking Robot Swarm. In *International Conference on Swarm Intelligence*. Springer, 290–297.
- [15] Daniel B Kearns. 2010. A field guide to bacterial swarming motility. *Nature Reviews Microbiology* 8, 9 (2010), 634–644. <https://doi.org/10.1038/nrmicro2405>
- [16] Jie Luo, Daan Zeeuwe, and Agoston E. Eiben. 2021. Gait-learning with morphologically evolving robots generated by L-system. In *arXiv*. <http://arxiv.org/abs/2107.08249>. arXiv:2107.08249
- [17] Viktor Makoviychuk, Lukasz Wawrzyniak, Yunrong Guo, Michelle Lu, Kier Storey, Miles Macklin, David Hoeller, Nikita Rudin, Arthur Allshire, Ankur Handa, et al. 2021. Isaac gym: High performance gpu-based physics simulation for robot learning. *arXiv preprint arXiv:2108.10470* (2021).
- [18] Andrew L Nelson, Gregory J Barlow, and Lefteris Doitsidis. 2009. Fitness functions in evolutionary robotics: A survey and analysis. *Robotics and Autonomous Systems* 57, 4 (2009), 345–370.
- [19] Randal S Olson, Arend Hintze, Fred C Dyer, David B Knoester, and Christoph Adami. 2013. Predator confusion is sufficient to evolve swarming behaviour. *Journal of The Royal Society Interface* 10, 85 (2013), 20130305. <https://doi.org/10.1098/rsif.2013.0305>
- [20] Sidhartha Panda and Narayana Prasad Padhy. 2008. Comparison of particle swarm optimization and genetic algorithm for FACTS-based controller design. *Applied Soft Computing Journal* 8, 4 (2008), 1418–1427. <https://doi.org/10.1016/j.asoc.2007.10.009>
- [21] Meh Pedersen. 2010. Good Parameters for Differential Evolution. *Evolution* (2010), 1–10. <http://www.hvass-labs.org/people/magnus/publications/pedersen10good-de.pdf>
- [22] Shanker G Radhakrishna Prabhu, Richard C Seals, Peter J Kyberd, and Jodie C Wetherall. 2018. A survey on evolutionary-aided design in robotics. *Robotica* 36, 12 (2018), 1804–1821.
- [23] James G Puckett, Aawaz R Pokhrel, and Julia A Giannini. 2018. Collective gradient sensing in fish schools. *Scientific reports* 8, 1 (2018), 1–11.
- [24] Rita Parada Ramos, Sancho Moura Oliveira, Susana Margarida Vieira, and Anders Lyhne Christensen. 2019. Evolving flocking in embodied agents based on local and global application of Reynolds' rules. *Plos one* 14, 10 (2019), e0224376.
- [25] Andreagiovanni Reina, Roman Miletich, Marco Dorigo, and Vito Trianni. 2015. A quantitative micro-macro link for collective decisions: the shortest path discovery/selection example. *Swarm Intelligence* 9, 2 (2015), 75–102.
- [26] Craig W Reynolds. 1987. Flocks, herds and schools: A distributed behavioral model. In *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*. 25–34.
- [27] Erol Şahin. 2004. Swarm robotics: From sources of inspiration to domains of application. In *International workshop on swarm robotics*. Springer, 10–20.
- [28] Thomas Schmickl, Franz Wotawa, Ronald Thenius, and Joshua Chierian Varughese. 2016. Fstaxis algorithm: Bio-inspired emergent gradient taxis. In *ALIFE 2016, the Fifteenth International Conference on the Synthesis and Simulation of Living Systems*. MIT Press, 330–337. <https://doi.org/10.1162/978-0-262-33936-0-ch055>
- [29] Mansoor Shaukat and Mandar Chitre. 2016. Adaptive behaviors in multi-agent source localization using passive sensing. *Adaptive behavior* 24, 6 (2016), 446–463. <https://doi.org/10.1177/1059712316664120>
- [30] Rainer Martin Storn. 1997. Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization* (1997), 131–141. <https://doi.org/10.1023/A>
- [31] Jakub M. Tomczak, Ewelina Weglarz-Tomczak, and Agoston E. Eiben. 2020. Differential Evolution with Reversible Linear Transformations. *GECCO 2020 Companion - Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion* 1 (2020), 205–206. <https://doi.org/10.1145/3377929.3389972> arXiv:2002.02869
- [32] Vito Trianni, Elio Tuci, Christos Ampatzis, and Marco Dorigo. 2014. Evolutionary swarm robotics: A theoretical and methodological itinerary from individual neuro-controllers to collective behaviours. *The horizons of evolutionary robotics* 153 (2014).
- [33] Gabriele Valentini, Davide Brambilla, Heiko Hamann, and Marco Dorigo. 2016. Collective perception of environmental features in a robot swarm. In *International Conference on Swarm Intelligence*. Springer, 65–76. https://doi.org/10.1007/978-3-319-44427-7_6
- [34] Gabriele Valentini, Eliseo Ferrante, and Marco Dorigo. 2017. The best-of-n problem in robot swarms: Formalization, state of the art, and novel perspectives. *Frontiers in Robotics and AI* 4 (2017), 9.
- [35] Joshua Chierian Varughese, Daniel Moser, Ronald Thenius, Franz Wotawa, and Thomas Schmickl. 2019. swarmfstaxis: Borrowing a swarm communication mechanism from fireflies and slime mold. In *Complex Adaptive Systems*. Springer, 213–222. https://doi.org/10.1007/978-3-030-20309-2_10
- [36] Mostafa Wahby, Julian Petzold, Catriona Eschke, Thomas Schmickl, and Heiko Hamann. 2019. Collective change detection: Adaptivity to dynamic swarm densities and light conditions in robot swarms. In *Artificial life conference proceedings*. MIT Press, 642–649. https://doi.org/10.1162/isal_a_00233
- [37] Warisa Wisittipanich and Voratas Kachitvichyanukul. 2011. Differential Evolution Algorithm for Job Shop Scheduling Problem. *Industrial Engineering and Management Systems* 10, 3 (2011), 203–208. <https://doi.org/10.7232/iems.2011.10.3.203>
- [38] Olaf Witkowski and Takashi Ikegami. 2016. Emergence of swarming behavior: foraging agents evolve collective motion based on signaling. *PloS one* 11, 4 (2016), e0152756.
- [39] Joe Wright and Ivan Jordanov. 2015. Intelligent Approaches in Locomotion - A Review. *Journal of Intelligent and Robotic Systems: Theory and Applications* 80, 2 (2015), 255–277. <https://doi.org/10.1007/s10846-014-0149-z>