

Emergence of Specialization in a Swarm of Robots

Ádám M. Halász, Yanting Liang, M. Ani Hsieh, and Hong-Jian Lai

Abstract. We investigate the emergence of specialized groups in a swarm of robots, using a simplified version of the stick-pulling problem [5], where the basic task requires the collaboration of two robots in asymmetric roles. We expand our analytical model [4] and identify conditions for optimal performance for a swarm with any number of species. We then implement a distributed adaptation algorithm based on autonomous performance evaluation and parameter adjustment of individual agents. While this algorithm reliably reaches optimal performance, it leads to unbounded parameter distributions. Results are improved by the introduction of a direct parameter exchange mechanism between selected high- and low-performing agents. The emerging parameter distributions are bounded and fluctuate between tight unimodal and bimodal profiles. Both the unbounded optimal and the bounded bimodal distributions represent partitions of the swarm into two specialized groups.

1 Introduction

In a robotic swarm, *heterogeneity* may be quantified in terms of *diversity*, or the variability of the properties of individual agents. Heterogeneity may also involve the *specialization* of individuals for certain tasks. This collective adaptation strategy is often seen in biology [6]. The design of heterogeneous swarms requires ways to quantify the degrees of heterogeneity and specialization as well as their impact on collective performance. Early work on heterogeneity and specialization in robot

Ádám Halász · Yanting Liang · Hong-Jian Lai
Department of Mathematics, P.O. Box 6310, West Virginia University,
Morgantown, WV 26506
e-mail: {halasz, lyt814, hjlai}@math.wvu.edu

M. Ani Hsieh
Mechanical Engineering and Mechanics Department, Drexel University,
3141 Chestnut Street, Philadelphia, PA 19104
e-mail: mhsieh1@drexel.edu

teams established methods for the composition of group level behaviors [7, 9] and proposed a measure for heterogeneity [1].

The stick-pulling problem was originally formulated [5] to explore the swarm intelligence paradigm [3] in a context where collaboration is realized through local interactions, with limited or no global communication. The basic task, finding and pulling randomly distributed sticks, requires two robots in asymmetric roles. A robot that finds a stick must wait for another one to help pull the stick. The *gripping* or *waiting time parameter* (WTP) [4,5] of a robot is the time it will wait for help before releasing a stick. In the original study [5], Ijspeert *et al.* found that this asymmetric task could benefit from specialization. Through experimentation and a two-level modeling approach, they identified an optimal WTP for a homogeneous swarm. For a heterogeneous swarm with two subgroups (*castes* or *species*) of agents, each with a different WTP, they found a family of high-performance pairs of WTP values. This type of heterogeneity led to better performance when the number of robots was less than the number of sticks and did not make a significant difference otherwise.

Li, Martinoli and Mustafa [8] investigated how specialization could be learned by the stick-pulling team. In their system, agents changed their WTP based on local or global reinforcement signals. Learning resulted in optimal performance accompanied by increases in information-theoretic measures of diversity and specialization. This work strengthened the correlation between group performance and diversity and provided an example of global performance improvement through individual adaptation. However, it left open the question whether distinct groups with specialized behaviors could emerge through individual adaptation.

We investigated the advantages of specialization in a slightly modified version of the stick pulling problem [4], using a methodology developed for task allocation [2]. The starting point of our modeling approach was similar to the probabilistic model of [5]. Our higher level of abstraction resulted in a concise and transparent analytical model and in the possibility of scaling simulations into the range of thousands of agents and millions of updates. We identified a maximal performance level that may not be exceeded for any WTP configuration, and showed that it could be reached in many different configurations. Comparing homogeneous and two-species configurations, we showed analytically and confirmed through simulations that the two-species swarm performed better under non-ideal circumstances than the homogeneous one (in the case with more sticks than robots). Echoing the results of [5], we found that specialization was advantageous.

In this work we expand the analysis of optimal configurations and explore *collective adaptation* based on *individual adjustment* of the agents. We investigate adaptation strategies from two perspectives: (1) convergence to optimal performance; (2) emergence of subgroups with specialized behaviors. We implement a distributed adaptation algorithm where robots randomly change their WTP with a frequency based on their own performance. In the second algorithm we add an exchange mechanism where WTPs of successful agents are assigned to underperforming ones. Both algorithms converge to configurations that ensure optimal performance. The WTP exchange mechanism increases the cohesion of the WTP distribution, causing the system to converge to bounded uni- or bimodal distributions.

2 Model and Analytical Results

2.1 The Stick Pulling Problem

N are robots tasked with pulling sticks from the ground. The S_T sticks are randomly distributed in the workspace. Two robots are required to pull a stick. Robot behaviors are sketched in Figure 1. Robots initially *wander* in search of sticks and can *discover* sticks in their immediate vicinity. When a robot finds a stick held by another robot, the robots pull the stick together. If a robot finds a free stick, it *holds* it waiting for another robot to come along, but will release it after a certain time. We model the discovery of sticks as a stochastic process, characterized by a *discovery rate* k_D , the same for all sticks, whether or not they are held by another robot. This rate accounts for all physical and technological constraints, such as: the physical density of sticks the size and accessibility of the area, the movement and detection capabilities of the robots. The numbers of sticks and robots are constant. The only element that can be chosen by design is the behavior of the robots upon discovery of a free stick. *Release* after waiting is described as a Poisson process whose characteristic time is the *waiting time parameter* (WTP) τ_i , set individually for each agent.

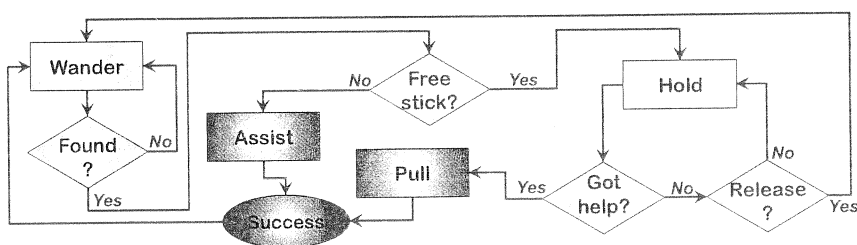


Fig. 1 Flow chart of robot behaviors in the stick pulling model

2.2 Equations of Motion

The N robots are subdivided into $p \leq N$ groups; N_i is the number of agents in group i . There are S_T sticks in total. At any time, a robot may be *free* (wandering), *c* holding a stick. We denote the number of free robots of type i with F_i , and by H_i the number of those holding a stick. The total number of free sticks and free and holding robots are denoted by S , F , and H . If total number robots of each type are fixed, we have:

$$\begin{aligned}
 N &= \sum_{i=1}^p N_i ; H = \sum_{i=1}^p H_i ; F = \sum_{i=1}^p F_i \\
 S_T &= S + H ; N = F + H ; N_i = F_i + H_i, \forall i \in \{1, \dots, p\}
 \end{aligned} \quad (1)$$

Robots in a group have the same WTP, τ_i , the *average* time a robot holds on to stick before *releasing* it. The release is controlled by a Poisson process whose rate

$\lambda_i = 1/\tau_i$. Similarly, the process of discovery of sticks by free robots is characterized by the *discovery rate* k_D . Due to (1), the state of the system is defined by the number of robots of each type holding a stick, $\{H_1, \dots, H_p\}$. We will write our equations in terms of these variables. Since we are interested in large swarms, we will adopt a *continuum approach* in describing the dynamics of the system [2].

There are three processes that contribute to the variation of H_i , *capture* (discovery), *pull*, and *release* of sticks. The *capture rate* $r_{capt}^{(i)}$ is proportional to the number of free robots of type i and the number of free sticks. The *pulling rate* $r_{pull}^{(i,j)}$ is proportional to the number of free robots of type i and the number of robots of type j holding a stick. These two processes have the same rate constant, k_D . Note that the pulling rate does *not* impact the number of free robots of the type of the second participant. By contrast, the robot that was holding the stick changes its state from holding to free. We denote by $r_{pull}^{(i)}$ the *total* rate of successful pulls of sticks held by robots of type i . The *release rate* $r_{release}^{(i)}$ of sticks by robots of type i is proportional to the number of robots of type i holding a stick, and the rate constant $\lambda_i = 1/\tau_i$.

$$\begin{aligned} r_{capt}^{(i)} &= k_D F_i S = k_D (N_i - H_i)(S_T - H) & ; & \quad r_{pull}^{(i,j)} = k_D F_i H_j = k_D (N_i - H_i) H_j \\ r_{release}^{(i)} &= \lambda_i H_i & ; & \quad r_{pull}^{(j)} = k_D (N - H) H_j \end{aligned} \quad (2)$$

The net rate of change in H_i is then:

$$\frac{dH_i}{dt} = k_D [(N_i - H_i)(S_T - H) - H_i(N - H)] - \lambda_i H_i \quad (3)$$

2.3 Steady State Analysis

We are interested in the steady-state(s) of (3). For any such configuration, the right-hand side of the equations of motion must vanish. Setting $\frac{dH_i}{dt} = 0$, we have:

$$k_D [N_i(S_T - H) - H_i(S_T + N - 2H)] = \lambda_i H_i \quad (4)$$

This equilibrium condition is more transparent in terms of dimensionless variables:

$$\frac{S_T}{N} \equiv \sigma ; \quad \frac{H}{N} \equiv \phi ; \quad \frac{H_i}{N_i} \equiv \varphi_i ; \quad \frac{N_i}{N} \equiv \rho_i \quad \longrightarrow \quad \varphi_i = \frac{\sigma - \phi}{\xi_i + (1 + \sigma - 2\phi)} \quad (5)$$

The *dimensionless time parameter* ξ_i is the ratio between the average time between two discoveries of the same stick by two robots, and the waiting time parameter τ_i :

$$\xi_i \equiv \frac{\lambda_i}{N k_D} = \frac{1}{N k_D \tau_i} = \frac{1/(N k_D)}{\tau_i} \quad (6)$$

The *occupancy fraction* ϕ is a weighted average of the individual *occupancies* φ_i . Substituting the individual equilibrium conditions, we arrive at a global condition:

$$\phi \equiv \frac{H}{N} = \sum_i \rho_i \varphi_i \longrightarrow \phi = \sum_i \frac{(\sigma - \phi) \rho_i}{\xi_i + (1 + \sigma - 2\phi)} = f(\phi) . \quad (7)$$

It can be shown that the equation $\phi = f(\phi)$ has a unique solution, which corresponds to a stable equilibrium of the equations of motion (3).

2.4 Pulling Rates and Optimality

The global *pulling rate*, (given N robots of which H are holding sticks) is

$$R_{pull} = k_D(N - H)H . \quad (8)$$

Since $0 \leq H \leq N$, R_{pull} is always positive, vanishes for $H = 0$ and $H = N$, and is maximal for $H = H^* \equiv \min(N/2, S_T)$. The *maximal pulling rate* is R_{pull}^* below:

$$R_{pull}^* = k_D(N - H^*)H^* ; R_{pull}^*(N, S_T) \leq R_{pull}^{max}(N) \equiv \frac{1}{4}k_D N^2 . \quad (9)$$

Here, R_{pull}^{max} is the maximal pulling rate *for N robots*; it may be achieved if there are enough sticks ($S_T > N/2$). If the number of robots N is larger than $2S_T$, the maximal pulling rate is limited to $k_D(N - S_T)S_T$. We will assume $N < 2S_T$, so $R_{pull}^* = R_{pull}^{max}$.

2.4.1 Optimal WTP Configurations

The objective of designing our swarm is to maintain the system performance as close to the ideal situation $H = N/2$ as possible. For a given configuration of groups and WTPs, we can calculate the equilibrium state of the system and the corresponding pulling rate, by solving the equilibrium condition (7) for the global occupancy ϕ , and use it to calculate the individual occupancies. We then specify conditions for optimality by requiring $\phi = 1/2$. A configuration of waiting time parameters that results in $\phi = 1/2$ is called *optimal* or *ideal*.

One species: If all agents have the same WTP τ , the equilibrium condition (9) reads

$$2\phi^2 - (2 + \sigma + \xi)\phi + \sigma = 0 . \quad (10)$$

Of the two solutions for ϕ , only one is in the $[0, 1]$ interval. The design problem here consists of determining the waiting time parameter τ (through the dimensionless time parameter $\xi = 1/Nk_D\tau$) so that optimal performance is achieved. We can calculate the value of the ideal $\xi = 1/Nk_D\tau$ by substituting $\phi = 1/2$:

$$\xi^* = \sigma - 1 \leftrightarrow \tau^* = \frac{1}{k_D(S_T - N)} . \quad (11)$$

Two types of robots: We have to design three quantities, the two WTPs τ_1, τ_2 , and the ratio ρ_1/ρ_2 between the sizes of the groups. Given $\{\xi_1, \xi_2, \rho_1, \rho_2\}$, the equilibrium configuration is uniquely defined. Choosing $\rho_1 = \rho_2 = 1/2$, we obtain the following constraint for the dimensionless time parameters $\{\xi_1, \xi_2\}$:

$$\frac{1}{\xi_1 + \sigma} + \frac{1}{\xi_2 + \sigma} = \frac{2}{2\sigma - 1}. \quad (12)$$

Equal size groups: Consider a number of p different species, each representing $1/p$ of the population.¹ The optimality condition is

$$\frac{p}{2\sigma - 1} = \sum_{i=1}^p \frac{1}{\xi_i + \sigma} \iff \frac{1}{p} \sum_{i=1}^p \frac{1}{\xi_i + \sigma} = \frac{1}{2\sigma - 1}. \quad (13)$$

For $p = 1$ we reobtain the condition for the ideal ξ . The second version of the condition can be interpreted as a requirement that the average of the quantities $1/(\xi_i + \sigma)$ match the ideal value $1/(2\sigma - 1)$.

2.4.2 Robustness Measures

The optimality requirement (13) represents a single algebraic constraint. With p equal sized groups, all but one of the WTPs $\{\tau_1, \dots, \tau_p\}$ may take any value over a semi-infinite interval. The corresponding configurations form a $p - 1$ dimensional manifold in the p -dimensional space of WTP configurations. We are interested in *additional* performance criteria to characterize these ideal configurations.

Performance under changing conditions: Consider the pulling rate of a system that is optimal for a stick/robot ratio of σ_0 , when faced with a different $\sigma \neq \sigma_0$. In Figure 2 we compare a one-group configuration with $\tau = \tau^*$ (11) and two configurations of two groups of equal size with WTP pairs $\{\tau_1, \tau_2\}$ that satisfy (12), for $\sigma = \sigma_0 = 10$. The larger τ_1 , the smaller τ_2 has to be. The factor $K = \tau_1/\tau^*$ is a measure of how far the $\{\tau_1, \tau_2\}$ pair is from the one-species case ($K = 1$). Theoretical predictions for the pulling rate for $K = 1, 10, 100$ are confirmed by simulation results as indicated. The loss of performance is the strongest for the one-group configuration ($K = 1$) and becomes milder as the ratio between τ_1 and τ^* increases.

Loss of agents: As a measure of how much of the optimality would be preserved by a subset of the agents in a given configuration, it is useful to compare the pulling efficiency *per agent* for a configuration where some agents are destroyed. This measure is relevant when comparing WTP configurations that result from randomized adaptation algorithms, where no two agents would likely have the same WTP. It provides a mechanism to penalize configurations that are “too heterogeneous”.

¹ This also applies to the situation when the robots are essentially independent, taking $p = N$.

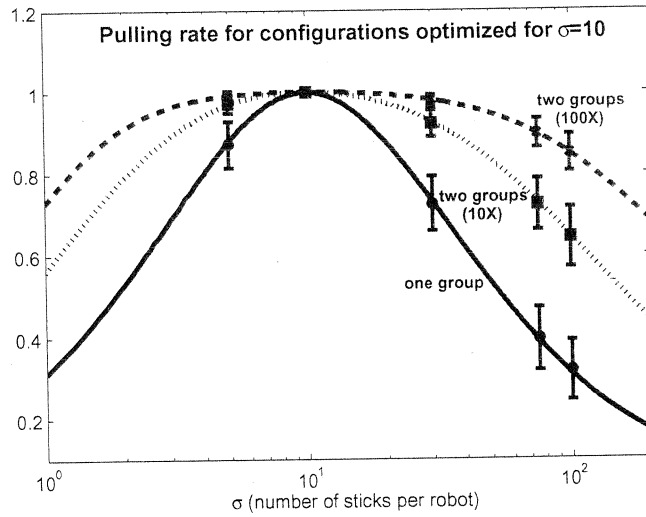
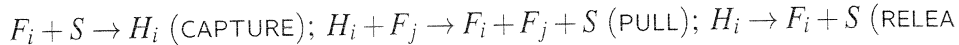


Fig. 2 Efficiency loss when conditions differ from ideal. Theoretical predictions and simulation results (points) for the pulling rate of a one-group and several two-configurations that are optimal for $\sigma = 10$. The WTP pairs $\{\tau_1, \tau_2\}$ of the two-group configurations satisfy (12), with the stated values of the ratio $K = \tau_1/\tau_2^*$ (10 and 100, respectively). The simulations (one per data point) had $N = 150$ robots and numbers of sticks that $\sigma = S_T/N = 5, 10, 30, 75, 100$.

3 Simulation Methods

3.1 Basic Simulation Algorithm

The main simulation algorithm is derived from the Gillespie algorithm used in chemistry and adapted to multi-robot systems in previous work [4]. The state system is defined by that of the individual agents. Each of the N agents can be free (F) or holding a stick (H). There are three possible transitions, corresponding to three processes discussed above:



In the CAPTURE(i) process, agent i goes from Free to Holding; the reverse is the RELEASE(i) process. In the PULL(i, j) process, agent i goes from holding to Free but the process requires another agent (j), whose state is not ultimately changed. Transitions are controlled by independent Poisson processes; the probability per unit time (or rate) for a specific transition is given by a time constant and the number of eligible partners, if applicable. For example, if both agents i and j are free, the probability per unit time for capturing a stick is the same for both of them, k_L for agent i while holding a stick is $\lambda_i = 1/\tau_i$.

In the Gillespie algorithm, simultaneous Poisson processes are simulated by generating *next event times* for each process, then implementing the state transition that corresponds to the smallest one of the next event times. When there are

possible transitions, one may calculate the cumulative transition rate for each *type* of transition, then choose a specific (pair of) agent(s) for the transition. This approach is correct for simultaneous Poisson processes and we do this for events that involve encounters between free robots and sticks. The additional computational cost due to updating the states of *individual agents* is almost negligible. Thus, the Poisson model for transitions allows us to have the equivalent of an agent-based simulation for the cost of a centralized one.

3.2 Individual Adaptation

We construct a self-evaluation measure or *satisfaction level* χ_i for agent i , as follows. Every time agent i participates in a successful pull (in either role), χ_i is incremented by 1. At every update, χ_i decreases exponentially with a characteristic time τ_{forget} : $\chi_i(t + \Delta t) = \chi_i(t) \exp(-\Delta t / \tau_{forget})$. Thus, agents have a memory of past successes, but their satisfaction level decreases as they go through a dry spell.

The satisfaction level defined here does not provide an absolute measure of an individual agent's effectiveness. There is no reference value for it, unless the agent knows what pulling rate it should expect. The maximal pulling rate can be computed from the number of sticks and agents; however, we are interested in an adaptive strategy that can find the optimum *without* relying on global knowledge.

In this algorithm, each agent changes their WTP *randomly*, at a rate proportional to *the inverse* of the agent's satisfaction level (lower satisfaction increases the rate of change). Adaptation is implemented as a Poisson process with time constant τ_{learn} / χ_i , that runs in parallel with the other transitions (but much slower). Every time this process fires, the respective agent changes its WTP with a small random quantity: $\tau_i' = \tau_i \exp((r - 1/2)\Delta)$ where r is a uniformly distributed random number between $[0, 1]$ and Δ is the *Monte-Carlo (MC) step size* (typically a small number). This algorithm results in a random walk in the space of $\log(\tau_i)$ biased by the satisfaction function. Our approach is simpler than the one used by Li *et al.*, but it also relies on a proper self-assessment of performance.

3.3 Swapping

As we discuss next, the individual adaptive strategy succeeds in optimizing the pulling rate, but generates configurations where the individual WTPs are spread over many orders of magnitude. In order to increase the coherence of the resulting WTP distributions, we introduced a *collective mechanism* to supplement individual adaptation. It consists of an additional WTP change, performed with a small (fixed) probability v , during normal WTP updates. This corresponds to an additional Poisson process, with propensity v / τ_{learn} . When this process fires, we select a pair of agents, a *donor* (with a high satisfaction level), and an *acceptor* (with a low satisfaction level), and change the WTP of the acceptor to that of the donor. This procedure is reminiscent of biologically inspired algorithms. While it requires some degree of

collective communication, it can be implemented in a way that ensures reasonable scaling as the number of agents increases. The key is in that the donor and acceptor agents can *self-select* and communicate using a pre-determined procedure of asynchronous communication to upload or download their WTPs.

4 Results and Discussion

4.1 Model Validation - Equilibration

The standard system in our simulations consists of $N = 150$ robots and $S_T = 2000$ sticks. We use time units where the discovery rate $k_D = 1$. Thus, the maximal pulling rate is $\frac{1}{4}N^2k_D = 5625$ pulls per unit time. The corresponding optimal waiting time is $\tau^* = 1/k_D(S_T - N) = 5.4 \times 10^{-4}$. The average time between two robot-stick encounters is $\tau_E = 1/(k_D S_T N) = 3.33 \times 10^{-6}$. The time between two consecutive updates in a simulation is on the order of 10^5 iterations per time unit. We performed simulations with various WTP configurations and verified that the system converges to the average occupancy fractions and pulling rates predicted by the continuum equations in Sec.2. The value of the equilibration time is comparable to the average time of 5×10^{-4} units it takes for one robot to find one of the 2000 sticks. Figure 2 shows theoretical and simulation results for the equilibrium pulling rate for configurations with one or two WTP groups, for the same number of robots ($N = 150$), but different numbers of sticks (simulations with $\sigma \equiv S_T/N = 5, 10, 30, 75, 100$). All configurations are ideal for $\sigma = 10$ ($S_T = 1500$ sticks). The simulation results confirm the analytical predictions given in Sec.2.4.1.

4.2 Individual Adaptation Algorithm

We implemented the individual adaptation algorithm described in Sec.3.2 on the $N = 150, S_T = 2000$ system, exploring parameter values around $\tau_{learn} = 1.0 \times 10^{-4}$, $\tau_{forget} = 0.1$ and a Monte-Carlo step size of $\Delta = 1.0 \times 10^{-2}$.

The evolution of the system with these parameter values is shown in Figure 3. The waiting time parameters are initially set to 50% of the optimal value τ^* . As the individual τ values change, the number of free robots evolves, reaching the optimum of 75 in approximately 400 time units. This τ -convergence time is significantly longer than the *equilibration time* of 5×10^{-4} it takes the number of free robots to reach the equilibrium value corresponding to a *fixed* WTP configuration. It is useful to visualize the time evolution of WTPs using the distribution of the $\log(\tau)$ values as in Figure 3. The results are qualitatively different from the one- or the two-group configurations described previously. As the simulation starts, the $\log(\tau)$ distribution spreads out and continues to do so over time, expanding into extremely large and small (positive) values, reaching widths of 10 orders of magnitude and higher after 10^8 steps. The $\log(\tau)$ distribution is close to a normal, whose standard deviation increases like the square root of the simulation time.

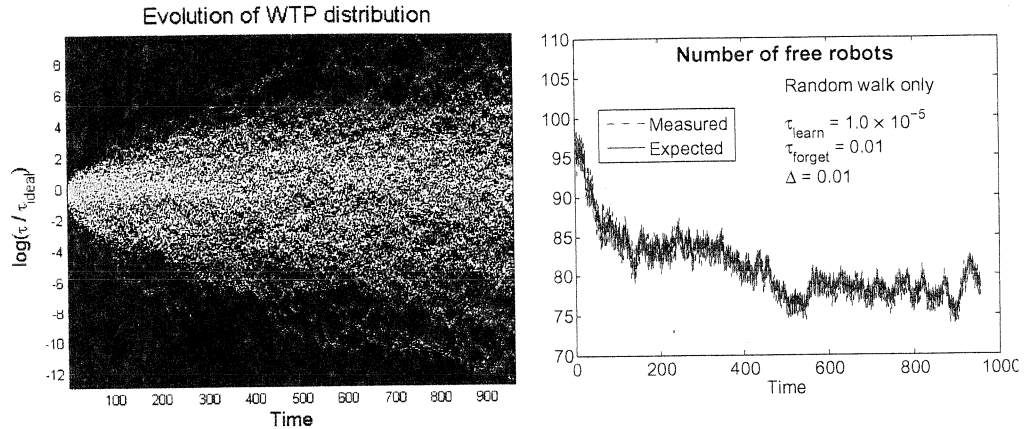


Fig. 3 (Left) Evolution of the distribution of WTPs in an adaptive simulation. Notice the similarity to a diffusion process. (Right) The number of free robots in the same simulation (expected = based on the current configuration of WTPs). Based on a single simulation with $N = 150$ robots, $S_T = 2000$ sticks, and 3×10^8 updates.

We performed a number of simulations to investigate the effect of changing the adaptation parameters on convergence. The results are presented in Figure 4. We define convergence for the purposes of these simulations as the state (after the initial equilibration) where the moving average over 10,000 iterations of the number of free robots is within 1 of the ideal value of 75. We limited the simulations to 10^7 iterations, and we plot both the time to convergence and the number of held sticks after the maximum number of iterations. For the converged simulations, the final number of sticks is very close to 75, and the convergence time varies. For the unconverged simulations, better adaptation corresponds to final sticks held counts that are closer to the ideal value of 75.

The dependence on the averaging time τ_{forget} is relatively weak. None of the simulations using this algorithm converged (within the iteration limit), due to the values for Δ and τ_{learn} . However, the final state approaches 75 as τ_{forget} is reduced by a factor of 10, and moves further away as τ_{forget} is increased. Increase in the frequency of WTP changes (decreased τ_{learn}) leads to marginal improvement. A 10-fold increase in the Monte-Carlo step size Δ improves the adaptation performance to the point where the system converges within the 10^7 iteration cutoff. Further increase of the step size leads to additional improvement in the convergence time. However, the configurations reached in this manner are increasingly incoherent, with a very wide WTP distribution.

4.3 Swapping Algorithm

The introduction of swapping leads to dramatically improved convergence, over all parameter values investigated. It is remarkable that WTP swapping, performed at a frequency corresponding to one swap per every 100 individual WTP changes, improves convergence this much. The parameter sensitivity results for this algorithm

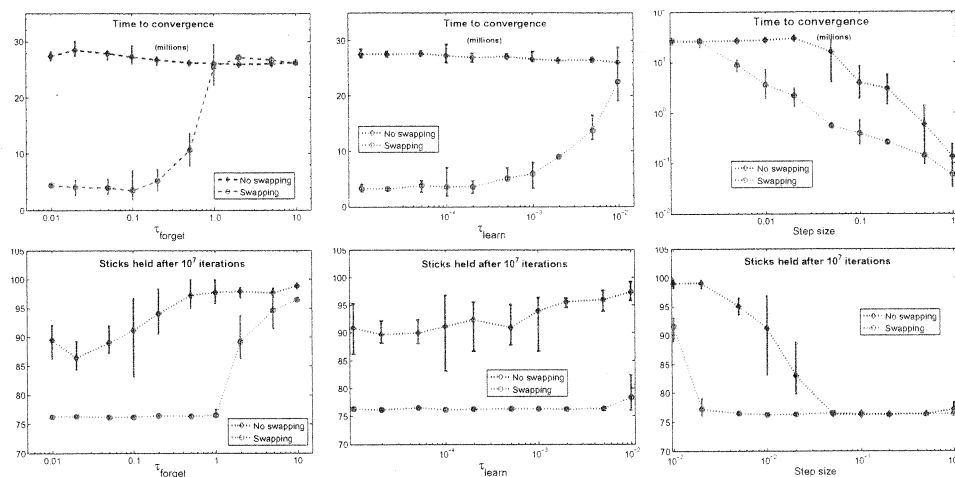


Fig. 4 Convergence time (*top*, in millions of iterations) and sticks held after 10^7 iterations (*bottom*) versus τ_{forget} , τ_{learn} , and MC step size Δ , in the individual adaptive algorithm (blue) and with swapping (red). Each point represents 5-20 simulations with $N = 150$ robots and $S_T = 2000$ sticks.

are also plotted in Figure 4. The effect of parameter changes is qualitatively similar in the two algorithms. The swapping result converges for all but the highest values of the averaging time τ_{forget} . The performance of the algorithm deteriorates as this parameter is increased, and convergence is lost as τ_{forget} goes from 0.1 to 1.0. Increased τ_{learn} also reduces the performance of the swapping algorithm.

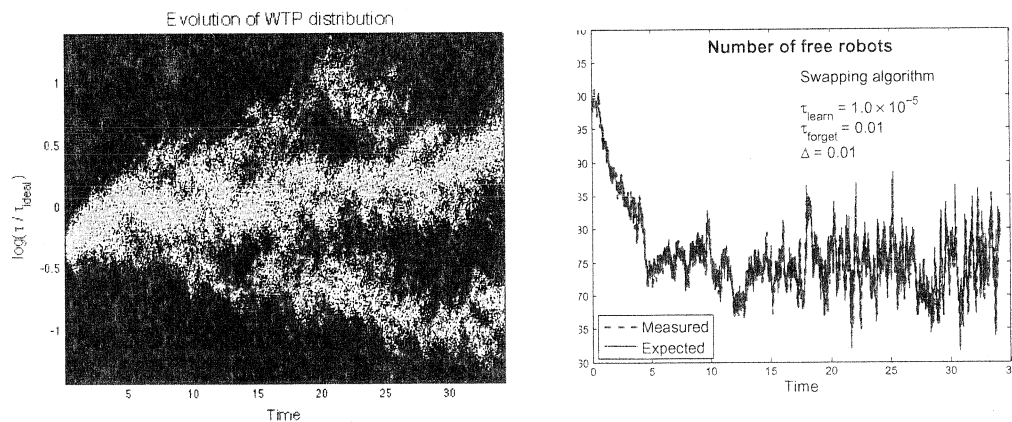


Fig. 5 Evolution of the distribution of waiting time parameters (left) and the number of free robots (right) for an adaptive simulation with swapping (expected = based on the current configuration of WTPs). Note the different time scale of convergence compared to the non-swapping simulation shown in Figure 3; the two simulations have the same adaptation parameters. Based on a single simulation with $N = 150$ robots, $S_T = 2000$ sticks, and 10^7 update steps.

The first algorithm was the most sensitive to the Monte-Carlo step size Δ . Increased Δ improved the convergence of both algorithms, and their performance became similar as $\Delta \approx 1$. A value of $\Delta = 1$ means that the random change of a WTP is comparable to the value of the WTP. With such large variation steps, the newly selected parameters have little to do with the previous ones. In this limit, the algorithm tends to become purely random selection of parameters rather than a search process (on the level of *individual* agents).

Swapping dramatically limits the expansion of the WTP distribution, as shown in Figures 5 and 6. For most parameter values (except very high Δ) the simulations resulted in bounded, unimodal distributions with a spread of little more than one order of magnitude, much less than in the individual adaptation case (compare Fig.6 and 3). In the longer term, some of the simulations exhibit transitions to bimodal distributions. The bimodal distributions we observed had narrow modes, with maxima separated by 1-2 orders of magnitude. While the bimodal distributions extended over almost three orders of magnitude, they remained bounded, and the system eventually transitioned back to the unimodal regime.

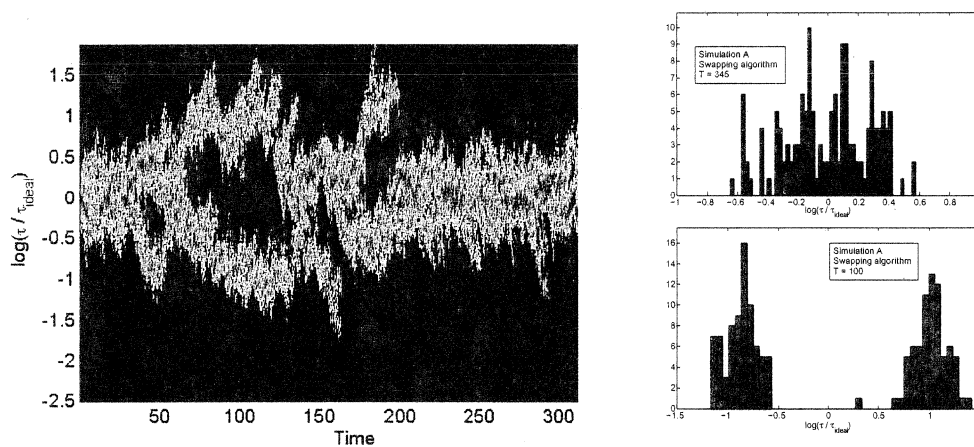


Fig. 6 Evolution of the distribution of waiting time parameters during a long simulation with swapping. Based on a single simulation with $N = 150$ robots, $S_T = 2000$ sticks, and 10^8 update steps.

4.4 Waiting Time Parameter Distributions

From a design and analysis perspective, configurations with one, two or a small number of distinct waiting time parameters seem more straightforward. By contrast, both adaptation algorithms result in configurations that can only be characterized by a continuous distribution of waiting time parameters, rather than one or a few distinct WTPs shared by groups of agents.

The evolution of the $\log(\tau)$ in the individual adaptation algorithm (Figure 3) is similar to pure diffusion, consistent with a random walk. This makes sense because each individual WTP change is a small variation taken from a distribution that is

symmetric in terms of $\log(\tau)$. This random walk is influenced through the satisfaction function and is practically confined to globally optimal configurations. The global optimality condition (13) corresponds to a single constraint on the N waiting time parameters. If the individual $\log(\tau)$ are allowed to increase or decrease indefinitely, most agents will have waiting times that are either much larger or much smaller than the ideal value. In this case, the $1/(\xi_i + \sigma)$ terms in Eq.(13) would approach $1/\sigma$ and 0, respectively. The optimality condition for a configuration with N_{high} agents with very high WTP ($\tau_i \gg 1 \rightarrow \xi_i \ll 1$) and the rest with very small WTP is simply

$$\frac{N_{high}}{N} = \frac{\sigma}{2\sigma - 1} = \frac{S}{2S - N} \quad (14)$$

This simple constraint on the values of the WTPs of the agents in the high and low groups ensures optimality for the late WTP configurations obtained in the individual adaptation algorithm. We will call these *divergent-optimal* or *DO* configurations. Presumably, these could also be obtained more easily, by a simple random search on the level of individual agents. A DO configuration can be interpreted as an example of specialization (castes with $\tau = \{0, \infty\}$).

5 Summary and Conclusions

We have expanded our analysis [4] of the stick-pulling problem and established that each WTP configuration corresponds to a unique equilibrium pulling rate which can be estimated analytically. We showed that there is a maximum possible or *optimal* pulling rate for a given number of sticks and robots (9). The optimality requirement can be formulated as a single algebraic condition (13) for the N parameters.

We designed and implemented two adaptive optimization strategies and showed that both converge to optimal configurations. The *individual adaptation algorithm* relies exclusively on the agents' own record of their performance, in the form of a *satisfaction function*. Robots change their WTP based on this function (low satisfaction \rightarrow higher change rate). Each change is a Monte-Carlo step in a random direction. The evolution of the WTP distribution in this algorithm is consistent with diffusion. The distribution of $\log(\tau)$ approaches a normal whose width increases indefinitely, while maintaining optimal performance. The long-term limit for this type of distribution, called *divergent-optimal* (DO), has WTPs that approach either zero or infinity. Optimality can be ensured by the appropriate ratio between the two groups (14). DO configurations can be regarded as extreme examples of emerging specialization. The $\tau \rightarrow \infty$ species specializes in discovering and holding sticks, and the $\tau \rightarrow 0$ specializes in assisting stick holders.

In the *swapping algorithm* we supplement individual adaptation with a mechanism that assigns the WTP of well performing agents to under-performing ones. While requiring a limited amount of global communication, this algorithm leads to dramatic improvement of the rate of convergence. It also limits the width of the

WTP distributions. Increased Monte-Carlo step size in the swapping algorithm leads to faster convergence but eventually results in the emergence of DO configurations.

Emergence of specialization can also be observed in the swapping algorithm, where long-term simulations fluctuate between *bounded* uni- and bimodal distributions with narrow modes. The bounded bimodal configurations are closer to the idea of specialized groups, each with a narrowly defined set of features (similar to biological phenotypes).

In conclusion, our results provide two mechanisms by which specialized groups of agents can emerge from an agent-based adaptation strategy. The more easily obtained DO configurations may not be satisfactory for a given application. Further refinements are necessary to stabilize the bounded bimodal configurations. This will require more sophisticated measures of performance, which can enforce our preference for one or another type of WTP distribution. We gave two possible examples of such measures that may be implemented in future applications. Finally, future work in this direction should also integrate results from machine learning and information theory.

References

1. Balch, T.: Hierarchic social entropy: An information theoretic measure of robot team diversity. *Autonomous Robots* 8(3), 209–238 (2000)
2. Berman, S., Halasz, A., Hsieh, M.A., Kumar, V.: Optimized stochastic policies for task allocation in swarms of robots. *IEEE Transactions on Robotics* 25(4), 927–937 (2009)
3. Bonabeau, E., Dorigo, M., Theraulaz, G.: *Swarm Intelligence: From natural to artificial systems*. Oxford University Press, New York (1999)
4. Hsieh, M.A., Halasz, A.M., Cubuk, E.D., Schoenholz, S., Martinoli, A.: Specialization as an optimal strategy under varying external conditions. In: *Proceedings of the International Conference on Robotics and Automation (ICRA)*, Kobe, Japan (2009)
5. Ijspeert, A., Martinoli, A., Billard, A., Gambardella, L.M.: Collaboration through the Exploitation of Local Interactions in Autonomous Collective Robotics: The Stick Pulling Experiment. *Autonomous Robots* 11(2), 149–171 (2001), doi:10.1023/A:1011227210047
6. Kussell, E., Leibler, S.: Phenotypic diversity, population growth, and information in fluctuating environments. *Science* 309(5743), 2075–2078 (2005)
7. Lerman, K., Galstyan, A., Martinoli, A., Ijspeert, A.J.: A Macroscopic Analytical Model of Collaboration in Distributed Robotic Systems. *Artificial Life* 7(4), 375–393 (2001), doi:10.1162/106454601317297013
8. Li, L., Martinoli, A., Abu-Mostafa, Y.: Learning and Measuring Specialization in Collaborative Swarm Systems. *Adaptive Behavior* 12(3-4), 199–212 (2004); Special issue on Mathematics and Algorithms of Social Interactions, Anderson, C., Balch, T. (eds.), doi:10.1177/105971230401200306
9. Mataric, M.J.: Designing and understanding adaptive group behavior. *Adaptive Behavior* 4, 50–81 (1995)