Improving collective decision accuracy via time-varying cross-inhibition

Mohamed S. Talamali, James A. R. Marshall, Thomas Bose, Andreagiovanni Reina

***Abstract*— We investigate decentralised decision-making, in which a robot swarm is tasked with selecting the best-quality option among a set of alternatives. Individual robots are simplistic as they only perform diffusive search, make local noisy estimates of the options' quality, and exchange information with near neighbours. We propose a decentralised algorithm, inspired by house-hunting honeybees, to efficiently aggregate noisy estimations. Individual robots, by varying over time a single decentralised parameter that modulates the interaction strength, balance exploration and agreement. In this way, the swarm first identifies the options under consideration, then rapidly converges on the best available option, even when outnumbered by lower quality options. We present stochastic analyses and swarm robotics simulations to compare the novel strategy with previous methods and to quantify the performance improvement. The proposed strategy limits the spreading of errors within the population and allows swarms of simple noisy units with minimal communication capabilities to make highly accurate collective decisions in predictable time.**

I. Introduction

Decision making is a key ability for any living organism or artificial system. Robot swarms are systems composed of a large number of simple and autonomous robots, and such systems are frequently required to make *collective* decisions in which all robots agree on one option among several available alternatives. Agreeing on a unique option allows the swarm to operate in unison and to express a coordinated response to external stimuli. For example, when the swarm needs to allocate all its resources to a single task which is localised in space, the swarm has first to decide at which location to perform the task, among the candidate spots. Consensus would guarantee an effective usage of the swarm's resources, while splitting would dilute the swarm's power and may hinder success.

While certain applications only require to reach a consensus for any option [1], [2], [3], in this study we ask the swarm to decide for the best quality option among n alternatives. This problem is known in the literature as the best-of-n decision problem [4], [5], [6], [7], [8], [9], [10]. In this work, we solve the best-of-n problem through a decentralised strategy inspired by house-hunting honeybees [11] and later adapted to artificial swarm systems [6], [12]. This study employs multi-scale modelling [13] to predict the system dynamics and tests the predictions through physics-based simulations. Previous swarm robotics studies relied on multi-scale modelling [14], [15], [16], [17], [18] and we believe that it is a powerful tool to engineer robot swarms. Most collective decision studies in swarm robotics limited their analysis to the binary case of $n = 2$ options [10], however it has been shown that decision models may qualitatively change their dynamics for $n > 2$, [19]. For this reason, here we investigate collective decisions with larger n.

The proposed strategy shows a considerable improvement of the decision accuracy and a constant decision speed for any tested number of options and problem difficulty. While this study focuses on the speed-accuracy analysis, previous research has shown that the investigated model can also be a mechanism for value-sensitive decisions [20], [21], [19]. The most relevant advantages of value-sensitive approaches are, in cases of equal-quality options, the ability to break or maintain decision deadlocks as a function of their quality, otherwise to select the uniquely best option when quality differences become large enough [20], [22].

In this study, n options are deployed in the environment and the robot swarm has no prior knowledge of the number, qualities, and locations of the available options. The swarm is tasked with searching for the options and selecting the highest-quality one (see problem description in Sec. II). While other methods may focus only on the exploration or the decision aspect, the investigated strategy includes both activities (see the robot behaviour in Sec. III). At the same time, it keeps the individual behaviour simple and limited to four probabilistic transitions (Sec. V). The proposed strategy is compared with other methods (Sec. IV) through stochastic analysis of the macroscopic model (Sec. V-C) and robot swarm simulations (Sec. VI). Limitations and possible extensions of the study are discussed in Sec. VII.

II. Decision problem formalisation

In this work, a swarm of S robots is required to solve the best-of-n decision problem, i.e. to reach consensus on the best-quality option out of n options available in the environment. Each option i (with $i \in \{1,...,n\}$) is located in space at position χ_i and has a quality $v_i \in [v_{min}, v_{max}]$ (v_{min}/v_{max} are the min/max quality that the robot can sense). The decision problem is to choose the option with highest value. The robots have no prior knowledge about the number of options, their location, or their qualities; robots can perceive an option only when they are in its proximity. The robots explore the environment to find the options, estimate their qualities, and collectively select the best-quality option. When a robot finds an option i, it can make an individual noisy estimate of its quality $\hat{v}_i$. We simulate noisy estimations

This work was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement number 647704).

The authors are with the Department of Computer Science, University of Sheffield, S1 4DP, UK (email: mstalamali1@sheffield.ac.uk; james.marshall@sheffield.ac.uk; t.bose@sheffield.ac.uk; a.reina@sheffield.ac.uk).

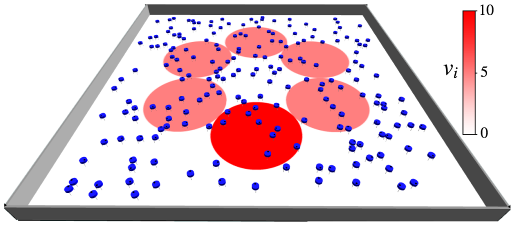


Fig. 1. Sample initial distribution of $S = 200$ simulated Kilobots in a scenario with $n = 6$ options and decision difficulty $\kappa = 0.5$. The red circles represent the areas (radius 25 cm) in which the options can be perceived by the robots via ARK. The colour intensity represents the option's quality $v_i \in [0, 10]$. The swarm is tasked to select the best-quality option.

as samples from a normal distribution $\mathcal{N}(v_i, \sigma^2)$ with mean v_i, variance σ^2, and reassignment to boundary values for samples out of range $[v_{min}, v_{max}]$.

Each robot has limited memory and can only store the location and quality of a single preferred option, which represents its commitment. In collective decision making, the decision is made when a quorum in favour of one option is reached [23], [24], [25], [26]. In our study, consensus is reached for option i when the number of robots committed to i reaches the quorum threshold $Q = 80\%$.

A. Experimental setup

Most studies focused on binary decision problems ($n = 2$) [20], [22], [8]; here we investigate the best-of-n problem with $n \geq 2$. We consider the scenario of one superior-quality option and $n-1$ inferior-quality distractors. This experimental scenario has been adopted in several empirical and theoretical studies in various domains as it allows to systematically vary the difficulty of the decision problem without increasing the number of experimental parameters [27], [28], [29], [30], [19]. The superior option has quality $v_{_H}$ while the other $(n-1)$ inferior equal-quality options have quality $v_{_L}$ (with $v_{_H}, v_{_L} \in [v_{min}, v_{max}] = [0, 10]$). The decision difficulty can be expressed as the ratio between inferior and superior qualities $\kappa = v_{_L}/v_{_H} \in [0, 1]$.

In this study, we simulate a robot swarm composed of $S = 200$ Kilobots [31] which are low-cost robots designed specifically to conduct large-scale swarm robotics experiments. Kilobots operate at a clock frequency of about 32 Hz which corresponds to a clock period $\delta_c \simeq 31$ ms. Kilobots are equipped with few sensors and actuators. They have differential-drive vibration motors to move on a flat surface at a speed of about 1 cm/s in straight motion and about 40 °/s in rotation. Kilobots are equipped with an IR transceiver to transmit 9-bytes messages to neighbours in a local range of ∼10 cm, Finally, Kilobots are equipped with an ambient light sensor, and an RGB LED to display their internal state. Despite their limited capabilities, Kilobots have been successfully employed in several swarm robotics studies, e.g. [32], [22], [33], [8], [34], [35]. Furthermore, Kilobot abilities can be enhanced via ARK, a form of 'augmented reality for Kilobots' [36], [37].

ARK allows Kilobots to use virtual sensors and actuators. In this work, we employ ARK to allow Kilobots to perceive an option i, estimate its quality $\hat{v}_i$, and geolocalise the option (i.e. compute χ_i). Robots within the option's sensing range (25 cm) receive via IR an ARK message which includes the option's location and its quality. Similarly, upon request to ARK, Kilobots can access to their GPS information (e.g. to navigate to the option's location). When a Kilobot lights up its red LED, ARK replies with its GPS information.

The $S = 200$ robots are initially uniformly distributed in a square environment 2 m × 2 m, as illustrated in Fig. 1. The n options are located at equal distance between each other on the vertex of a regular polygon with n edges and radius 50 cm (see an example with $n = 6$ options in Fig. 1). The superior-quality option is placed at a random vertex in each simulation run.

III. Robot behaviour

We simulated the Kilobot swarm behaviour with the physics-based simulator ARGoS [38], which allows quick simulation of large-scale swarms through its multi-threaded architecture. ARGoS supports the simulation of Kilobots through a dedicated plugin which is particularly convenient as it offers the possibility to run the same identical control code in simulation and on the real robots [39]. ARGoS also supports experiments with a simulated ARK infrastructure which allows quick conversion of ARK experiments between simulation and reality.

We implemented a generic Kilobot behaviour to plug in different decision strategies and test their performance. The behaviour is composed of three concurrent actions:

1) Environment exploration: Robots have no prior knowledge about the decision problem (number, location, and quality of the options), hence robots need to explore the environment in order to find the available options and estimate their qualities. A simple and effective way to search for options in an unknown environment is to perform a diffusive isotropic random walk [34]. In this study, the Kilobots alternate straight motion for approximately 10 s and rotations in a random direction for a random number of seconds chosen from a uniform distribution $U(0, 5)$ s. Beside allowing environment exploration, the random walk allows robots to encounter new peers and sample different information within the swarm. When a Kilobot perceives an option i, it stores the option's location χ_i and estimates its quality $\hat{v}_i$ which is used to update the internal commitment state.

It is important to note that the robots do not resample the same options multiple times to average qualities over time because we are not interested in individual strategies to attenuate the noise on individual estimates. We assume that in any real application scenario, the robot would conduct the necessary sampling operations to obtain the most accurate possible quality estimate. This estimate would anyway be subject to a certain level of noise, here modelled through a normal distribution $\mathcal{N}(v_i, \sigma^2)$. This study investigates collective strategies to efficiently aggregate noisy estimations.

2) Social interactions: While exploring the environment, robots interact with neighbours within a local range of about 10 cm. Each robot broadcasts a message every second with information on its commitment state and, if committed to an option, the option's location χ_i and (possibly) the estimated option quality $\hat{v}_i$. The receiving robots use this information to update their commitment state.

3) Commitment updates: Each robot r has two possible individual commitment states $\{U,C\}$; either committed (C) to an option i, or uncommitted (U). When committed, the Kilobots keep record of the option's location χ_i and the estimated quality $\hat{v}_i$. All robots start without any prior knowledge, therefore in the uncommitted state. Upon discovery of an option or incoming messages from other robots, each Kilobot updates its individual commitment state either by changing state or by remaining committed and only modifying the option i. We implemented and compared various update strategies which we describe in Sec. IV and V.

IV. Direct comparison strategy

A simple and naive strategy to collectively reach consensus on the best-quality option is the *direct comparison strategy* (DC) [40]. The DC strategy requires the Kilobots to share their estimated quality $\hat{v}_i$ with each other, and to update their commitment state as follows. An uncommitted robot exposed to information of an option i (either via independent discovery or via another robot's message) changes its commitment state to C and stores location χ_i and estimated quality $\hat{v}_i$. Instead, committed robots update their commitment to option i only when the new option j, received from another robot, is different (i.e. $i \neq j$) and has a better quality $\hat{v}_j > \hat{v}_i$. If $\hat{v}_j$ is higher, the robot forgets the previous option and stores the new received information χ_j and $\hat{v}_j$. When the two options have equal estimated quality $\hat{v}_i = \hat{v}_j$, the robot randomly selects option i or j. This random selection may allow the swarm to break the decision deadlock in case of equal-quality options.

As discussed in Sec. III, robots do not perform multiple sampling of the same option, instead performing an individual estimate of an option and let the information spread within the swarm. This simple strategy has the advantage to quickly reach consensus for one option. Although, at the same high speed, errors spread. In fact, an individual overestimation of an option is quickly accepted by other robots which, in turn, use it in their messages. In other words, robots use *second-hand* quality estimates (received from a neighbour) to recruit themselves other robots. The effect of noise on this decision strategy can be appreciated in Fig. 2(a) which shows ARGoS simulations results for a scenario with six options. We varied the problem difficulty $\kappa = v_L/v_H \in [0.5, 1.0]$ and the noise strength $\sigma^2 \in [0,5]$. Decision accuracy quickly decreases as the problem difficulty or noise strength increases.

V. Collective Decisions through Cross-inhibition

To overcome the poor performance of DC (Fig. 2(a)), we propose an extension of the Collective Decisions through

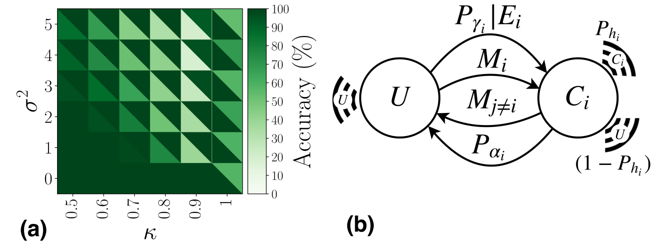


Fig. 2. (a) 200-Kilobot swarm results (100 simulations each condition) showing the effect of noise strength $\sigma^2 \in [0,5]$ on the decision accuracy in the best-of-6 problem with difficulty $\kappa = v_L/v_H \in [0.5,1]$ ($v_H = 10$). We compare the accuracy of the DC strategy of Sec. IV (bottom-left triangles) with the accuracy of the time-varying strategy $r_{step}(t)$ (with $\tau_0 = 50$) of Sec. V-C (top-right triangles). While DC is highly sensitive to noise, the proposed strategy shows remarkably high performance ($\geq 93\%$) for any tested noise level and difficulty κ up to 0.9. In case of equal-quality options ($\kappa = 1$), the quick dynamics of DC allows to break the symmetry within 2 hours more often than the proposed strategy with a suboptimal parameterisation of τ_0 (see more details in Fig. 5(b)). (b) The PFSM controlling the robot's individual decision state. Robots update their commitment state using the probabilities P_{γ_i} and P_{α_i} of Eq. (1) or upon receiving a message M_i from a robot committed to i. The symbol $|$ on the arrow for the discovery transition indicates conditional probability on the occurrence of the event E_i of encountering option i. The transmission symbols indicate that a robot in state C_i sends an interaction message (for recruitment and cross-inhibition) with probability P_{h_i}.

Cross-Inhibition strategy (CDCI) proposed in [6]. This strategy has been inspired by the house-hunting process of European honeybees [11] and been applied in a variety of swarm robotics experiments [41], [22]. This work extends the previous CDCI strategy by removing any need to share quality estimates between the agents and by introducing time varying interactions. We show through various comparisons an improvement in both decision speed and accuracy.

A. The basic CDCI strategy

Each robot's commitment state is controlled by the probabilistic finite state machine (PFSM) shown in Fig. 2(b). The active state of the PFSM corresponds to the individual commitment state of the robot: the state U is active when the robot is uncommitted while the state C is active when the robot is committed to an option. In Fig. 2(b), the option commitment is indicated by index i, with $i \in \{1,\ldots,n\}$, indicating that the robots has stored the option i's information (χ_i and $\hat{v}_i$) and is committed to it. As indicated in the CDCI design pattern and explained below, changes in commitment depend upon the local estimate of the distribution of other agents' commitments. Therefore, we designed an update time $\delta_u = 50$ clock cycles (i.e. ~ 1.5 s) to allow the robot to gather a local sample of its neighbourhood. Every δ_u, the robot updates its commitment state as follows. If an uncommitted robot satisfies the condition E_i by encountering the option i during the last δ_u clock cycles, it may become committed to i (i.e. the robot *discovers* option i) with probability P_{γ_i}. A robot committed to option i may spontaneously *abandon* its commitment to i and become uncommitted with probability P_{α_i}. If an uncommitted robot satisfies the condition M_i by receiving a message from a robot committed to option i during the last δ_u clock cycles, it *gets recruited* and commits to option i. If a robot committed to option i satisfies the condition $M_{j\neq i}$ by receiving a message from a robot committed to a different option j (with $i \neq j$) during the last δ_u clock cycles, it *gets cross-inhibited* and reverts to the uncommitted state. At each broadcast tick, a robot committed to option i probabilistically decides either to *interact* with peers and share the location χ_i of option i with probability P_{h_i} or to not interact and appear as uncommitted (with probability $1-P_{h_i}$).

Following the CDCI guidelines [6], the robot makes interaction transitions (recruitment and cross-inhibition) with probability function of the distribution of commitment in its neighbourhood. These probabilistic transitions can be reduced to the selection at random of one message M (among all the received) which is used to conditionally trigger transitions (see Fig. 2(b) and [6] for more details). In this study, the robot only keeps the last received message by overwriting the information each time. This choice allows the robot to minimise the memory usage and to keep the most up-to-date information (compared to older messages). After each commitment update, the last message is deleted.

As robots do not exchange quality estimates, after recruitment, the recruited robot is required to visit the option's location in order to self-estimate the option's quality. This is similar to the behaviour observed in honeybees [42] and ants [43] during nest-site selection. Once the robot estimates the quality, it resumes interactions. While individual estimates by each robot is a (time-consuming) necessary component of the CDCI strategy, it can also prevent the spreading of inaccurate estimates. In fact, from the analysis of the DC method, we understood that the reuse of second-hand information (i.e. the received quality) can lead to the spreading of inaccurate estimates. We conducted analyses to estimate the impact of self-estimates on the DC performance. A modified DC strategy with robots sampling the quality after each recruitment increases the performance to values similar to CDCI for $\kappa \leq 0.9$, although it cannot break symmetry for $\kappa \approx 1$ even after several hours (results not shown). In terms of speed, the DC strategy shows quicker dynamics than CDCI although it is important to consider that the CDCI requires lower cognitive abilities of the individual robots who do not need to share qualities in a common range.

To allow the swarm to converge to consensus for the best-quality option, the individual robots modulate their behaviour as a function of the option's quality, exhibiting more frequent behaviours in support of better quality options. In this study, the probabilities of the robot's behaviours follows the parameterisation introduced in [19] which preserves the value-sensitive decision-making characteristics when the number of options is greater than two:

$$P_{\gamma_i} = k\hat{v}_i\Delta, \quad P_{\alpha_i} = k\hat{v}_i^{-1}\Delta, \quad P_{h_i} = h\hat{v}_i\Delta\,, i \in \{1,2,\ldots,n\} \quad (1)$$

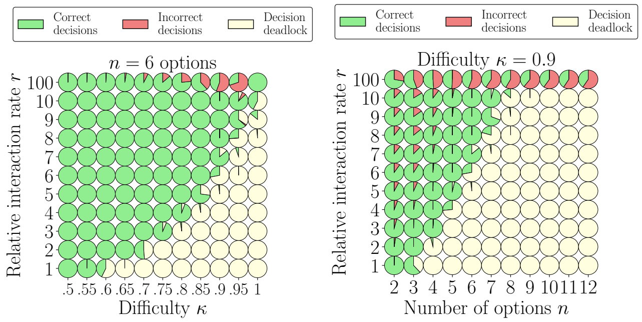
where $\hat{v}_i$ is the estimated quality of option i, while h and k are parameters to control the frequency at which the robots send interaction messages and perform individual behaviours, respectively. The ratio $r = h/k$ represents the relative interaction rate. Following [6], the parameter Δ is required to scale probabilities within the valid range $[0,1]$ and guarantee a match between microscopic and macroscopic description of the process. $\Delta = \delta_u\delta_c\delta_s$ is determined by three components: the number of Kilobot clock cycles between two updates ($\delta_u = 50$), the Kilobot clock period ($\delta_c \simeq 31\,\text{ms}$) and the temporal scaling factor $\delta_s = 0.000594$ which controls the process speed. As shown in [19], the key parameter in the swarm decision dynamics is the relative interaction rate r.

B. Stochastic analysis of the basic CDCI strategy

We investigate the effect of the relative interaction rate r on the decision outcome through stochastic analysis with the goal of identifying the best r for our swarm robotics system. The CDCI strategy can be described in the form of a master equation [11], [6]. This description form allows us to investigate the macroscopic system dynamics with random fluctuations proportional to our system size $S = 200$. We approximate the solution of the master equation through 1,000 runs of the stochastic simulation algorithm (SSA) [44].

For values of $r \in [1,100]$ with $k = 1$, we investigated the effects of varying the decision difficulty $\kappa \in [0.5,1]$ keeping constant the number of options $n = 6$ (Fig. 3(a)), and of varying number of options $n \in [2,12]$ keeping constant the decision difficulty $\kappa = 0.9$ (Fig. 3(b)). Each run terminates either when the maximum decision time $T_{max} = 10$ is reached or when a sub-population committed to a single option reaches the decision quorum $Q = 0.8$ (i.e. at least 160 robots are committed to the same option). The maximum decision time has been selected to be safely above the average decision time to avoid premature terminations. The results of the analysis are shown in Fig. 3 where each coloured pie represents a tested condition and indicates the percentage of runs terminating in a decision deadlock (yellow), a decision for the best quality option (green), or a decision for any of the $(n-1)$ inferior equal-quality options (red). The obtained results are in agreement with previous deterministic mean-field analyses that show the need of high positive/negative feedback to break the decision deadlock [19], [45] and with stochastic analyses of decision-making models of ants and slime moulds [46], that show that a high positive feedback reduces decision accuracy.

Figure 3(a) shows the existence of a dilemma: On the one hand, low values of r guarantee accurate decisions when there is a superior-quality option ($v_H \gg v_L$) but lead to a decision deadlock when all options are similar ($v_H \approx v_L$). On the other hand, high values of r guarantee deadlock-breaking but the system is very sensitive to initial random fluctuations which lead to inaccurate decisions (i.e. the first discovered option has higher probability to be selected even of inferior quality). Additionally, Fig. 3(b) shows that the minimum r necessary to break deadlock increases quadratically with the number of options. Therefore, the interaction rate r that maximises accuracy varies as a function of the decision problem (number and quality of options) which is normally unknown to the swarm. Hence any fixed value of r has drawbacks and causes poor performance in certain scenarios.



(a) Varying decision difficulty κ (b) Varying number of options n

Fig. 3. Results of the SSA showing the influence of the interaction ratio $r = h/k$ with $k = 1$ (of Eq. (1)) for various best-of-n problems in the case of $S = 200$ robots. The pie-charts indicate the percentage of 1,000 runs terminating in a decision deadlock, i.e. below quorum $Q = 0.8$, after $T_{max} = 10$ (yellow), a decision for the best-quality option $v_H = 10$ (green), or a decision for any $n-1$ inferior-quality distractor $v_L = \kappa \cdot v_H$ (red). (a) Sensitivity of CDCI to the ratio r for various problem difficulties $\kappa \in [0.5, 1]$ in case of $n = 6$; the minimum r necessary to break decision deadlock grows quadratically with κ. (b) Sensitivity of CDCI to the ratio r for various number of options $n \in [2, 12]$ in case of $\kappa = 0.9$; the minimum r necessary to break decision deadlock grows quadratically with n. Sufficiently high values of interaction rate r, always lead to a decision but accuracy rapidly decreases with increasing n or κ.

C. The time-varying CDCI strategy

To solve the previous dilemma without prior knowledge of the decision problem (n and κ), we propose a novel decentralised strategy which consists of beginning with low interaction rate r (to limit initial random fluctuations) and then increasing the interaction rate over time to reach consensus. The initial low r corresponds to relatively sporadic interactions to prevent commitment to the first discovered options (which may have inferior quality v_L) propagating through the swarm. The increase of interactions over time, instead, has the function to break decision deadlocks and to allow the swarm to build up consensus for the best discovered option. Additionally, the robots will modulate the increase speed of interactions as a function of the estimated quality; that is, they will start to recruit (and cross-inhibit) earlier for better quality options. We expect that this quality-sensitive increase of interactions could result in improved accuracy and decision speed compared with the standard CDCI.

We investigate two variants of the proposed time-varying strategy: a gradual increase of the interaction rate r (function $r_{ramp}(t)$ in Fig. 4(a)) and a jump of the interaction rate from low to high values (function $r_{step}(t)$ in Fig. 4(b)):

$$r_{ramp}(t) = \frac{h_{ramp}(t)}{k},\; h_{ramp}(t) = \begin{cases} \dfrac{H_{max}}{\tau(\hat{v}_i)} t & \text{if } t < \tau(\hat{v}_i) \\ H_{max} & \text{if } t \geq \tau(\hat{v}_i) \end{cases}, \quad (2)$$

$$r_{step}(t) = \frac{h_{step}(t)}{k},\; h_{step}(t) = \begin{cases} 0 & \text{if } t < \tau(\hat{v}_i) \\ H_{max} & \text{if } t \geq \tau(\hat{v}_i) \end{cases}, \quad (3)$$

In both functions, the individual transitions strength k remains constant at $k = 1$, while the interaction strength $h(t)$ increases over time until a maximum H_{max}. The function

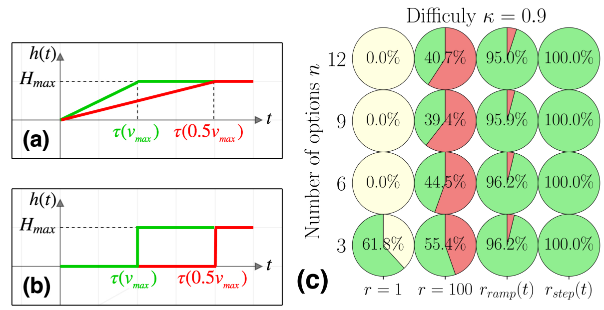


Fig. 4. (Left) Two forms for the time-varying interaction $h(t)$: $h_{ramp}(t)$ of Eq. (2) in panel (a) and $h_{step}(t)$ of Eq. (3) in panel (b). With the ramp function, the robot constantly increases the interaction strength $h_{ramp}(t)$ with slope proportional to the estimated quality $\hat{v}_i$; instead, with the step function, the robot does not interact $h_{step}(t) = 0$ until a time $\tau(\hat{v}_i)$ that is inversely proportional to the estimated quality $\hat{v}_i$. (c) Results of the SSA (1,000 runs each condition and same colour code of Fig. 3) predicting the decision outcome of a 200-robot swarm for various best-of-n problems ($n \in \{3, 6, 9, 12\}$) and difficulty $\kappa = 0.9$. The $r_{ramp}(t)$ and $r_{step}(t)$ from Eqs. (2)-(3) with parameters $\tau_0 = 5$ and $H_{max} = 100$ show a considerable improvement in accuracy (reported on each pie-chart).

$\tau(\hat{v}_i) = \tau_0 v_{max}/\hat{v}_i$ describes the time at which a robot committed to option i reaches H_{max}, therefore it determines the slope of the ramp and the jumping time of the step function. If the estimated quality is the maximum value, i.e. $\hat{v}_i = v_{max}$, the maximum interaction strength is reached at τ_0, otherwise it happens later. The functions of Eqs. (2)-(3) can be implemented in a decentralised fashion by asking the robots to modify the strength of their interactions over time, in particular by increasing the probability of sending recruitment and cross-inhibition messages. In other words, the robots vary the probabilities of Eq. (1) with the time-varying term h from Eqs. (2)-(3).

We assess the performance of the proposed strategy in its two variants through master equation analysis. The original version of SSA proposed by Gillespie only considers constant transition rates, therefore, to take into account the time-varying probabilities of the individual behavioural rules of Eq. (1), we used a modified version of SSA, as proposed in [47]. Figure 4(c) compares the decision outcome of time-invariant interaction rate r and the two proposed time-varying interaction strategies for various problems with $n \in \{3, 6, 9, 12\}$ and $\kappa = 0.9$. The results of each tested condition are obtained via 1,000 SSA runs and follow the same colour code of Fig. 3. A swarm with low time-invariant relative interaction rate $r = 1$ has dynamics dominated by the individual behaviours of discovery and abandonment, which are not sufficient to break the deadlock between more than three similar-quality alternatives. Using high time-invariant $r = 100$, the dynamics are dictated by the interaction behaviours of recruitment and cross-inhibition, which act respectively as positive and negative feedbacks on the committed sub-populations. In this case, initial random fluctuations quickly spread through the system and the probability to select the inferior options is high and increases with n. Both functions of time-varying $r(t)$ show a considerable performance improvement in terms of decision accuracy, in particular $r_{step}(t)$ has 100% accurate outcomes.

VI. Robot swarm simulations

We simulate through ARGoS each proposed strategy in 100 runs on a swarm of $S = 200$ Kilobots[1]. We impose a time limit of $T_{max} = 2$ hours within which we expect the swarm to reach an agreement in favour of one option (decision quorum $Q = 80\%$). We use T_{max} as a cutoff to compute the swarm decision/indecision. Figure 5(a) shows the decision speed and accuracy for $n \in \{3, 6\}$ options in a difficult decision problem of $\kappa = 0.9$ ($v_H = 10$ and $v_L = 9$), with a quality estimation noise strength $\sigma^2 = 1$. As correctly predicted by the stochastic analysis of Sec. V-B, low interaction rate (i.e. $r = 1$) gives slow dynamics and leaves the swarm undecided, unable to break the decision deadlock within the cutoff time. A high interaction rate (i.e. $r = 100$) speeds up the convergence dynamics at the cost of low accuracy. The time-varying strategy $r_{ramp}(t)$ (with $\tau_0 = 50$ min) shows an improvement in both speed and accuracy. The accuracy performance is further improved by the $r_{step}(t)$ strategy (with $\tau_0 = 50$ min) which interestingly also has a highly consistent and predictable decision time at few minutes after τ_0. Finally, DC has the highest speed but a very low accuracy due to the quick spreading of noisy quality overestimates. SSA gives a good prediction of the expected dynamics although it does not (and is not supposed to) give the exact same swarm dynamics because the swarm robotics process and the master equation model are different. The main differences are caused by the local discovery and interactions. Discovery transitions are conditional on the event of robots encountering the localised option in space and Kilobots' local communication happens only between nearby neighbours with slow motion dynamics. Therefore, the robot swarm is subject to correlated interactions while SSA is computed under the assumption of uncorrelated interactions.

VII. Discussion

This study investigates novel strategies for consensus-decision making in decentralised systems in the context of swarm robotics. Reaching an agreement among all group members, in our case all robots, can be very challenging, especially when candidate options have similar qualities [48]. We propose a decentralised strategy that considerably improves decision accuracy and that we tested through computational analysis of master equations and swarm robotics simulations. The solution, inspired by honeybee house-hunting behaviour [11], consists of increasing the strength of interactions among robots over time and is based on principled understanding and formal analysis of the system dynamics from previous research [20], [6], [19]. While this work is limited to simulation, we plan to test the proposed system on a large-scale Kilobot swarm interfaced with ARK [36].

Most research on decision algorithms in swarm robotics limits its analysis to binary choice experiments [10]. In the swarm robotics literature, only a few exceptions investigated

[1]The robot control software is available online at https://github.com/DiODeProject/Time-Varying-CDCI

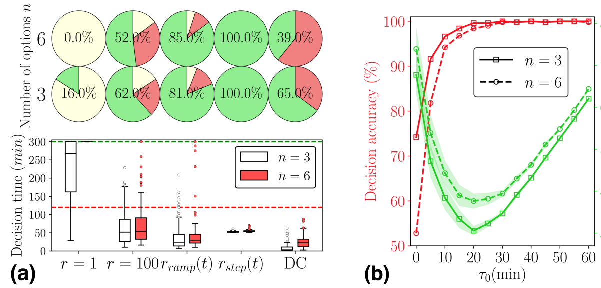


Fig. 5. (a) 200-Kilobot swarm results (100 simulations for each condition) for different decision strategies (in each column) in case of $n \in \{3, 6\}$ options with difficulty $\kappa = 0.9$ and noise strength $\sigma^2 = 1$. Top pie-charts show the decision accuracy (same colour code of Figs. 3-4(c)). Bottom boxplots show the decision time; the horizontal red line (at 2 hours) is the cutoff time to compute the decision outcome (e.g. indecision vs decision). We let the simulation run a maximum of 5 hours to display the complete decision time dynamics. Low interaction rate ($r = 1$) shows low convergence rate and frequent deadlocks. High interaction rate ($r = 100$) shows low accuracy. Time-varying $r_{ramp}(t)$ shows an improvement in accuracy, which is further improved by $r_{step}(t)$ (both time-varying strategies use $\tau_0 = 50$ min). The DC (of Sec. IV) shows low accuracy due to the spreading of noisy estimates. (b) Speed (green lines with 95% confidence shades and right y-axis) and accuracy (red lines and left y-axis) of the swarm robotics system for varying interaction speed $\tau_0 \in [0, 60]$min for the $r_{step}(t)$ strategy. An inaccurate tuning of τ_0 may lead to sub-optimal performance.

scenarios with more than two options, e.g. [49], [50]. However, previous theoretical analysis showed that increasing the number of options can considerably change the swarm dynamics [19]. We therefore performed our analyses and experiments in a genuinely best-of-n setup. The proposed time-varying strategy is able to consistently show remarkably high accuracy performance and a predictable decision speed, for any tested number of options.

The time-variant strategy with the best performance separates the collective decision into two phases: an initial phase of environment exploration based on independent behaviour and absence of interactions, and a second phase of information exchange based on frequent robot interactions to reach consensus. Similar analyses have shown how systems of collective motion [51], [52], [53], or foraging [54], [55], can benefit from modulating the strength of individuals' interaction in relation to environmental features. While our strategy shows good performance, we note that the time to move from the exploration to the interaction phase (i.e. τ_0) should be accurately tuned to the speed of the decision process and, likewise, to environmental features. In Fig. 5(b) we show how speed and accuracy vary as a function of τ_0 for $r_{step}(t)$. These results show that inaccurate tuning of τ_0 may reduce the system performance. For instance, if the swarm starts interacting before any option is discovered it nullifies the positive effects of the strategy; conversely a long exploration phase could unnecessarily delay consensus and reduce decision speed. Instead of accurately tuning this key parameter, in future work we plan to allow individual robots to autonomously estimate when to move from exploration to the interaction phase. In this way, the swarm should be able to adapt its response to the decision problem without requiring any speed tuning.

REFERENCES

[1] S. Garnier, J. Gautrais, M. Asadpour, C. Jost, and G. Theraulaz, "Self-organized aggregation triggers collective decision making in a group of cockroach-like robots," *Adaptive Behavior*, vol. 17, no. 2, pp. 109–133, 2009.

[2] H. Hamann, T. Schmickl, H. Wörn, and K. Crailsheim, "Analysis of emergent symmetry breaking in collective decision making," *Neural Computing and Applications*, vol. 21, no. 2, pp. 207–218, 2012.

[3] A. Özdemir, M. Gauci, S. Bonnet, and R. Groß, "Finding consensus without computation," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1346–1353, 2018.

[4] C. A. C. Parker and H. Zhang, "Cooperative decision-making in decentralized multiple-robot systems: the best-of-n problem," *IEEE Transactions on Mechatronics*, vol. 14, no. 2, pp. 240–251, 2009.

[5] Á. Gutiérrez, A. Campo, F. Monasterio-Huelin, L. Magdalena, and M. Dorigo, "Collective decision-making based on social odometry," *Neural Computation*, vol. 19, no. 6, pp. 807–823, 2010.

[6] A. Reina, G. Valentini, C. Fernández-Oto, M. Dorigo, and V. Trianni, "A design pattern for decentralised decision making," *PLoS ONE*, vol. 10, no. 10, p. e0140950, 2015.

[7] A. Scheidler, A. Brutschy, E. Ferrante, and M. Dorigo, "The k-unanimity rule for self-organized decision-making in swarms of robots," *IEEE Transactions on Cybernetics*, vol. 46, no. 5, pp. 1175–1188, 2016.

[8] G. Valentini, E. Ferrante, H. Hamann, and M. Dorigo, "Collective decision with 100 Kilobots: speed versus accuracy in binary discrimination problems," *Autonomous Agents and Multi-Agent Systems*, vol. 30, no. 3, pp. 553–580, 2016.

[9] G. Valentini, *Achieving Consensus in Robot Swarms*, ser. Studies in Computational Intelligence. Springer International Publishing, 2017, vol. 706.

[10] G. Valentini, E. Ferrante, and M. Dorigo, "The Best-of-n Problem in Robot Swarms: Formalization, State of the Art, and Novel Perspectives," *Frontiers in Robotics and AI*, vol. 4, p. 9, 2017.

[11] T. D. Seeley, P. K. Visscher, T. Schlegel, P. M. Hogan, N. R. Franks, and J. A. R. Marshall, "Stop signals provide cross inhibition in collective decision-making by honeybee swarms." *Science*, vol. 335, no. 6064, pp. 108–11, 2012.

[12] R. Gray, A. Franci, V. Srivastava, and N. E. Leonard, "An agent-based framework for bio-inspired, value-sensitive decision-making," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8238–8243, 2017.

[13] N. Correll and H. Hamann, *Springer Handbook of Computational Intelligence*. Springer, 2015, ch. Probabilistic Modelling of Swarming Systems, pp. 1423–1432.

[14] K. Lerman, A. Martinoli, and A. Galstyan, "A review of probabilistic macroscopic models for swarm robotic systems," in *Swarm Robotics*, ser. LNCS. Springer, 2005, vol. 3342, pp. 143–152.

[15] H. Hamann and H. Wörn, "A framework of spacetime continuous models for algorithm design in swarm robotics," *Swarm Intelligence*, vol. 2, no. 2-4, pp. 209–239, 2008.

[16] M. A. Hsieh, A. Halász, S. Berman, and V. Kumar, "Biologically inspired redistribution of a swarm of robots among multiple sites," *Swarm Intelligence*, vol. 2, no. 2-4, pp. 121–141, 2008.

[17] S. Berman, A. Halasz, M. A. Hsieh, and V. Kumar, "Optimized Stochastic Policies for Task Allocation in Swarms of Robots," *IEEE Transactions on Robotics*, vol. 25, no. 4, pp. 927–937, 2009.

[18] M. Vigelius, B. Meyer, and G. Pascoe, "Multiscale modelling and analysis of collective decision making in swarm robotics," *PLoS ONE*, vol. 9, no. 11, p. e111542, 2014.

[19] A. Reina, J. A. R. Marshall, V. Trianni, and T. Bose, "Model of the best-of-N nest-site selection process in honeybees," *Physical Review E*, vol. 95, no. 5, p. 052411, 2017.

[20] D. Pais, P. M. Hogan, T. Schlegel, N. R. Franks, N. E. Leonard, and J. A. R. Marshall, "A mechanism for value-sensitive decision-making," *PLoS ONE*, vol. 8, no. 9, 2013.

[21] T. Bose, A. Reina, and J. A. R. Marshall, "Collective decision-making," *Current Opinion in Behavioral Sciences*, vol. 6, pp. 30–34, 2017.

[22] A. Reina, T. Bose, V. Trianni, and J. A. R. Marshall, "Effects of spatiality on value-sensitive decisions made by robot swarms," in *Proceedings of the 13th International Symposium on Distributed Autonomous Robotic Systems (DARS2016)*, ser. STAR. Springer, 2018, vol. 6, pp. 461–473.

[23] S. C. Pratt, E. Mallon, D. J. T. Sumpter, and N. Franks, "Quorum sensing, recruitment, and collective decision-making during colony emigration by the ant Leptothorax albipennis," *Behavioral Ecology and Sociobiology*, vol. 52, no. 2, pp. 117–127, 2002.

[24] T. D. Seeley and P. K. Visscher, "Quorum sensing during nest-site selection by honeybee swarms," *Behavioral Ecology and Sociobiology*, vol. 56, no. 6, pp. 594–601, 2004.

[25] C. A. C. Parker and H. Zhang, "Collective unary decision-making by decentralized multiple-robot systems applied to the task-sequencing problem," *Swarm Intelligence*, vol. 4, no. 4, pp. 199–220, 2010.

[26] J. R. Cody and J. A. Adams, "An evaluation of quorum sensing mechanisms in collective value-sensitive site selection," in *Proceedings of the 2017 International Symposium on Multi-Robot and Multi-Agent Systems (MRS)*. IEEE, 2017, pp. 40–47.

[27] T. D. Seeley and S. C. Buhrman, "Nest-site selection in honey bees: how well do swarms implement the "best-of-n" decision rule?" *Behavioral Ecology and Sociobiology*, vol. 49, no. 5, pp. 416–427, 2001.

[28] N. R. Franks, A. Dornhaus, C. S. Best, and E. L. Jones, "Decision making by small and large house-hunting ant colonies: one size fits all," *Animal Behaviour*, vol. 72, no. 3, pp. 611–616, 2006.

[29] E. J. H. Robinson, N. R. Franks, S. Ellis, S. Okuda, and J. A. R. Marshall, "A simple threshold rule is sufficient to explain sophisticated collective decision-making," *PLoS ONE*, vol. 6, no. 5, p. e19981, 2011.

[30] J. D. Schall and D. P. Hanes, "Neural basis of saccade target selection in frontal eye field during visual search," *Nature*, vol. 366, no. 6454, pp. 467–469, 1993.

[31] M. Rubenstein, C. Ahler, N. Hoff, A. Cabrera, and R. Nagpal, "Kilobot: A low cost robot with scalable operations designed for collective behaviors," *Robotics and Autonomous Systems*, vol. 62, no. 7, pp. 966–975, 2014.

[32] M. Rubenstein, A. Cornejo, and R. Nagpal, "Programmable self-assembly in a thousand-robot swarm," *Science*, vol. 345, no. 6198, pp. 795–799, 2014.

[33] Y. K. Lopes, S. M. Trenkwalder, A. B. Leal, T. J. Dodd, and R. Groß, ""supervisory control theory applied to swarm robotics"," *Swarm Intelligence*, vol. 10, no. 1, pp. 65–97, 2016.

[34] C. Dimidov, G. Oriolo, and V. Trianni, "Random walks in swarm robotics: an experiment with Kilobots," in *Swarm Intelligence (ANTS 2016)*, ser. LNCS. Springer, 2016, vol. 9882, pp. 185–196.

[35] V. Trianni, D. De Simone, A. Reina, and A. Baronchelli, "Emergence of consensus in a multi-robot network: from abstract models to empirical validation," *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 348–353, 2016.

[36] A. Reina, A. J. Cope, E. Nikolaidis, J. A. R. Marshall, and C. Sabo, "ARK: Augmented Reality for Kilobots," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1755–1761, 2017.

[37] A. Font Llenas, M. S. Talamali, X. Xu, J. A. R. Marshall, and A. Reina, "Quality-sensitive foraging by a robot swarm through virtual pheromone trails," in *Swarm Intelligence (ANTS2018)*, ser. LNCS. Springer, 2018, vol. 11172, pp. 135–149.

[38] C. Pinciroli, V. Trianni, R. O'Grady, G. Pini, A. Brutschy, M. Brambilla, N. Mathews, E. Ferrante, G. Di Caro, F. Ducatelle, M. Birattari, L. M. Gambardella, and M. Dorigo, "ARGoS: a modular, parallel, multi-engine simulator for multi-robot systems," *Swarm Intelligence*, vol. 6, no. 4, pp. 271–295, 2012.

[39] C. Pinciroli, M. S. Talamali, A. Reina, J. A. R. Marshall, and V. Trianni, "Simulating kilobots within argos: models and experimental validation," in *Swarm Intelligence (ANTS2018)*, ser. LNCS. Springer, 2018, vol. 11172, pp. 176–187.

[40] G. Valentini, D. Brambilla, H. Hamann, and M. Dorigo, "Collective perception of environmental features in a robot swarm," in *Swarm Intelligence (ANTS2016)*, ser. LNCS. Springer, 2016, pp. 65–76.

[41] A. Reina, R. Miletitch, M. Dorigo, and V. Trianni, "A quantitative micro-macro link for collective decisions: The shortest path discovery/selection example," *Swarm Intelligence*, vol. 9, no. 2–3, pp. 75–102, 2015.

[42] N. F. Britton, N. R. Franks, S. C. Pratt, and T. D. Seeley, "Deciding on a new home: how do honeybees agree?" *Proceedings of the Royal*

Society of London. Series B: Biological Sciences, vol. 269, no. 1498, pp. 1383–1388, 2002.

[43] R. Kaur, K. Anoop, and A. Sumana, "Leaders follow leaders to reunite the colony: Relocation dynamics of an Indian queenless ant in its natural habitat," *Animal Behaviour*, vol. 83, no. 6, pp. 1345–1353, 2012.

[44] D. T. Gillespie, "A general method for numerically simulation the stochastic time evolution of coupled chemical reactions," *Journal of Computational Physics*, vol. 22, no. 4, pp. 403–434, 1976.

[45] R. Gray, A. Franci, V. Srivastava, and N. E. Leonard, "Multiagent decision-making dynamics inspired by honeybees," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 2, pp. 793–806, 2018.

[46] S. C. Nicolis, N. Zabzina, T. Latty, and D. J. T. Sumpter, "Collective irrationality and positive feedback," *PLoS ONE*, vol. 6, no. 4, p. e18901, 2011.

[47] R. R. P. Purtan and A. Udrea, "A modified stochastic simulation algorithm for time-dependent intensity rates," in *Proceedings of the 19th International Conference on Control Systems and Computer Scienc (CSCS2013)*. IEEE, 2013, pp. 365–369.

[48] A. Baronchelli, "The emergence of consensus: a primer," *Royal Society Open Science*, vol. 5, no. 2, 2018.

[49] C. A. C. Parker and H. Zhang, "Active versus passive expression of preference in the control of multiple-robot decision-making," in *IEEE International Conference on Intelligent Robots and Systems (IROS2005)*. IEEE, 2005, pp. 3706–3711.

[50] S. Garnier, M. Combe, C. Jost, and G. Theraulaz, "Do ants need to estimate the geometrical properties of trail bifurcations to find an efficient route? a swarm robotics test bed," *PLoS Computational Biology*, vol. 9, no. 3, p. e1002903, 2013.

[51] C. J. Torney, Z. Neufeld, and I. D. Couzin, "Context-dependent interaction leads to emergent search behavior in social aggregates," *Proceedings of the National Academy of Sciences*, vol. 106, no. 52, pp. 22 055–22 060, 2009.

[52] A. Shklarsh, G. Ariel, E. Schneidman, and E. Ben-Jacob, "Smart swarms of bacteria-inspired agents with performance adaptable interactions," *PLoS Computational Biology*, vol. 7, no. 9, 2011.

[53] Y. Khaluf, I. Rausch, and P. Simoens, "The impact of interaction models on the coherence of collective decision-making: a case study with simulated locusts," in *Swarm Intelligence (ANTS2018)*, ser. LNCS. Springer, 2018, vol. 11172, pp. 252–263.

[54] R. Pagliara, D. M. Gordon, and N. E. Leonard, "Regulation of harvester ant foraging as a closed-loop excitable system," *PLoS Computational Biology*, vol. 14, no. 12, p. e1006200, 2018.

[55] L. Pitonakova, R. Crowder, and S. Bullock, "The Information-Cost-Reward framework for understanding robot swarm foraging," *Swarm Intelligence*, vol. 12, no. 1, pp. 71–96, 2018.