The Power of Social Information in Distributed Consensus in Ant-Colonies: Model and Analysis

by

Jiajia Zhao

S.B. Computer Science and Engineering, Massachusetts Institute of Technology, 2016

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2021

Author . . . .
Department of Electrical Engineering and Computer Science
September 10, 2021

Certified by . . . .
Nancy Lynch
Professor
Thesis Supervisor

Accepted by . . . .
Katrina LaCurts
Chair, Master of Engineering Thesis Committee

The Power of Social Information in Distributed Consensus in Ant-Colonies: Model and Analysis

by

Jiajia Zhao

Submitted to the Department of Electrical Engineering and Computer Science
on September 10, 2021, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

The decentralized cognition of animal groups is both a challenging biological problem and a potential basis for bio-inspired engineering design. The understanding of these systems and their application can benefit from modeling and analysis of the underlying algorithms. In Chapter 2, we define a modeling framework that can be used to formally represent all components of such algorithms. As an example application of the framework, we adapt to it the much-studied house-hunting algorithm used by emigrating colonies of *Temnothorax* ants to reach consensus on a new nest. We provide a Python simulator that encodes accurate individual behavior rules and produces simulated behaviors consistent with empirical observations, on both the individual and group levels. We use the simulator to make predictions about several aspects of collective emigration behavior, some with empirical support and some are new predictions. Critically, our results highlight the value of individual sensitivity to site population in ensuring consensus, and suggest its empirical measurement.

Though the above model captures a wide range of observed phenomenon and make new predictions, our work and previous work have mostly focused on experimental or modeling work, and lack rigorous mathematical justification. Building a theoretical understanding of the key mechanisms in the house-hunting process is necessary for the designs of novel distributed consensus algorithms. In order to do so, in this chapter we further simplified the model introduced in Chapter 2 and investigated the marginal benefits of the quorum sensing mechanism. We show theoretical confirmation of the hypothesized evolutionary advantage of quorum sensing in helping consensus. In addition, the desirable values of the quorum size from our theoretical results have been observed empirically.

It is our hope that the scientific insights and the modeling and mathematical tools can inspire further research from both the biology and computer science community.

Thesis Supervisor: Nancy Lynch
Title: Professor

Acknowledgments

I want to thank all my family and friends for supporting me throughout my thesis research. I would like to thank Dr. Nancy Lynch for her support and supervision for all aspects of my research. I want to thank Dr. Stephen C. Pratt for his guidance on biological insights on Chapter 2 of my thesis; Dr. Lili Su for her guidance on mathematical analysis on Chapter 3 of my thesis; Emily Zhang, Brabeeba Wang, Sabrina Drammis and Chi-Ning Chou for valuable feedback on my research results. Chapter 2 is joint work with Drs. Nancy Lynch and Stephen C. Pratt, and Chapter 3 is joint work with Drs. Lili Su and Nancy Lynch. Finally, I want to thank Dr. Nancy Lynch again for providing me such an amazing opportunity to conduct important research with a wonderful group of people!

Contents

List of Figures

List of Tables

Chapter 1

Thesis Overview

Animal groups are capable of remarkable displays of highly coordinated behavior. Fish schools collectively choose foraging sites [69], locusts self-organize into orderly swarms [70], oceanic fish assemble in vast migratory shoals [34], and social insects perform a host of collective actions including group foraging, construction of complex nests, and adaptive allocation of tasks across the labor force [4, 7, 8, 41, 43, 62]. Although well-informed leaders may play a role, group organization is typically very decentralized [7, 17, 62]. Coordination emerges from interactions among large numbers of animals acting on limited local information with appropriate decision rules.

However, how these group actions result from individual behavior remains a major research challenge. Understanding the connection between individual behavior and group outcomes is too much for unaided intuition, hence mathematical models and agent-based simulations have become useful tools for understanding. In this thesis, we develop general modeling and analytical tools to understand a notable example of decentralized decision-making: nest site selection by ants of the genus *Temnothorax.*

Models, in combination with experimental studies, have already revealed much about these ants, making them a leading model system for collective decision-making [51]. A *Temnothorax* ant colony is composed of **adult workers** and brood items (immature ants), each group making up 40% to 60% of colony members. Adults are roughly equally divided between active workers, who organize and execute emigrations, and passive workers, who, like brood items, are typically transported to the new nest by active workers and do not themselves recruit nestmates [53, 15, 67]. Colonies live in pre-formed cavities such as rock crevices or hollow nuts; if their home is damaged, they are adept at finding candidate new homes, evaluating each site's quality, and moving the entire colony to the best one. Their decision emerges from the separate efforts of many scouts, each independently recruiting nestmates to the site it has found. Because recruitment is quality-dependent, better sites accumulate ants more rapidly [35]. These differences are amplified by a quorum rule under which scouts accelerate recruitment with transporting instead of tandem running to a site, once its adult population crosses a threshold; the winner of the race to attain a quorum becomes the colony's choice [46]. An agent-based model has shown that this algorithm helps the colony reach consensus on the best site [53]. Other models have shown how a colony can make a good choice even when no individual directly compares sites [38, 56], and how individual behavioral strategies optimize speed/accuracy tradeoffs at the colony level [25, 37, 36, 44, 52, 65].

Although successful, models of this process have been limited to the simple challenge of choosing between two distinct and equidistant nests in a controlled laboratory environment. Real colonies face more complex scenarios, such as selecting among several sites of varying quality, avoiding splits when candidate nest sites are identical, and resolving colony splits when they occur [10, 60]. It also remains unclear how the colony maintains high performance with noisy and heterogeneous individuals, and how individuals modify their behavior to account for changes in context or colony state. In addition, a large body of experimental work has uncovered new collective colony behavior that has yet to be explained. These include the more complex scenarios mentioned above, as well as effects on decision-making of colony size and emigration distance, colony reconnaissance of potential new homes, and the emergence of group-level rationality despite individual-level irrationality [21, 59, 64]. To better capture the complexities of nest-site selection, in Chapter 2 we develop a new, flexible, general model for the analysis and exploration of these questions.

Additionally, previous work has mostly focused on experimental studies and modeling work, and lacks rigorous mathematical justification of the models that have been developed based on empirical findings. Although data shown from both experiments and simulations reveal much about the various mechanisms involved in the house hunting process, rigorous mathematical theorems about the process is essential to advancing our understanding the benefits of each of the key mechanism at work in this process. These understandings might in turn generalize to other biological processes, or help us design artificial systems such as robot swarms. In Chapter 3, we use analytical methods to investigate the marginal benefits of one of the key mechanisms in this process: quorum sensing.

Thus, in this thesis we add to the growing body of literature that shed light on the house-hunting process of *Temnothorax* ants by addressing both of the aspects above.

Chapter 2

The Power of Social Information: A Computational Modeling Approach

2.1 Introduction

In this chapter, we first define a modeling framework that can be used to formally represent all components of discrete-time agent-based algorithms where each agent is a state machine that transitions at most once in a round, and transitions involve the state change of either a single agent or two agents. As an example application of the framework, we adapt to it the much-studied house-hunting algorithm used by emigrating colonies of *Temnothorax* ants to reach consensus on a new nest. Using this house-hunting model, we investigate a range of emergent properties. Some of these properties have empirical support and some are important predictions that suggest further experimental or theoretical studies.

We first demonstrate the value of the model by reproducing results of earlier models showing how the ants' algorithm can account for decision-making and speed/accuracy tradeoffs in simple one- and two-choice experiments [52, 53]. We then extend the model to account for more recent empirical observations, including robust decision-making among larger numbers of options and rational colony decisions about decision speed [60, 61]. After that, we test out a proposed novel but critical role for social information, in which ants directly incorporate nestmate presence into their assessment of nest site quality. We use the model to test the effect of such information on a colony's ability to decide between two identical nests, a context that poses a particular challenge to consensus formation. Finally, we make predictions about the relationship between quorum size and the speed/accuracy tradeoff, and about the ability of a colony to re-unify after dispersal of its members among multiple competing sites. Our results indicate a more complex relationship between individual behavior and the speed/accuracy trade-off than seems to have been previously appreciated. The model proved relatively weak at resolving colony divisions among multiple sites, suggesting either limits to the ants' ability to reach consensus, or an aspect of their behavior not captured in our model. It is our hope that these insights and predictions can inspire further research from both the biology and computer science community.

Our model touches on several aspects of the emergence of collective intelligence in the house hunting process, but many more are yet to be explored. Therefore, an additional goal is to provide a versatile, easy-to-use and maintainable modeling tool that can be used to quantitatively test hypotheses beyond those included in this chapter. We achieve this goal with our Python simulator. Last but not least, from a theoretical perspective, our model can serve as a starting point for simpler models that allow rigorous proofs on convergence speed and accuracy.

Chapter Organization The rest of the chapter is organized as follows: Section 2.2 defines a general framework that we believe will be useful not only for this algorithm but for other agent-based distributed algorithms as well. Section 2.3 applies this framework, with designs and interpretations specific to the house hunting context. Section 2.4 describes our Python implementation of the house hunting simulator, instructions on running simulations, and their scoring goals and metrics. Section 2.5 validates our model by comparing its results with those obtained by empirical studies on individual behaviors and the collective properties of the colony. These validations give us confidence in the accuracy of our model as we proceed into further confirmations of newly observed colony behaviors with limited experimental results, as listed in Section 2.6. And Section 2.7 showcases simulation results that establish the power of the new social information rule in the colonies' ability to reach consensus. Lastly, Section 2.8 includes preliminary and exploratory predictions that suggest additional interesting ways to use our simulator.

2.2 Modeling Framework

In this section, we introduce a general modeling "language" that has the potential to be useful for a wide range of applications. In Section 2.3 we instantiate this language in the context of the house hunting process in ant colonies.

2.2.1 Agent-based Model

Formally, the components below define the entities in the system and their static capabilities. More explanatory text follows after the list.

- **agent-ids**, a set of ids for agents. Each *agent-id* uniquely identifies an agent. We also define **agent-ids**$'$ to be **agent-ids** $\cup\{\perp\}$ where $\perp$ is a placeholder for "no agent". In general, we add $'$ to a set name to denote the original set with the addition of a default element $\{\perp\}$.

- **external-states**, a set of external states an agent might be in. Each element in the set is an *external-state*. In addition, **all-externals** is the set of all mappings from **agent-ids** to **external-states**. Each element of the set is an *all-external*.

- **internal-states**, a set of internal states an agent might be in. Each element in the set is an *internal-state*.

- **env-states**, a set of states that the agents' environment might take on. Each element in the set is a *env-state*.

- **action-types**, a set of the types of actions agents might perform. Each element in the set is an *action-type*.

- **env-choices**, a set of values an agent can access in the environment. Each element in the set is an *env-choice.*

- **actions**, a set of quadruples of the form (*action-type, agent-id, agent-id′, env-choice*) ∈ **action-types** × **agent-ids** × **agent-ids**′× **env-choices**. Each element in the set is an *action.*

- **select-action**(*agent-id, state, env-state, all-external*): A *state* is a pair of (*external-state, internal-state*) ∈ **external-states** × **internal-states**. Each (*agent-id, state, env-state, all-external*) quadruple is mapped to a probability distribution over the sample space of **actions**, for which the second component is equal to the input argument *agent-id* and the third component is not equal to it. The function then outputs this probability distribution.

- **transition**(*agent-id, state, all-external, action*): A *state* is a pair of (*external-state, internal-state*) ∈ **external-states** × **internal-states**. Each (*agent-id, state, all-external, action*) quadruple determines a *state* as the resulting state of the agent identified by the input argument *agent-id.* The function outputs the resulting *state.*

Each agent has a unique *agent-id* ∈ **agent-ids**, and is modeled by a state machine. Agents can transition from one *state* to another. A *state* is a pair: an *external-state* ∈ **external-states** that is visible to other agents, and an *internal-state* ∈ **internal-states** that is invisible to other agents.

We define **all-externals** to be the set of all mappings from **agent-ids** to **external-states**. Each element of the set is an *all-external* and represents a particular mapping from **agent-ids** to **external-states** where each *agent-id* is mapped to an *external-state.*

The set **env-states** represents the set of states that the agents' environment might take on. In this chapter, we will assume that the environment is fixed. That is, the *env-state* does not change during the execution of the system. The reason we use a set here is to enable us to model the same set of agents operating in different environments.

Agents can also access values in the environment, and each value is called an *env-choice*. The set **env-choices** is the set of all possible values for *env-choice*.

An agent can transition from one *state* to another by taking an *action* $\in$ **actions**. Each *action* consists of an *action-type* $\in$ **action-types**, the id of the initiating agent *agent-id* $\in$ **agent-ids**, the id of the (optional) received agent *agent-id*$'$ $\in$ **agent-ids**$'$, and *env-choice* $\in$ **env-choices**.

The function **select-action**(*agent-id, state, env-state, all-external*) is intended to select an *action* for the agent with the given *agent-id*, who is the initiating agent in the *action*. The function outputs a probability distribution over the sample space **actions**. However the sample space limits its elements to have the second component equal to the input argument *agent-id*, and the third component not equal to it. Thus, any sampled *action* will have *agent-id* being the initiating agent's id, and the (optional) receiving agent necessarily has a different id.

The function **transition**(*agent-id, state, all-external, action*) represents a transition to be performed by the agent identified by the input argument *agent-id*. Given the input arguments, the function deterministically outputs the resulting *state* of the transition.

2.2.2 Timing and Execution Model

In this section, we introduce the dynamic aspects of our model, including the discrete and synchronous timing model, and how different components in the system interact with each other at different points during the execution of the algorithm.

Our system configuration contains 1) an environment state, called *env-state*, and 2) each agent's *state*, which is a pair (*external-state, internal-state*), independent of *env-state*. Agents receive inputs from and react to the environment during the execution of the system. In this chapter, we will assume that the environment is fixed. That is, the *env-state* does not change during the execution of the system.

Incorporating some theoretical ideas from [26, 54], we divide the total time into *rounds*. Each round is a discrete time-step, and times are the points between rounds. At any time t, there is a corresponding system configuration t. The initial time is time 0, and the first round is round 1, taking the system from configuration 0 at time 0 to configuration 1 at time 1. In general, round t starts with system configuration $(t-1)$. During round t, agents can perform various **transition**'s, which take the system from configuration $(t-1)$ at time $(t-1)$ to configuration t at time t.

We now describe the execution of an arbitrary round t. At any point in the execution of round t, each agent x is mapped to a *state*, *state_x*, which is visible to agent x itself. However, to other agents, only agent x's *external-state*, *external_x* is visible. We denote *all-external* $\in$ **all-externals** to be the mapping from every *agent-id* $\in$ **agent-ids** to the corresponding *external-state* $\in$ **external-states** in round t. These mappings can be updated during the execution.

Accounting for the randomness of the order of execution for all the agents, a randomly chosen permutation of **agent-ids** is generated at the beginning of round t, serving as the order of execution for the agents in the round. We also instantiate a set $Trans = \emptyset$ at the beginning of the round. An agent is prevented from changing its *state* further in the round once it adds its *agent-id* to $Trans$, which can happen during its turn (even if there is no resulting state change) or when it performs a **transition** during another agent's turn. As a result, each agent can change its state at most once in the round. After all agents are in the set *Trans*, round t is over, and all agents enter round $t+1$ synchronously.

The rest of this section describes all possible operations during one agent x's turn in round t. When an agent with *agent-id* x (a.k.a. agent x) gets its turn to execute, it first checks whether $x \in Trans$. If so, agent x does nothing and ends its turn here.

Otherwise, agent x has not yet transitioned in round t. Let *state_x* denote the *state* of agent x. Agent x calls the function **select-action**(*x, state_x, env-state, all-external*). The function outputs a probability distribution over the sample space of a subspace of **actions**, for which the second component is x, and the third component is not x. Agent x randomly selects an *action*, *act* $= (a, x, x', e)$, according to this probability distribution.

Agent x then calls **transition**(*x, state_x, all-external, act*), to determine the resulting *state*, *new_state_x*, for agent x. As the initiating agent, x also gets added to $Trans$. Next, in the case where $x' \neq \perp$, agent x' also calls **transition**(x', *state*_x', *all-external, act*) where *state*_x' is the current *state* of agent x', maps itself to the function output, and updates its entry in *all-external.* Note that x' is added to $Trans$ if the function output is different than *state*_x' in any way. This is the end of agent x's **transition** call. Agent x then maps itself to the resulting state *new*_*state*_x, and updates its entry in *all-external.* Agent x finally ends its turn here.

2.2.3 Discussion

Although our model keeps track of the *external-state* of all the agents in *all-external*, when performing a transition, an agent can only access *local* information in it. Locality here is flexible to the context, i.e. local to the location of the agent initiating an action.

Agent-based models are especially powerful for simulating and analyzing collective behaviors given their natural compatibility with object-oriented programming methodologies and their flexibility for allowing individual differences in realized state transition probabilities among the agents [6, 66, 39, 53].

2.3 House Hunting Model

This section uses the framework defined in Section 2.2 to describe the house-hunting process.

2.3.1 Informal Description

Now we give an informal description of how ants perform an emigration, consistent with but more intuitive than the formal model definition in Section 2.3.2 later.

There are four distinct phases for an active worker in the house-hunting process. In the first, the **Exploration** phase, the ant randomly starts to explore her surroundings for a suitable new nest. If she finds a candidate site, she enters the **Assessment** phase, where she individually assesses the site's quality according to various metrics

[29, 22, 47]. If she judges the site to be satisfactory, the ant accepts it and enters the **Canvassing** phase, in which she returns to the old nest to recruit other ants to the site by leading **forward tandem runs** (FTR). In a FTR, the recruiter slowly leads a single follower (another active worker) from the old nest to the new [40, 55, 68]. Upon arriving at the nest, the follower ant goes directly into the Assessment phase and evaluates the nest's quality independently of the leader ant. If she finds the nest satisfactory, she will transition to the Canvassing phase and start leading FTRs to the nest. A canvasser continues leading FTRs until she perceives that the new nest's population has exceeded a threshold, or quorum [48]. At this point, she enters the **Transport** phase, in which she fully commits to the new nest as the colony's home. She ceases FTRs and instead switches to picking up and carrying nestmates from the old to the new nest. These transports are faster than FTRs, and they are largely directed at the passive workers and brood items, hence they serve to quickly move the entire colony to the new nest [46, 53]. Previous models and experiments indicate that the quorum rule helps the colony to reach consensus rather than splitting among multiple sites [46, 19, 23]. Splitting becomes a danger if ants at different sites, each ignorant of their nestmate's discoveries, launch FTRs to their respective sites. The quorum rule makes it likely that whichever site first hits the threshold will quickly end up with all or most of the colony, due to the speediness of transport.

Although experimental evidence is equivocal, we assume that the quorum size is correlated with the number of adult workers in the colony [13, 19]. We also assume that passive workers can contribute to quorum attainment. Once the quorum is met, the switch to Transport phase is irreversible: an ant continues transporting nestmates to her new home nest even if the nest population later drops below the quorum size [48]. However, transporters do sometimes interrupt transport to search for and assess alternative nest sites. If the search yields a new site that is better than the ant's current nest, then she exits the Transport phase and enters the Assessment phase with the new site as her candidate nest.

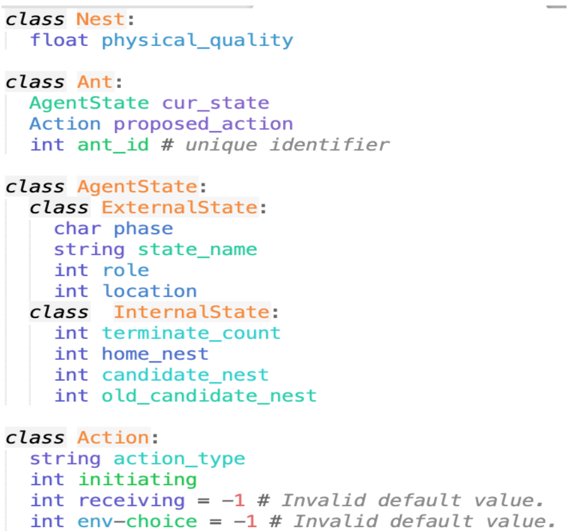
An ant in the Canvassing or Transport phase does not recruit indefinitely. Once the site from which she is recruiting is empty, she returns to her home nest and transitions back to the **Exploration** phase. However, this happens only upon meeting a “termination" condition consisting of ten occurrences of either of the following events: 1) the worker tries to lead a FTR where the solicited follower is also trying to lead her own FTR, and 2) the worker tries to carry another worker who is also in the Transport phase. This condition is based on frequent observation of these events at recently emptied nests. We hypothesize that an ant’s requirement of several such events is a means of ensuring thorough exploration of the old nest so that no nest-mates are left behind. We do not have a precise measure of how many such events are required, but chose the number 10 as an upper-bound estimate.

The emigration is completed when all ants in the colony are relocated to the new nest, except possibly for a few active scouts [46].

2.3.2 Formal Model

Model components

In this section, we show how each component in our modeling framework (Section 2.2.1) is defined in the house hunting algorithm context.



```
class Nest:
  float physical_quality

class Ant:
  AgentState cur_state
  Action proposed_action
  int ant_id # unique identifier

class AgentState:
  class ExternalState:
    char phase
    string state_name
    int role
    int location
  class  InternalState:
    int terminate_count
    int home_nest
    int candidate_nest
    int old_candidate_nest

class Action:
  string action_type
  int initiating
  int receiving = -1 # Invalid default value.
  int env-choice = -1 # Invalid default value.
```

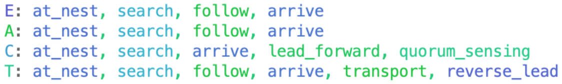
Figure 2-1: Native data structures that define different entities in the distributed system.

Fig. 2-1 shows our native data structures as used by various components in the system: *Nest* objects, an array which constitutes an *env-state*; *Ant* objects, each corresponding to an agent; *State = (ExternalState, InternalState)* objects, each corresponding to a *state = (external-state, internal-state)*, and *Action* objects, each corresponding to an *action.* Each of the data structures contains a set of variables, as seen in Fig. 2-1. Note that we consider all variables belonging to either the class *ExternalState* or the class *InternalState* to belong to the class *State* as well. Throughout the rest of the chapter, we use the notation *object.variable* to denote the value of a *variable* belonging to a class *object.* Using these data structures as building blocks, we now show all possible values for the components in the framework presented in Section 2.2.1. Note that for consistency with our implementation in Section 2.4, we use -1 or an empty string “” to represent any invalid default integer or string values represented by $\perp$ in Section 2.2.

- **agent-ids**, the set containing all integers in the range [0, num_ants), where num_ants is the total number of ants in the colony. In addition, **agent-ids**$'$ = **agent-ids** $\cup$ -1. Each *Ant* is initialized with its corresponding *ant_id*, which corresponds to a *agent-id.*

- **external-states**, the set containing all possible values for an *ExternalState* class object, each corresponding to an *external-state.* We designed these variables to be in the *external-state* because these contain information that influences other ants' activities. Therefore, it is biologically plausible that individuals have access to this information about one another.

  In any *ExternalState* class object, *phase* has one of four possibilities - Exploration (searching for new nests), Assessment (assessing new nests), Canvassing (leading other active workers on FTRs to her accepted candidate nest), and Transport (committing to the new nest and rapidly carrying other ants to it). Note we abbreviate the four phases to names “E", “A", “C" and “T", respectively. The initialization of an *Ant*'s *phase* and *state_name* can be found in Section 2.3.2. For each *phase*, the variable *state_name* take values from a different set, as follows:



```
E: at_nest, search, follow, arrive
A: at_nest, search, follow, arrive
C: at_nest, search, arrive, lead_forward, quorum_sensing
T: at_nest, search, follow, arrive, transport, reverse_lead
```

The variable *role* can be one of (0,1,2) representing (active ant, passive ant, brood), and each *Ant* is initialized with the appropriate value. The variable *location* can be any integer in the range [0, num_nests) where num_nests is the total number of nests in the environment, with 0 representing the original home nest. In addition, recall that **all-externals** is the set of all possible mappings from **agent-ids** to **external-states**. Each element of the set is an *all-external*.

- **internal-states**, the set containing all possible values for an *InternalState* class object, each corresponding to an *internal-state*. The set of fields we designed for the *InternalState* class represent information that should only be accessed and modified by an ant's internal memory. Each of *home_nest* (initial value = 0), *candidate_nest* (initial value = -1), and *old_candidate_nest* (initial value = -1) can take any integer in the range [0, num_nests), where num_nests is the total number of nests. Lastly, *terminate_count* (initial value = 0) takes any value in the range [0, 10].

- **env-states**, a set of arrays, each being an array of the *Nest* class objects. Each array corresponds to an *env-state*. For an *env-state*, the *Nest* at index 0 represents the original home nest and has *physical_quality* 0. All other *Nest*'s have *physical_quality* in range [0, 4]. The maximum quality 4 here is arbitrary. Recall that the array does not change throughout the execution of the system, and the array is read from a configuration file introduced in Section 2.4.1.

- **action-types**, the set of the types of actions includes: “search", “no_action", “find", “follow_find", “get_lost", “reject", “no_reject", “accept", “recruit", “quorum_met", “quorum_not_met", “stop_trans", “delay", “terminate", “lead", "carry". *Action-type* is initialized to "no_action". Each item in the set above is an *action-type*.

- **env-choices**, the set of integers in [0, num_nests) ∪ -1 where num_nests is the number of nests in the environment. Each element in the set is an *env-choice* and is an integer representing an index into *env-state*. An *env-choice* has initial value -1.
- **actions**, the same set as defined in Section 2.2.1. Note that not all actions require a receiving agent, and not all actions require an *env-choice*. In case that they are not needed, they take the invalid default value -1.
- **select-action**(*agent-id, state, env-state, all-external*): the same function as defined in Section 2.2.1. Refer to Section 2.3.2 for details.
- **transition**(*agent-id, state, all-external, action*): the same function as defined in Section 2.2.1. Refer to Section 2.3.2 for details.

The select-action function

The function **select-action**(*x, state_x, env-state, all-external*) outputs a probability distribution over the sample space of **actions**, for which the second component is equal to the input argument *agent-id* and the third component is not equal to it. Let any *action* in the sample space be denoted by (a, x, x', ec), where the second component is fixed. We now list out the probability distribution on other components for each possible value of the *state_name* variable in *state_x*, as it is the only variable in *state_x* that affects the output probability distribution. The boldface words are parameters that we can tune and whose values are read from a configuration file, introduced in Section 2.4.1.

- For *search*, the probabilities of choosing a to be "find" and "no_action" are **search_find** and 1-**search_find** respectively, and all other *action-type*'s have 0 probability. Both variables x' and ec take the invalid default value -1 with probability 1.

- For $follow$, the probabilities of choosing a to be “follow_find" and “get_lost" are **follow_find** and 1-**follow_find** respectively, and all other *action-type*’s have 0 probability. Both variables x' and ec take the invalid default value -1 with probability 1.

- For $reverse\_lead$, the probabilities of choosing a to be“delay" and “no_action" are **transport** and 1-**transport** respectively, and all other *action-type*’s have 0 probability. Both variables x' and ec take the invalid default value -1 with probability 1.

- For $quorum\_sensing$, let the set $\widetilde{X}$ be the set containing id’s of all agents with *external-state* having $role \in \{0, 1\}$ and $location = state\_x.location$. If the set size $|\widetilde{X}| \geq$ **quorum_threshold**, the probabilities of choosing a to be “quorum_met" and “quorum_not_met" are 1 and 0 respectively, and are 0 and 1 otherwise, and all other *action-type*’s have 0 probability. Both variables x' and ec take the invalid default value -1 with probability 1.

- For $lead\_forward$, let $\widetilde{X}$ be the set containing id’s of the agents that are not x, and whose *external-state* has $role = 0$ and $location = state\_x.location$. The function selects an action $\widetilde{act} = (\widetilde{a}, x, x', ec)$ according to the following probability distribution. In case $terminate\_count < 10$, $\widetilde{a}$ is chosen among “lead" and “get_lost" with probabilities **lead_forward** and 1-**lead_forward** respectively, and all other *action-type*’s have probability 0. In case $terminate\_count \geq 10$, $\widetilde{a}$ is “terminate" with probability 1. The variable ec is equal to $\{state\_x.candidate\_nest\}$ with probability 1. The distribution of x' depends on $\widetilde{a}$, as follows:
  - For “lead", if $\widetilde{X} \neq \emptyset$, the variable x' is uniformly selected from $\widetilde{X}$, and all other values in *agent-id*$'$ have 0 probability; otherwise, $x' = -1$ with probability 1.
  - For “get_lost", $x' = -1$ with probability 1.
  - For “terminate", $x' = -1$ with probability 1.

- For *transport*, let $\widetilde{X}$ be the set containing id's of all agents that are not x, and whose *external-state* has $location = state\_x.location$. In addition, let $\widetilde{X'}$ be the subset of $\widetilde{X}$ containing agents that have $role \in \{1, 2\}$. The function first selects an action $\widetilde{act} = (\widetilde{a}, x, x', ec)$ according to the following probability distribution. In case $terminate\_count < 10$, $\widetilde{a}$ is chosen among "carry" and "stop_trans" with probabilities **transport** and 1-**transport** respectively, and all other *action-type*s have probability 0. In case $terminate\_count \geq 10$, $\widetilde{a}$ is "terminate" with probability 1. The variable ec is equal to $\{state\_x.home\_nest\}$ with probability 1. The distribution of x' depends on $\widetilde{a}$, as follows:

  - For "carry", if $\widetilde{X'} \neq \emptyset$, x' is uniformly sampled from $\widetilde{X'}$, and all other values in *agent-id*$'$ have 0 probability. Otherwise if $\widetilde{X'} = \emptyset \cap \widetilde{X} \neq \emptyset$, x' is uniformly sampled from $\widetilde{X}$, and all other values in *agent-id*$'$ have 0 probability. Otherwise, $x' = -1$ with probability 1.

  - For "stop_trans", $x' = -1$ with probability 1.

  - For "terminate", $x' = -1$ with probability 1.

- For *at_nest*, the probability of choosing a to be "search" is $1 - p(x)$, where x is the quality of the nest option under assessment (Figure. 2.1) and $p(x)$ defined in Equation 2.2. There are always two possible actions for a *state* with $state\_name = at\_nest$, and the one that is not "search" naturally has probability $p(x)$. All other *action-type*'s have 0 probability. Both variables x' and ec take the invalid default value -1 with probability 1. To determine $p(x)$, an ant is required to assess the quality of a nest in the environment. The assessment of the quality of a nest includes both its physical qualities [2, 57] and the nest population [49, 13]. Therefore, we use a simple linear combination of these two values to denote the final nest quality with a new parameter called **pop_coeff** as the coefficient of the population effect. In other words, the final nest quality of a nest with *physical_quality* q and population *pop* (obtained from *all-external*) is

$$\frac{q}{4} + \mathbf{pop\_coeff} \times \frac{pop}{num\_ants}, \tag{2.1}$$

where 4 is the maximum value of nest qualities, and num_ants is the total colony size. We further define the following sigmoidal function (Fig. 2-2)

$$p(x) = \frac{1}{1 + e^{-\lambda \times x}} \tag{2.2}$$

where λ is a parameter that controls how "steep" the sigmoidal function is, and x is the above defined nest quality. Higher λ values correspond to lower individual noise level, and bring $p(x)$ closer to a step function.

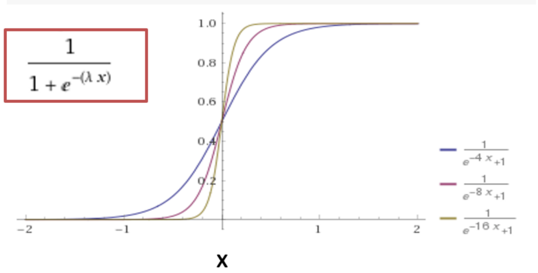


Figure 2-2: Sigmoidal function with $\lambda = 4, 8, 16$.

- For *arrive*, the probabilities of choosing a to be "reject" and "no_reject" are $1 - p(x)$ and $p(x)$ respectively, where x is the difference in quality of the candidate nest compared to the home nest (Equation 2.3) and $p(x)$ defined in Equation 2.2. All other *action-type*'s have 0 probability. The variable x' take the invalid default value -1 with probability 1. The variable ec take the invalid default value -1 with probability 1. To determine $p(x)$, an ant x is required to compare the quality of its $candidate\_nest$ (with physical_quality q_1 and population pop_1) and its $home\_nest$ (with physical_quality q_0 and population pop_0). We still use the sigmoidal function in Equation 2.2, with the change that the input x to the function now is

$$\frac{q_1 - q_0}{4} + \mathbf{pop\_coeff} \times \frac{pop_1 - pop_0}{num\_ants} \tag{2.3}$$

where 4 is the maximum value of nest qualities, and num_ants is the total colony size.

The transition function

Passive Workers and Brood Items Active worker scouts are defined as those who engage in the emigration process by independently discovering the new nests (entering without carrying or being carried) or by carrying brood items or other adult ants to the new nest or both. Passive workers remain in the old nest until they are carried to the new nest. Brood items are similar to passive workers but do not contribute to quorum attainment [46, 15].

We use sn_p to denote a *state* with a certain $state\_name = sn$ and $phase = p$. Passive workers and brood items together form the passive majority population in the colony. Their behavior pattern is thus very simple — they only have one $state\_name_{phase}$, at_nest_E, available to them. They can only allow one *action* with *action-type* “carry" and themselves as the receiving agent. The action results in the *location* variable in their *state* set to the last component of the action, *env-choice*, and no other variables in their *state*'s can change throughout the execution. Therefore, the rest of the section focus on the *state* transitions of **active** workers only, including any initiating and receiving ants involved.

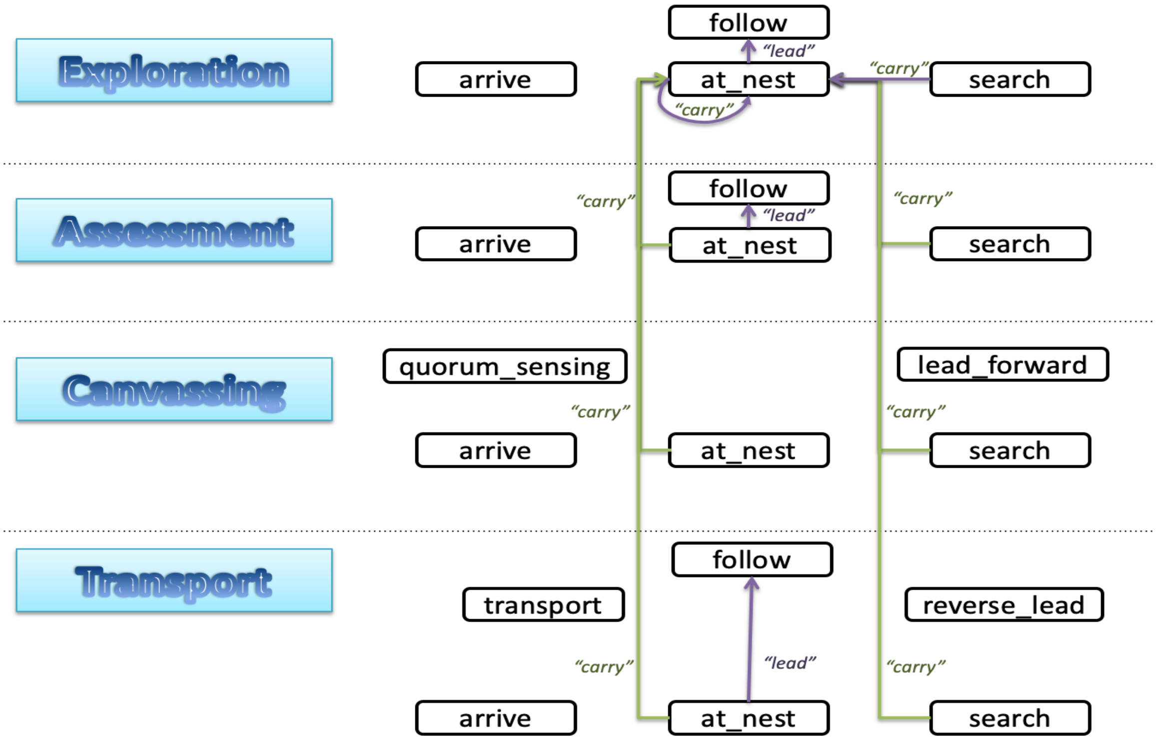
Initiation and Termination of Emigration All ants start in at_nest_E. Their *role* variable values are assigned the corresponding numbers, and $home\_nest$, *location* are both initiated with 0, the original home nest. The variables $candidate\_nest$ and $old\_candidate\_nest$ are set to -1 as the default invalid value. And $terminate\_count$ starts with 0.

We do not designate a separate “termination state" that disables an ant from exploring further, but at the termination of the emigration process, we expect most active workers to be in at_nest_E. This is enforced softly through the population effect introduced in Section 2.3.2 - if an agent in at_nest_E is in a nest with both a high physical quality and a high nest population it is highly likely that she is happy staying put in this nest and stabilizes in the state at_nest_E. As a result, the more agents stabilizes in the same nest, the more likely that they will stay stable and that new agents will stabilize as well. In the house hunting algorithm, the conditions that trigger this "termination" behavior contains two cases, as mentioned in Section 2.3.1. The details of this special "termination" case handling is discussed in the next paragraph.

Special and General Cases In the house-hunting algorithm, there are some special cases that the **transition** function handles before outputting the resulting *state*. To facilitate, we define a set **allowed-in**(*external-state*) to be a mapping from **external-states** to subsets of **action-types**. Consider an *external-state* s, and the allowed subset is then **allowed-in**(s), representing the set of actions the agent in the *external-state* s is allowed to receive. The four variables s contains (as shown in the *ExternalState* class) each affects **allowed-in**(s) in the following way. *location* has no influence. If *role* is 1 (passive) or 2 (brood), **allowed-in**(s) = "carry". Otherwise, $role = 0$. Let $state\_name_{phase}$ denote the *state_name* and *phase* variables in s. For at_nest_E, at_nest_A, and at_nest_T, **allowed-in**(s) = "lead", "carry". For $search_E$, $search_A$, $search_C$, $search_T$, and at_nest_C, **allowed-in**(s) = "carry". For all other cases, **allowed-in**(s) = $\emptyset$.

We now list out how the function **transition**(*agent-id, state, all-external, action*) handles each of the special cases, and also the general case. Let the input argument *action* be expanded to the quadruple ($act = a$, x, x', ec). Also recall that the set $Trans$ is a set containing the id's of all the agents that have completed a *state* change in the round (Section 2.2.2).

- The first special case is if the input argument $agent\text{-}id = x'$. This case only happens when agent $x' \neq -1$ invokes (in agent x's turn) a **transition**(x', *state,*

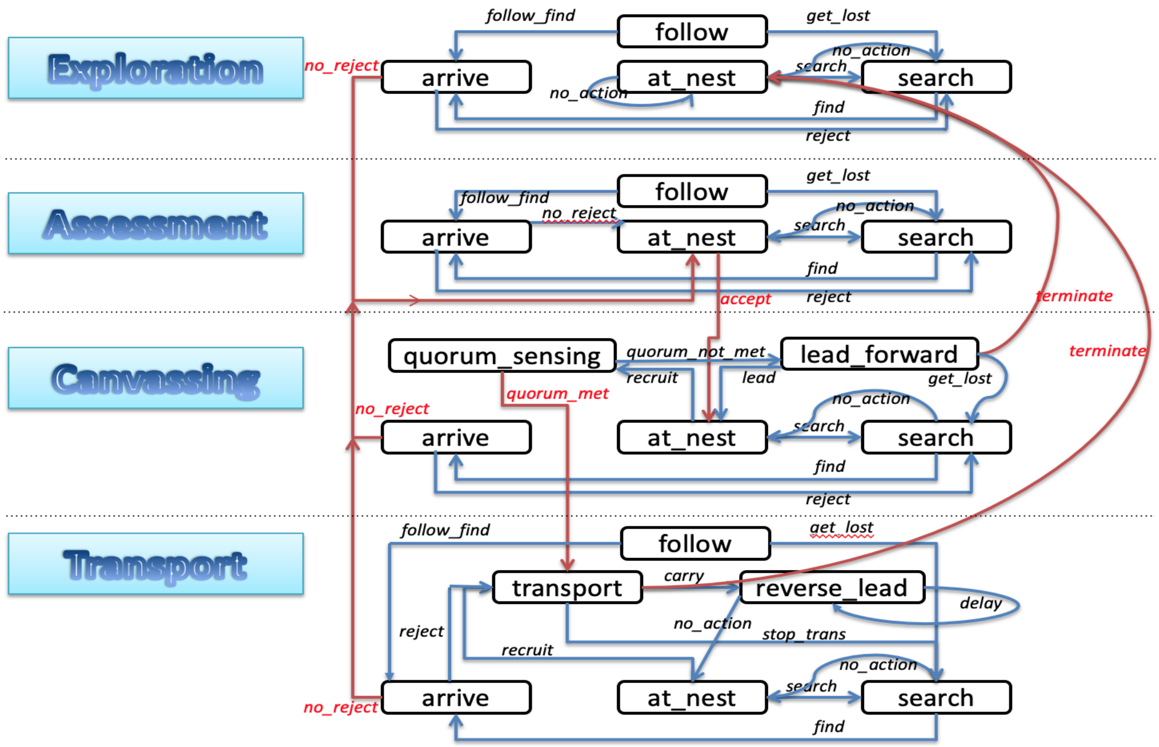


(a) *Action-types* an active ant can receive from another ant and the corresponding *state_name* and *phase* transitions.

Figure 2-3: States and actions modeling the behavior of active ants responsible for organizing colony emigrations. As described in Section 2.3.1, the four distinct phases are in different boxes: Exploration, Assessment, Canvassing, and Transport.

all-external, act), where $state = (external', internal')$ is the current *state* of agent x'. If $x' \in Trans$ or if $a \notin$ **allowed-in**$(external')$, the function simply ends by returning the input argument *state*. Otherwise, the function adds x' to $Trans$. It then finds the black text box corresponding to *state.phase* and *state.state_name* in Fig. 2-3a, and the black text box that a leads to contains the *phase* and *state_name* of the resulting *state*. The rest of the variables in *state* are modified as well, and the details are listed for each possible value of the (*phase*, *state_name*) pair at the end of the section. The function then outputs the resulting *state*.

- The second special case is if *act* satisfies the termination condition mentioned earlier in this Section. Specifically, the cases are when $agent\text{-}id = x$, and *act*



(b) *Action-type*'s an active ant can initiate and the corresponding *state_name* and *phase* transitions.

Figure 2-3: (Cont.) States and actions modeling the behavior of active ants responsible for organizing colony emigrations. As described in Section 2.3.1, the four distinct phases are in different boxes - Exploration, Assessment, Canvassing, and Transport.

is either 1) $(lead\_forward, x, x', state\_x.candidate\_nest)$ and $x' \neq -1$ has an *external-state* with $state\_name = lead\_forward$, or 2) $(transport, x, x', state\_x.home\_nest)$ and $x' \neq -1$ has an *external-state* with $state\_name = transport$. We call these the "termination conditions". When *act* satisfied either clauses, after adding x to $Trans$, the function ends its execution by outputting a resulting *state* that only differs from the input argument *state* by adding 1 to the $terminate\_count$ variable.

- The third special case is if $agent\text{-}id = x$ and *act* does not satisfy the termination conditions, but $x' \neq -1$ and either of the following is true: 1) $x' \in Trans$, or 2) $a \notin$ **allowed-in**$(external')$. Note the second case here excludes cases that satisfy our termination conditions stated in the last bullet point. In other words,

Arrow Color	Init/Recv	Phase Change
blue	Initiating	No
red	Initiating	Yes
purple	Receiving	No
green	Receiving	Yes

Table 2.1: Color coding of arrows representing *action-type*'s in Fig. 2-3a and Fig. 2-3b.

the second special case has priority over this third special case. In this third special case, the function adds x to $Trans$, and ends its execution by outputting the original input argument, *state*.

- Lastly, in the general case where none of the above special cases applies, the function first adds x to $Trans$. Then it finds the black text box corresponding to *state.phase* and *state.state_name* in Fig. 2-3b, and the black text box that a leads to contains the *phase* and *state_name* of the resulting *state*. The rest of the variables in *state* are modified as well, and the details are listed for each possible value of the (*phase*, *state_name*) pair at the end of the section. The function then outputs the resulting *state*.

In Fig. 2-3a and Fig. 2-3b, *action-type*'s are color-coded as shown in Table 2.1. We walk through in Appendix A all possible transitions of an ant and the associated changes in the internal and external states. An example is an ant in at_nest_E has four possible actions. First, she can perform "no_action" and remain in the current nest. Second, she can perform "search" and go into the state $search_E$. Third, she can receive a "lead" by another ant to follow a FTR to a destination nest, $ec \in$ **env-choices**, in which case she sets *old_candidate_nest* to the value of *candidate_nest*, and sets *candidate_nest* to *ec*. Then she transitions to the state $follow_E$. Finally, she can receive a "carry" by another active worker ant to a destination nest $ec \in$ **env-choices**, in which case her *location* and *candidate_nest* are changed to *ec*, and she stays in at_nest_E.

2.4 Model Simulation and Metrics

In this section we describe the configuration file that contains all the parameters of the model in Section 2.3. We also quantitatively define the speed and accuracy measures for our simulation runs. A detailed description of the Python simulator implementation is in Appendix A.2, and our simulator code is open sourced at https://github.com/snowbabyjia/Collective-Decision-Making-HH.

2.4.1 Configuration Parameters

There are three kinds of parameters: environment, algorithm, and settings.

Environment parameters are controlled by the environment and not considered changable or tune-able. These include the number of ants in the colony, and the number and physical qualities of the nests as potential new nest options.

Algorithm parameters are parameters that we can manipulate in order to change the *select-action* function and hence the outcomes of our simulations. These include the λ for the sigmoid function in Equation 2.2, the pop_coeff value, parameters related to quorum sizes, the probability of finding a new nest in the environment, the probabilities of following and leading a FTR without getting lost, and the probability of continued transports instead of stopping transportation. Related explanations are in Section 2.3.2 and 2.3.2.

Settings parameters control plotting features and also convergence criteria. These include the option to generate a plot, the total number of runs for every environment/algorithm setting, the maximum number of rounds per simulation run, the percentage of ants needed in a nest to declare **convergence**, and the number of rounds the convergence needs to persist to declare **persisted convergence** which marks the end of the simulation run. An example is in the supplemental materials.

Baseline Default Parameter Values Compared to the agent-based model in [53], our model places less emphasis on assigning specific observed values to a large number of parameters, but rather on a simple and elegant model that is more agile in representing a wide range of possible behaviors. For that reason, some parameters cannot be directly drawn from existing empirical data. We estimate these parameter values in a trial-and-error fashion until simulation results match well with the empirical results in [53]. These baseline values are used as a default from Section 2.5 to 2.8, unless otherwise specified.

The sources for determining the parameter values are listed in Table 2.2. In particular, the values of **lambda_sigmoid** (range: 1 to 16) and **pop_coeff** (range: 0 to 1) are picked by trial-and-error to model individual sensitivity to nest qualities, and the significance of colony information versus individual judgements. The quorum size (**quorum_thre** × (num_active+num_passive) + **quorum_offset**) is observed to have a median value between 4 and 18 ants for worker populations from 24 to 150, with the quorum size having a significant positive correlation with the number of adult ants [46, 24]. Therefore, with a colony of 200 members (including 100 adult workers), we use a **quorum_thre** of 15% and set **quorum_offset** to 0, estimating a quorum size of 15. The value of **search_find** (range: 0 to 1) is determined experimentally by trial-and-error. This parameter can be influenced by many other factors such as the spatial geometry of the nests and the experience level of the individual. These nuances are not captured in our model in the interest of simplicity. But they can significantly affect the simulation outcomes, and are an important future extension of our work. The parameter **follow_find** denotes the success rate of a tandem run without the follower getting lost and starting a new search. A successful tandem run requires that both ants reach the target nest. Empirical observations suggest large variation in tandem success, with observed success ranging from 30% to over 90% [50, 27]. However, even lost followers enjoy a significantly increased chance of finding the target nest on their own [50]. We thus chose a high FTR success rate of 0.9 to capture both these direct and indirect effects of tandem following on nest discovery. The parameter **lead_forward** (range: 0 to 1) is the probability that an ant performs an FTR when in the *lead_forward* state. The alternative option, *get_lost*, is designed to model the slower speed of an FTR, and is determined experimentally in a trial-and-error fashion. The parameter **transport** is the probability that an ant keeps

Parameter	Value	Source
lambda_sigmoid	8	trial-and-error, Sec. 2.7.1
pop_coeff	0.35	trial-and-error, Sec. 2.7.1
quorum_thre	0.15	[46, 24]
quorum_offset	0	[46]
search_find	0.005	trial-and-error
follow_find	0.9	[27, 50]
lead_forward	0.6	trial-and-error
transport	0.7	[53]

Table 2.2: Default parameter values and the sources that helped determine these values.

transporting instead of stopping to resume search for additional sites. The stopping probability is observed to be between 0.06 and 0.44, meaning our **transport** should take values between 0.56 and 0.94. We chose 0.7 as our baseline value.

An average colony size of 200 with 50 active workers, 50 passive workers, and 100 brood items is within the range of real colony compositions [24]. One round approximately translates to 0.5-1 minutes, though this is a very rough estimate. A simulation with 2000 rounds thus translates to 16-32 hours, and one with 4000 rounds translates to 32-64 hours. The values for variables **percent_conv** and **persist_rounds** are determined by trial-and-error and rough estimates from past empirical observations.

2.4.2 Speed and Accuracy Measures

We define the speed and accuracy metrics below for the *whole emigration process* until either convergence or the end of simulation, including cases resulting in splitting.

Convergence Score as Speed The final goal of the house hunting algorithm is to achieve fast convergence in any given environment and stabilize at that convergence. To assess how well this was achieved, we calculate a **convergence score** as the inverse of the round number when a persistent convergence started. If no persistent convergence was reached before the end of the simulation, the convergence score is 0. Each simulation run has a convergence score.

Accuracy Another important metric is **accuracy**, which is defined for a group of simulation runs. This metric tells us how good the colony is at selecting the best choice in the environment. Thus, each of the nest options in the configuration has an empirical probability of the colony converging to it, called the nest's **convergence probability**. Note that we also get a probability of splitting. To calculate the final accuracy, we also normalize the nests' physical qualities, such that the best nest has quality 1 after normalization, and the worst nest, which is the home nest, has quality 0. The **accuracy** of the configuration is then

$$\sum_{i \in nests} p_i \times q_i$$

where p_i is nest i's convergence probability, and q_i is its normalized physical quality. If no convergence is reached (splitting), the physical quality corresponding to that probability will be 0, thus not contributing to the summation above.

2.5 Model Validation

We begin by validating our model against the same empirical data that were successfully accounted for by two earlier models [53, 52]. First, we examine a simple scenario where colonies have only one candidate nest in the environment. Then we consider a decision between two nests that clearly differ in quality. Finally, we investigate how colonies trade off speed and accuracy depending on the urgency of their move. For all scenarios, we simulate the same data explored by the earlier models, and compare our results, at both individual and colony level, to the empirical observations. All simulations for the rest of the chapter default to the configuration file described in Section 2.4.1, unless specified otherwise.

2.5.1 Single-Nest Emigrations

The first question we ask is: does our model accurately reproduce statistics on individual recruitment acts in single-nest emigrations? Previous empirical work showed

Colony	Active	Passive	Brood	Total
A4	70	28	228	326
A6	59	74	111	244
A8	62	95	106	263
A14	67	42	192	301
A16	53	88	61	202
A17	73	101	173	347

Table 2.3: Compositions of colonies used in two-nest emigrations for model validation as shown in [53].

the distributions across ants of key behaviors contributing to the collective outcome [53]. These include the number of recruitment acts per ant, the number of ants performing each recruitment type, and the number of ants arriving at the new site by different routes. We asked whether our model could replicate the empirical distributions. To answer the question, we simulated the single-nest experiments conducted in [53], on the six colonies with compositions detailed in Table 2.3. We used default parameter values, except we increased *search_find* to 0.05. This increase accounts for the presence of only one new nest, hence all “find" actions after the first one are re-discoveries of this nest, which we assume has a higher probability than finding a previously unknown site [53]. In future work, this variable should be expanded to depend on other factors, such as the number of nests in the environment or the spatial geometry. We ran 500 simulations for each colony.

Results We compared the statistics of the model output to the same statistics reported in [53] (Fig. 2-4). Fig. 2-4(a) shows the histograms of individual workers grouped by the number of recruitment acts. More than half of the simulated workers never recruited, consistent with the empirical finding of about 60% non-recruiting active workers. The other bins show similar mean and variance to the empirical data. Fig. 2-4(b) classifies ants by their recruitment behavior, and the breakdowns are again consistent with the experimental observations. Fig. 2-4(c) categorizes workers by their routes to discovery of the candidate nest, and is again consistent with the findings in [53], at least when the experimental data are pooled over six emigrations by three colonies. However, the distributions across the three different routes vary strongly across emigrations. Indeed, the results in [53] notably differ from those in [45]. While our model does not account for this variation, we conclude that it does adequately reproduce key distributions in recruitment behavior in single-nest emigrations.

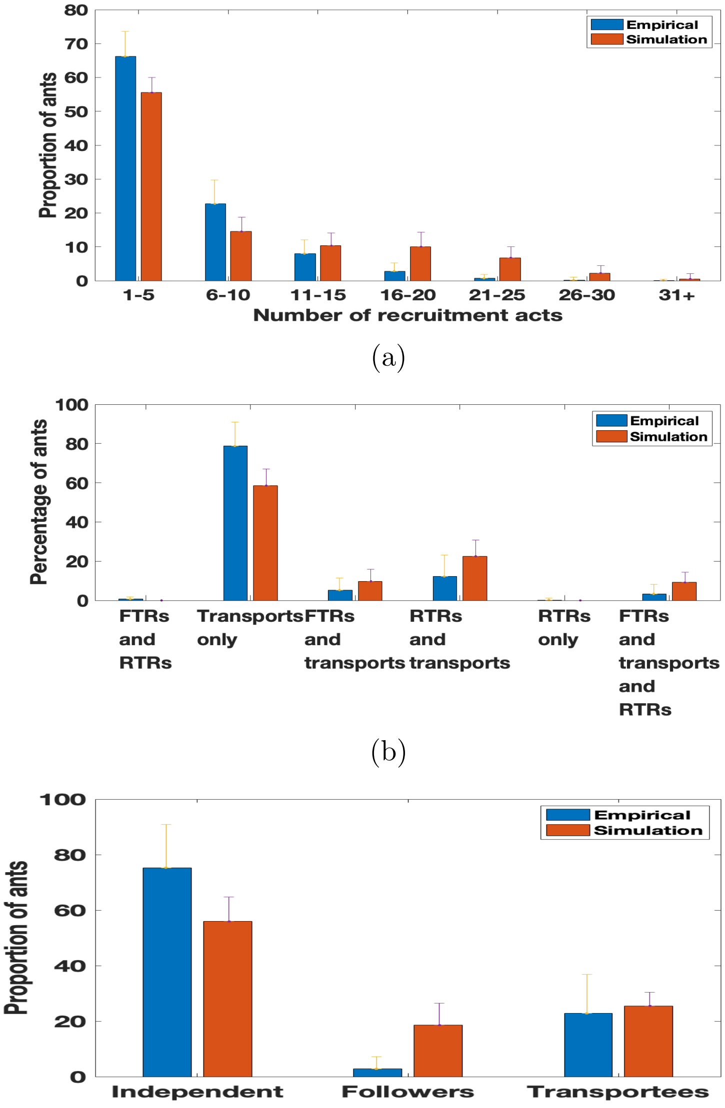
2.5.2 Two Unequal Nest: Splits

The second question we ask is: does our model account for the degree of splitting in two-nest emigrations with unequal qualities? In these circumstances, colonies do not always make a unanimous choice, but may temporarily split between the sites before eventually coalescing on a single one. We focus on splitting because it is a primary hindrance to consensus. We measure splitting as the percentage of brood items in the better candidate nest at the time when the last ant has been moved from the home nest.

We replicated the two-nest emigrations in [53], with six colonies whose member compositions are listed in Table 2.3. We set **nest_qualities** = [0,1,2], representing a destroyed old nest and two candidate nests of mediocre and good quality, respectively. The rest of the configuration parameters were left at the default values.

We ran 500 simulations for each colony, and for each colony we recorded the average percentage of brood items in the better nest at the time the home nest became empty. To compare the simulations with empirical data, we measured for each colony the proportion of simulations departing as far or farther from the colony average as did the experimental value. Twice this proportion gave the p-value for a test of the null hypothesis that the observed value was drawn from the same probability distribution as the simulated values.

Results The results show no significant difference between experiment and simulation for five of six colonies (Table 2.4). This outcome validates our model's ability to reproduce observed patterns of splitting in two-nest emigrations for a variety of colony compositions.



(a)

(b)

(c)

Figure 2-4: (a): Histogram of workers grouped by the number of recruitment acts performed. (b): Histogram of workers who performed different types of recruitment acts. (c): Histogram of workers grouped by the route by which workers arrived at the new nest. Blue bars are results from [53]. Orange bars show our simulation results. Bar values are averaged over 500 simulations. Error bars show standard deviations.

Colony	% Pred	%Observed	P
A4	59 ± 17	61	0.86
A6	61 ± 28	80	0.56
A8	63 ± 30	99	0.36
A14	59 ± 20	98	0.1
A16	61 ± 34	100	0.5
A17	60 ± 25	2	0.02

Table 2.4: Percentage of brood in the better nest for each of the six colonies, predicted vs observed. The last column is the p-value, with $P < 0.05$ indicating a significant difference between predicted and observed percentages.

2.5.3 Two Unequal Nests: Speed-Accuracy Trade-off

The third question we ask is: does our model reflect the way that colonies trade off speed and accuracy when the urgency of the emigration changes? In “urgent" situations, ants face a critical need for immediate new shelter and thus benefit from moving out of the old nest location as fast as possible. In contrast, in less urgent situations they can deliberate longer among alternatives to increase the likelihood of moving directly to the best site [52, 14]. We simulated experiments that adjusted urgency by offering colonies a choice between a mediocre and a good nest under two circumstances: their old nest has just been destroyed (high urgency) or their old nest is of acceptable quality but worse than either of the new candidates (low urgency). Our simulations followed the same tactic by tuning the physical quality of the home nest to adjust urgency. We ran 300 simulations each for eleven home nest qualities in range [0,1], with candidate nest qualities of [1,2]. We used the default parameter values, except **lambda_sigmoid**, which was set to 16 in order to increase the ants' sensitivity to home nest quality differences, and **pop_coeff**, which was set to 0 in order to better match the model assumptions in [52]. As in [52], we measured the duration of emigration as the time in rounds at which the old site was completely abandoned, and we measured the accuracy of decision-making as the proportion of the colony's members inside the good site at the time of old nest abandonment.

Results The results show that time taken to complete an emigration decreased as urgency increased (i.e., as old nest quality decreased) (Fig. 2-5b). This is consistent with the empirical observation that higher urgency induces faster emigrations (Fig. 2-5a). Furthermore, the simulations show that higher urgency (lower old nest quality) reduces the likelihood of the colony achieving consensus on the better site. This also matches the empirical results, which show that higher urgency leads to lower accuracy [52].

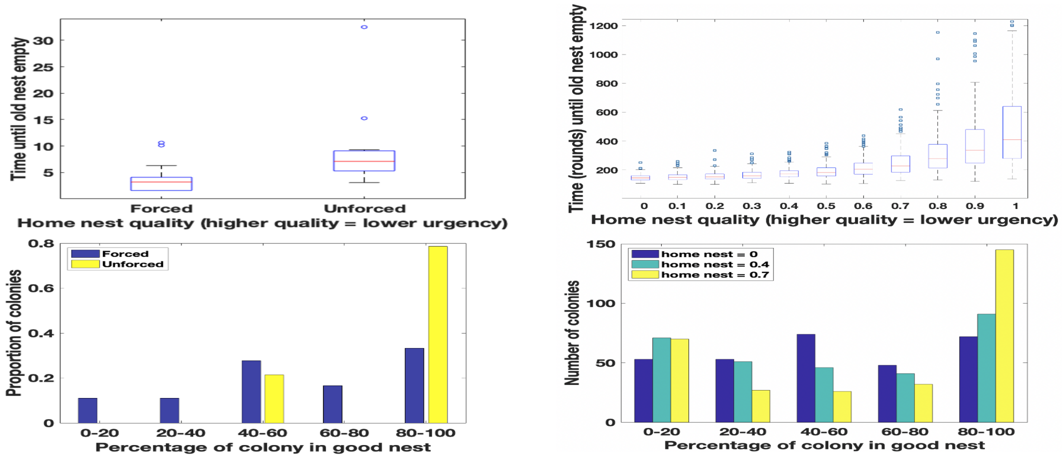
These results confirm that our model can account for the empirically observed speed and accuracy trade-off up to old nest abandonment. However, it is worth noting that real colonies in the low urgency situation were better able to reach consensus than our simulated colonies. This might suggest the existence of other mechanisms at work that this simulation failed to capture. One such mechanism could be sensitivity to the presence of nestmates when assessing the quality of a nest. This could enhance consensus by amplifying the differential treatment of competing nests, as the better one's population increases and makes it still more attractive. Our model can capture this phenomenon by using a non-zero *pop_coeff*, a possibility explored further in the next section.

2.6 Confirmation of New Experiments

In this section we consider more complex scenarios where the link between colony patterns and individual behavior has not previously been modeled. For scenarios that have been explored empirically, we determine how well our model can account for observed results. Section 2.6.1 examines a colony's ability to choose well when faced with larger option arrays; and Section 2.6.2 focuses on how colonies make rational decision time investments depending on nest quality differences.

2.6.1 Colonies Have High Cognitive Capacity

How well do colonies perform when selecting from many nests? A previous study [60] showed that colonies are quite good at selecting a single good nest from a set of eight

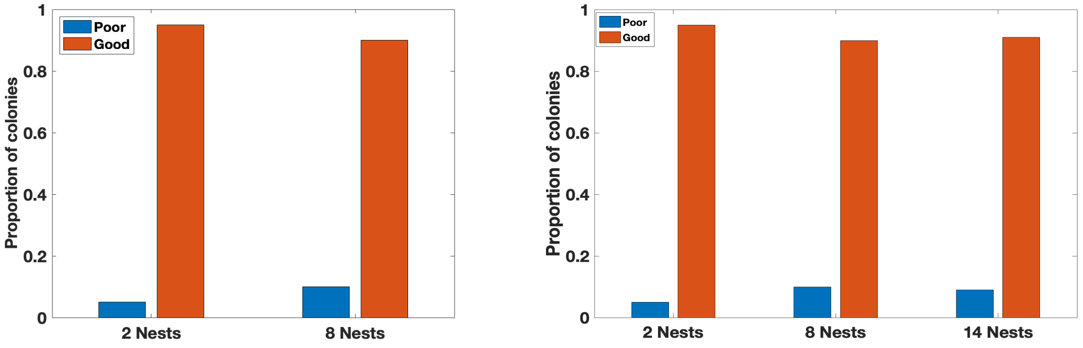


(a) (b)

Figure 2-5: (a): From [52]: Speed and accuracy of decision-making in high-urgency and low-urgency emigrations. (Top) Time until the old nest is empty for each treatment. The ends of each box mark the upper and lower percentiles, and the horizontal line inside the box gives the median. The brackets show the data range, and circles are outliers. (Bottom) Histograms of the degree to which colonies split between the good and mediocre new nest sites, under high urgency (blue bars) and low urgency (yellow bars). (b): Simulation results for high and low urgency emigrations. Top and bottom panels correspond to those in (a). Simulations were run for more urgency levels than in the experiments.

nests, four of which are good and four of which are mediocre. This is in contrast to individual ants, who are as likely to choose a mediocre as a good nest when faced with the same scenario. The colony advantage has been hypothesized to result from sharing the burden of nest assessment: very few of the scouts ever visit more than one or two nests, whereas a lone ant visits several, potentially overwhelming her ability to process information about them successfully. We simulate this experiment to determine whether we can reproduce both the colony's ability and the observed distribution of nest visits across scouts.

We designed a simulated experiment with multiple nests in the environment, half of which are mediocre (physical_quality 1.0) and the rest of which are good (physical_quality 2.0). We considered three environments with 2, 8, and 14 nests, respectively. For each environment, We ran 600 simulations with a fixed colony size 200, containing 50 active and passive ants each, and 100 brood items.



(a) (b)

Figure 2-6: The proportions of colonies that eventually moved into poor or good nests. (a): Empirical results in 2-nest and 8-nest settings [60]. (b): Simulation results from our model in 2-nest, 8-nest, and 14 nest settings.

Results First, we found that simulated colonies reached consensus on a good nest with high probability, matching that seen in empirical data (Fig. 2-6). This was true even when the number of nests was increased to 14.

Next, we verified that the high cognitive capacity of colonies is associated with a low number of nests visited by each scout. The proportion of ants visiting only one or two nest was similar in the simulations and experiments [60]: over 80% of individual ants visited only one or two nests in the course of the emigration. Fig. 2-7 shows similar pattern is seen for the number of transports: that is, if we focus only on the ants who contributed to the emigration by transporting nestmates, over 80% visited only one or two nests. Thus, ants that access many nests have a minor role in the transportation process, supporting the hypothesis that colonies' high cognitive capacity results from avoiding the overloading of individual ants.

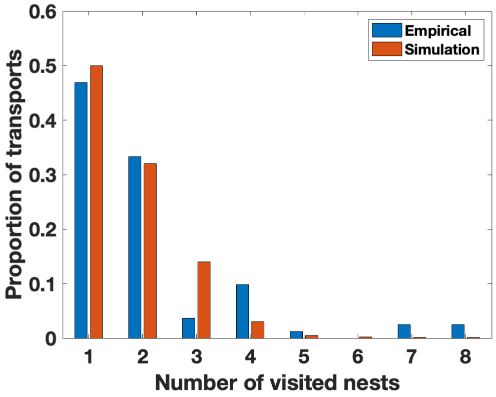


Figure 2-7: Proportions of transport efforts as a function of the number of candidate sites visited by each ant. The blue bars show the percentage of transports done by ants that visited a given number of nests [60], and the dark orange bars show the same for simulated ants. Colonies choose among eight nests (four good and four mediocre) in both simulations and experiments [60].

2.6.2 Colonies Make Rational Choices about Decision Speed

For choices between two nests, how does the difference between the nests affect the speed of decision-making? Counter-intuitively, a previous study [61] found that colonies move more quickly when site qualities are more similar. But this behavior accords with decision theory predictions that decision-makers should take less time if the consequences of their choice are small; that is, since the nests are similar in quality, the opportunity cost of making a wrong decision is small, so it's rational to save time costs by taking on a higher risk of choosing the wrong nest.

We simulate this scenario to determine if we can reproduce the same pattern, but we also explore a broader range of quality differences to better describe the relation between quality difference and decision time. We designed an environment with two candidate nests, one good and the other mediocre. The good nest has physical_quality 2 in all simulations, but the physical_quality of the mediocre nest varies across simulations from 0.2 to 1.7. We asked whether the quality of the mediocre nest is correlated with the convergence score (a measure of decision speed). We ran 300 simulations for each environment with a colony of size 200, consisting of 50 active workers, 50 passive workers, and 100 brood items. We repeated this set of simulations for five different values of **lambda_sigmoid** values: [8,10,12,14,16].

Results If our model reproduces the rational time investment choices of colonies [61], then we expect the convergence score to increase as the mediocre nest quality increases, thus becoming more similar to the good nest. Our results partially match this prediction, with convergence score increasing as the mediocre nest quality goes from 0.2 to about 1 (Fig. 2-8). However, at higher mediocre nest qualities, the pattern reverses and convergence score declines. This basic pattern is seen for all tested values of **lambda_sigmoid**.

We propose that the nest qualities studied in [61] came from the region below the peak score that saw an increase of speed with decreasing quality difference. But from our more granular simulations, we predict that as the quality difference gets still smaller, the convergence score will start decreasing, meaning colonies will start investing more time.

Why might this happen? Recent studies have explained the behavioral difference between individuals and colonies via two different decision models: the tug-of-war model describes individual behavior, while colony behavior is better accounted for by the horse race model [31]. The tug-of-war correctly predicts the irrational behavior of individual ants, in that their decision-making slows down for options that are more similar. The horse race, in contrast, correctly predicts colonies' rational acceleration of decision making for similar options. We hypothesize that the applicability of these models to the colony's behavior changes as the quality difference changes. More specifically in Fig. 2-8, before the peak score is reached, the colony may effectively distribute its decision-making across many ants with limited information, the situation envisioned in the horse-race model. After the peak score is reached, the colony may come to depend more on individual comparisons between nest sites made by a few well-informed ants, and thus to show the irrational slow-down predicted by the tug-of-war model. It could also be the case that more transports are performed between the two candidate nests as the likelihood of the mediocre nest achieving quorum attainment increases.

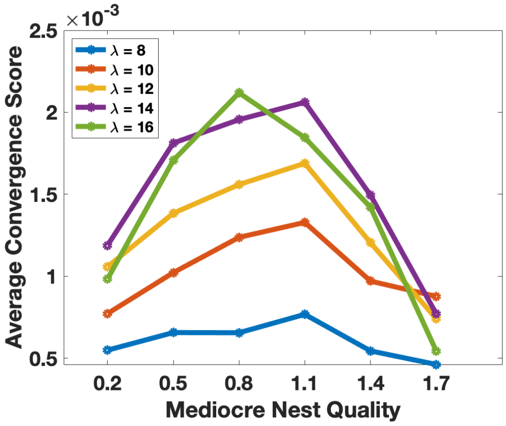


Figure 2-8: Average convergence score as a function of the physical_quality of the mediocre nest. The physical_quality of the good nest is 2, and that of the home nest is 0.

2.7 The Power of Social Information

In this section, we explore the influence of social information on migration speed, accuracy, and cohesion. Section 2.7.1 explores correlations between **pop_coeff** and the degree of randomness in individual decision-making; and Section 2.7.2 reveals how **pop_coeff** decreases splitting by colonies facing two equal options.

2.7.1 Balancing Personal and Social Information

Individual ants are capable of directly comparing nests and choosing the better one, but their discriminatory ability is less than that of whole colonies. This may be seen as a kind of “wisdom of crowds," in which the estimations of many noisy individuals are integrated into a more precise group perception. Ants do this via positive feedback loops based on recruitment, which can amplify small differences in site quality [58]. They also use social information via the quorum rule, under which full commitment to a site is conditioned on a minimum number of nestmates “voting" for it by spending time there. The quorum rule inspired us to consider another way that ants might use social information to improve decision-making: by taking population into account when assessing a site’s quality. We do this via the parameter **pop_coeff**, which controls the degree to which the presence of nestmates increases a site’s perceived value. We propose that this population sensitivity might be able to complement the noisy perception of individual ants, modeled by the parameter λ in the Eq. 2.2. We hypothesize that ants may adapt to different values of **lambda_sigmoid** by changing the value of **pop_coeff**. In particular, we sought evidence for a correlation between the values of **lambda_sigmoid** and **pop_coeff** needed to achieve the best convergence score.

To investigate this question, we ran simulations for different combinations of **pop_coeff** (ranging from 0.002 to 0.8) and **lambda_sigmoid** (ranging from 2 to 16). We ran simulations for an environment containing two identical new nests [0,1,1]. For each combination of **pop_coeff** and **lambda_sigmoid**, we ran 500 simulations with a colony of size 200, consisting of 50 active workers, 50 passive workers, and 100 brood items.

Results The results show evidence for an inverse relation between **pop_coeff** and **lambda_sigmoid** (Fig. 2-9). For each value of **lambda_sigmoid** in the range [2,16], there is a value of **pop_coeff** that maximizes the convergence score, and this value increases as **lambda_sigmoid** decreases. Thus, when an individual ant makes noisy local decisions (modeled with lower values of **lambda_sigmoid**), she can counteract this deficiency by relying more on the input of her peers through a higher value of **pop_coeff**.

However, a high value also has risks. Here we interpret the advantages and disadvantages of *increasing* the value of **pop_coeff**:

Advantages

- Higher momentum in the system. This can promote the colony to accumulate population at a certain nest more quickly, and thus achieve faster convergence.
- Better prevention of splits. Multiple candidate nests may reach the quorum, especially when the nests have similar physical qualities. This can lead to the colony splitting between more than one site. Social information via **pop_coeff**

might help to break ties, by amplifying small random differences in the populations of competing sites.

Disadvantages

- Slower error correction. Since we are dealing with a randomized algorithm, there is always a chance that the colony will collectively make a “bad" temporary decision, even if individuals have low noise levels. The higher momentum will then make the wrong decision more “sticky" by accumulating more ants at a mediocre nest even if a better one is available. The colony would then have to move later to the better nest, adding costs in time and risk. In this way, high **pop_coeff** can cause slower convergence, and lead to “madness of the crowd".

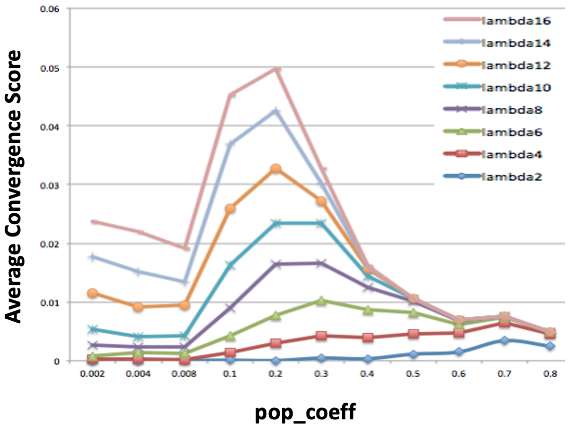


Figure 2-9: Average convergence score (across 500 simulations) as a function of **pop_coeff**. Different colored curves represent different **lambda_sigmoid** values as described in the individual decision model (Fig. 2-2). This shows that the optimal value of **pop_coeff** increases as **lambda_sigmoid** decreases.

These trade-offs suggest that there is an optimal value of **pop_coeff** for a given **lambda_sigmoid** as seen in Fig. 2-9. This predicts that colonies may tune **pop_coeff** according to the uncertainty of individual behavior in order to achieve the highest convergence score for a given environment.

2.7.2 Avoiding Splits Between Two Equal Nests

In this section we further explore how social information can help colonies to reach consensus when faced with two identical nests. Many social insects have highly non-linear recruitment mechanisms that lead to symmetry breaking when faced with two identical resources. For example, ant species that recruit via trail pheromones will choose one of two identical food sources rather than forming trails to both. This is because the attractiveness of a trail is a sigmoidal function of the amount of pheromone it contains, which leads to rapid amplification of small random differences in the strengths of competing trails [1, 42]. However, similar experiments on *Temnothorax* ants found that they do not always break symmetry, instead exploiting both feeders equally, a result that has been attributed to the linear relationship between tandem running effort and recruitment success [63].

An open question is whether this lack of symmetry breaking also holds for nest site selection. When presented with identical nests, do colonies choose only one or split between them? If they can reach consensus, then how do they do so? One possibility is that the quorum rule provides sufficient non-linearity to amplify small random differences in site population, thus ensuring that the colony does not split. Another possibility is that colonies have some other as of now unrecognized mechanism of avoiding splits. A good candidate for such a mechanism is incorporation of site population into each scout's assessment of site quality, as discussed in Section 2.7.1. This would allow amplification of early random differences in population, by increasing the likelihood of recruitment to the nest with more ants. We explore this question by simulating emigrations in which a colony is presented with two identical nest sites. We assess how well they reach consensus on a single one. We also vary the degree of scout sensitivity to site population by considering different values of **pop_coeff**.

We ran 200 simulations each for **pop_coeff** = [0, 0.1, 0.2, 0.3, 0.4], in an environment with **nest_qualities** = [0,2,2]. We set **lambda_sigmoid** to 16 in order to be more sensitive to temporal differences in nest populations. From an initial set

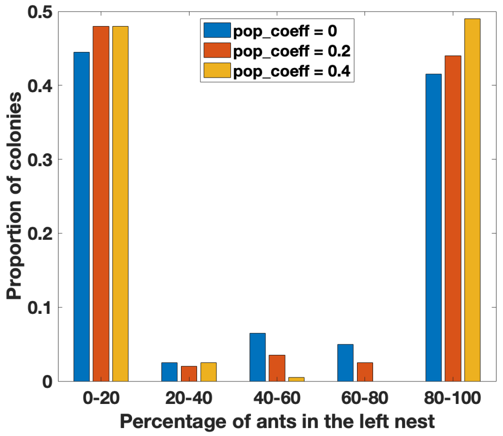


Figure 2-10: Simulation results for colonies choosing between two identical nests. The histograms show the distribution of the percentage of the colony occupying the left nest, for three different values of **pop_coeff**.

of simulations, we observed that almost all simulations converged within the default value of **num_rounds**. Therefore, in order to gain more insight into the effect of **pop_coeff** on the degree of splitting before convergence, we set it to a smaller value 1000. The rest of the parameters take the default values.

Results The simulation results show strong symmetry breaking (Fig. 2-10). That is, a large majority of simulations ended with 80% to 100% of the colony in one of the two nests. When consensus was reached, it was roughly equally likely to be in nest 1 or nest 2, producing the distinctive U-shaped distribution seen in Fig. 2-10. This pattern was true regardless of the value of **pop_coeff**, suggesting that the quorum rule is enough to generate symmetry breaking in this case. However, as the value of **pop_coeff** increases, the histograms also aggregate more towards the two end bins, meaning there are fewer split cases. Thus we confirm the positive effect of **pop_coeff** in reducing splits, either by prevention or by facilitating later re-unification. These mechanisms can have significant effects in more challenging environments with more candidate nests.

2.8 New Predictions

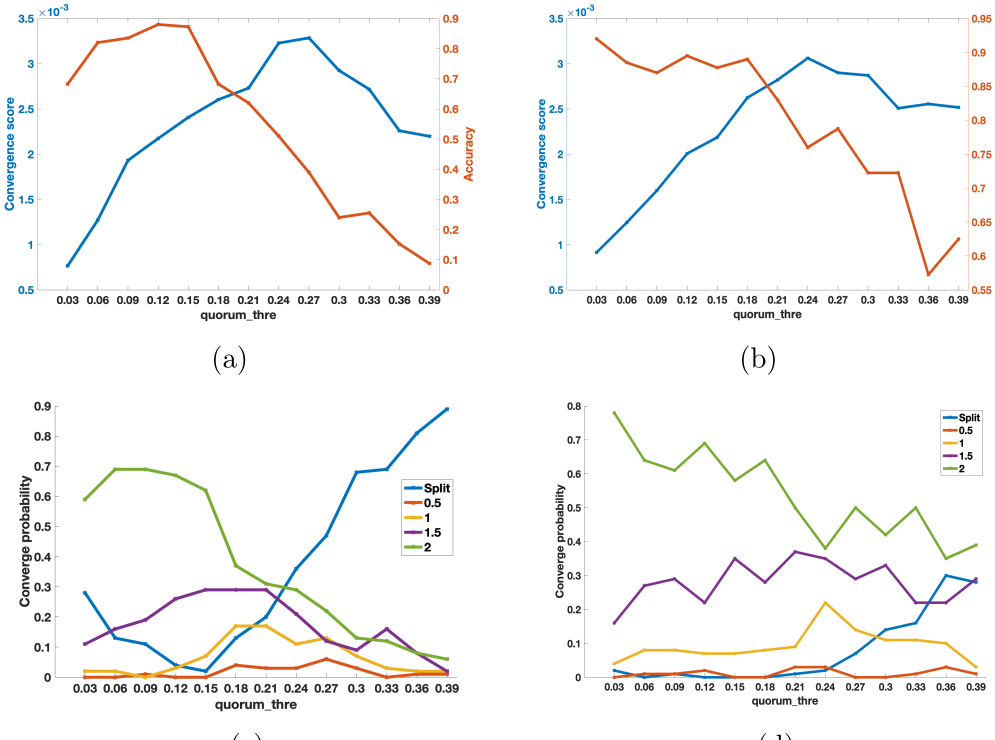
We also use our model to develop new hypotheses and predictions for future experimental study. Section 2.8.1 gives simulated evidence for a surprising speed-accuracy trade-off for the entire emigration process, tuned by the quorum size; and Section 2.8.2 discusses colony re-unification after splits with an increasing level of difficulty.

2.8.1 Quorum Size and the Speed/Accuracy Trade-off

Temnothorax colonies can adjust their behavior to adaptively trade off the speed and accuracy of decision-making [52, 14] (Section 2.5.3). One of the behavioral tools implicated in this adjustment is the quorum rule. Colonies lower the quorum in more urgent situations, increasing their reliance on individual judgement. This allows them to make rapid decisions and quickly move the colony out of the old nest, at the cost of an increased probability of splitting or choosing an inferior nest [20, 18].

When considering speed, previous studies focused on the time to move out of the old nest, but the completion of an emigration often requires more than that. A fast “first" decision does not always mean a fast emigration. In fact, a low quorum and hence a fast “first" decision could lead to slower emigrations [18] since it could cause more splitting, which the colony must subsequently resolve in a second phase of movement. Here, we explore the effect of quorum size on the speed and accuracy as we have defined them for the whole process (Section 2.4.2). Within the accuracy measure, we pay special attention to the rate of splitting, which is the percentage of emigrations that do not reach a persistent convergence within the given number of rounds. A natural question arises: is there a speed-accuracy trade-off if we define “speed" as (the inverse of) the time taken to the final completion of the emigration? In other words, do the convergence score and accuracy have inverse correlations with **quorum_thre**, and are these relationships affected by splitting rate?

We simulated an environment with candidate nests [0.5, 1, 1.5, 2] and a home nest with quality 0 as usual. We used a colony of size 200, consisting of 50 active workers, 50 passive workers, and 100 brood items. Quorum size is assumed to be proportional to the total number of adults in the colony, and is set to **quorum_thre** × **num_adults**. We varied **quorum_thre** from 0.03 to 0.39, and set **pop_coeff** to either 0 or 0.35. We set *num_rounds* to 2000 and ran 100 simulations for each unique combination of **quorum_thre** and **pop_coeff**.



(a) (b)

(c) (d)

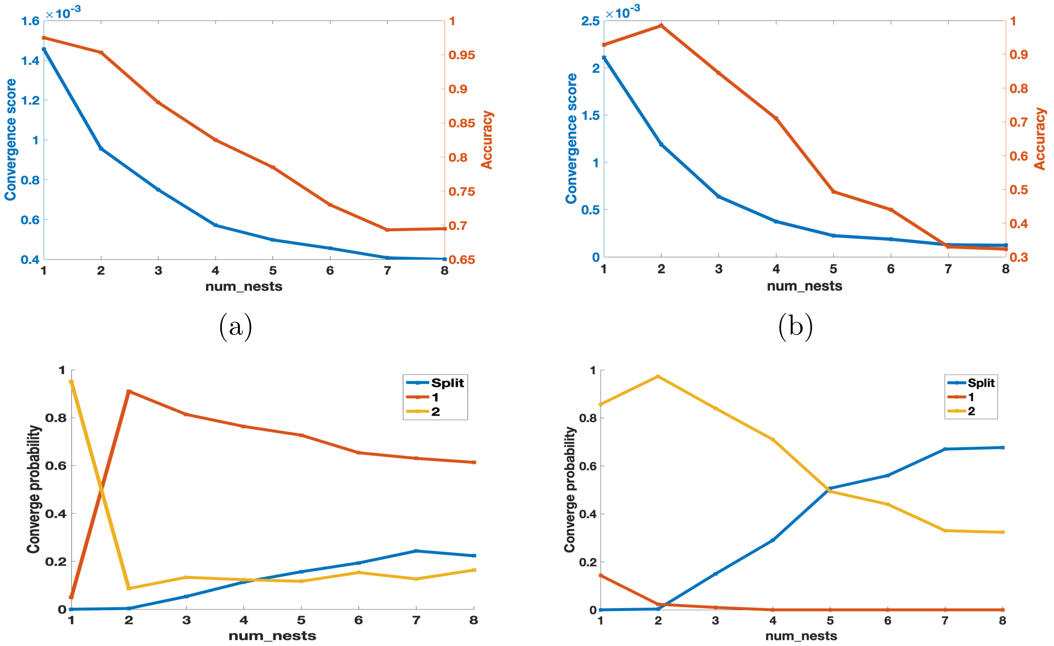
Figure 2-11: (a), (b): Convergence score and Accuracy as a function of **quorum_thre**, with **pop_coeff** = 0 and 0.35 respectively. (c), (d): Probabilities of converging to each nest (or splitting) as a function of **quorum_thre**, with **pop_coeff** = 0 and 0.35 respectively.

Results The simulation results show that the convergence score generally has a reverse-U shape that peaks at **quorum_thre** = 0.24-0.27 (Fig. 2-11a, Fig. 2-11b). In addition, the accuracy measure has a similar shape, but peaks roughly at **quorum_thre** = 0.1-0.15. The split rate, in contrast, has a U-shape, with a trough around **quorum_thre** = 0.15 to 0.18 (Fig. 2-11c and 2-11d).

The above results indicate a surprising speed-accuracy trade-off in the segments where the two lines form an “X" shape in Fig. 2-11(a) and (b): the increase of **quorum_thre** is accompanied by a decrease in accuracy and an increase in speed.

This is the opposite of our findings in Section 2.5.3 and in the related experimental work [52, 14]. However, it is important to note that the current definitions of speed and accuracy differ from those used in the prior work, which defined both quantities only up to the point where the old nest is empty. The results on splitting rate could give more insight into the conflicting results - if repairing splits is costly, lowering the probability of splits by increasing the quorum would indeed significantly increase the average convergence score. But another factor is that setting the quorum too high to reach will also delay convergence. These results point to the need for better understanding of how colonies reunite after splits, as well as the costs of reunification relative to other components of the emigration.

2.8.2 Reunification after Splitting



(a) (b)

(c) (d)

Figure 2-12: Convergence scores and splitting for environments with different numbers and qualities of nests. (a) Convergence score and accuracy as a function of the number of nests with quality 1 in the environment. (b) Convergence score and accuracy as a function of the number of nests with quality 2 in the environment. (c), (d) Same environments as (a) and (b), respectively, but plotting convergence probabilities to different nests (or splitting) on the y-axis.

Finally, we touch on another aspect of the robustness of the house hunting algorithm — reunification after splitting. Experimental studies on the speed-accuracy trade-off showed that colonies often split in urgent emigrations, but they also noted that split colonies were eventually able to reunite [20, 18]. Later studies [10, 9, 11] showed that artificially divided colonies readily re-unite, using the same behavioral tools as in emigrations, but relying more on the efforts of a small group of active workers. These findings suggest that emigrations depend on a mixture of individual and colony-level decision making. In this section, we explore how well our model achieves convergence after an arbitrary division among multiple nests. What can we learn about the mechanisms that achieve re-unification?

We ran simulations in which colonies were randomly divided among 2 to 9 nests. At the start of a simulation, each ant's *location* variable in her *ExternalState* was sampled uniformly at random from all *env-choices*. We ran one set of simulations in which one nest was of quality 2 and the rest were of quality 1, and another set in which one nest was of quality 1 and the rest were of quality 2. We ran 300 simulations for each environment with a colony of size 200, consisting of 50 active workers, 50 passive workers, and 100 brood items.

Results As the number of equal quality nests increases, the reunification task becomes increasingly difficult. Additional candidate nests have a negative effect on the convergence score and accuracy of reunification even when they are significantly worse than the best nest in the environment, possibly due to more distractions during evaluations of all nests. But the marginal effect of each additional nest diminishes (Fig. 2-12). As a result, the convergence score eventually stabilizes.

However, we see that adding nests of quality 2 (highest quality in the environment) makes reunification much harder since split rate increases quickly. Intuitively speaking, having multiple nests that are the highest quality nest in the environment can greatly intensify competition among them. But this hypothesis needs additional quantitative analyses and empirical confirmation.

In these simulations we randomized the location at the start of our simulations, but not other variables in the internal and external states of individual ants. In reality, when ants are distributed among multiple nests, they most likely have a variety of values for these other variables. We further hypothesize that 1) randomizing the other variables may help with reunification, and/or 2) colonies may have mechanisms to prevent splitting to this extent during the emigration. However, further investigation is needed to test these hypotheses.

2.9 Discussion and Future Work

In this chapter, we introduced an agent-based modeling framework that can be used to formally represent a variety of distributed algorithms used by animal groups to understand the emergence of collective intelligence. We used the framework to examine the collective nest site selection process in colonies of *Temnothorax* ants. To test against existing experimental data and to make predictions, we built a convenient Python simulator (Section 2.4 and Appendix A.2) that can easily be extended to add extra features. Our model reveals several underlying mechanisms behind the flexible, noise-tolerant and efficient collective house-hunting process in ant colonies. These include the mixed applicability of the horse-race model and the tug-of-war model to the house-hunting case; the colony's emergent ability to tolerate noise and prevent splits by balancing social information and personal information; and the impact of splitting on the speed-accuracy trade-off of the emigration. Moreover, the generalizable modeling framework that we present can be used to investigate many other collective behaviors in biological systems.

The model successfully replicated published colony-level outcomes, suggesting that it accurately accounts for underlying individual behavior. It replicated the statistics on individual recruitment acts in single nest emigrations and on emigration accuracy in two-nest emigrations as well as an earlier model [53], but did so with a more concise set of individual decision rules. The new model was also able to replicate more realistic and challenging emigration contexts that have not previously been modeled. Thus, on top of matching previous studies of simple one- and two- choice environments, our model successfully and quantitatively predicts colony behavior in environments with many more choices. Additionally, with varying degrees of experimental support, our model suggests the quantitative relationships between multiple parameters and their effects on the colony's collective behavior. The most important highlight is a novel role of social information in the speed, accuracy, and the rate of splitting during emigrations. The model and the accompanying software are versatile and easy to extend to additional investigations on unexplored scenarios such as emigrations in a changing environment. However, modeled colonies proved relatively weak at resolving splits in more challenging environments (i.e., with the colony split among several similar high-quality sites). This may reflect real limitations of the ants' algorithm, but it seems more likely that the model underestimates their reunification abilities. This topic has so far received little attention; further development of this model could help to guide new experimental work.

In the rest of this section, we discuss several specific directions for future research. While our model captures many aspects of individual behavior, it leaves out some important features, including many that affect timing. These include 1) effects of the spatial distribution of nests, which has already gathered some interests [3], 2) effects of individual experience on recruitment probability and speed, and 3) actions that may last a variable duration such as the evaluation of a new nest. Adding these to the model would allow it to explore a broader range of colony abilities and to reveal as yet unknown components of individual behavior and how they interact with known aspects. For example, more realistic models of timing would undoubtedly affect the discovery behavior currently captured by the single variable *search_find*.

On the simulation data analysis side, there are many directions for further research. First, we note the link between the effects of different quorum sizes and the horse-race and tug-of-war models that have been successfully used to describe group and individual decision-making, respectively, in these ants [31]. Our model finds that group decision-making may be better captured by the tug-of-war model when a colony is choosing between two very similar options. If so, this suggests that colonies can change their relative reliance on individual decision-making according to the decision context. This indicates the value of developing a more quantitative model that combines the tug-of-war and the horse-race models, based on the same factors that affect how a colony chooses the most beneficial quorum size.

Additionally, our model shows the potential utility of individual ants taking account of site population when assessing a site's quality. Whether real ants use social information in this way has not yet been experimentally tested. Our results suggest that it may be important for preventing and repairing split decisions. However, the amount of social information that individuals should rely on is an intricate balance, as we described in Section 2.7.1. It would be highly valuable to quantify the relationship between the frequency and degree of splitting to the quorum size and to **pop_coeff**. A related research direction is to find other factors that allow colonies to robustly reunify in split cases. However, the runtime of our program over hundreds of simulations can be substantial, making it difficult to investigate the system dynamics and performance in all possible parameter settings. Overcoming this challenge may require software optimization techniques such as code parallelization and possibly further model simplifications.

On the theoretical analysis side, our model serves as a stepping stone for more rigorous mathematical formulations and proofs of guaranteed bounds on any metrics of interest. Starting with simpler environments, our model can be reduced to analytically derive the goals different mechanisms can and cannot achieve. These results can then potentially provide insights on why certain collective behaviors have emerged through evolution, as well as on engineering artificial distributed systems subject to similar limitations to reach consensus. In fact, Chapter 3 is one such analytical study. Compared to the model in this chapter, the model we analyze in Chapter 3 focuses only on the marginal benefits of the quorum sensing mechanism. Therefore, in Chapter 3 we were able to further reduce the number of states and actions, while only preserving those that are necessary to capture the population dynamics that can trigger the quorum sensing mechanism.

Finally, we emphasize that our modeling framework can be flexibly adapted to other distributed algorithms inspired by animal groups. One compelling example is that honeybee colonies use a very similar algorithm in their nest-selection process [32], and can be easily modeled by our framework. Comparing it to our ant colony model can reveal commonalities and differences in how different animal groups achieve various goals and organize potentially conflicting priorities.

Chapter 3

Quorum Sensing Helps Consensus: An Analytical Study

3.1 Introduction

In Chapter 2, we saw that *Temnothorax* ant colonies have many biological constraints: individuals with limited memory and computational power, limited communication, and no central control. Despite so, colonies as a whole can reach various global goals such as nest-site selections and foraging. Their remarkable collective intelligence is not only an interesting problem for biologists, but also inspiring for the computer science community. In particular, from the distributed computing perspective, the collective house-hunting behavior is closely related to the fundamental problem of consensus. Building a theoretical understanding of the key mechanisms in the house-hunting process can thus shed light on the designs of novel distributed consensus algorithms. In order to do so, in this chapter we simplify the model introduced in Section 2.3 and investigate the marginal benefits of the quorum sensing mechanism. The results and theoretical tools used in this chapter are especially meaningful because previous work is mostly experimental or modeling work, and lacks rigorous mathematical justification that we task ourselves with in this chapter.

First we recall the emigration process. Colonies consist of active ants who move the remaining passive workers, the queen, and brood items (immature ants) [53, 15].

At the beginning of the emigration, each individual active ant independently searches for new nest sites. If she finds one, she evaluates the site's quality according to various metrics [29, 22]. Quality evaluation is relative to the old home nest [12]. If she is not satisfied with the site, she keeps searching. Otherwise, if she is satisfied with the site, she returns to the home nest after some time interval that is inversely related to the new nest site quality; during this interval she might continue searching for other new potential nest sites [35, 46]. If she returns to the old nest, she recruits another active ant to the site by leading a slow *tandem runs* from the old nest to the new site [40, 55]. Upon arriving at the nest, the follower ant also evaluates the nest's quality independently of the leader ant. Both the leader and the follower then continues monitoring the quality of the nest and repeats the process of quality estimation, wait interval/continued search, and further recruitment [57].

A leader ant continues leading tandem runs until she perceives that the new nest's population has exceeded a threshold, or quorum [48]. At this point, she ceases tandem runs and instead picks up and carries nest-mates from the old to the new nest. Once she starts transporting, she is considered a "committed" ant to the new nest. These transports are much faster than tandem runs, and they are largely directed at the passive workers and brood items, hence they serve to quickly move the entire colony to the new nest [46, 53]. The transporter rarely drops out of transporting other ants, and hence is considered fully committed to the new nest as the colony's home [57].

Both tandem runs and transports are forms of recruitment to accelerate the emigration process, but the marginal benefits of transports remain relatively poorly understood. Previous studies have attributed the quorum sensing mechanism as a way to tune the speed-accuracy trade-off [25, 37, 36, 44, 52, 65], generally equating accuracy with consensus or cohesion [18, 20]. As a result, emigrations that end in splits are much less investigated [52, 14], where the colony splits among multiple sites that all reach quorum. Although reunification is possible, it generally requires a second emigration process, incurring additional costs in time and risk of exposure [9, 10, 48]. Therefore, the quorum size has to be carefully calibrated in order to avoid the risk of splitting. However, apart from its role in tuning the speed-accuracy trade-off, whether the quorum sensing mechanism helps with reaching consensus remains an open question. Therefore, in this chapter, we test the hypothesis that another evolutionary advantage of the quorum sensing mechanism is that it helps emigrations reach consensus, when the quorum size take on desirable values.

To test this hypothesis, we conducted a theoretical analysis of the marginal benefits of the quorum sensing mechanism and the effects of different quorum sizes. The emigrations that serve as the control group follow a model that does not have the quorum sensing mechanism (Section 3.4), and those that serve as experimental group follow a model that incorporates the quorum sensing mechanism with a quorum size (Section 3.5). We compare the two sets of emigrations in terms of their ability to reach consensus. Our results show that with high probability, emigrations without quorums cannot reach consensus. In addition, *average case emigrations* that do use quorums can reach consensus when the quorum size take specific values. However, we note that *average case emigrations* do not represent *all* emigrations. Therefore, to fully test the hypothesis, our model and analyses in Section 3.5 has to be extended to a wider range of emigrations. Overall, our results confirm the hypothesis, and suggest useful mathematical methodologies for future work and for analyzing other agent-based algorithms.

Chapter Organization In Section 3.2, we present our simplified model of individual ants, of the entire colony, and of an execution. In Section 3.3, we formally state the definition of consensus, and the metrics to measure a model's performance in terms of consensus. In Section 3.4, we consider the performance of individual models without quorum sensing, i.e., the only form of recruitment is tandem running. We show that with high probability, emigrations cannot eventually reach consensus without quorum sensing. In Section 3.5, with the addition of the quorum sensing mechanism, we focus on characterizing the effects of different quorum sizes in the *average* case emigration. Specifically, we show a necessary condition and a sufficient condition for consensus, both conditions being values of the quorum size. The desirable values of the quorum size has also been observed in empirical studies. However, as noted before, these results for the average case require further developments to compare to empirical results as well as to fully test our hypothesis on the evolutionary advantage of the quorum sensing mechanism. Lastly, in Section 3.6 we note that our results provide theoretical evidence of a hypothesized evolutionary advantage of using the quorum rule to reach consensus by *Temnothorax* ant colonies in their house-hunting process. The analysis technique can be applied to other social insects such as honeybees [5, 33]. Additionally, our results suggest the application of the quorum rule in other artificially noisy, dynamic and distributed systems such as robot swarms.

3.2 Model

Under a discrete timing and execution model, we represent individual active ants as coupled random processes. Due to the noisy nature of the distributed individual decisions and the dynamical dependency introduced by tandem runs and transports, the transitions of an ant are non-Markovian. To accommodate this challenge, we consider the dynamics of the entire ant colony which turn out to be a Markov chain. Analyzing this Markov chain is highly non-trivial because the state is quite involved and the state space is huge; as can be seen later, the state space of this Markov chain contains all the possible partitions of the active ants into three groups, with each group representing one nest as the state of an individual ant.

3.2.1 Timing Model and the Environment

We divide time into discrete rounds. Individual active ants are modeled as identical probabilistic finite state machines and their dynamics are coupled through recruitment actions, as described later in Section 3.2.2. Let N denote the total number of active ants in the colony. Note that brood items can only be transported and have no states, and are thus not considered individual agents. Each ant starts a round with its own state. During each round, ants can perform various state transitions and have new states, before all entering the next round at the same time. Throughout this chapter, the state of an ant at round t refers to her state at the end of round t.

The environment contains the original home nest n_0 and two new nests n_1 and n_2. The new nests n_1 and n_2 have qualities q_1 and q_2 respectively, *relative* to the home nest quality. For convenience of our analysis, we let $0 < q_2 < q_1 \leq 1$, where a higher value corresponds to a better nest. Each nest is also associated with a *population* that changes from round to round. We use $x_0(t)N$, $x_1(t)N$ and $x_2(t)N$, where $x_0(t)+x_1(t)+x_2(t) = 1$, to denote active ant populations in nest n_0, n_1 and n_2 respectively at the beginning of round t. Initially, individual ants have no information on q_1 and q_2.

3.2.2 Model of Individual Ants

Individual active ants are modeled as identical probabilistic finite state machines that interact with each other. In Fig. 3-1 below, we describe the dynamics of an individual active ant without quorums (a.k.a. without performing state transitions based on seeing a quorum). Note that these dynamics are also accurate before she sees a quorum for the first time at either nest. Though these dynamics are not Markovian as the state transition of an ant is influenced by other ants during tandem runs and transports, in Section 3.4 we prove that after a finite time, the state transitions of different ants become independent. Then, in Fig. 3-2, we show the dynamics of an individual active ant after she sees a quorum for the first time either at n_1 or n_2. With a quorum size, an active ant starts her transitions according to Fig. 3-1 before she sees a quorum at either nest for the first time, and afterwards she performs her transitions according to Fig. 3-2.

Let $\mathcal{S} = \{n_0, n_1, n_2\}$ denote the collection of possible states of an active ant. Each state n_i refers to the ant being at nest n_i, and thus throughout the chapter we use "in state n_i" and "in nest n_i" interchangeably. Denote the state of ant a_i at the end of time t as $\boldsymbol{s}_i(t)$ with $\boldsymbol{s}_i(0) = n_0$ for all a_i, i.e., initially all ants locate at the home nest n_0.

An active ant in n_0 can get recruited by either following a tandem run or getting transported to either n_1 or n_2 (shown in Fig. 3-1 transitioning from n_0 to n_1 or n_2, Eq. (3)-(4)). Recruitment takes priority over her performing a probabilistic state transition to either n_1 or n_2 through independent discovery. In the model shown in Fig. 3-1, for each ant a_i and round t, if she is in n_0, let $TR_1(t), TR_2(t)$ be the event that any ant a_i gets recruited to n_1 and n_2 respectively during round t. Let $\tau_1(t), \tau_2(t)$ represent their respective probabilities during round t. Note that for any ant a_i, the two events are mutually exclusive, and $\tau_1(t) + \tau_2(t) \leq 1$.

Also included in the transitions from n_0 to n_1 or n_2 in Fig. 3-1 (Eq. (3)-(4)) is any ant a_i's independent discovery of either new nest. Specifically, conditioned on not getting recruited, an arbitrary ant a_i at n_0 discovers nest n_1 or n_2 for the first time through independent discovery with probability α for either nest and a total discovery probability of 2α. Note that the biological meaning of the parameter α is that it encodes the home nest quality - the higher the home nest quality, the less likely a_i is to search for a new nest during any round t and the smaller α is.

Thus, if an ant a_i is at state n_0 at time t, we denote the probability of her transitioning to n_1 in the nest round with $\alpha_1(t)$, which sums up the probability of her getting recruited to n_1 ($\tau_1(t)$) and the probability of independent discovery of n_1 in the case that she does not get recruited to either n_1 or n_2 ($\alpha(1 - \tau_1(t) - \tau_2(t))$, as shown in Fig. 3-1 and Eq. (3). Similarly, we define $\alpha_2(t)$ as the probability of her transitioning to n_2 during round $t + 1$ (Eq. (4)).

Each time an active ant a_i in either state n_1 or n_2 chooses to stay in the same state, she tries to recruit another active ant from state n_0 through either tandem run if the quorum is not reached ((Fig. 3-1, Eq. (6) and (8)), or through transports otherwise (Fig. 3-2, Eq. (10)). Whether the recruitment is successful or not has no effect on a_i's own state transitions during this round. Otherwise, if she does not recruit, she searches her environment and discovers the new nest she is not currently at (Eq. (7) and (9)).

In addition, the parameter $\lambda > 0$ represents the noise level of individual decision making to evaluate the quality of a nest n_k for $k \in \{1, 2\}$. Note that this evaluation only happens every time that a_i is in n_k before she detects quorum attainment at n_k: a larger λ means a higher probability (represented by u_1, u_2 to nests n_1, n_2 respectively in Fig. 3-1) of an ant a_i remaining at n_1 or n_2 and leading a tandem run. Once she detects that a quorum is reached at a nest n_k, however, she becomes "committed" to n_k. We model an ant's commitment by disallowing her to transition out of n_k. This means an active ant a_i at n_k has to perform a transport state and stay in the n_k state at any round after n_k's quorum attainment, even if n_k's population drops below the quorum later on. As a result, once a nest reaches the quorum, it never drops out of the quorum and every ant that transitions to that nest gets "stuck" in that nest. We thus model a "committed" ant with a separate Markov chain that essentially only has one possible state, as shown in Fig. 3-2.

Also worth noting is that during any given round t, if an ant a_i at n_0 does not get recruited, her transitions are Markovian and independent (Fig. 3-1, Fig. 3-2).

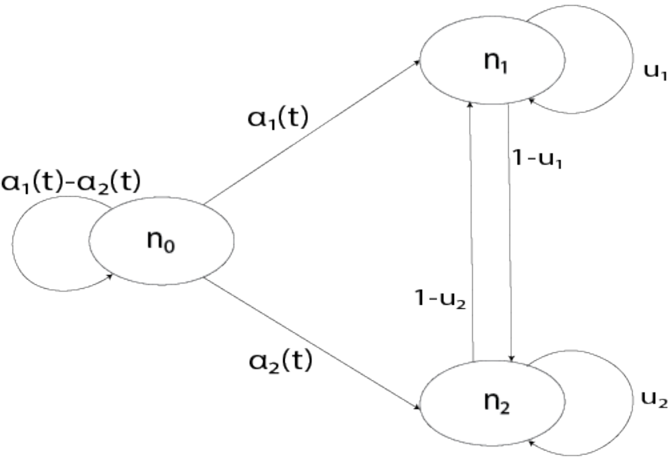


Figure 3-1: State transition diagram for individual active ants, at any round t before/without quorum attainment. $\alpha_1(t)$ and $\alpha_2(t)$ are composite functions each including the probabilities of an active ant taking different paths (independent discovery, tandem running, transports) to transition out of n_0 into n_1 and n_2, respectively.

Individual Model Without Quorums We model the transition probabilities for individual ants *without quorums*, a.k.a. without any state transitions based on seeing a quorum, as in Fig. 3-1 and in Eq. (1)-(9). Among these equations, only Eq. (1)-(5) are potentially affected by events $TR_1(t)$ and $TR_2(t)$ for any t. We do not model an ant transitioning from n_1 to n_2 or vice versa as a recruitment since recruited ants usually come from n_0 [40].

$$\mathbb{P}\{TR_1(t)\} = \tau_1(t) \tag{3.1}$$

$$\mathbb{P}\{TR_2(t)\} = \tau_2(t) \tag{3.2}$$

$$\mathbb{P}\{\boldsymbol{s}_i(t+1) = n_1 \mid \boldsymbol{s}_i(t) = n_0\} = \alpha_1(t) \tag{3.3}$$

$$\mathbb{P}\{\boldsymbol{s}_i(t+1) = n_2 \mid \boldsymbol{s}_i(t) = n_0\} = \alpha_2(t) \tag{3.4}$$

$$\mathbb{P}\{\boldsymbol{s}_i(t+1) = n_0 \mid \boldsymbol{s}_i(t) = n_0\} = 1 - \alpha_1(t) - \alpha_2(t) \tag{3.5}$$

$$\mathbb{P}\{\boldsymbol{s}_i(t+1) = n_1 \mid \boldsymbol{s}_i(t) = n_1\} = u_1 \tag{3.6}$$

$$\mathbb{P}\{\boldsymbol{s}_i(t+1) = n_2 \mid \boldsymbol{s}_i(t) = n_1\} = 1 - u_1 \tag{3.7}$$

$$\mathbb{P}\{\boldsymbol{s}_i(t+1) = n_2 \mid \boldsymbol{s}_i(t) = n_2\} = u_2 \tag{3.8}$$

$$\mathbb{P}\{\boldsymbol{s}_i(t+1) = n_1 \mid \boldsymbol{s}_i(t) = n_2\} = 1 - u_2, \tag{3.9}$$

where

$$\begin{aligned}
\alpha_1(t) &\triangleq \mathbb{P}\{TR_1(t)\} + \mathbb{P}\{\boldsymbol{s}_i(t+1) = n_1 \mid (\boldsymbol{s}_i(t) = n_0 \wedge \neg(TR_1(t) \vee TR_2(t)))\} \\
&\quad \times \mathbb{P}\{\neg(TR_1(t) \vee TR_2(t)) \mid \boldsymbol{s}_i(t) = n_0\} \\
&= \tau_1(t) + \alpha(1 - \tau_1(t) - \tau_2(t)), \\
\alpha_2(t) &\triangleq \tau_2(t) + \alpha(1 - \tau_1(t) - \tau_2(t))), \\
u_1 &\triangleq \frac{1}{1 + \exp(-\lambda q_1)}, \\
u_2 &\triangleq \frac{1}{1 + \exp(-\lambda q_2)}.
\end{aligned}$$

Recall that $\alpha \in (0, \frac{1}{2}]$ is the probability of independent discovery of n_1 (or n_2) in the case that any ant a_i does not get recruited to either n_1 or n_2 during any round t.

Formally,

$$\alpha \triangleq \mathbb{P}\left\{\boldsymbol{s}_i(t+1) = n_1 \mid (\boldsymbol{s}_i(t) = n_0 \wedge \neg(TR_1(t) \vee TR_2(t)))\right\}$$
$$= \mathbb{P}\left\{\boldsymbol{s}_i(t+1) = n_2 \mid (\boldsymbol{s}_i(t) = n_0 \wedge \neg(TR_1(t) \vee TR_2(t)))\right\}.$$

Also recall that $\tau_1(t) \in [0,1]$ and $\tau_2(t) \in [0,1]$ are constrained by $\tau_1(t)+\tau_2(t) \leq 1$. The exact expressions for $\tau_1(t)$ and $\tau_2(t)$ are not important for the proofs we have for the rest of this chapter. Note that $u_1, u_2 \in [0.5, 1]$ and $u_1 > u_2$.

Individual Model After Quorum Attainment As mentioned earlier, the state transitions of a “committed" ant (i.e. after quorum attainment at a new nest) is different and are shown by a separate Markov chain (Fig. 3-2). For any committed ant a_i, let n_k be the nest that she is committed to where $k \in \{1, 2\}$. Then the other new nest she is not committed to is n_{3-k}. The transition probability equations for individual committed ants are thus as follows:

$$\mathbb{P}\left\{\boldsymbol{s}_i(t+1) = n_k \mid \boldsymbol{s}_i(t) = n_k\right\} = 1 \tag{3.10}$$
$$\mathbb{P}\left\{\boldsymbol{s}_i(t+1) = n_{3-k} \mid \boldsymbol{s}_i(t) = n_k\right\} = 0 \tag{3.11}$$
$$\mathbb{P}\left\{\boldsymbol{s}_i(t+1) = n_k \mid \boldsymbol{s}_i(t) = n_{3-k}\right\} = 0 \tag{3.12}$$
$$\mathbb{P}\left\{\boldsymbol{s}_i(t+1) = n_{3-k} \mid \boldsymbol{s}_i(t) = n_{3-k}\right\} = 0. \tag{3.13}$$

Figure 3-2: State transition diagram for individual active ants committed to n_1 and n_2, on the left and right, respectively.

Individual Model With Quorums With a quorum size, an active ant a_i starts her transitions according to Fig. 3-1 before she sees a quorum at either nest for the first time. After she sees the population at either nest reaching the quorum, she commits to that nest and performs her transitions according to Fig. 3-2. We show the full model in Fig. 3-3.

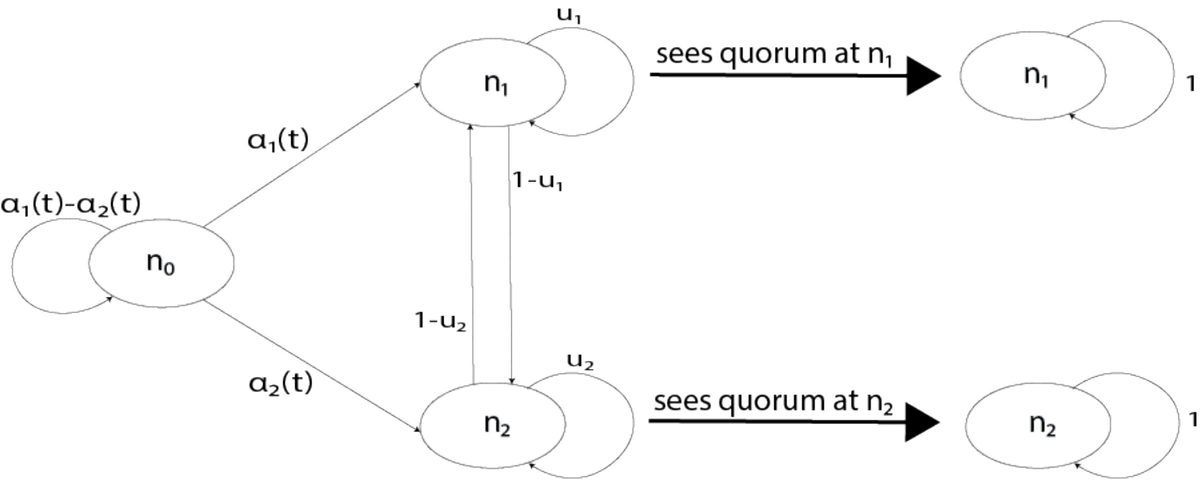


Figure 3-3: State transition diagram for individual active ants with the quorum sensing mechanism. An active ant first starts transitioning according to the left part of the figure, identical to Fig. 3-1. Then once she see a quorum at either n_1 and n_2 (but not both), she commits to that nest and can only stay in that nest, as shown on the right part of the figure, identical to Fig. 3-2.

3.2.3 Dynamics of Entire Colony

We now describe what happens in an arbitrary execution, or emigration. Throughout this chapter, we use “an execution" and “an emigration" interchangeably.

Let $\boldsymbol{s} = \{\boldsymbol{s}_1, \cdots, \boldsymbol{s}_N\}$ be the Markov chain of entire colony state, represented by a vector of dimension N that stacks the states of individual ants in the colony.

An emigration starts from round 1, at the beginning of which all ants are in nest n_0, and ants initially know the value of N but not q_1 or q_2. During each round, each ant performs one state transition in a random order.

At the beginning of a round t, each ant has her own state $\boldsymbol{s}_i(t-1)$ and the colony has state $\boldsymbol{s}(t-1)$. The value of q_1 or q_2 is available during this round to an ant a_i only if at the beginning of round t she is in nest n_1 or n_2, respectively. In this case, the population at that nest at the beginning of round t is also available to a_i. During a round t, each individual ant performs one state transition according to the individual models in Section 3.2.2, which results in a transition of the colony state as well during this round. At the end of round t, each ant has a new state $\boldsymbol{s}_i(t)$ and the colony has state $\boldsymbol{s}(t)$. All ants then enter the next round $t+1$ with their new states.

The colony state $\boldsymbol{s}(t)$ is a Markov chain and its transitions are Markovian throughout the execution. Recall that the colony state $\boldsymbol{s}(t)$ is composed of individual states $[\boldsymbol{s}_i(t)]_{i=1}^N$ of all active ants a_i at the end of round t. As shown in Fig. 3-1 and Fig. 3-2, the state $\boldsymbol{s}_i(t)$ of any individual ant a_i during round t have dependencies on 1) her own state in the previous round $\boldsymbol{s}_i(t-1)$, and 2) the states of other ants at the end of this round $[\boldsymbol{s}_j(t)]$ for any $j \neq i$, which in turn depends on the individual states of all ants in the last round $[\boldsymbol{s}_j(t-1)]_{j=1}^N$. Both of these dependencies are confined to either the colony state at round t or at round $t-1$, and not on the individual or colony states of any prior round. In other words, given the colony state at round $t-1$, $\boldsymbol{s}(t-1)$, the new colony state at round t, $\boldsymbol{s}(t)$, does not depend on any history of the colony state prior to $\boldsymbol{s}(t-1)$, or in other words, the colony state $\boldsymbol{s}(t)$ is a Markov chain.

3.3 The Consensus Problem

Here we define what it means for an emigration to reach consensus. We assume there is a *consensus error margin* $\Delta \in [0, \frac{1}{2}]$ that can be arbitrarily small. We say that an emigration has reached consensus if there exists some time t and a nest $k \in \{1, 2\}$, such that at any round from round t onward, the proportion of population at that nest n_k is greater than or equal to $1-\Delta$, i.e., $x_k(t) \geq (1-\Delta)$.

The metric to evaluate a model's performance is the *consensus probability* C, which is the probability that an emigration reaches consensus as defined above.

3.4 Consensus Without Quorum Sensing

In this section, we explore colony emigration behavior *only* with individual transition rules and tandem runs defined above (Eq. (1)-(9)) without quorum sensing. Equivalently, we consider the case where the quorum size is at least N, so that the quorum sensing mechanism never has any effect. We show an upper bound on the consensus probability C for a given Δ and colony size N. This upper bound, in turn, can be arbitrarily small if N is arbitrarily large. This motivates the desire for extra mechanisms to reach consensus with higher probabilities, which we will discuss in Section 3.5.

In this section, we analytically show that despite the fact that the emigration behaviors of individual active ants are *interactive*, the dynamics of any individual active ant are *asymptotically independent* of other ants (i.e. the only dependence is through quorum sensing), shortly after she leaves the original home nests either through discovery or through recruitment. Moreover, we show that this "asymptoticity" manifests itself in a non-trivially way after a few rounds – suggesting that a large portion of the active ants quickly rely only on individual intelligence. We show that this "asymptotic" independence is harmful to realizing social cohesion. Our results are derived with probability tools such as conditional probability, concentration bounds and Markov mixing time. Both the results on asymptotically independence and its negative impact can be extended to the general multiple-new-nest environments.

3.4.1 Model and Notation

In this section we use the model shown in Fig. 3-1 and Eq. (1)-(9). Additionally, after Proposition 6, we focus on only the part of the model defined by Eq. (6)-(9), for any ant that has transitioned out of n_0 at some point and afterwards can only be in either n_1 or n_2. For the convenience of our analysis, we show this partial model separately in Fig. 3-4.

We define the following constants:

- H, a 2×2 transition matrix of an arbitrary ant a_i's state $\boldsymbol{s}_i$ after she transitions out of n_0, as specified in Eq. (6)-(9).

- $\pi^* = \frac{1}{2-u_1-u_2}\left[1-u_2, 1-u_1\right]$, a vector representing the limiting distribution of an arbitrary ant a_i. The first element is the limiting distribution of state n_1 and the second is that of state n_2 (details in Eq. 3.16).

3.4.2 Main Theorem

Recall that Δ is the *consensus error margin*, representing the proportion of ant that can be *exempted* from the consensus requirement. We can see that the smaller Δ is (lowest value is 0), the larger $(1-\Delta)N$ is, and hence the more ants it is required for an emigration to reach consensus. In other words, the smaller Δ is, the more “strict" the consensus metric is and the more challenging it is for an emigration to reach consensus.

Theorem 1. *For any* $\Delta \in [0, 1-\pi^*(n_1)]$, *let* $\epsilon_0 = \frac{1-\pi^*(n_1)-\Delta}{2} > 0$. *Then it holds that*

$$\mathbb{P}\left\{\sum_{i=1}^{N} \mathbb{1}\{\boldsymbol{s}_i(t) = n_1\} \geq \left(\pi^*(n_1) + 2\epsilon_0\right) N = (1-\Delta)N\right\} \leq 2\exp\left(-\frac{\epsilon_0^2 N}{2}\right),$$

for any $t > \left(\frac{1}{\ln(1-2\alpha)} + \frac{1}{\ln(1-R(H))}\right) \ln\frac{\epsilon_0}{2}$, *where* $R(H)$ *is the Dobrushin's coefficient of ergodicity [28, Chapter 6.2] of* H.

Theorem 1 is stated for n_1. Similar result holds for n_2. Theorem 1 says that for any t greater than $\left(\frac{1}{\ln\beta} + \frac{1}{\ln(1-R(H))}\right)\ln\frac{\epsilon_0}{2}$, the probability of $x_1(t)$ reaching $(1-\Delta)$ is upper bounded by $2\exp\left(-\frac{\epsilon_0^2 N}{2}\right) < 1$. Thus, the total consensus probability is upper bounded by $4\exp\left(-\frac{\epsilon_0^2 N}{2}\right) < 1$.

Real ant colonies often need Δ to be very small or even zero for survival. From the theorem expression, we can see that the smaller Δ is, the lower is the upper bound of the consensus probability. Furthermore, we can see that for N arbitrarily large, C can be arbitrarily close to 0, for any given Δ. Therefore, this calls for extra mechanisms to help the emigration reach consensus, such as the quorum rule.

We also claim that the proofs in this section and related results can easily extend to environments with more than 2 nests, where one can add a new state for each new nest, each similar to the n_1 and n_2 in our model and all new nests can exchange ants and receive ants from n_0 through recruitment or discovery. After adjusting the constants of the model, H and π^*, one can derive results similar to Theorem 1: without quorum sensing, the probability of consensus can be arbitrarily low.

3.4.3 Proof of Theorem 1

Definition 2. For each $i \in [N]$, define random variable $T_i^1 \triangleq \inf\{t : \boldsymbol{s}_i(t) \neq n_0\}$ as the first round at the beginning of which ant a_i has transitioned out of the n_0 state in any arbitrary execution of the emigration.

Remark 3. It can be shown that T_i^1 is finite with probability 1 (Appendix B.1). Also from Definition 2, we have that for any ant a_i,

$$\mathbb{P}\left\{\boldsymbol{s}_i(t) = n_0 \mid t > T_i^1\right\} = 0.$$

Consider N random indicator variables $\mathbb{1}\{T_i^1 > t\}$ for any t, each variable taking values in the range $\{0, 1\}$. Using Hoeffding's inequality [30] (also Appendix B.1), we show an upper bound on the number of ants still in n_0 at round t.

Proposition 4. *For $t \geq 1$ and any number $d \in [0, 1]$, it holds that*

$$\mathbb{P}\left\{\sum_{i=1}^{N} \mathbb{1}\{T_i^1 > t\} < N\left(\beta^t + d\right)\right\} > 1 - \exp\left(-2Nd^2\right), \quad \textit{where } \beta \triangleq 1 - 2\alpha,$$

i.e., with probability at least $(1 - \exp(-2Nd^2))$, the number of active ants staying at home nest beyond time t is at most $N(\beta^t + d)$.

Proof. Notably the N indicator random variables are not independent. Hence, we prove this proposition via coupling. For each ant a_i, we introduce an auxiliary random variable $\widehat{T}_i^1$ whose value is determined by a sequence of independent coin tosses one for each round. Each state transition of ant a_i during a round t is coupled with a coin toss as follows. The coin toss happens at the end of a round t. For any of the cases below, if the coin toss shows "HEAD", set $\widehat{T}_i^1 = t$.

- If ant a_i is not in n_0 at the beginning of t and $\widehat{T}_i^1 > t-1$, then toss a biased coin whose "HEAD" probability is 2α.

- If ant a_i is in n_0 at the beginning of round t and gets recruited during round t, then toss a biased coin whose "HEAD" probability is 2α.

- Otherwise, a.k.a. if ant a_i is in n_0 at the beginning of round t and does not get recruited during round t, the conditional probability of ant a_i leaves n_0 through independent discovery is 2α during round t. If ant a_i spontaneously leaves n_0, toss a biased coin whose "HEAD" probability is 1; otherwise, toss a biased coin whose "HEAD" probability is 0.

Since initially all ants stay at home nest n_0, under the above construction, it holds that $\widehat{T}_i^1 \geq T_i^1$. Therefore, if $T_i^1 > t$, then $\widehat{T}_j^1 > t$.

Summing up the cases listed above, for each a_i and each round t, if she has not transitioned out of n_0 by the end of round $t-1$, the probability of her coin toss showing a head is thus 2α, i.e. ,

$$\mathbb{P}\left\{\widehat{T}_i^1 = t \mid \widehat{T}_i^1 \geq t\right\} = \begin{cases} 1, & \text{with probability } 2\alpha, \\ 0, & \text{otherwise.} \end{cases}$$

Moreover, let $E_{i,t} \triangleq \mathbb{1}\left\{\widehat{T}_i^1 = t \mid \widehat{T}_i^1 \geq t\right\}$. From the above construction, we know for fixed a_i, $\{E_{i,t}\}_{t=0}^{\infty}$ are independent. Also, $\{E_{i,t}\}_{t=0}^{\infty}$ are *i.i.d.* across all ants.

We further define N independent random indicator variables $\mathbb{1}\left\{\widehat{T}_i^1 > t\right\}$ for any t:

$$\begin{aligned}
\widehat{T}_i^1 > t &= \widehat{T}_i^1 > t \mid \widehat{T}_i^1 \geq t-1, \text{by definition} \\
&= \neg(\widehat{T}_i^1 = t-1 \mid \widehat{T}_i^1 \geq t-1) \\
&= \neg E_{i,t-1}.
\end{aligned}$$

Hence, for any t, the variables $\mathbb{1}\left\{\widehat{T}_i^1 > t\right\}$ for each a_i are independent, each taking values in the range $\{0, 1\}$ and takes value 1 with probability $1 - 2\alpha$.

Since $\mathbb{1}\{T_i^1 > t\} = 1$ implies $\mathbb{1}\left\{\widehat{T}_i^1 > t\right\} = 1$, it holds that

$$\sum_{i=1}^{N} \mathbb{1}\{T_i^1 > t\} \leq \sum_{i=1}^{N} \mathbb{1}\left\{\widehat{T}_i^1 > t\right\}.$$

Combined with Hoeffding's inequality [30] (also Appendix B.1), we are ready to prove the Proposition statement.

$$\begin{aligned}
&\mathbb{P}\left\{\sum_{i=1}^{N} \mathbb{1}\{T_i^1 > t\} < N\left((1-2\alpha)^t + d\right) = N\left(\beta^t + d\right)\right\} \\
&\geq \mathbb{P}\left\{\sum_{i=1}^{N} \mathbb{1}\left\{\widehat{T}_i^1 > t\right\} < N\left(\beta^t + d\right)\right\} \\
&> 1 - \exp\left(-2Nd^2\right)
\end{aligned}$$

□

Corollary 5. *For any given $\epsilon \in (0, 1)$, for any $t \geq \log_\beta(\frac{\epsilon}{2})$,*

$$\mathbb{P}\left\{\sum_{i=1}^{N} \mathbb{1}\{T_i^1 > t\} < \epsilon N\right\} > 1 - \exp\left(-N\epsilon^2/2\right),$$

where $\beta = 1 - 2\alpha$. In other words, with probability at least $(1 - \exp(-N\epsilon^2/2))$, at most ϵN ants remain in the home nest n_0 after round $\log_\beta(\frac{\epsilon}{2})$.

Proof. For any given $\epsilon \in (0, 1)$, we can set $d = \frac{\epsilon}{2}$ and apply Proposition 4 and thus proves the statement. □

Next, we show that ant a_i's state transitions become independent of other ants after T_i^1, the time that a_i leaves n_0. Note that this is different than the two types of independence proved in Proposition 4, in which we show the independence of E_i, t for any individual ant a_i across time t and also the independence of E_i, t across all ants.

These two types of independence target at dynamics of ants until they transition out of n_0. However our next Proposition shows independence across different ants after they transition out of n_0.

Proposition 6. *For every $i, j \in [N], i \neq j$ and every $t > T_i^1$, the state transitions of ant a_i are independent from those of $s_j(t)$, i.e.,*

$$\begin{aligned} &\mathbb{P}\left\{\boldsymbol{s}_i(t+1) = s_1' \mid (\boldsymbol{s}_i(t) = s_1) \wedge (\boldsymbol{s}_j(t) = s_2) \wedge (t > T_i^1)\right\} \\ &= \mathbb{P}\left\{\boldsymbol{s}_i(t+1) = s_1' \mid (\boldsymbol{s}_i(t) = s_1) \wedge (t > T_i^1)\right\}, \end{aligned}$$

where $s_1, s_2, s_1' \in \mathcal{S}$ and $s_1' \neq n_0$.

Proof of Proposition 6. For ant a_i, at the beginning of round $T_i^1 + 1$, her state is either n_1 or n_2. As can be seen from the state transition probabilities for each ant, once ant a_i jumps out of n_0, i.e. $t > T_i^1$, she can never return to n_0 according to Eq.(6)-(9).

Therefore, after T_i^1, a_i's state transition is independent of other ants, that is, the transition probabilities in Eq.(6)-(9) are not affected by the states of other ants. □

Nest, we show that every ant a_i that has transitioned out of n_0 has the same limiting distribution. Furthermore, we show that all ants eventually transition out of n_0 and thus all active ants share the same limiting distribution.

Let

$$Q(t) \triangleq \{a_i : \boldsymbol{s}_i(t) \neq n_0\} \tag{3.14}$$

be a random variable representing the set of ants that have transitioned out of n_0 by the end of round t, in an arbitrary emigration. $Q(t)$ is thus a function of an execution. It is easy to see that w.r.t. this emigration, $Q(t-1) \subseteq Q(t)$ for any $t \geq 1$. From Proposition 6 we know that for any $t \geq 1$, the state transitions of the ants in $Q(t-1)$ are independent of other ants. For any ant $a_i \in Q(t)$, her state transitions are thus modeled as in Fig. 3-4 and Eq. (6)-(9). Additionally, from Corollary 5, we know that for a large enough t, with high probability, a large portion of the active ants will have been transitioned out of n_0. In fact, it can be shown that within finite time, all ants transition out of n_0 and follow independent Markov transitions (Appendix B.1). An immediate consequence of this is that the Markov chain representing the entire colony is irreducible and aperiodic. Hence, the stationary distribution of the colony state $\boldsymbol{s}$ is well-defined and is unique. Denote this distribution by π. Note that π is a joint distributions of the states $\boldsymbol{s}_1, \cdots, \boldsymbol{s}_N$ of all ants. Let π_i be the marginal distribution of π on the state of ant a_i. We show below that $\pi_i = \pi_j$ for any $i \neq j$, and thus for ease of exposition we define $\pi^* = \pi_i$ for every i.

Lemma 7. *Let* $\pi^* \triangleq \frac{1}{2-u_1-u_2}[1-u_2, 1-u_1]$. *For any* i, $\pi_i = \pi^*$, *where* π_i *is the marginal distribution of* π *on the state of ant* a_i.

Proof. From Definition 2 and Proposition 6 we know that π_i is supported on $\{n_1, n_2\}$, and that $\pi_i = \pi_j$, where H is the transition matrix of state $\boldsymbol{s}_i$ specified in Eq.(6)-(9), as shown below.

$$H = \begin{bmatrix} u_1 & 1-u_1 \\ 1-u_2 & u_2 \end{bmatrix}. \tag{3.15}$$

Solving the equation system $\pi_i = \pi_i H$, we get

$$\pi_i = \frac{1}{2-u_1-u_2}\left[1-u_2, 1-u_1\right]. \tag{3.16}$$

Eq.(3.16) show that π_i does not depend on i, thus it is easy to see that $\pi_i = \pi_j = \pi^*$ for any ants $a_j \neq a_i$.

□

From Eq.(3.16), we derive the probability ratio

$$\frac{\pi^*(n_1)}{\pi^*(n_2)} = \exp\left(\lambda(q_1 - q_2)\right),$$

i.e., it is very sensitive to the nest quality gap $(q_1 - q_2)$ and λ.

Recall that $u_1 > u_2$ and both take values in range $[\frac{1}{2}, 1]$. The Dobrushin's coefficient of ergodicity [28, Chapter 6.2] of H is

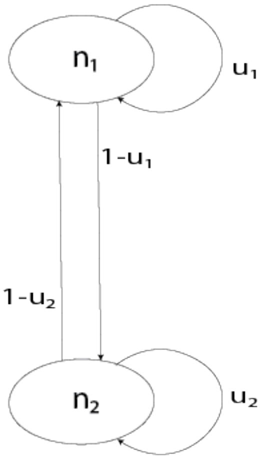


Figure 3-4: State transition diagram for individual active ants after they leave n_0, before/without quorum attainment.

$$R(H) = 2 - u_1 - u_2 \tag{3.17}$$

Next, we show that for t large enough, any ant that has transitioned out of n_0 has state distributions "close" to the stationary distribution π^*.

Lemma 8. *For any ant i, let $\pi_{i,t}$ denote the probability distribution of a_i's state over the possible states depicted in Fig.3-4 at time $t \geq T_i^1$.*

Then for any number of rounds $\ell > 0$, it holds that

$$\|\pi_{i,T_i^1+\ell} - \pi^*\|_1 \leq 2\left(1 - R(H)\right)^{\ell}.$$

Proof. First note that at the beginning of round T_i^1, the probability distribution of a_i's state is such that the probability of state $s_i(T_i^1)$ (either nest n_1 or n_2) is 1 and of the new nest state she is not at is 0. By Proposition 6 we know that after T_i^1, the dynamics of $\boldsymbol{s}_i$ is Markovian. Hence we know that $\pi_{i,T_i^1+\ell} = \pi_{i,T_i^1} H^{\ell}$.

It follows from [28, Proposition 6.5] that $\|\pi_{i,T_i^1} H^\ell - \pi^*\|_1 \leq 2\left(1 - R(H)\right)^\ell$, proving the lemma. $\square$

Using Lemma 8, the following corollary immediately follows:

Corollary 9. *Fix any* $\delta \in (0,1)$. *For any ant* a_i *and* $t > T_i^1 + \ell$, *where* $\ell \triangleq \log_{(1-R(H))} \frac{\delta}{2}$, *it holds that*

$$\|\pi_{i,t} - \pi^*\|_1 \leq \delta.$$

Combined with Corollary 5, we are now ready to prove Theorem 1.

Proof of Theorem 1. We first give the intuition and a proof sketch to show an upper bound on the probability of the active ant population at n_1 being higher than a certain number C_0, for t large enough. We break down the problem into two cases, first if by a certain round k_1 when we expect most ants to be out of n_0, the actual number of ants that have transitioned out of n_0 is low; and second is if that number is high. Intuitively speaking, k_1 should be higher if Δ is lower, i.e. the consensus has a “stricter" population requirement. Now, the first case should have a pretty low probability. The second case has a high probability. From Corollary 9 we know that after a certain amount of rounds k_2 these most of the ants that are out of n_0 will reach the limiting distribution π^*. Thus, the number of ants in n_1 should be close to $\pi^*(n_1)$. In other words, after $k_1 + k_2$ rounds the probability of n_1's population being much higher than $\pi^*(n_1)$ should be quite low. Summing up the probabilities of the first and second cases gives us an overall probability as an upper bound, proving the theorem.

For ease of exposition, let $B_i(t) = \mathbb{1}\{\boldsymbol{s}_i(t) = n_1\}$ for each $i \in [N]$ and $t \geq 0$. Let C_0 be an arbitrary positive number, $C_0 \in [0, N]$. Let $C_1 = (1 - \epsilon_0)N$. Recall that

$\beta = 1 - 2\alpha$.

$$\mathbb{P}\left\{\sum_{i=1}^{N} \mathbb{1}\{\boldsymbol{s}_i(t) = n_1\} \geq C_0\right\} = \mathbb{P}\left\{\sum_{i=1}^{N} B_i(t) \geq C_0\right\}$$
$$= \mathbb{P}\left\{\sum_{i=1}^{N} B_i(t) \geq C_0 \mid \left|Q(\log_\beta \frac{\epsilon_0}{2})\right| < C_1\right\} \mathbb{P}\left\{\left|Q(\log_\beta \frac{\epsilon_0}{2})\right| < C_1\right\}$$
$$+ \mathbb{P}\left\{\sum_{i=1}^{N} B_i(t) \geq C_0 \mid \left|Q(\log_\beta \frac{\epsilon_0}{2})\right| \geq C_1\right\} \mathbb{P}\left\{\left|Q(\log_\beta \frac{\epsilon_0}{2})\right| \geq C_1\right\}$$
$$\leq \mathbb{P}\left\{\left|Q(\log_\beta \frac{\epsilon_0}{2})\right| < C_1\right\} + \mathbb{P}\left\{\sum_{i=1}^{N} B_i(t) \geq C_0 \mid \left|Q(\log_\beta \frac{\epsilon_0}{2})\right| \geq C_1\right\}. \tag{3.18}$$

We bound the two terms in the RHS of Eq.(3.18) separately.

Bounding the 1st term: For any $t \geq \log_\beta \frac{\epsilon_0}{2}$, we have

$$\begin{aligned}\mathbb{P}\left\{|Q(t)| < C_1\right\} &= \mathbb{P}\left\{|Q(t)| < (1-\epsilon_0)N\right\} \\ &= \mathbb{P}\left\{\sum_{i=1}^{N} \mathbb{1}\{\boldsymbol{s}_i(t) \neq n_0\} < (1-\epsilon_0)N\right\} \\ &= \mathbb{P}\left\{\sum_{i=1}^{N} \mathbb{1}\left\{T_i^1 < t\right\} < (1-\epsilon_0)N\right\} \\ &= \mathbb{P}\left\{\sum_{i=1}^{N} \mathbb{1}\left\{T_i^1 > t\right\} > \epsilon_0 N\right\} \\ &\leq \exp\left(-\frac{\epsilon_0^2 N}{2}\right),\end{aligned}$$

where the last inequality follows from Corollary 5.

Bounding the 2nd term: Note that

$$\sum_{i=1}^{N} B_i(t) = \sum_{a_i \in Q(t)} B_i(t) + \sum_{a_i \notin Q(t)} B_i(t).$$

It is easy to see that

$$\sum_{a_i \notin Q(t)} B_i(t) = 0. \tag{3.19}$$

In addition, we have

$$\begin{aligned}
&\mathbb{P}\left\{ \sum_{a_i \in Q(t)} B_i(t) - \sum_{a_i \in Q(t)} \mathbb{E}\left[B_i(t)\right] \geq \epsilon_0 \left|Q(t)\right| \quad \mid \quad \left|Q(t)\right| \geq (1-\epsilon_0)N \right\} \\
&= \mathbb{P}\left\{ \sum_{a_i \in Q(t)} B_i(t) - \sum_{a_i \in Q(t)} \pi_{i,t}(n_1) \geq \epsilon_0 \left|Q(t)\right| \quad \mid \quad \left|Q(t)\right| \geq (1-\epsilon_0)N \right\} \\
&\leq \exp\left(-2\left|Q(t)\right|\epsilon_0^2\right) \\
&\leq \exp\left(-2(1-\epsilon_0)\epsilon_0^2 N\right).
\end{aligned}$$

Conditioning on $\left|Q(\log_\beta \frac{\epsilon_0}{2}))\right| \geq (1-\epsilon_0)N$, from Corollary 9, we know that for each $a_i \in Q(\log_\beta \frac{\epsilon_0}{2})$, for any $t > \log_\beta \frac{\epsilon_0}{2} + \ell$, where $\ell = \log_{(1-R(H))} \frac{\epsilon_0}{2}$, it holds that $\pi_{i,t}(n_1) \leq \pi^*(n_1) + \epsilon_0$. Hence we get

$$\begin{aligned}
\sum_{a_i \in Q(t)} \pi_{i,t}(n_1) + \epsilon_0 \left|Q(t)\right| &\leq \left(\pi^*(n_1) + \epsilon_0\right)\left|Q(t)\right| + \epsilon_0\left|Q(t)\right| \\
&\leq \left(\pi^*(n_1) + 2\epsilon_0\right) N.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\mathbb{P}\left\{ \sum_{a_i \in Q(t)} B_i(t) \geq \left(\pi^*(n_1) + 2\epsilon_0\right) N \right\} \\
&\leq \mathbb{P}\left\{ \sum_{a_i \in Q(t)} B_i(t) - \sum_{a_i \in Q(t)} \mathbb{E}\left[B_i(t)\right] \geq \epsilon_0 \left|Q(t)\right| \quad \mid \quad \left|Q(t)\right| \geq (1-\epsilon_0)N \right\} \\
&\leq \exp\left(-2(1-\epsilon_0)\epsilon_0^2 N\right).
\end{aligned} \tag{3.20}$$

Combining Eq.(3.19) and (3.20), we conclude that

$$\mathbb{P}\left\{\sum_{i=1}^{N} B_i(t) \geq \left(\pi^*(n_1) + 2\epsilon_0\right) N \mid |Q(t)| \geq (1-\epsilon_0)N\right\} \leq \exp\left(-2(1-\epsilon_0)\epsilon_0^2 N\right).$$

Combining the probability bounds on the first and second terms of Thm.(1), we have

$$\begin{aligned}\mathbb{P}\left\{\sum_{i=1}^{N} \mathbb{1}\{\boldsymbol{s}_i(t) = n_1\} \geq \left(\pi^*(n_1) + 2\epsilon_0\right) N\right\} &\leq \exp\left(-2(1-\epsilon_0)\epsilon_0^2 N\right) + \exp\left(-\frac{\epsilon_0^2 N}{2}\right) \\ &\leq 2\exp\left(-\frac{\epsilon_0^2 N}{2}\right) \quad \text{as } \epsilon_0 \in (0, 1/2),\end{aligned}$$

proving Theorem 1.

□

3.5 Consensus with Quorum Sensing

In this section, we analyze the addition of the quorum sensing mechanism to an average case emigration, which is defined as any execution that follows the population dynamics described in Eq.(21)-(26), i.e., the average case dynamics. The distribution of the active ants population among the three nests changes deterministically over time in the average case as defined in Section 3.5.1. We discuss how different quorum sizes influence whether an average emigration reaches consensus. Let $QS \times N$ denote the quorum size: $QS \in [0, 0.5]$.

Recall that the definition of consensus for an emigration allows an error margin denoted by $\Delta \geq 0$, representing the proportion of ants *exempted* from the consensus requirement. For example, when $\Delta = 0$, i.e., no ant is exempted from the consensus requirement, *all* ants have to be in one nest from some time t onward for an emigration to reach consensus. As discussed in Section 3.4, without quorum sensing, for any N, a smaller Δ means a smaller upper bound on the consensus probability C, and thus a higher need of extra mechanisms to reach consensus. For simplicity, our analysis in this section thus assumes that $\Delta = 0$.

Our main results are four values that are important for consensus: 1) a **necessary** condition for consensus - a value $QS_- \leq 0.5$ such that for $QS \leq \max\{QS_-, \alpha\}$, an average case does not reach consensus, and 2) a **sufficient** condition for consensus - a value QS_+ such that for $QS \in [0, 0.5] \wedge QS > \max\{QS_+, \alpha\}$, an average case emigration does reach consensus. However, both conditions have complicated expressions and use constants whose numeric value we cannot directly evaluate. To provide more interpretable results, we also derive: 3) a corollary as a necessary condition - a value $QS'_- \leq 0.5$ whose expression only uses model parameters, such that $QS > \max\left\{QS'_-, \alpha\right\}$ is a necessary condition for consensus, and 4) a strong conjecture suggested by simulation results as a sufficient condition - a value QS'_+ whose expression only uses model parameters, such that $QS \in [0, 0.5] \wedge QS > \max\left\{QS'_+, \alpha\right\}$ is a sufficient condition for consensus. These two bounds on QS aren't as tight as those in the first two results, but they have simpler forms.

Studying the average case dynamics is meaningful because similar results and/or proof techniques may apply to population dynamics that are “close" to the average case dynamics. Although we do not yet have a clear metric of “closeness" between two population dynamics, we hope that our insights on the average case serves as an important milestone towards quantifying the relationship between the quorum size QS and the consensus probability C under this model where individuals use the quorum rule.

3.5.1 Model with Quorums

First, we revisit the individual model with quorums (Fig. 3-3) that we defined in Section 3.2.2. Then, in order to make our model more bio-plausible, we describe an execution model which is a modification of the execution model described in Section 3.2.3. This modification enforces that if a nest reaches quorum for the first time at some point in the middle of a round, any ant that in that nest that has not transitioned yet also sees the quorum, and transitions according to Fig. 3-2 instead of Fig. 3-1. This is more bio-plausible because a real ant does not wait until the an arbitrary end of round time to evaluate the nest population again, but instead her evaluation can happen at any time when she is in the nest through encounter rates. We also define a special case using this modified execution model to define an important concept, *simultaneous split*, which happens when two nests reach consensus in the same round, as opposed to a "separate" split where nests reach consensus in different rounds. Following that, we describe informally the definition of the average case dynamics. Then, we introduce the recruitment dynamics followed by the full population dynamics that formally define the average case dynamics.

Individual Model with Quorums

Each individual ant is modeled by Fig. 3-3, which uses both Fig. 3-1 and Fig. 3-2, divided by her seeing a quorum at either new nest.

We determine that a nest n_k has "reached quorum" or "hit quorum" at round t if by the end of round t, $x_k(t) \geq QS$. Recall that we model the quorum rule at n_k by disallowing any active ant in state n_k to transition out of n_k (Eq. (10)-(13)) once n_k hits quorum. Active ants in n_k also stop leading tandem runs from n_0 but still stay in the state n_k, and start transporting passive and brood items [46]. That is, once a nest n_k hits the quorum, it never drops below the quorum. This also implies that a *split* where the populations at both nests are at least QS is irreversible, and results in the emigration failing to reach consensus.

Execution Model with Quorums

In order to make our execution model more bio-plausible, we apply a modification to the execution model introduced in Section 3.2.3, specifically *only* the rounds in which either n_1 or n_2 hits quorum for the first time. Recall that as described in Section 3.2.3, during a round, individual transitions happen in a random order, and each ant performs exactly one transition. However, for the execution model with quorums, for any round t at which a nest hits quorum QS *for the first time*, we only allow a proportion of the individual state transitions that would have happened according to Section 3.2.3 to be successful. Specifically, let $x'_k(t) \geq QS$ ($k \in \{1, 2\}$) represent the population at n_k at the end of round t in the execution model in Section 3.2.3, where n_k hits quorum for the first time at round t. Let $m = \frac{QS - x_k(t-1)}{x'_k(t) - x_k(t-1)}$ denote the proportion of each individual transitions in Eq. (1)-(9) that are allowed to happen in this round. For each of the state transition equations, exactly which ants are picked to be in the successful proportion is random. We then halt the execution in the middle of round t exactly after this successful proportion of individual transitions to happen. Any individual state transitions that would have happened in this round after the halt do not happen. That is, some ants do not get to perform a transition during this round.

If both nests hit quorum for the first time in the same round t, the n_k used for the successful proportion m is the nest with the higher $x'_k(t)$. In the modified model where only a proportion of ants join either nests in round t, $x_k(t)$ would be exactly QS, but it is still possible for the other nest to hit quorum during this round t. In these cases we say that the emigration has resulted in a **simultaneous split** and therefore failed to reach consensus.

An ant a_i's state at the end of this round, $\boldsymbol{s}_i(t)$, still depends only on $\boldsymbol{s}_i(t-1)$ and $\boldsymbol{s}_j(t)$ for any $j \neq i$ which in turn depends on $[\boldsymbol{s}_j(t-1)]_{j=1}^{N}$, both still confined to the colony state at either round t or round $t-1$, and not on the individual or colony states of any prior round. Thus the Markov properties of the entire colony still hold.

The purpose of this execution model is to enforce that by end of round t, $x_k(t)$ is exactly equal to QS, where x_k ($k \in \{1, 2\}$) is the nest that hits quorum for the first time at t.

The Average Case Dynamics

Below we define the **average case dynamics**. We start with an intuitive overview of the average case dynamics during any t that is not the first round that a nest hits quorum, followed by the dynamics during a t that is. Then, more formally, we quantify the recruitment dynamics at any round t. Finally, we quantify the full dynamics during the execution at any round t. An average case emigration is thus any execution that follows the population dynamics defined in Eq. (21)-(26).

Recall that for any execution, during each round t, the colony state $\boldsymbol{s}$ has a new state due to the state transitions of individual ants during t, where individual transitions are carried out in a random order within the round. The population partition among the three states shown in Fig. 3-1 at the end of round t is $\{x_0(t)N, x_1(t)N, x_2(t)N\}$. In the average case dynamics, during every round t, the proportion of ants transitioning between the three different states are as specified in Eq. (1)-(9) for an uncommitted ant, and Eq. (10)-(13) for committed ants. For example, if by the end of round t, n_1's population $x_1(t)$ has not reached the quorum, then during round $t+1$ the number of ants transitioning from n_1 to n_2 is $(1-u_1)x_1(t)$ (Eq. (7)), and the number of ants that perform tandem runs towards n_1 is thus $u_1x_1(t)$ (Eq. (6)). If n_1 already reached quorum at round t, then all ants in n_1 transport brood items to n_1, without influencing the dynamics of active ants at n_0. The same applies to all other states and transitions.

Later we prove that in our construction, n_1 reaches quorum first, at least as early as n_2 does (Proposition 14). Intuitively, this is because $q_1 > q_2$, and thus through the tandem run positive feedback loops, on average, n_1 accumulate population faster than n_2. Let t^* be the round at the end of which n_1 hits quorum for the first time. Recall that $x'_1(t^*) \geq QS$ represent the population at n_k at the end of round t^* using the execution model *without* quorums (Section 3.2.3). Later we show in Prop. 15 that as long as $QS > \alpha$, an average case emigration does not result in a *simultaneous split.*

If a *simultaneous split* does not occur, then at round t^*, only n_1 hits quorum and n_2 has not. As described earlier, in the average case dynamics, the proportion of ants that perform transitions are only partial during round t^*: a proportion $m = \frac{QS-x_1(t^*-1)}{x'_1(t^*)-x_1(t^*-1)}$ of the number of ants that would have performed the transition listed in Eq.(1)-(9), while the rest of the ants do not perform any transitions during round t^*. Then the next round, t^*+1, starts with this “snapshot" distribution of ants among all the nests. Using the previous example, during round t^*, the number of ants transitioning from n_1 to n_2 is $(1-u_1)x_1(t^*-1)m$ (Eq. (7)), and the number of ants attempting tandem runs and returning to n_1 is thus $u_1x_1(t^*-1)m$ (Eq. (6)). The same applies to all other states and transitions.

Recruitment Dynamics The recruitment dynamics in the average case dynamics are defined below. For transports, since they are targeted at passive ants and broods, once an active ant becomes committed to a nest, she does not transport other active ants from n_0 anymore. For tandem runs, we consider all tandem runs during a round $t+1$ to be successful as long as there are at least as many active ants as (potential) tandem run followers in n_0 as tandem run recruiters from n_1 and n_2 combined, except during round t^*.

Let $\widehat{x}_1(t+1), \widehat{x}_2(t+1)$ denote the number of tandem followers from n_0 to n_1 and n_2, respectively, during round $t+1$. Then by the end of round t,

- If $t+1 < t^*$: If neither nest has reached quorum,
  - If $x_0(t) \geq u_1x_1(t) + u_2x_2(t)$, then all tandem runs are successful during round $t+1$, i.e. $\widehat{x}_1(t+1) = u_1x_1(t), \widehat{x}_2(t+1) = u_2x_2(t)$.
  - Otherwise, $\widehat{x}_1(t+1) = x_0(t)\frac{u_1x_1(t)}{u_1x_1(t)+u_2x_2(t)}$, and $\widehat{x}_2(t+1) = x_0(t)\frac{u_2x_2(t)}{u_1x_1(t)+u_2x_2(t)}$. Then starting from the beginning of round $t+2$, n_0 is empty (has no active ants).
- If $t+1 = t^*$: let $x'_1(t^*) = u_1x_1(t) + \widehat{x}_1(t) + \alpha(x_0(t) - \widehat{x}_1(t) - \widehat{x}_2(t)) + (1-u_2)x_2(t)$, representing n_1's population at the end of t^* without the modification to the execution model. Then $\widehat{x}_k(t+1) = \widehat{x}_k(t+1)\frac{QS - x_1(t)}{x'_1(t+1) - x_1(t)}$ for $k \in \{1, 2\}$.
- If $t+1 > t^*$ and n_1 has reached quorum but n_2 has not: then only ants in n_1 would still be leading tandem runs from n_0. $\widehat{x}_1(t+1) = \min\{x_0(t), u_1x_1(t)\}$, and $\widehat{x}_2(t+1) = 0$.
- If $t+1 > t^*$ and both n_1 and n_2 have reached quorum: then no more tandem runs from n_0 to either nests during round $t+1$ (or any round afterwards). $\widehat{x}_1(t+1) = \widehat{x}_2(t+1) = 0$.

Full Dynamics Let $Y_1 \triangleq \mathbb{1}\{x_1(t) < QS\}$, and $Y_2 \triangleq \mathbb{1}\{x_2(t) < QS\}$, by the end of an arbitrary round $t \neq t^* - 1$. The following set of equations define the average case dynamics during round $t+1$.

$$x_0(t+1) = (x_0(t) - Y_1(t)\widehat{x}_1(t+1) - Y_2(t)\widehat{x}_2(t+1))(1-2\alpha) \tag{3.21}$$

$$\begin{aligned} x_1(t+1) &= u_1 x_1(t) + (1 - Y_1(t))(1-u_1)x_1(t) + Y_1(t)\widehat{x}_1(t+1) \\ &\quad + \alpha(x_0(t) - Y_1(t)\widehat{x}_1(t+1) - Y_2(t)\widehat{x}_2(t+1)) + Y_2(t)(1-u_2)x_2(t) \end{aligned} \tag{3.22}$$

$$\begin{aligned} x_2(t+1) &= u_2 x_2(t) + (1 - Y_2(t))(1-u_2)x_2(t) + Y_2(t)\widehat{x}_2(t+1) \\ &\quad + \alpha(x_0(t) - Y_1(t)\widehat{x}_1(t+1) - Y_2(t)\widehat{x}_2(t+1)) + Y_1(t)(1-u_1)x_1(t) \end{aligned} \tag{3.23}$$

For $t = t^* - 1$, the dynamics during round t^* are instead as follows. Recall that $x'_k(t^*)$ denote the proportion of ants at nest $k \in \{0, 1, 2\}$ without the execution model modification, i.e. according to Eq.(21)-(23) above for $t = t^* - 1$. Let that proportion be denoted by $m = \frac{QS - x_1(t^*-1)}{x'_1(t^*) - x_1(t^*-1)}$. Using Eq. (25)-(26) below, we show that $x'_2(t^*) \leq x'_1(t^*)$ (Appendix B.2, Proposition 26). We thus define formally that a **simultaneous split** event occurs when $x_2(t^*) \geq QS$. Note that in the average case, for any $t \neq t^*$, the population dynamics are exactly the same under the execution model with and without modification in Section 3.5.1.

$$x_0(t^*) = x_0(t^* - 1) + (x'_0(t^*) - x_0(t^* - 1))m \tag{3.24}$$

$$x_1(t^*) = x_1(t^* - 1) + (x'_1(t^*) - x_1(t^* - 1))m \tag{3.25}$$

$$x_2(t^*) = x_2(t^* - 1) + (x'_2(t^*) - x_2(t^* - 1))m, \tag{3.26}$$

where

$$\begin{aligned} x'_0(t^*) &= (x_0(t^*-1) - \widehat{x}_1(t^*-1) - \widehat{x}_2(t^*-1))(1-2\alpha) \\ x'_1(t^*) &= u_1 x_1(t^*-1) + \widehat{x}_1(t^*-1) + \alpha(x_0(t^*-1) - \widehat{x}_1(t^*-1) - \widehat{x}_2(t^*-1)) \\ &\quad + (1-u_2)x_2(t^*-1) \\ x'_2(t^*) &= u_2 x_2(t^*-1) + \widehat{x}_2(t^*-1) + \alpha(x_0(t^*-1) - \widehat{x}_1(t^*-1) - \widehat{x}_2(t^*-1)) \\ &\quad + (1-u_1)x_1(t^*-1), \end{aligned}$$

represent the populations at n_0, n_1, n_2 *if* we did not apply our modification to the execution model.

3.5.2 Notations

Below we define two concepts used in our main results: a constant t^*, and a function γ of the quorum size QS as well as model parameters α, u_1, u_2.

- t^*, an integer representing the first round at the end of which $x_1(t)$ reaches QS for the first time, i.e. $x_t(t^* - 1) < QS$ and $x_1(t^*) = QS$.
- $\gamma(QS, \alpha, u_1, u_2) \triangleq \frac{x_2(t^*)}{QS}$, where γ is a function of QS, α, u_1 and u_2.

3.5.3 Main Results

We first explain our range of consideration for QS, $[0, 0.5]$. When $QS = 0.5$, the emigration will reach consensus in finite time after all ants have moved out of n_0 (proof in Appendix B.1) and all ants have to be in either n_1 or n_2. Thus one of the two nests will have at least $0.5N$ ants, and hitting the quorum since $QS \leq 0.5$. Moreover, a *simultaneous split* is impossible when $QS = 0.5 > \alpha$ (Appendix B.2, Proposition 15). Hence, when $QS = 0.5$, only one nest will hit quorum and the emigration eventually achieves consensus. It is possible for the colony to have a quorum larger than $0.5N$ and reach consensus; it is strictly worse than having $QS = 0.5$ from a biological standpoint, since the larger the quorum size, the slower it is to reach it and hence worse for the colony's survival. Therefore we do not consider cases where $QS > 0.5$ for the purpose of the quorum attainment mechanism.

With the addition of quorum sensing with a specific quorum size QS, once the quorum is hit by a nest, that nest never drops out of quorum. With our construction and choice of $QS \leq 0.5$, as discussed in the previous paragraph, at least one nest will hit quorum. Thus, the outcome of an emigration is either that one nest hits quorum and eventually all active ants move there, or both nests hit quorum and the emigration ends in a split. Therefore, we treat the first outcome where all ants are in one nest as consensus, with a 0 error margin (hence smaller than or equal to any given $\Delta \in [0, 1 - \pi^*(n_1)]$).

We are now ready to present our results on the effects of different values of QS within the range of consideration. Our first result (Theorem 10) shows a value $QS_- \leq 0.5$ such that for $QS \leq \max\{QS_-, \alpha\}$, an average case does not reach consensus, i.e. $QS > \max\{QS_-, \alpha\}$ is a necessary condition for consensus. The α term in the expression is because in order to reach consensus, it is required that a *simultaneous split* does not occur, which we later prove is equivalent to $QS > \alpha$ (Appendix B.2, Proposition 15). However, the expression of QS_- is complicated and involves the γ function, whose numeric value is deterministic but we cannot derive directly. This makes it difficult to interpret the expression of QS_- and subsequently to compare $\max\{QS_-, \alpha\}$ to real quorum sizes observed empirically. To that end, in Corollary 11 we derive a smaller but more interpretable value, $QS'_- \leq 0.5$ that does not use the γ function in its expression, such that $QS > \max\left\{QS'_-, \alpha\right\}$ is a necessary condition for consensus. Furthermore, in Theorem 12 we show our third result: a value QS_+ such that for $QS \in [0, 0.5] \wedge QS > \max\{QS_+, \alpha\}$, an average case emigration does reach consensus, i.e. $QS \in [0, 0.5] \wedge QS > \max\{QS_+, \alpha\}$ is a sufficient condition for consensus. Last but not least, similar to the necessary conditions, in Strong Conjecture 13 we show an interpretable but larger value, QS'_+, such that $QS \in [0, 0.5] \wedge QS > \max\left\{QS'_+, \alpha\right\}$ is a sufficient condition for consensus. QS'_+ does not use the γ function in its expression. Also note that the conjecture is based on simulation results, instead of analytical results.

It is worth noting that α, the independent discovery rate of new nests, is quite low in the environments that *Temnothorax* real ant colonies usually live in, i.e. it is very challenging for individual ants to find new nests independently. Therefore, if applied to real environments, QS_-, QS'_-, QS_+, QS'_+ derived in our results below should likely be much higher than α, but here we still present our formal results in the fullest form.

Theorem 10. *For $\Delta = 0$, a necessary condition for consensus is $QS > \max\{QS_-, \alpha\}$, where*

$$QS_- = \frac{\min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}}{1 + \min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}(1+u_2) + \gamma(QS, \alpha, u_1, u2)\left(\min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\} - 1\right)},$$

i.e. an average emigration could reach consensus only if QS *is larger than* $\max\{QS_-, \alpha\}$.

Since the above expression of QS_- uses the γ function, we obtained a lower bound on γ (Lemma 19) that resulted in the following Corollary. We interpret the following results through a few examples. Let $\alpha = 0.01$ in the examples below.

- If $u_2 \geq \frac{2}{3}$ and u_1 in the range $(\frac{2}{3}, 1]$: it is necessary that $QS > \frac{1}{4}$ for an average emigration to reach consensus. This case includes many environments that are bio-plausible, and the value $\frac{1}{4}$ as the quorum size has been observed empirically in 2-nest environments [J: TODO add reference].
- If $u_2 < \frac{2}{3}$: for example, when $u_2 = 0.6$, and $u_1 = 1 > u_2$, it is necessary that $QS > 0.217$ for an average emigration to reach consensus.

Corollary 11. *For* $\Delta = 0$*, a necessary condition for consensus is* $QS > \max\left\{QS'_-, \alpha\right\}$*, where*

$$QS'_- \triangleq \begin{cases} \frac{1}{4} & \textit{if } u_2 \geq \frac{2}{3} \\ \frac{4-\frac{2}{u_2}}{2+4u_2-(\frac{1}{u_1}+1)(\frac{1}{u_2}-1)} & o.w. \end{cases},$$

i.e. an average emigration could reach consensus only if QS *is larger than* $\max\left\{QS'_-, \alpha\right\}$.

Next, we present the formal statements for our sufficient condition results.

Theorem 12. *For* $\Delta = 0$*, a sufficient condition for consensus is* $QS \in [0, 0.5] \wedge QS > \max\{QS_+, \alpha\}$*, where*

$$QS_+ = \frac{\max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}}{\frac{1}{2u_2} + \max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\} + \gamma(QS, \alpha, u_1, u_2)\left(\max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\} - 1\right)},$$

i.e. an average emigration reaches consensus if QS *is in the range of consideration and* $QS > \max\{QS_+, \alpha\}$.

Similar to Theorem 10, it is also hard to interpret Theorem 12 since it uses the γ function in its expression. To that end, we used simulation results on the full range of model parameters to show an upper bound $\frac{1-u_1}{1-u_2}$ of the γ function. Then we obtained the Strong Conjecture below. We interpret the results as follows:

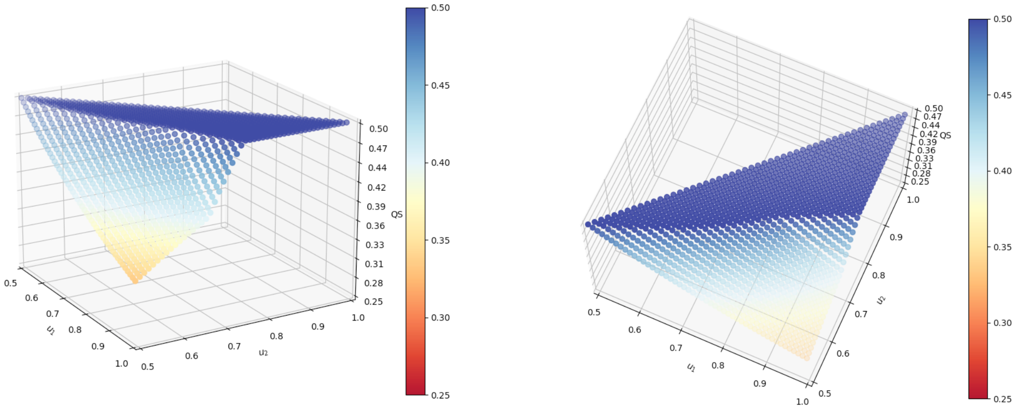
- If $\alpha \leq \frac{1}{3}$: the results are demonstrated in the simulation results shown in Fig. 3-5. For example, when $u_1 = 1, u_2 = 0.5$, it is sufficient for consensus that $QS > QS'_+ = \frac{1}{3}$. Recall that α is usually quite small in real environments, much smaller than $\frac{1}{3}$. So the results shown in Fig. 3-5 applies to a wide range of real environments.

- Otherwise: the sufficient condition for consensus is that $QS > \max\left\{QS'_+, \alpha\right\}$.

Strong Conjecture 13. *For* $\Delta = 0$*, a sufficient condition for consensus is* $QS \in [0, 0.5] \wedge QS > \max\left\{QS'_+, \alpha\right\}$*, where*

$$QS'_+ \triangleq \begin{cases} \frac{4u_2-2}{4u_2+2u_1-3} & \textit{if } u_2 \geq \frac{2}{3} \\ \frac{1}{\frac{1}{u_2}+1-\frac{1-u_1}{1-u_2}} & o.w. \end{cases},$$

i.e. an average emigration reaches consensus if QS *is in the range of consideration and larger than* $\max\left\{QS'_+, \alpha\right\}$.



(a)

(b)

Figure 3-5: 3D plots demonstrating the sufficient condition in our Strong Conjecture 13, when $\alpha \leq \frac{1}{3}$. Views from two angles.

3.5.4 Proof of Main Results

Preliminary Results

First, we give some intuitions on how we defined t^*, γ and why they are important quantities.

Proposition 14. *Let t_2 denote the round that n_2 hits quorum for the first time; let $t_2 = \infty$ if n_2 never hits quorum. In an average emigration, it holds that*

$$t_2 \geq t^*,$$

i.e., n_1 reaches quorum first, at least as early as n_2 does, if n_2 does at all.

Proof. Because we are looking at only an average case emigration and $q_1 > q_2$ and hence $u_1 > u_2$, it can be shown that up until the first round that either nest hits quorum, the population at n_1 is always larger than or equal to that at n_2, i.e. $\frac{x_2(t)}{x_1(t)} < 1$ for any $t \leq t^* - 1$ (Appendix B.2). As a result, for any given t, if $x_1(t) < QS$, then $x_2(t) < x_1(t) < QS$. In other words, if n_1 does not hit the quorum at round t, then n_2 does not either. Therefore, n_2 cannot hits quorum for the first time at round t_2 without n_1 already reaching quorum as well at a round $t^* \leq t_2$. Therefore, since at least one nest hits quorum throughout the emigration, n_1 has to hit quorum at some point t^* and n_2 may or may not hit quorum throughout the emigration. $\square$

Additionally, below we show that n_2 also reaches quorum at t^* if and only if $QS \leq \alpha$. Otherwise, if $QS > \alpha$, n_2 cannot reach the quorum at the end of round t^* (Proposition 15), but may reach it at a later round. That is, if and only if $QS > \alpha$, $x_2(t^*) < QS \Rightarrow \gamma(QS, \alpha, u_1, u_2) < 1$.

Proposition 15. *If and only if $QS > \alpha$, $\gamma(QS, \alpha, u_1, u_2) < 1$, i.e. the occurrence of the special case, simultaneous split, is equivalent to $QS \leq \alpha$.*

Proof. During round t^*, in an average emigration, the population changes at each nest is as specified in Eq. (24)-(26). In order to prove the proposition, we need to have $\frac{x_2(t^*)}{x_1(t^*)} < 1$.

If $QS \leq \alpha$, both nests will have the same number of ants that is exactly QS in round 1. Therefore, $\gamma(QS, \alpha, u_1, u_2) = 1$ and the emigration ends in a *simultaneous split.*

If $QS > \alpha$, $t = 1, x_2(t) = x_1(t) = \alpha < QS$, thus no quorum is hit in the first round. From Proposition 26, we know it holds that $\frac{x_2(t^*-1)}{x_1(t^*-1)} < 1$. Denote the population discount factor by $m = \frac{QS - x_1(t^*-1)}{x_1'(t^*) - x_1(t^*-1)}$. Proposition 26 also shows that $x_2'(t^*) < x_1'(t^*)$. Thus,

$$\begin{aligned}
\frac{x_2(t^*)}{x_1(t^*)} &= \frac{x_2(t^*-1) + (x_2'(t^*) - x_2(t^*-1))m}{x_1(t^*-1) + (x_1'(t^*) - x_1(t^*-1))m} \\
&\leq \max\left\{\frac{(1-m)x_2(t^*-1)}{(1-m)x_1(t^*-1)}, \frac{x_2'(t^*)}{x_1'(t^*)}\right\}, \text{ derived from Lemma 25} \\
&= \max\left\{\frac{x_2(t^*-1)}{x_1(t^*-1)}, \frac{x_2'(t^*)}{x_1'(t^*)}\right\} \\
&< 1.
\end{aligned}$$

This proves the proposition. □

Next, we define another important milestone time which is the time that n_0 becomes empty, in the case that $x_2(t^*) < QS$, i.e. $QS > \alpha$.

Definition 16. Let $t_0' \geq t^*$ denote the round that n_0 first becomes empty, after n_1 reaches quorum, i.e. $x_0(t_0' - 1) > 0$ and $x_0(t_0') = 0$.

Proposition 17. *In an average case emigration, it holds that* $t^* \leq t_0'$.

Proof. When n_0 becomes empty, the entire active ant population would be distributed between n_1 and n_2. We prove by contradiction that n_1 would definitely have hit the quorum by the time n_0 becomes empty. Let t be the first round that n_0 becomes empty. If by round t, n_1 does not hit quorum, then neither nests hits quorum. Since $\frac{x_2(t)}{x_1(t)} < 1$, $x_1(t) > 0.5$. Further since $QS < 0.5 < x_1(t)$, $x_1(t)$ does reach QS before t. Contradiction.

Therefore, in our model, n_1 definitely hits quorum for the first time at a round t^* before n_0 first becomes empty at time $t_0' \geq t^*$. □

From Propositions 14, 15 and 17, we know that the emigration reaches a consensus if $QS > x_2(t)$ for any $t \geq t^*$, and a split otherwise. Furthermore, since $x_2(t)$ monotonically decreases for any $t > t'_0$ (Appendix B.2, Proposition 28), we need to look only at the period $[t^*, t'_0]$ to determine whether a consensus is reached or not. Thus we can characterize the effects of a given QS in the below two cases.

- If $x_2(t)$ reaches QS in the period $[t^*, t'_0]$: the emigration ends in a **split**.
- If $x_2(t)$ never reaches QS in the period $[t^*, t'_0]$: the emigration reaches **consensus**, since after t'_0, $x_2(t)$ monotonically decreases by losing ants to n_1 during every round.

Proof of Necessary Conditions

Below we derive a value $QS_- \leq 0.5$ such that for $QS \leq \max\{QS_-, \alpha\}$, an average case does not reach consensus. Recall there are two criteria for consensus: 1) the lack of *simultaneous split*, i.e. n_2 does not reach quorum during t^*, and 2) if a *simultaneous split* does not occur, n_2 does not reach quorum after t^*. Since the first criterion is equivalent to $QS > \alpha$ (Proposition 15), in this proof we thus focus on understanding what the second criterion requires of QS.

We derive the maximum value b that n_2's (proportional) population is *guaranteed* to grow to in the period $[t^*, t'_0]$. If QS is lower than or equal to b, n_2's population is thus guaranteed to reach QS and the emigration ends in a split. In other words, it is necessary for consensus that $QS > b$. Then, combined with the *simultaneous split* case, the overall requirement on QS to avoid emigration splits is $QS > \max\{b, \alpha\}$.

Intuitively, a lower QS is generally more likely to be hit by n_2 after t^*. Therefore, the main question here is: what is the highest value that $x_2(t)$ is guaranteed to reach throughout the emigration? The necessary condition criterion is thus that QS is larger than this value.

In the case that $QS > \alpha$, it can be shown that $x_2(t)$ monotonically increases for any $t < t'_0$ and decreases for any $t \geq t'_0$ (Appendix B.2). Note that $x_2(t'_0)$ could be larger or smaller than $x_2(t'_0 - 1)$. However, in either case, it is necessary that

$x_2(t'_0 - 1) < QS$ in order to avoid splits. To that end, in Proposition 18 below we compare the value of $x_2(t'_0 - 1)$ to QS. If $t^* + 1 \leq t'_0 - 1 \Rightarrow t^* \leq t'_0 - 2$, i.e. there are at least two rounds between the first round that n_1 hits quorum and the round that n_0 first becomes empty, we investigate how $x_2(t)$ grows for $t \in [t^* + 1, t'_0 - 1]$. Otherwise, $x_2(t'_0 - 1) \leq \gamma(QS, \alpha, u_1, u2)QS < QS$ according to Proposition 15, so the necessary condition criterion is already met.

Proposition 18. *If a simultaneous split does not occur,*

$$\frac{x_2(t) - x_2(t^*)}{x_0(t^*) - x_0(t)} \in \left[\min\left\{\frac{2u_2 - 1}{u_2}, \frac{1}{2}\right\} \max\left\{\frac{2u_2 - 1}{u_2}, \frac{1}{2}\right\}\right],$$

for any $t \in [t^* + 1, t'_0 - 1]$.

Proof. We prove this by induction where the induction variable is t. Each iteration examines the value of $\frac{x_2(t)-x_2(t^*)}{x_0(t^*)-x_0(t)}$ for $t \in [t^* + 1, t'_0 - 1]$.

Base case: $t = t^* + 1$

$$\begin{aligned} \frac{x_2(t^* + 1) - x_2(t^*)}{x_0(t^*) - x_0(t^* + 1)} &= \frac{(2u_2 - 1)x_2(t^*) + \alpha(x_0(t^*) - u_2 x_2(t^*))}{u_2 x_2(t^*) + 2\alpha(x_0(t^*) - u_2 x_2(t^*))} \\ &\in \left[\min\left\{\frac{2u_2 - 1}{u_2}, \frac{1}{2}\right\}, \max\left\{\frac{2u_2 - 1}{u_2}, \frac{1}{2}\right\}\right], \end{aligned}$$

where the last step derives from Appendix B.2, Lemma 25.

Inductive Hypothesis: Assume $\frac{x_2(t)-x_2(t^*)}{x_0(t^*)-x_0(t)} \in \left[\min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}, \max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}\right]$ for some $t \in [t^* + 1, t'_0 - 2]$.

Induction Step: Show the theorem holds for round $t + 1$

$$\begin{aligned} \frac{x_2(t + 1) - x_2(t^*)}{x_0(t^*) - x_0(t + 1)} &= \frac{2u_2 x_2(t) + \alpha(x_0(t) - u_2 x_2(t)) - x_2(t^*)}{x_0(t^*) - x_0(t) + x_0(t) - x_0(t + 1)} \\ &= \frac{2u_2 x_2(t) - x_2(t^*) + \alpha(x_0(t) - u_2 x_2(t))}{x_0(t^*) - x_0(t) + u_2 x_2(t) + 2\alpha(x_0(t) - u_2 x_2(t))} \\ &\in \left[\min\left\{\frac{x_2(t) - x_2(t^*)}{x_0(t^*) - x_0(t)}, \frac{(2u_2 - 1)x_2(t)}{u_2 x_2(t)}, \frac{1}{2}\right\},\right. \\ &\qquad \left.\max\left\{\frac{x_2(t) - x_2(t^*)}{x_0(t^*) - x_0(t)}, \frac{(2u_2 - 1)x_2(t)}{u_2 x_2(t)}, \frac{1}{2}\right\}\right] \end{aligned}$$

Since $\frac{x_2(t)-x_2(t^*)}{x_0(t^*)-x_0(t)} \geq \min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}$,

$$\min\left\{\frac{x_2(t)-x_2(t^*)}{x_0(t^*)-x_0(t)}, \frac{(2u_2-1)x_2(t)}{u_2x_2(t)}, \frac{1}{2}\right\} = \min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}.$$

Similarly, since $\frac{x_2(t)-x_2(t^*)}{x_0(t^*)-x_0(t)} \leq \max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}$,

$$\max\left\{\frac{x_2(t)-x_2(t^*)}{x_0(t^*)-x_0(t)}, \frac{(2u_2-1)x_2(t)}{u_2x_2(t)}, \frac{1}{2}\right\} = \max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}.$$

Thus, the above together proves that

$$\frac{x_2(t+1)-x_2(t^*)}{x_0(t^*)-x_0(t+1)} \in \left[\min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\} \max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}\right],$$

hence proving the proposition using induction. □

Following Proposition 18, we derive a necessary bound for consensus (Theorem 10) as below.

Proof of Theorem 10. First we look at the case that a *simultaneous split* does not occur, i.e. $QS > \alpha$.

Recall that $x_2(t^*) = \gamma(QS, \alpha, u_1, u2)QS$. Because at round t'_0, n_0 becomes empty, we know that at the beginning of round t'_0, the number of potentially tandem followers must be at most equal to the number of leaders from n_2, i.e. $x_0(t'_0-1) \leq u_2x_2(t'_0-1)$. Applying Proposition 18, we obtain a lower bound for $x_2(t'_0-1)$ as below:

$$\min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\} \leq \frac{x_2(t'_0-1)-x_2(t^*)}{x_0(t^*)-x_0(t'_0-1)} \leq \frac{x_2(t'_0-1)-x_2(t^*)}{x_0(t^*)-u_2x_2(t'_0-1)}$$

Rearranging the terms, we obtain that

$$
\begin{aligned}
x_2(t'_0-1) &\geq \frac{\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}x_0(t^*)+x_2(t^*)}{1+u_2\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}}\\
&= \frac{\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}(1-(1+\gamma(QS,\alpha,u_1,u2))QS)+\gamma(QS,\alpha,u_1,u2)QS}{1+u_2\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}}
\end{aligned}
$$

Therefore, $x_2(t'_0-1)$ definitely reaches the above bound. Thus, if the above bound is bigger than or equal to QS, the emigration will definitely not reach consensus but end in a split instead. In other words, a necessary condition for consensus is:

$$
\begin{aligned}
&\frac{\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}(1-(1+\gamma(QS,\alpha,u_1,u2))QS)+\gamma(QS,\alpha,u_1,u2)QS}{1+u_2\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}} < QS\\
&\Rightarrow QS > QS_- = \frac{\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}}{1+\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}(1+u_2)+\gamma(QS,\alpha,u_1,u2)\left(\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}-1\right)}
\end{aligned}
$$

Further, we prove that the above bound is less than 0.5. Note that it always holds that $\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}\leq\frac{1}{2}$. Then,

$$
\begin{aligned}
&\frac{\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}}{1+\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}(1+u_2)+\gamma(QS,\alpha,u_1,u2)\left(\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}-1\right)} < 0.5\\
&\iff\\
&\gamma(QS,\alpha,u_1,u2) < \frac{\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}(1-u_2)-1}{\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}-1} = 1+\frac{u_2\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}}{1-\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}} \qquad \iff\\
&\gamma(QS,\alpha,u_1,u2) < 1,\ \text{since}\ \min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}\in[0,\frac{1}{2}).
\end{aligned}
$$

The last statement is from Proposition 15 in Appendix B.2. We thus finished the theorem proof.

Note that $\min\left\{\frac{2u_2-1}{u_2},\frac{1}{2}\right\}-1\leq\frac{1}{2}-1=-\frac{1}{2}<0$. Thus, the lower $\gamma(QS,\alpha,u_1,u2)$

is, the lower the value of QS_- is.

Lastly, combined with the no *simultaneous split* criterion where $QS > \alpha$, the overall requirement on QS to avoid emigration splits is $QS > \max\{\alpha, QS_-\}$. We thus finished proving the theorem. $\square$

The above expression of QS_- is quite complex. It involves the value of the γ function, whose numeric value is deterministic but we cannot derive directly. In order to obtain a cleaner and more interpretable expression, we show a lower bound on $\gamma(QS, \alpha, u_1, u_2)$.

Lemma 19. *If* $QS > \alpha$, $\gamma(QS, \alpha, u_1, u_2) > \frac{1-u_1}{2u_1}$.

Proof. Consider any $t < t^*$.

$$\begin{aligned}
&0 < x_2(t+1) < x_1(t+1) \\
&\Rightarrow 0 < 2u_2x_2(t) + \alpha(x_0(t) - u_1x_1(t) - u_2x_2(t)) + (1-u_1)x_1(t) \\
&\qquad < 2u_1x_1(t) + \alpha(x_0(t) - u_1x_1(t) - u_2x_2(t)) + (1-u_2)x_2(t) \\
&\Rightarrow 0 < 2u_2x_2(t) + (1-u_1)x_1(t) < 2u_1x_1(t) + (1-u_2)x_2(t) \\
&\Rightarrow \frac{2u_2x_2(t) + (1-u_1)x_1(t)}{2u_1x_1(t) + (1-u_2)x_2(t)} < 1
\end{aligned}$$

Thus, for any $t < t^*$

$$\begin{aligned}
\frac{x_2(t+1)}{x_1(t+1)} &= \frac{2u_2x_2(t) + \alpha(x_0(t) - u_1x_1(t) - u_2x_2(t)) + (1-u_1)x_1(t)}{2u_1x_1(t) + \alpha(x_0(t) - u_1x_1(t) - u_2x_2(t)) + (1-u_2)x_2(t)} \\
&> \min\left\{\frac{2u_2x_2(t) + (1-u_1)x_1(t)}{2u_1x_1(t) + (1-u_2)x_2(t)}, 1\right\}, \text{from Appendix B.2} \\
&= \frac{2u_2x_2(t) + (1-u_1)x_1(t)}{2u_1x_1(t) + (1-u_2)x_2(t)} \\
&= \frac{2u_2\frac{x_2(t)}{x_1(t)} + (1-u_1)}{2u_1 + (1-u_2)\frac{x_2(t)}{x_1(t)}}.
\end{aligned}$$

Define $f(x) = \frac{2u_2x + (1-u_1)}{2u_1 + (1-u_2)x} = \frac{2u_2}{1-u_2} + \frac{1-u_1-\frac{4u_1u_2}{1-u_2}}{2u_1+(1-u_2)x}$, for $x \in (0,1)$.

Since $1 - u_2 > 0$, and $1 - u_1 - \frac{4u_1u_2}{1-u_2} < 1 - 0.5 - \frac{4*0.5*0.5}{1-0.5} = -1.5 < 0$, we conclude that $f(x)$ monotonically increases for $x \in (0,1)$. Hence, since $\frac{x_2(t)}{x_1(t)} \in (0,1)$ and

$x_1(t^*) = QS$,

$$\gamma(QS, \alpha, u_1, u_2) = \frac{x_2(t^*)}{QS} = \frac{x_2(t^*)}{x_1(t^*)} > \frac{2u_2 \times 0 + (1-u_1)}{2u_1 + (1-u_2) \times 0} = \frac{1-u_1}{2u_1}$$

$\square$

With a lower bound on $\gamma(QS, \alpha, u_1, u_2)$, we can now apply it to Theorem 10 and derive a cleaner expression QS'_- analogous to QS_-. Below we show the technical proof for Corollary 11.

Proof of Corollary 11. Again we first look at the case where a *simultaneous split* does not occur.

When $u_1 > u_2 \geq \frac{2}{3}$, $\min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\} = \frac{1}{2}$. Thus, plugging in the lower bound of $\gamma(QS, \alpha, u_1, u2) = \frac{1-u_1}{2u_1}$, we obtain that

$$\begin{aligned} QS_- =& \frac{\min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}}{1 + \min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}(1+u_2) + \gamma(QS, \alpha, u_1, u2)\left(\min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\} - 1\right)} \\ &\geq \frac{2}{7 + 2u_2 - \frac{1}{2u_1}} \\ &> \frac{1}{4}, \text{ since } u_2 < u_1 < 1. \end{aligned}$$

When $u_2 < \frac{2}{3}$, $\min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\} = \frac{2u_2-1}{u_2}$. $\frac{2u_2-1}{u_2} - 1 = \frac{u_2-1}{u_2} < 0$. Similar to above, plugging in the lower bound of $\gamma(QS, \alpha, u_1, u2) = \frac{1-u_1}{2u_1}$, we obtain that

$$\begin{aligned} QS_- =& \frac{\min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}}{1 + \min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}(1+u_2) + \gamma(QS, \alpha, u_1, u2)\left(\min\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\} - 1\right)} \\ &> \frac{4 - \frac{2}{u_2}}{2 + 4u_2 - (\frac{1}{u_1} + 1)(\frac{1}{u_2} - 1)}. \end{aligned}$$

To combine the above two cases, we define that

$$QS'_{-} \triangleq \begin{cases} \frac{1}{4} & \text{if } u_2 \geq \frac{2}{3} \\ \frac{4-\frac{2}{u_2}}{2+4u_2-(\frac{1}{u_1}+1)(\frac{1}{u_2}-1)} & \text{o.w. } . \end{cases}$$

Since $QS_{-} > QS'_{-}$, it holds that if $QS \leq \max\left\{QS'_{-}, \alpha\right\}$, then $QS \leq \max\left\{QS_{-}, \alpha\right\}$ and the emigration ends in a split according to Theorem 10. Therefore, the emigration cannot reach consensus if $QS \leq \max\left\{QS'_{-}, \alpha\right\}$. In other words, $QS > \max\left\{QS'_{-}, \alpha\right\}$ is a necessary condition for consensus. We thus proved the corollary.

□

Sufficient Condition for Consensus

Below we derive a value QS_{+} such that for $QS \in [0, 0.5] \wedge QS > \max\left\{QS_{+}, \alpha\right\}$, an average case does reach consensus. Recall there are two criteria for consensus: 1) the lack of *simultaneous split*, i.e. n_2 does not reach quorum during t^*, and 2) if a *simultaneous split* does not occur, n_2 does not reach quorum after t^*. Since the first criterion is equivalent to $QS > \alpha$ (Proposition 15), in this proof we thus focus on understanding what the second criterion requires of QS.

The main question here is: what is maximum value b that $x_2(t)$ could grow to for $t \in [t^*, t'_0]$? Once we have this value b, one sufficient condition criterion is that $QS > b$. Then, combined with the *simultaneous split* case, the overall requirement on QS to reach consensus is $QS > \max\left\{b, \alpha\right\}$. Recall that this is different from the necessary condition proof logic, where we obtained the maximum value that $x_2(t)$ is *guaranteed* to grow to in the period $[t^*, t'_0]$. In the sufficient condition proof, $x_2(t)$ may or may not reach the expression of b we will derive below - b is an upper bound of $x_2(t)$.

Proof of Theorem 12. We first focus on the case where $QS > \alpha$ and there is no *simultaneous split*.

Since $x_0(t'_0 - 1) \leq u_2 x_2(t'_0 - 1)$, we obtain that

$$x_2(t'_0) \leq u_2 x_2(t'_0 - 1) + x_0(t'_0 - 1) \leq 2u_2 x_2(t'_0 - 1).$$

Thus, $\max(x_2(t))$ for any $t \in [1, t'_0]$ is equal to $\max(x_2(t'_0), x_2(t'_0 - 1)) \leq 2u_2 x_2(t'_0 - 1)$.

Now we derive an upper bound for $x_2(t'_0 - 1)$ as below:

$$\begin{aligned}
\max\left\{\frac{2u_2 - 1}{u_2}, \frac{1}{2}\right\} &\geq \frac{x_2(t'_0 - 1) - x_2(t^*)}{x_0(t^*) - x_0(t'_0 - 1)} \\
&\geq \frac{x_2(t'_0 - 1) - x_2(t^*)}{x_0(t^*)} \\
&= \frac{x_2(t'_0 - 1) - \gamma(QS, \alpha, u_1, u_2)QS}{(1 - (1 + \gamma(QS, \alpha, u_1, u_2))QS)}
\end{aligned}$$

Rearranging the terms, we obtain the below:

$$x_2(t'_0 - 1) \leq \max\left\{\frac{2u_2 - 1}{u_2}, \frac{1}{2}\right\}(1 - (1 + \gamma(QS, \alpha, u_1, u_2))QS) + \gamma(QS, \alpha, u_1, u_2)QS$$

Therefore, the maximum value that $x_2(t)$ can reach for any $t \in [1, t'_0]$ is thus upper bounded by

$$2u_2 x_2(t'_0 - 1) \leq 2u_2\left(\max\left\{\frac{2u_2 - 1}{u_2}, \frac{1}{2}\right\}(1 - (1 + \gamma(QS, \alpha, u_1, u_2))QS) + \gamma(QS, \alpha, u_1, u_2)QS\right)$$

Therefore, in order to reach consensus, it suffices that the right hand side of the above inequality is less than QS. After rearranging the terms again, we obtain the following:

$$QS > QS_+ = \frac{\max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}}{\frac{1}{2u_2} + \max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\} + \gamma(QS, \alpha, u_1, u_2)\left(\max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\} - 1\right)}.$$

Then, combined with the no *simultaneous split* criterion, i.e., $QS > \alpha$, to achieve consensus, QS has to be larger than $\max\{QS_+, \alpha\}$. Lastly, recall that QS has to fall within the range of consideration $[0, 0.5]$. Thus the final necessary condition for consensus is that $QS \in [0, 0.5] \wedge QS > \max\{QS_+, \alpha\}$. We thus finished proving the theorem. $\square$

The above expression is very complicated. Therefore we want to derive a cleaner and more interpretable result, similar to the forms in Corollary 11. Although we do not have the analytical proof right now, our simulation results show that $\gamma(QS, \alpha, u_1, u_2) \leq \frac{1-u1}{1-u2}$ for all range of the model parameters (α, u_1, u_2), and for all range of QS that satisfy the necessary condition. To that end, we obtain our Strong Conjecture 13 as below.

Proof of Strong Conjecture 13. To obtain the expressions of QS'_+ shown in Strong Conjecture 13, we first plug in the upper bound $\frac{1-u_1}{1-u_2}$ for $\gamma(QS, \alpha, u_1, u_2)$ into the bound derived in Theorem 12:

- If $u_2 \geq \frac{2}{3}$: $QS_+ = \frac{\max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}}{\frac{1}{2u_2}+\max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}+\gamma(QS,\alpha,u_1,u_2)\left(\max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}-1\right)} \leq \frac{4u_2-2}{4u_2+2u_1-3}$
- Otherwise: $QS_+ = \frac{\max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}}{\frac{1}{2u_2}+\max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}+\gamma(QS,\alpha,u_1,u_2)\left(\max\left\{\frac{2u_2-1}{u_2}, \frac{1}{2}\right\}-1\right)} \leq \frac{1}{\frac{1}{u_2}+1-\frac{1-u_1}{1-u_2}}$

To combine the above two cases, we define that

$$QS'_+ \triangleq \begin{cases} \frac{4u_2-2}{4u_2+2u_1-3} & \text{if } u_2 \geq \frac{2}{3} \\ \frac{1}{\frac{1}{u_2}+1-\frac{1-u_1}{1-u_2}} & \text{o.w. } . \end{cases}$$

Since $QS_+ \leq QS'_+$, it holds that if $QS \in [0, 0.5] \wedge QS > \max\left\{QS'_+, \alpha\right\}$, then $QS \in [0, 0.5] \wedge QS > \max\{QS_+, \alpha\}$ and the emigration ends in a consensus according to Theorem 12. Therefore, $QS \in [0, 0.5] \wedge QS > \max\left\{QS'_+, \alpha\right\}$ is a sufficient condition for consensus. $\square$

3.6 Discussion and Future Work

In this chapter, we used analytical tools to test the hypothesis that the quorum sensing mechanism observed in the collective nest site selection process by *Temnothorax* ants helps emigrations reach consensus, when the quorum size take certain desirable values. Our theoretical results support this hypothesis. We first analyze the control group: emigrations that do not use quorums. Without quorum sensing, the only form of recruitment, tandem runs, does speed up the emigration process, but our results show that emigrations have a high probability of splitting among multiple new sites, imposing significant risks to the colony's survival. We analyze a model of a two-new-nest environment, but the above result extends easily to environments with more nests. We then analyze the experimental group: emigrations that do use quorums. We investigate whether *average case emigrations* reach consensus with different values of quorum sizes. We thus derive a necessary condition and a sufficient condition for consensus for an average case emigration. Both conditions are values of the quorum size within its value range of consideration, $[0, 0.5]$. The desirable values of the quorum size show general consistency with experimental findings of the observed quorum size employed by *Temnothorax* ant colonies [24, 46], however more analysis on non-average-case emigrations are needed to fully test our hypothesis and to better compare the theoretical to observed results. Overall, our results show confirmation of our hypothesis, and provide insights into the importance of the quorum sensing mechanism in an unpredictable environment with multiple nests. Though the desirable values of the quorum size only apply to the average case emigration and dynamics, any emigration with population dynamics "close" to the average case likely share similar results. However, we do not yet have a clear metric on the closeness between two emigrations or dynamics, and as such it remains a particularly important possibility for future work.

Additionally, another future work direction is to make our model more bio-plausible. Specifically, our model does not consider the very small probability that committed ants "drop out" of the nest they are committed to, and go back to searching. Adding this into the model could make it more biologically realistic.

One more way to strengthen our theoretical results is by adding a time bound metric to our consensus problem. Our current consensus metric, the consensus probability C, only requires that at least $(1 - \Delta)N$ ants keep staying at either n_1 or n_2 after a finite number of rounds. By adding a time bound metric as well, we would be able to better characterize the consensus probability (even if lower than a given C) of an emigration by a certain time t.

Finally, our results and mathematical methodologies can extend to environments with more than 2 new nests. Applying similar methods to the general environment can give us more insights on how the number of nests and their qualities might influence the desirable values for the quorum size, with the goal to avoid splits, or to ensure consensus, or with an objective involving a specific degree or probability of consensus.

Chapter 4

Concluding Remarks

This thesis contains two main projects. The first project in Chapter 2 focuses on developing a comprehensive and biologically-plausible model that is still robust enough to enable analytical studies. This model is based on many empirical measurements, and can perform as well as its predecessor [53, 52] on known phenomenon. In addition, this model is highly tractable, and can be easily used to test new hypothesis. Among many interesting examples we surveyed in this chapter, we highlight a particularly intriguing one: a hypothesized way for ants to perceive a nest's quality integrates the peer opinion about that nest, represented by the proportion of ants at that nest at that time. We also develop a modeling framework that can be adapted to other biologically distributed algorithms. The *Temnothorax* ants house-hunting algorithm can thus be seen as one example of this modeling framework. Additionally, a recent modeling and data analysis study that extended our house-hunting model to include geospatial features [3] has also demonstrated the value of this framework and the ease of use of our simulator.

On the theoretical analysis side, the model in Chapter 2 is set up to enable analytical studies. Chapter 3 is one of these studies, first simplifying the model and then using analytical tools such as conditional probability, concentration bounds and Markov mixing time to rigorously prove our results. We test the hypothesis that appropriate quorum sizes not only tunes the speed-accuracy trade-off as shown in many previous literature and our Chapter 2, but also plays an essential role in helping emigrations reach consensus. We show theoretical results on 1) a low probability of consensus without quorum sensing, and 2) values of the quorum size that are necessary or sufficient for consensus. Our results thus shows confirmation of the main hypothesis. Additionally, another recent theoretical work on the effect of quorums in single-nest emigrations [16] also demonstrates the value of the general house-hunting model in Chapter 2.

Both projects suggest many interesting extensions, as discussed in detail in Section 2.9 and Section 3.6. Both projects highlight the elegant ways that noisy and limited individuals utilizes social information to “offset" their own imperfections. As such, these results have direct implications for both the biology community, and the math/computer science community. It is our hope that the work presented in this thesis can provide valuable tools and insights for the science community at large.

Appendix A

Supplemental Materials for Chapter 2

A.1 State Transition Details

Exploration

- An ant in at_nest_E has four possible actions. First, she can perform “no_action" and remain in the current nest. Second, she can perform “search" and go into the state $search_E$. Third, she can receive a “lead" by another ant to follow a FTR to a destination nest, $ec \in$ **env-choices**, in which case she sets *old_candidate_nest* to the value of *candidate_nest*, and sets *candidate_nest* to *ec*. Then she transitions to the state $follow_E$. Finally, she can receive a “carry" by another active worker ant to a destination nest $ec \in$ **env-choices**, in which case her *location* and *candidate_nest* are changed to *ec*, and she stays in at_nest_E.

- An ant is in the state $follow_E$ if she is in the middle of following an FTR, and has two possible actions. First, she can successfully follow the FTR to the destination nest (“follow_find") and change her *location* to her *candidate_nest*, which results in the state $arrive_E$. Otherwise, she may lose contact with her tandem leader (“get_lost"), and then enters the state $search_E$ and assigns the value of *old_candidate_nest* to *candidate_nest*.

- An ant in the state $search_E$ has three possible actions. First, she can have "no_action" and transition to at_nest_E by staying at her last known *location*. Second, she can "find" a new nest, $ec \in$ **env-choices**, in this round, assign the value of *candidate_nest* to *old_candidate_nest* and assign *ec* to both *location* and *candidate_nest*, and transition into $arrive_E$ state to evaluate it further. Third, she can receive an action, "carry", and the results are the same as receiving the "carry" action in at_nest_E.

- An ant in the state $arrive_E$ has two action options. First, she can "reject" the nest she just arrived at. She then assigns the value of *candidate_nest* to *location* and then that of *old_candidate_nest* to *candidate_nest* go into the $search_E$ state. Otherwise, if she performs "no_reject", she transitions into the state at_nest_A and assigns the value of *candidate_nest* to *location*.

Assessment

- An ant in the state at_nest_A is assessing a new nest and is currently located at that nest. From here, three actions are available. First, she can "accept" the nest if she deems it high quality, which results in her transitioning to at_nest_C. Second, she may perform "search" to get into the $search_A$ state. Third, she can receive a "lead" by another ant to follow a FTR to a destination nest, in which case she assigns the value of *candidate_nest* to *old_candidate_nest* and assigns the destination nest $ec \in$ **env-choices** to *candidate_nest*, and then she transitions to the state $follow_A$. Finally, she can receive a "carry" by another active worker ant to a destination nest $ec \in$ **env-choices**, in which case her *location* and *candidate_nest* are changed to *ec* and transitions back to at_nest_E.

- An ant in the states $follow_A$ or $search_A$ has the same options and variable changes as in $follow_E$ or $search_E$ respectively, but the resulting state subscripted with A except the "carry" action.

- An ant in $arrive_A$ state has the same options and variable changes as in $arrive_E$, but with "reject" action leading to $search_C$.

Canvassing

- An ant in at_nest_C state has three available actions. First, she can decide to "recruit" and go into $quorum\_sensing_C$ state. Second, she can decide to "search" more and result in $search_C$ state. Third, she may receive a "carry" by another active worker ant to a destination nest, in which case her *location* and $candidate\_nest$ are changed to that nest and results back to at_nest_E.
- An ant in $quorum\_sensing_C$ state is at a nest different than her home nest, and has two options. If she estimates the current nest population to be higher than the quorum threshold, she performs "quorum_met", swap the values of $home\_nest$ and $candidate\_nest$, and enters the state $transport_T$. Otherwise, she performs "quorum_not_met" and enters $lead\_forward_C$ state.
- An ant in $lead\_forward_C$ state has three actions available to her. First, she can "lead" another active worker and lead her on an FTR from the original home nest to the candidate new nest. She changes her *location* to the value of $candidate\_nest$, and enters at_nest_C state. Second, she can "get_lost" in the process if she loses contact with the follower, and enters $search_C$ state. Lastly, she can "terminate" her emigration if the termination conditions are met, namely if she has repeated attempts to call other active workers who are also in $lead\_forward_C$ state. In this case, she changes her *location* to her $home\_nest$, resets $terminate\_count$ to 0, and enters state at_nest_E.
- An ant in $search_C$ state has the same options and variable changes as in $search_E$ with the resulting state sub-scripted with C.
- An ant in $arrive_C$ state has the same options and variable changes as in $arrive_E$, but with "reject" action leading to $search_C$.

Transport

- An ant in at_nest_T state has the same options and variable changes as in at_nest_C with the resulting state sub-scripted with T, except that a "recruit" action results in $transport_T$, and that it can receive one additional action “lead", in which case she assigns the value of *candidate_nest* to *old_candidate_nest*, assigns the destination nest $ec \in$ **env-choices** to *candidate_nest*, and transitions to $follow_T$.

- An ant in $transport_T$ state has three available actions. First, she can decide to carry another ant, active, passive, or brood, to her newly committed nest. This results in her entering $reverse\_lead_T$ mode, meaning she can lead a reverse tandem run (RTR). These are tandem runs lead from the newly committed nest to the old home nest or another nest. Second, she can decide to “stop_trans" and stops her transport to go into the state $search_T$. Third, similar to the state $lead\_forward_C$, there is a “terminate" action when the termination condition is met, namely if she has repeated attempts to carry other active workers who are also in $transport_T$ state. In this case, she changes her *location* to her *home_nest*, resets *terminate_count* to 0, and enters state at_nest_E.

- An ant in $reverse\_lead_T$ only has two actions as her options. First, she may perform *no_action* and returns to at_nest_T state. Second, she may experience “delay" in her tandem runs, and will stay in $reverse\_lead_T$ state. There's no conclusion on the purpose of RTRs at this point in the research community, so we model it as a round-trip from an agent's candidate nest to the original home nest and back, eventually ending up with no state changes.

- An ant in the states $follow_T$ or $search_T$ has the same options and variable changes as in $follow_E$ or $search_E$ respectively, but the resulting state are subscripted with T except the “carry" action.

- An ant in $arrive_T$ state has the same options and variable changes as in $arrive_A$, but with “reject" action leading to $transport_T$.

A.2 Simulation Details

A.2.1 Sample Configuration File

```
[ENVIRONMENT]
num_ants = 200
nest_qualities = 0,1,2

[ALGO]
lambda_sigmoid = 8
pop_coeff = 0.35
quorum_thre = 0.15
quorum_offset = 0
search_find = 0.005
follow_find = 0.9
lead_forward = 0.6
transport = 0.7

[SETTINGS]
plot = 0
total_runs_per_setup = 500
num_rounds = 4000
percent_conv = 0.9
persist_rounds = 200
```

A.2.2 Data Structures and Global Variables

We define four native data structures, as shown in Fig. 2-1. The global variables include 1) the transition tables defined in Fig. 2-3, 2) *Nests*, the array of all nests including the home nest which by default has quality 0 and id 0, and 3) *Ants*, the array of all ants in the colony.

A.2.3 Simulation Overview

We describe our algorithm implementation in details below. Our executable software and instructions are available upon request.

Consider a colony of size num_ants where all the ants start the house-hunting task synchronously. We divide the total time to completion into *round*s, with a maximum round number of $total\_runs\_per\_setup$.

At the beginning of round t, no ant has transitioned yet (instantiate $Trans = \emptyset$). Then a random permutation of all ant_ids is generated as the order of execution.

When an ant gets her turn during this round, she first checks if her *ant_id* is in *Trans*. If so, she does nothing. Otherwise, knowing its id and current state, she chooses an action for this round according to the probability distribution defined in the **select-action** function.

The action picked by an ant x has an *action_type*, a receiving ant id, and a nest id. Please note here that in real ant colonies, an action can involve either just a single ant, or a pair of ants (tandem run and carry). In the single ant action case, the receiving ant's id is assigned value -1. In the pair ant action case, the action includes the valid *ant_id* of the receiving ant y. Similarly, not all actions require a nest, in which case the nest id for the action is -1.

By looking up the *Ants* array, x can also get the current external state of all ants including the receiving ant y, if any, of the picked action. These values are enough for x to call the **transition** function, and adds its own id to *Trans*. The special case handling is detailed in Section 2.3.2, including the case where y might also call a **transition** function and adds itself to *Trans*.

When one round finishes, each ant has had one chance to initiate or receive an action, and potentially has a new state. Repeat rounds like the above until the criteria is met for convergence with persistence, or until the program reaches the maximum number of rounds specified in the configuration file.

Appendix B

Supplemental Materials for Chapter 3

B.1 Proofs for Section 3.4

Theorem 20. *(Hoeffding's Inequality) If* $X_1, X_2, ..., X_N$ *are independent random variables with finite first and second moments, and* $0 \leq X_i \leq 1$ *for* $i = 1, ..., n$, *then for all* $D \geq 0$,

$$\mathbb{P}\left\{\frac{1}{N}\sum_{i=1}^{N}(X_i - \mathbb{E}\left[X_i\right]) \geq D\right\} \leq \exp\left(-2ND^2\right).$$

Proposition 21. *For any round* $t \geq 1$ *and ant* a_i, *it holds that*

$$\mathbb{P}\left\{T_i^1 > t\right\} < (1 - 2\alpha)^t = \beta^t,$$

where $\beta = 1 - 2\alpha$.

Proof. Let F^t represent the event that a_i is in state n_0 and does not transition out of n_0 during round t. That this could happen only if a_i has not jumped out of n_0 during any round previous to t as well. In turn, this requires that every coin toss that a_i has done

Further, F^t is influenced by the set of events $\{F^l\}$ for $\forall l \leq t - 1$ only through

F^{t-1}. Thus we have the following:

$$\begin{aligned}&\mathbb{P}\left\{F^t \mid F^{t-1} \wedge F^{t-2} \wedge \ldots \wedge F^0\right\}\\&= \mathbb{P}\left\{F^t \mid F^{t-1}\right\}\mathbb{P}\left\{F^{t-1} \mid F^{t-2}\right\}\mathbb{P}\left\{F^{t-2} \mid F^{t-3}\right\}\ldots\mathbb{P}\left\{F^2 \mid F^1\right\}\mathbb{P}\left\{F^1\right\}\end{aligned}$$

For any given round $t \geq 2$, conditioned on F^{t-1} is true, the probability of any ant a_i still *not* transitioning out of n_0 during round t is $1-2\alpha(1-\tau_1(t)-\tau_2(t))-\tau_1(t)-\tau_2(t) < 1-2\alpha$. Furthermore, $\mathbb{P}\{F^1\} = 1-2\alpha$ since there are no ants leading tandem runs yet and $\tau_1(1) = \tau_2(1) = 0$. Therefore,

Therefore,

$$\begin{aligned}\mathbb{P}\left\{T_i^1 > t\right\} &= \mathbb{P}\left\{F^t \mid F^{t-1} \wedge F^{t-2} \wedge \ldots \wedge F^0\right\}\\&= \mathbb{P}\left\{F^t \mid F^{t-1}\right\}\mathbb{P}\left\{F^{t-1} \mid F^{t-2}\right\}\mathbb{P}\left\{F^{t-2} \mid F^{t-3}\right\}\ldots\mathbb{P}\left\{F^2 \mid F^1\right\}\mathbb{P}\left\{F^1\right\}\\&= \mathbb{P}\left\{F^1\right\}\prod_{l=2}^{t}\mathbb{P}\left\{F^l \mid F^{l-1}\right\}\\&= \prod_{l=2}^{t}\mathbb{P}\left\{F^l \mid F^{l-1}\right\}(1-2\alpha)\\&< (1-2\alpha)^{t-1}(1-2\alpha)\\&= (1-2\alpha)^t.\end{aligned}$$

□

Proposition 22. *Let T_i^1 be defined as in Definition 2. With probability 1, ant a_i transitions out of the state n_0 in finite time, i.e.*

$$\mathbb{P}\left\{T_i^1 < \infty\right\} = 1.$$

Proof. By definition, we know that $\{T_i^1 < \infty\} = \cup_{t=1}^{\infty}\{T_i^1 \leq t\}$. Thus, it holds that

$$\left\{T_i^1 = \infty\right\} = \left\{T_i^1 < \infty\right\}^c = \left(\cup_{t=1}^{\infty}\left\{T_i^1 \leq t\right\}\right)^c = \cap_{t=1}^{\infty}\left\{T_i^1 > t\right\}.$$

It is easy to see that $\{T_i^1 > 1\} \supset \{T_i^1 > 2\} \supset \cdots \supset \{T_i^1 > t\} \supset \{T_i^1 > t+1\} \supset \cdots$.

By continuity of probability [28, Lemma 1.1], we know that

$$\lim_{t\to\infty} \mathbb{P}\left\{T_i^1 > t\right\} = \mathbb{P}\left\{\cap_{t=1}^{\infty}\left\{T_i^1 > t\right\}\right\}.$$

In addition, from Proposition 21, we have

$$\mathbb{P}\left\{T_i^1 > t\right\} \le (1-2\alpha)^t.$$

Therefore, we have that

$$\left\{T_i^1 = \infty\right\} = \mathbb{P}\left\{\cap_{t=1}^{\infty}\left\{T_i^1 > t\right\}\right\} = \lim_t \mathbb{P}\left\{T_i^1 > t\right\} \le \lim_t (1-2\alpha)^t = 0,$$

proving the proposition.

$\square$

Corollary 23. *With probability 1, all ants jump out of n_0 in finite time, i.e.,*

$$\mathbb{P}\left\{\max_{i\in[N]} T_i^1 < \infty\right\} = 1.$$

Proof.

$$\mathbb{P}\left\{\left\{\max_{i\in[N]} T_i^1 < \infty\right\}^c\right\} = \mathbb{P}\left\{\exists\ i \in [N] \text{ such that } T_i^1 = \infty\right\} = \sum_{i=1}^{N} \mathbb{P}\left\{T_i^1 = \infty\right\} = N \times 0 = 0.$$

$\square$

Corollary 24 follows immediately from Corollary 23 and Proposition 6.

Corollary 24. *For any $t \ge \max_{i\in[N]} T_i^1 + 1$, the state transitions $\boldsymbol{s}_i(t) \to \boldsymbol{s}_i(t+1)$ is independent of other ants. Moreover, for all $t \ge \max_{i\in[N]} T_i^1$,*

$$\mathbb{P}\left\{\exists i \in [N] \ \ \textit{such that}\ \boldsymbol{s}_i(t) = n_0\right\} = 0.$$

B.2 Proofs for Section 3.5

Lemma 25. *If a, b, c, d are all positive, the following holds*

$$\min\left\{\frac{a}{c}, \frac{b}{d}\right\} \leq \frac{a+b}{c+d} \leq \max\left\{\frac{a}{c}, \frac{b}{d}\right\},$$

where equality only holds if $\frac{a}{c} = \frac{b}{d}$.

Proof. If $\frac{a}{c} \geq \frac{b}{d}$, then $\frac{a+b}{c+d} \geq \frac{\frac{bc}{d}+b}{c+d} = \frac{b}{d}$; and $\frac{a+b}{c+d} > \frac{a}{c}$ otherwise.

If $\frac{a}{c} \geq \frac{b}{d}$, then $\frac{a+b}{c+d} \leq \frac{a+\frac{ad}{c}}{c+d} = \frac{a}{c}$; and $\frac{a+b}{c+d} < \frac{b}{d}$ otherwise. □

Proposition 26. *$x_2(t)/x_1(t) < 1$ for $\forall t \in [2, t^* - 1]$. Additionally, $x_2'(t^*) \leq x_1'(t^*)$ where equality only holds if $QS \leq \alpha$.*

Proof. We prove by induction. Preliminary: $t = 1, x_2(t) = x_1(t) = \alpha$.

Before either nests hit quorum,

Base case: $t = 1$.

$$\begin{aligned} x_2(t) &= 2u_2\alpha + \alpha(1 - u_1\alpha - u_2\alpha) + (1 - u_1)\alpha \\ &= (2u_2 - u_1 + 1)\alpha + \alpha(1 - u_1\alpha - u_2\alpha) \\ &< (2u_1 - u_2 + 1)\alpha + \alpha(1 - u_1\alpha - u_2\alpha) \\ &= x_1(t) \end{aligned}$$

Inductive Step: Assume $x_2(t) < x_1(t)$ for some $t \in [2, t^*-1)$, we have the following,

$$
\begin{aligned}
& x_2(t+1) - x_1(t+1) \\
& = 2u_2x_2(t) + \alpha(x_0(t) - u_1x_1(t) - u_2x_2(t)) + (1-u_1)x_1(t) \\
& \qquad - (2u_1x_1(t) + \alpha(x_0(t) - u_1x_1(t) - u_2x_2(t)) + (1-u_2)x_2(t)) \\
& = (3u_2 - 1)x_2(t) - (3u_1 - 1)x_1(t) \\
& < (3u_1 - 1)(x_2(t) - x_1(t)), \text{since } u_1 > u_2 > \frac{1}{3}, \\
& < 0, \text{since } u_1 > \frac{1}{2} \wedge x_2(t) < x_1(t).
\end{aligned}
$$

This completes the inductive proof and thus proves the first claim in the proposition.

Extending the above proof one more round, we prove that if $QS > \alpha$, i.e. no nest hits quorum in the first round, $x'_2(t^*) < x'_1(t^*)$ as defined in Eq. (25)-(26).

$$
\begin{aligned}
& x'_2(t^*) - x'_1(t^*) \\
& = (2u_2 - 1 + k)x_2(t^* - 1) - (2u_1 - 1 + k)x_1(t^* - 1) \\
& < (2u_1 - 1 + k)(x_2(t^* - 1) - x_1(t^* - 1)), \text{since } u_1 > u_2 > \frac{1}{2}, \\
& < 0, \text{since } 2u_1 - 1 + k > 0 \wedge x_2(t) < x_1(t),
\end{aligned}
$$

where $k = \frac{u_1x_1(t^*-1)}{u_1x_1(t^*-1)+u_2x_2(t^*-1)}$ if $x_0(t^*-1) < u_1x_1(t^*-1) + u_2x_2(t^*-1)$ and $k = 1$ otherwise.

Furthermore, if $QS \leq \alpha$, Eq. (25)-(26) shows that $x'_2(t^*) = x_2(2) = x'_1(t^*) = x_1(2) = QS$.

Hence we have finished proving the second part of proposition: $x'_2(t^*) \leq x'_1(t^*)$ and equality only holds if $QS < \alpha$. □

Proposition 27. *$x_2(t)$ monotonically increases for $\forall t < t'_0$.*

Proof.

- For $t < t^* - 1$: n_0 is not empty by time $t + 1$. $x_2(t+1) = 2u_2x_2(t) + \alpha(x_0(t) - \widehat{x}_1(t) - \widehat{x}_2(t)) + (1-u_1)x_1(t) > x_2(t)$, since $u_2 > 0.5$.
- For $t = t^* - 1$: similar to the above, we get $x'_2(t^*) > x_2(t^* - 1)$. Thus $x_2(t^*) -$

$x_2(t^* - 1) = (x'_2(t^*) - x_2(t^* - 1))d > 0$, where $d = \frac{x_1(t^*)-x_1(t^*-1)}{x'_1(t^*)-x_1(t^*-1)} > 0$.

- For $t \in [t^*, t'_0)$: $x_2(t+1) = 2u_2 x_2(t) + \alpha(x_0(t) - \widehat{x}_2(t)) > x_2(t)$.

$\square$

Proposition 28. *$x_2(t)$ monotonically decreases for $\forall t \geq t'_0$ if n_2 doesn't hit the quorum by t'_0.*

Proof. After n_0 becomes empty, n_2 has no “incoming" flux and only loses ants every round since n_1 has already hit quorum, according to Fig. 3-1 and Fig. 3-2 $\square$

Bibliography

[1] R. Beckers, Jean-Louis Deneubourg, and J. M. Pasteels. Collective decision-making through food recruitment. *Insectes Sociaux*, 37(3):258–267, 1990.

[2] Dominic D.R. Burns, Ana B. Sendova-Franks, and Nigel R. Franks. The effect of social information on the collective choices of ant colonies. *Behavioral Ecology*, 27(4):1033–1040, 02 2016.

[3] Grace Cai, Wendy Wu, Wayne Zhao, Jiajia Zhao, and Nancy Lynch. A spatially dependent probabilistic model for house hunting in ant colonies, 2021.

[4] Daniel Charbonneau and Anna Dornhaus. When doing nothing is something. how task allocation strategies compromise between flexibility, efficiency, and inactive agents. *Journal of Bioeconomics*, 17(3):217–242, 2015.

[5] Lars Chittka, Adrian G. Dyer, Fiola Bock, and Anna Dornhaus. Bees trade off foraging speed for accuracy. *Nature*, 424(6947):388–388, 2003.

[6] H. De Vries and J.C. Biesmeijer. Modelling collective foraging by means of individual behaviour rules in honey-bees. *Behavioral Ecology and Sociobiology*, 44(2):109–124, 1998. cited By 92.

[7] Claire Detrain and Jean-Louis Deneubourg. Collective decision-making and foraging patterns in ants and honeybees. volume 35 of *Advances in Insect Physiology*, pages 123 – 173. Academic Press, 2008.

[8] Claire Detrain, Hugo Pereira, and Vincent Fourcassié. Differential responses to chemical cues correlate with task performance in ant foragers. *Behavioral Ecology and Sociobiology*, 73(8):107, 2019.

[9] G. N. Doering and S. C. Pratt. Queen location and nest site preference influence colony reunification by the ant temnothorax rugatulus. *Insectes Sociaux*, 63(4):585–591, 2016.

[10] Grant Navid Doering and Stephen C. Pratt. Symmetry breaking and pivotal individuals during the reunification of ant colonies. *Journal of Experimental Biology*, 222(5), 2019.

[11] Grant Navid Doering, Kirsten A. Sheehy, James B. Barnett, and Jonathan N. Pruitt. Colony size and initial conditions combine to shape colony reunification dynamics. *Behavioural Processes*, 170:103994, 2020.

[12] Carolina Doran, Zac F. Newham, Ben B. Phillips, and Nigel R. Franks. Commitment time depends on both current and target nest value in temnothorax albipennis ant colonies. *Behavioral Ecology and Sociobiology*, 69(7):1183–1190, 2015.

[13] A. Dornhaus and N. R. Franks. Colony size affects collective decision-making in the ant temnothorax albipennis. *Insectes Sociaux*, 53(4):420–427, 2006.

[14] A. Dornhaus, N.R. Franks, R.M. Hawkins, and H.N.S. Shere. Ants move to improve: colonies of leptothorax albipennis emigrate whenever they find a superior nest site. *Animal Behaviour*, 67(5):959 – 963, 2004.

[15] Anna Dornhaus, Jo-Anne Holley, Victoria G. Pook, Gemma Worswick, and Nigel R. Franks. Why do not all workers work? colony size and workload during emigrations in the ant temnothorax albipennis. *Behavioral Ecology and Sociobiology*, 63(1):43–51, 2008.

[16] Nancy Lynch Emily Y. Zhang, Jiajia Zhao. A distributed consensus model for house-hunting in *Temnothorax* ant colonies. unpublished, 2021.

[17] Ofer Feinerman and Amos Korman. Individual versus collective cognition in social insects. *Journal of Experimental Biology*, 220(1):73–82, 2017.

[18] Nigel R Franks, François-Xavier Dechaume-Moncharmont, Emma Hanmore, and Jocelyn K Reynolds. Speed versus accuracy in decision-making ants: expediting politics and policy implementation. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1518):845–852, 03 2009.

[19] Nigel R. Franks, Anna Dornhaus, Charlotte S. Best, and Elizabeth L. Jones. Decision making by small and large house-hunting ant colonies: one size fits all. *Animal Behaviour*, 72(3):611 – 616, 2006.

[20] Nigel R. Franks, Anna Dornhaus, Jon P. Fitzsimmons, and Martin Stevens. Speed versus accuracy in collective decision making. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1532):2457–2463, 2003.

[21] Nigel R. Franks, Katherine A. Hardcastle, Sophie Collins, Faith D. Smith, Kathryn M.E. Sullivan, Elva J.H. Robinson, and Ana B. Sendova-Franks. Can ant colonies choose a far-and-away better nest over an in-the-way poor one? *Animal Behaviour*, 76(2):323 – 334, 2008.

[22] Nigel R Franks, Eamonn B Mallon, Helen E Bray, Mathew J Hamilton, and Thomas C Mischler. Strategies for choosing between alternatives with different attributes: exemplified by house-hunting ants. *Animal Behaviour*, 65(1):215 – 223, 2003.

[23] Nigel R Franks, Thomas O Richardson, Nathalie Stroeymeyt, R W Kirby, W M D Amos, P M Hogan, James A. R. Marshall, and T Schlegel. Speed-cohesion trade-offs in collective decision making in ants and the concept of precision in animal behaviour. *Animal Behaviour*, 85:1233–1244, April 2013.

[24] Nigel R. Franks, Jonathan P. Stuttard, Carolina Doran, Julian C. Esposito, Maximillian C. Master, Ana B. Sendova-Franks, Naoki Masuda, and Nicholas F. Britton. How ants use quorum sensing to estimate the average quality of a fluctuating resource. *Scientific Reports*, 5(1):11890, 2015.

[25] N.R. Franks, T.O. Richardson, N. Stroeymeyt, R.W. Kirby, W.M.D. Amos, P.M. Hogan, J.A.R. Marshall, and T. Schlegel. Speed–cohesion trade-offs in collective decision making in ants and the concept of precision in animal behaviour. *Animal Behaviour*, 85(6):1233 – 1244, 2013.

[26] Mohsen Ghaffari, Cameron Musco, Tsvetomira Radeva, and Nancy Lynch. Distributed house-hunting in ant colonies, 2015.

[27] Simone M Glaser and Christoph Grüter. Ants (temnothorax nylanderi) adjust tandem running when food source distance exposes them to greater risks. *Behavioral Ecology and Sociobiology*, 72(3):40, 2018.

[28] Bruce Hajek. *Random processes for engineers.* Cambridge university press, 2015.

[29] Christiane I. M. Healey and Stephen C Pratt. The effect of prior experience on nest site evaluation by the ant temnothorax curvispinosus. *Animal Behaviour*, 76:893–899, 2008.

[30] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[31] Alex Kacelnik, Marco Vasconcelos, Tiago Monteiro, and Justine Aw. Darwin’s “tug-of-war”vs. starlings”“horse-racing”: how adaptations for sequential encounters drive simultaneous choice. *Behavioral Ecology and Sociobiology*, 65(3):547–558, 2011.

[32] Teeraphan Laomettachit, Teerasit Termsaithong, Anuwat Sae-Tang, and Orawan Duangphakdee. Stop-signaling reduces split decisions without impairing accuracy in the honeybee nest-site selection process. *Journal of Insect Behavior*, 29(5):557–577, 2016.

[33] Teeraphan Laomettachit, Teerasit Termsaithong, Anuwat Sae-Tang, and Orawan Duangphakdee. Stop-signaling reduces split decisions without impairing accuracy in the honeybee nest-site selection process. *Journal of Insect Behavior*, 29(5):557–577, 2016.

[34] Nicholas C. Makris, Purnima Ratilal, Srinivasan Jagannathan, Zheng Gong, Mark Andrews, Ioannis Bertsatos, Olav Rune Godø, Redwood W. Nero, and

J. Michael Jech. Critical population density triggers rapid formation of vast oceanic fish shoals. *Science*, 323(5922):1734–1737, 2009.

[35] E. Mallon, S. Pratt, and N. Franks. Individual and collective decision-making during nest site selection by the ant leptothorax albipennis. *Behavioral Ecology and Sociobiology*, 50(4):352 – 359, 2001.

[36] James A. R. Marshall, Rafal Bogacz, Anna Dornhaus, Robert Planqué, Tim Kovacs, and Nigel R. Franks. On optimal decision-making in brains and social insect colonies. *Journal of The Royal Society Interface*, 6(40):1065–1074, 2009.

[37] James A.R Marshall, Anna Dornhaus, Nigel R Franks, and Tim Kovacs. Noise, cost and speed-accuracy trade-offs: decision-making in a decentralized system. *Journal of The Royal Society Interface*, 3(7):243–254, 2006.

[38] Naoki Masuda, Thomas A. O'shea-Wheller, Carolina Doran, and Nigel R. Franks. Computational model of collective nest selection by ants with heterogeneous acceptance thresholds. *Royal Society Open Science*, 2(6):140533, 2015.

[39] Naoki Masuda, Thomas A. O'shea-Wheller, Carolina Doran, and Nigel R. Franks. Computational model of collective nest selection by ants with heterogeneous acceptance thresholds. *Royal Society Open Science*, 2(6):140533, 2015.

[40] Michael H. J. Moglich. Social organization of nest emigration in leptothorax (hym., form.). 1978.

[41] Benjamin P. Oldroyd and Stephen C. Pratt. Chapter four - comb architecture of the eusocial bees arises from simple rules used during cell building. volume 49 of *Advances in Insect Physiology*, pages 101 – 121. Academic Press, 2015.

[42] Andrea Perna, Boris Granovskiy, Simon Garnier, Stamatios C Nicolis, Marjorie Labédan, Guy Theraulaz, Vincent Fourcassié, and David J T Sumpter. Individual rules for trail pattern formation in Argentine ants (*Linepithema humile*). *PLoS Computational Biology*, 8(7):e1002592, July 2012.

[43] Andrea Perna and Guy Theraulaz. When social behaviour is moulded in clay: on growth and form of social insect nests. *Journal of Experimental Biology*, 220(1):83–91, 2017.

[44] Robert Planqué, Anna Dornhaus, Nigel R. Franks, Tim Kovacs, and James A. R. Marshall. Weighting waiting in collective decision-making. *Behavioral Ecology and Sociobiology*, 61(3):347–356, 2007.

[45] S. C. Pratt. Behavioral mechanisms of collective nest-site choice by the ant temnothorax curvispinosus. *Insectes Sociaux*, 52(4):383–392, 2005.

[46] S.C. Pratt, E.B. Mallon, D.J. Sumpter, and et al. Quorum sensing, recruitment, and collective decision-making during colony emigration by the ant leptothorax albipennis. *Behavioral Ecology and Sociobiology*, 52(2):117–127, 2002.

[47] Stephen C Pratt. Behavioral mechanisms of collective nest-site choice by the ant temnothorax curvispinosus. *Insectes Sociaux*, 52:383–392, 2005.

[48] Stephen C. Pratt. Quorum sensing by encounter rates in the ant Temnothorax albipennis. *Behavioral Ecology*, 16(2):488–496, 01 2005.

[49] Stephen C. Pratt. Quorum sensing by encounter rates in the ant Temnothorax albipennis. *Behavioral Ecology*, 16(2):488–496, 01 2005.

[50] Stephen C. Pratt. Efficiency and regulation of recruitment during colony emigration by the ant temnothorax curvispinosus. *Behavioral Ecology and Sociobiology*, 62(8):1369–1376, 2008.

[51] Stephen C. Pratt. Nest site choice in social insects. In Jae Chun Choe, editor, *Encyclopedia of Animal Behavior (Second Edition)*, pages 766 – 774. Academic Press, Oxford, second edition edition, 2019.

[52] Stephen C. Pratt and David J. T. Sumpter. A tunable algorithm for collective decision-making. *Proceedings of the National Academy of Sciences*, 103(43):15906–15910, 2006.

[53] Stephen C. Pratt, David J.T. Sumpter, Eamonn B. Mallon, and Nigel R. Franks. An agent-based model of collective nest choice by the ant temnothorax albipennis. *Animal Behaviour*, 70(5):1023 – 1036, 2005.

[54] Tsvetomira Radeva. *A Symbiotic Perspective on Distributed Algorithms and Social Insects*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA 02139, June 2017.

[55] Thomas O Richardson, Philippa A Sleeman, John M Mcnamara, Alasdair I Houston, and Nigel R. Franks. Teaching with evaluation in ants. *Current Biology*, 17(17):1520–1526, September 2007.

[56] Elva J. H. Robinson, Nigel R. Franks, Samuel Ellis, Saki Okuda, and James A. R. Marshall. A simple threshold rule is sufficient to explain sophisticated collective decision-making. *PLOS ONE*, 6(5):1–11, 05 2011.

[57] Takao Sasaki, Blake Colling, Anne Sonnenschein, May M. Boggess, and Stephen C. Pratt. Flexibility of collective decision making during house hunting in temnothorax ants. *Behavioral Ecology and Sociobiology*, 69(5):707–714, 2015.

[58] Takao Sasaki, Boris Granovskiy, Richard P. Mann, David J. T. Sumpter, and Stephen C. Pratt. Ant colonies outperform individuals when a sensory discrimination task is difficult but not when it is easy. *Proceedings of the National Academy of Sciences*, 110(34):13769–13773, 2013.

[59] Takao Sasaki and Stephen C. Pratt. Emergence of group rationality from irrational individuals. *Behavioral Ecology*, 22(2):276–281, 01 2011.

[60] Takao Sasaki and Stephen C. Pratt. Groups have a larger cognitive capacity than individuals. *Current Biology*, 22(19):R827 – R829, 2012.

[61] Takao Sasaki, Benjamin Stott, and Stephen C. Pratt. Rational time investment during collective decision making in <i>temnothorax</i> ants. *Biology Letters*, 15(10):20190542, 2019.

[62] T.D. Seeley. *The Wisdom of the Hive.* Harvard University Press, 1995.

[63] Zachary Shaffer, Takao Sasaki, and Stephen C. Pratt. Linear recruitment leads to allocation and flexibility in collective foraging by ants. *Animal Behaviour*, 86:967–975, September 2013.

[64] Nathalie Stroeymeyt, Martin Giurfa, and Nigel R. Franks. Improving decision speed, accuracy and group cohesion through early information gathering in house-hunting ants. *PLOS ONE*, 5(9):1–10, 09 2010.

[65] David J.T Sumpter and Stephen C Pratt. Quorum responses and consensus decision making. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1518):743–753, 2009.

[66] D.J.T. Sumpter, G.B. Blanchard, and D.S. Broomhead. Ants and agents: A process algebra approach to modelling ant colony behaviour. *Bulletin of Mathematical Biology*, 63(5):951–980, 2001. cited By 35.

[67] Gabriele Valentini, Naoki Masuda, Zachary Shaffer, Jake R Hanson, Takao Sasaki, Sara Imari Walker, Theodore P Pavlic, and Stephen C Pratt. Division of labour promotes the spread of information in colony emigrations by the ant *Temnothorax rugatulus*. *Proceedings of the Royal Society B: Biological Sciences*, 287:20192950, 2020.

[68] Gabriele Valentini, Nobuaki Mizumoto, Stephen C. Pratt, Theodore P. Pavlic, and Sara Imari Walker. Revealing the structure of information flows discriminates similar animal social behaviors. *eLife*, 9:e55395, 2020.

[69] Ashley J. W. Ward, Jens Krause, and David J. T. Sumpter. Quorum decision-making in foraging fish shoals. *PLOS ONE*, 7(3):1–8, 03 2012.

[70] Christian A. Yates, Radek Erban, Carlos Escudero, Iain D. Couzin, Jerome Buhl, Ioannis G. Kevrekidis, Philip K. Maini, and David J. T. Sumpter. Inherent noise can facilitate coherence in collective swarm motion. *Proceedings of the National Academy of Sciences*, 106(14):5464–5469, 2009.