



More Is Faster: Why Population Size Matters in Biological Search

JANNATUL FERDOUS,¹ GEORGE MATTHEW FRICKE,^{1,2} and MELANIE E. MOSES^{1,3,4}

ABSTRACT

Many biological scenarios have multiple cooperating searchers, and the timing of the initial first contact between any one of those searchers and its target is critically important. However, we are unaware of biological models that predict how long it takes for the first of many searchers to discover a target. We present a novel mathematical model that predicts initial first contact times between searchers and targets distributed at random in a volume. We compare this model with the extreme first passage time approach in physics that assumes an infinite number of searchers all initially positioned at the same location. We explore how the number of searchers, the distribution of searchers and targets, and the initial distances between searchers and targets affect initial first contact times. Given a constant density of uniformly distributed searchers and targets, the initial first contact time decreases linearly with both search volume and the number of searchers. However, given only a single target and searchers placed at the same starting location, the relationship between the initial first contact time and the number of searchers shifts from a linear decrease to a logarithmic decrease as the number of searchers grows very large. More generally, we show that initial first contact times can be dramatically faster than the average first contact times and that the initial first contact times decrease with the number of searchers, while the average search times are independent of the number of searchers. We suggest that this is an underappreciated phenomenon in biology and other collective search problems.

Keywords: collective search, extreme first passage time, foraging, immunity, initial first contact time.

¹Department of Computer Science, The University of New Mexico, Albuquerque, New Mexico, USA.

²Center for Advanced Research Computing, The University of New Mexico, Albuquerque, New Mexico, USA.

³Department of Biology, The University of New Mexico, Albuquerque, New Mexico, USA.

⁴Santa Fe Institute, Santa Fe, New Mexico, USA.

1. INTRODUCTION

IN MANY BIOLOGICAL SCENARIOS, a large number of searchers simultaneously search for one or more targets. Sometimes, the first target discovered by the first lucky searcher is particularly important because it results in a cascade of downstream events. For example, in ant colonies, many ants forage for food; when the first one is successful, it may use pheromones to recruit other searchers. Thus, the time for the first ant to discover food may have a disproportionate impact on subsequent resource collection rates; similarly, naive T cells search for cognate antigen in lymph nodes. When T cells bind with target antigen, they activate and replicate in an exponentially growing population of cells. Thus, the timing of the first contact by the first T cell that finds its target is particularly important in initiating the adaptive immune response.

In these examples, the time it takes for the initial discovery of a target (what we call the initial first contact time) may be more important than the average time of all of the searchers to find their first targets. In this study, we develop and analyze mathematical models and computer simulations to understand which factors affect the initial first contact times.

In immune system models, it is common to consider the mean or median search time for a typical T cell to find its target (Celli et al., 2012; Ferdous et al., 2022; Krummel et al., 2016; Perelson and Wiegel, 2009; Wiegel and Perelson, 2004). An analysis of initial first contact times in three-dimensional (3D) volumes has only recently been investigated in a series of publications (Lawley, 2020b; Lawley, 2020c; Lawley and Madrid, 2020), but these models do not apply to the T cell search that motivated our work. The derivations in these other studies apply to searchers that start in a single location and search for a single target, and they rely on population sizes that approach infinity. We call the predictions of initial first contact times from such models t_∞ .

In this work, we derive initial first contact time predictions for finite populations of searchers, t_{fin} . This model initially assumes a number of searchers and targets distributed at uniform random within a volume, and we predict how t_{fin} scales as the number of searchers increases. We then test different search scenarios, assess when assumptions made by the t_{fin} and t_∞ models hold, and compare simulated initial first contact times predicted by t_{fin} and t_∞ .

While we were motivated to develop t_{fin} to predict how quickly the first T cell would find its target antigen in a lymph node, we also examine whether t_{fin} can explain other search scenarios relevant for social insects and other collective foragers. Group size is a key feature of a colony (Dornhaus et al., 2012) and many studies have investigated how colony size affects social insect foraging, finding that larger colonies are more complex (Anderson and McShea, 2001), mass recruitment is more effective in larger ant colonies (Beekman et al., 2001), but waggle dances are not necessarily more effective in larger bee colonies (Dornhaus et al., 2006). Popp and Dornhaus (2023) found larger colonies searched an area faster and Adler and Gordon (1992) found larger colonies found more events. However, none of these studies modeled whether the first event that triggered subsequent recruitment was found faster in larger colonies.

Some studies suggest this may be the case. For example, a field study by Donaldson-Matasci et al. (2013) found that larger honeybee colonies found resources faster, and as a result, overall foraging rates were better for larger colonies. Our previous field study (Flanagan et al., 2012) also found that larger seed harvester ant colonies found piles of seeds faster; however, neither of these studies had sufficiently detailed estimates of forager population size paired with first discovery times to establish a quantitative relationship between the two. That is the focus of the models we build and analyze here.

In this study, we develop a model that predicts the initial first contact time for a finite number of searchers (t_{fin}). We simulate search in different scenarios to validate conditions under which our theoretical assumptions do, and do not, hold. Our goal in this article is not to test model predictions with biological data (which we leave to future work), but we parameterize our simulations to reflect the numbers of T cell searchers, antigen-carrying dendritic cell targets, and lymph node volumes across mammals. We show that t_{fin} depends on the number of searchers (N). In scenarios where multiple searchers are involved, such as foraging ants, bees, or immune cells, a larger number of searchers can increase the likelihood of discovering a target quickly.

We predict that more searchers exploring the environment increase the chances that one will stumble upon the target quickly, whether that target is a seed, a flower, or an antigen. Our goal is to provide quantitative predictions and simulations that test this qualitative expectation, given the different numbers and placements of searchers and targets in different search volumes. In this article, we develop a general model of initial first contact times, considering various placements of searchers and targets, and we compare our analysis for finite N with physics models of initial first contact times that assume an infinite number of searchers all starting in the same location. This work is a step toward a more general biological theory of first contact times.

While we are unaware of quantitative models or data of initial first contact times for a finite number of searchers t_{fin} , physicists have developed an extensive literature analyzing extreme first passage time (t_{∞}) (Basnayake et al., 2019; Benichov et al., 2005; Lawley, 2020a; Redner, 2001; Weiss et al., 1983), an approach that predicts the time it takes for the fastest among an infinite number of searchers to find its target. However, t_{∞} generally only considers search problems in which searchers all start at the same location and search for a single target, the number of searchers is assumed to approach infinity, and most work considers searchers on a one-dimensional (1D) line. To address these departures from biological reality, we propose a mathematical model designed to predict first contacts in the immune system. Motivated by Ferdous et al. (2024), t_{fin} is an idealized model of the initial first contact times between T cells and dendritic cells within the lymph node.

This first contact is the critical step that initiates adaptive immunity and is a key factor in determining how long it takes T cells to control a viral infection. We consider a finite number of T cells as searchers and dendritic cells as targets, with both distributed uniformly in a bounded space representing the lymph node. We show in simulations that t_{fin} is proportional to the volume of the search space and inversely proportional to both the number of searchers and the targets. We investigate the influence of two factors that can reduce search time when there are more searchers: (1) the initial distance between the searchers and the targets and (2) the chance that a searcher happens to take a direct path to the target.

We expect that both the shortest distance to a target and the most direct path to the target will decrease when there are more searchers because both of these fortuitous events are more likely to occur at least once when there are more searchers. The major contributions of this article are that we (1) develop a model (t_{fin}) that predicts the initial first contact times for a finite number of randomly distributed searchers and targets, (2) compare t_{fin} with the t_{∞} model that assumes an infinite number of searchers all starting in one location, (3) analyze the models in four scenarios (cases 0–3) with varying searcher and target distributions to provide insights into the applicability of t_{fin} and t_{∞} in different biologically relevant scenarios, and (4) simulate idealized versions of those scenarios to examine how well the analytical predictions correspond to those scenarios.

We consider the following cases: In Case 0, we consider a random distribution of multiple searchers and multiple targets. We test whether t_{fin} assumptions are met and whether simulations are consistent with t_{fin} predictions. We then compare t_{fin} with t_{∞} , which assumes only a single target. Thus, for all other cases (1–3), we consider a single target located at the center of the search space ($T = 1$), and all searchers start at the same fixed point at some specified distance from the target (Fig. 1). The different predictions of the t_{∞} and t_{fin} approaches are summarized for each case in Table 1.

2. PREDICTIONS AND COMPUTATIONAL METHODS

2.1. Summary of scenarios

We show four scenarios in Figure 1. The t_{fin} and t_{∞} predictions of initial first contact times differ because each approach rests on different assumptions. The predictions are summarized in Table 2.

- Case 0: For multiple searchers and multiple targets, both distributed at uniform random and with a constant density of searchers and targets, t_{fin} is expected to be inversely proportional to the volume of the search space and the number of searchers (N) [Eq. (5)]. One of the primary assumptions for t_{∞} is that all the searchers start from the same location; since that assumption does not hold in Case 0, t_{∞} does not make a prediction for Case 0.
- Case 1: For a fixed initial distance between all the searchers and a single target, t_{fin} is expected to decrease linearly with the number of searchers, according to Equation (7) (the same as Case 0). In contrast, t_{∞} predicts the initial first contact times to decrease logarithmically with N [Eq. (13)].
- Case 2: For a fixed number of searchers, t_{fin} is expected to increase proportionally to the cube of the distance between all the searchers starting from the same location and the target according to Equation (8). Our analysis of t_{∞} predicts that t_{∞} scales quadratically with the distance in Equation (14).
- Case 3: If the density of searchers remains constant within a given volume, but all searchers initiate their search from a common starting point, t_{fin} is expected to be constant in Equation (9). t_{∞} is expected to scale with length squared divided by N in Equation (15).

While t_{fin} and its underlying assumptions were originally formulated for Case 0, with multiple searchers and targets distributed randomly in a volume, we extend the examination of t_{fin} to Cases 1–3 to assess

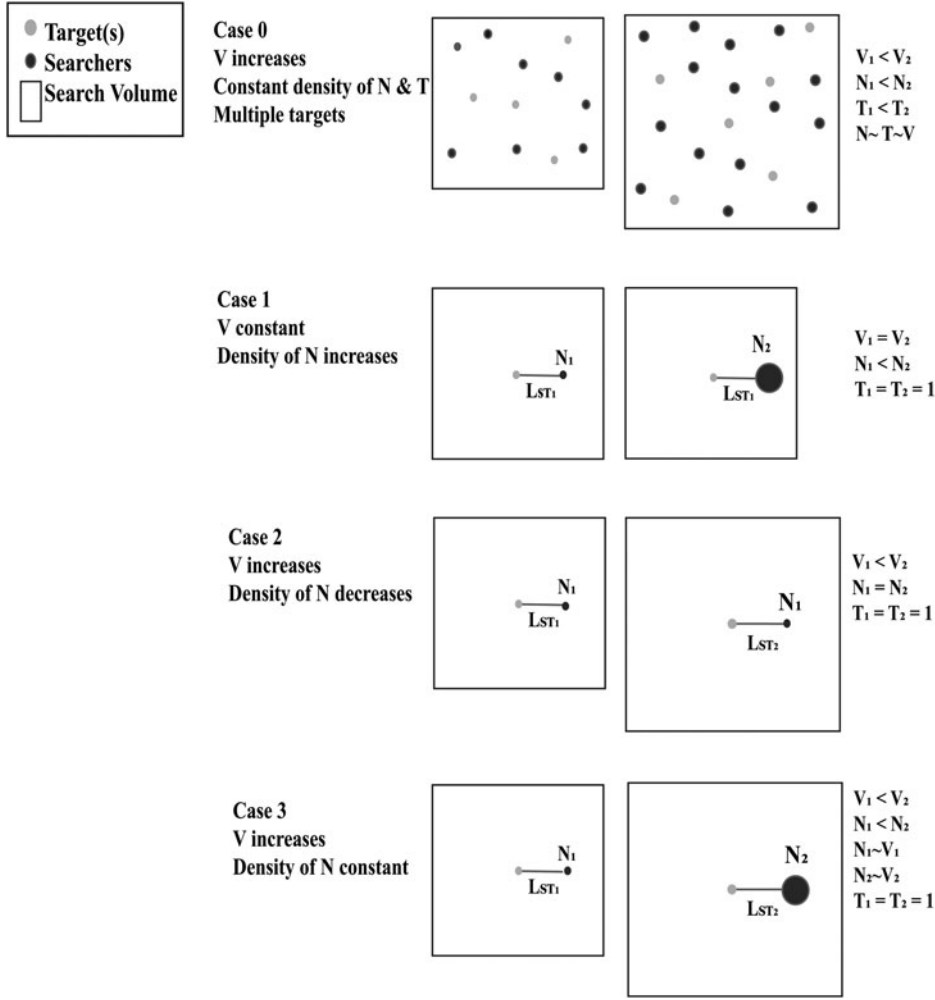


FIG. 1. Graphical representation of the simulated cases. In all cases there are multiple searchers (N). Here, V is the search space, T is the number of searchers, N is the number of searchers, L_{ST} is the length of the distance from the searchers to the target. In Case 0, there are multiple targets and the targets and searchers are distributed at uniform random in the volume, and their density is constant as the volume increases. In Cases 1–3, there is a single target and the searchers are all located at the same place at distance $L_{ST} = 0.25 L$ from the target so that searchers are equidistant from the target and the boundary. In Case 1, the number of searchers N increases (indicated by the larger circle on the right), while L_{ST} and V are constant. In Case 2, N is constant, and L_{ST} and V increase (indicated by the larger box on the right). In Case 3, N , L_{ST} , and V all increase (indicated by the larger box size and larger circle on the right). We assume L_{ST} is proportional to $V^{1/3}$.

TABLE 1. COMPARISON OF t_{∞} AND t_{fin}

Experimental variation in N , T , and V	t_{∞} prediction	t_{fin} prediction
Case 0 (N , T , and V increase; multiple searchers and targets are distributed at uniform random)	No prediction	$E[t_{\text{fin}}] \propto N^{-1}$
Case 1 (N increases, V is constant; searchers start at one location; single target)	$E[t_{\infty}] \propto \frac{1}{\ln(N)}$	$E[t_{\text{fin}}] \propto \frac{1}{N}$
Case 2 (N is constant, V increases; searchers start at one location; single target)	$E[t_{\infty}] \propto L^2$	$E[t_{\text{fin}}] \propto L^3$
Case 3 (N and V increase; searchers start at one location; single target)	$E[t_{\infty}] \propto \frac{L^2}{\ln N}$	$E[t_{\text{fin}}] \propto 1$

We assumed an exponential distribution of search times and a finite number of searchers in our model of t_{fin} . In contrast, t_{∞} assumes an infinite number of searchers and that the fastest searcher will follow the shortest (and therefore, most direct) path to the target.

TABLE 2. STATISTICAL COMPARISONS OF THE GAUSSIAN AND EXPONENTIAL DISTRIBUTION FIT TO THE INITIAL FIRST CONTACT TIMES FOR CASE 0

	<i>Number of searchers</i>				
	20	63	200	630	2000
Exponential					
SSE	$8.1.2 \times 10^{-8}$	4.6×10^{-10}	4.6×10^{-11}	4.6×10^{-12}	7×10^{-13}
AIC	2600	2800	3000	3200	3400
Gaussian					
SSE	1.3×10^{-8}	6.7×10^{-10}	1.2×10^{-10}	3.7×10^{-11}	1.9×10^{-11}
AIC	2600	3000	3300	3900	4300

The exponential distribution has lower mean SSE and AIC values than the Gaussian distribution, indicating a better fit to the data. One hundred simulations were replicated for each number of searchers. Numbers are given to two significant digits.

AIC, Akaike information criteria; SSE, sum of squares error.

whether it applies. Similarly, t_∞ and its assumptions were initially designed for Case 1 with the number of searchers approaching infinity. We explore how well the predictions of t_{fin} and t_∞ correspond to simulations given that the original assumptions may not hold in all cases. We note that our predictions for t_{fin} are based on an assumption that each searcher is independent, resulting in an exponential distribution of search times (Celli et al., 2012). We show that the assumption holds in Case 0 (the original scenario that t_{fin} was developed for), but it may not hold for other cases.

2.2. Analysis of t_{fin}

We initially developed t_{fin} to predict the initial first contact time given randomly distributed searchers and targets in a volume. We assume that the contact times between the searchers and targets are memoryless random processes, such that the first contact times of each searcher follow an exponential distribution. This is shown in Case 0 as a bounded cube with N searchers looking for a number of targets (T). If the probability that a searcher finds the target in time t is $P(T_t)$, then the probability $P(T_t)$ that at least one searcher encounters a target between time 0 and time t that can be modeled with an exponential distribution,

$$P(T < t) = \frac{NT}{\lambda} e^{-\frac{NT}{\lambda}} \quad (1)$$

Here, λ is the mean contact time between a single searcher and a single target within a volume V . Celli et al. (2012) showed that

$$\lambda \propto V \quad (2)$$

To obtain the expected encounter time, we integrate over all time,

$$E[t_{\text{fin}}] = \int_0^\infty \frac{tNT}{\lambda} e^{-\frac{NT}{\lambda}} dt = \frac{\lambda}{NT} \quad (3)$$

By replacing λ with V from Equation (2) in Equation (3) we get,

$$E[t_{\text{fin}}] \propto \frac{V}{NT} \quad (4)$$

- Case 0 (N , T , and L all increase, multiple targets):

Here we assume there are multiple searchers and targets, and that the density of searchers and targets is constant, $N \propto T \propto V$, then from Equation (4) we get,

$$E[t_{\text{fin}}] \propto V^{-1} \propto N^{-1} \quad (5)$$

That is, given a constant density of searchers and targets, the time for the initial first contact between a searcher and a target is inversely proportional to the search volume and the number of searchers. In this scenario, the initial first contact predicted t_{fin} speeds up linearly with N .

In Figure 3, we show that simulations of Case 0 generate exponentially distributed contact times among all of the searchers in a simulation, consistent with the primary assumption in t_{fin} . In Figure 4, we show that the initial first contact times produced by the simulations are consistent with the (t_{fin}) prediction. Next, to compare our t_{fin} prediction with the t_{∞} prediction, we consider three other cases with searchers starting at the same location and searching for a single target. Thus we set the number of targets $T=1$, simplifying Equation (4) to

$$E[t_{\text{fin}}] \propto \frac{V}{N} \quad (6)$$

In addition to allowing us to compare the predictions of the t_{fin} and t_{∞} approaches, these scenarios also give us insights into the contribution of two factors that determine t_{fin} : (1) the initial distance between the lucky first searcher and the nearest target versus (2) the directness of the path of the lucky first searcher toward the target because in these scenarios we fix (1) the distance between all searchers and the target.

- Case 1 (N increases, V is constant, single target):

Since V is constant and N increases, Equation (6) is simplified to:

$$E[t_{\text{fin}}] \propto \frac{V}{N} \propto \frac{1}{N} \quad (7)$$

This case with a single target in a fixed volume gives the same prediction as Case 0 for multiple randomly distributed searchers and targets in an increasing volume: as N increases, the expected initial first contact time decreases linearly.

- Case 2 (N is constant, V increases, single target):

Since $V=L^3$ and N is constant, Equation (6) becomes,

$$E[t_{\text{fin}}] \propto \frac{V}{N} \propto L^3 \quad (8)$$

In this scenario, the expected initial first contact time increases linearly with the volume. As the search volume (and therefore L) increases, it takes more time for the searchers to cover the increased space required to find the target.

- Case 3 (N and V both increase, single target)

In this scenario, we set $N \propto V$, so that Equation (6) predicts,

$$E[t_{\text{fin}}] \propto \frac{V}{N} \propto 1 \quad (9)$$

Thus, if density is constant, t_{fin} is predicted to be constant across volumes.

2.3. Analysis of t_{∞}

The t_{∞} prediction was developed by physicists to describe the time for the first of an infinite number of searchers to find a target in an idealized search, originally on a 1D line (Redner, 2001; Weiss et al., 1983). The theory has recently been expanded to more dimensions and more complex searches (Basnayake et al., 2019; Lawley, 2020a; Lawley, 2020c; Ro and Kim, 2017), but most formulations consider only very large N that approaches infinity. t_{∞} predicts that,

$$E[t_{\infty}] \propto \frac{L_{ST}^2}{4D \ln(N)}, N \rightarrow \infty \quad (10)$$

where N is the number of searchers, D is a diffusivity, and L_{ST} is the distance between the identical starting location of all searchers and the target. The model assumes searchers move via a random walk. The time for

the first searcher to reach the target is, in essence, the rate of spread of the furthest of the diffusing searchers. That diffusion creates a Gaussian distribution of distances from the start location. The logarithmic term in the denominator comes from determining the time it takes for the tail of the distribution to reach the target.

The full derivation of Equation (10) is in Lawley (2020c) [Eqs. (5–7), Section 2.1]. t_∞ predicts that the distance traveled by the fastest searcher increases slowly. As the number of searchers increases, the time required for the first searcher to reach the target decreases as a slow logarithmic function of N . Note that this formulation of t_∞ does not specifically address how the volume of the search domain affects search time; it considers only the distance between the searcher and target. This is because the approach assumes that the fastest searcher will be the one that has taken the most direct path from the start location to the target; any searcher that meandered through the full volume of the search space would not be the first searcher to arrive, and therefore would not be relevant to the initial first contact time. This argument only holds for large N that approaches infinity.

However, many biological search processes have a finite number of searchers within a bounded volume, for example, immune cells search in the volume of a lymph node or an infected organ, and foraging animals search within a two-dimensional (2D) or 3D territory. We consider only the 3D case here and set L_{ST} to scale with the side L of the volume V of the search domain such that,

$$L_{ST} \propto L \propto V^{1/3} \quad (11)$$

For simplicity, we consider a constant rate of diffusivity (D). By replacing L_{ST} with L and D with a constant, Equation (10) gives us a simplified expression for t_∞ in a bounded volume of length L :

$$E[t_\infty] \propto \frac{L^2}{\ln(N)}, N \rightarrow \infty \quad (12)$$

We consider three cases for Equation (12).

- Case 1 (N increases, L is constant, single target):

In this case, we assume that L_{ST} remains fixed. As the number of searchers N increases, we reformulate Equation (12) by ignoring scaling constants to obtain a prediction from t_∞ ,

$$E[t_\infty] \propto \frac{L^2}{\ln(N)} \propto \frac{1}{\ln(N)} \quad N \rightarrow \infty \quad (13)$$

This equation predicts that as N grows, the expected time for finding the target decreases logarithmically with N . The diminishing returns of adding more searchers are evident, indicating that large groups of searchers only slightly reduce the search time compared with smaller groups when the distance (L) is fixed.

- Case 2 (N is constant, L increases, single target):

In this case, we keep N constant, while V (and therefore L_{ST}) increases. Then, Equation (12) becomes,

$$E[t_\infty] \propto L^2 \quad (14)$$

In this scenario, t_∞ is predicted to increase with L^2 , while volume increases with L^3 .

- Case 3 (N and L increase, single target):

In this case, N increases proportional to V , leading to:

$$E[t_\infty] \propto \frac{L^2}{\ln(N)}, \text{ where } N \propto V \propto L^3 \quad (15)$$

The different predictions from t_∞ and t_{fin} are summarized in Table 1.

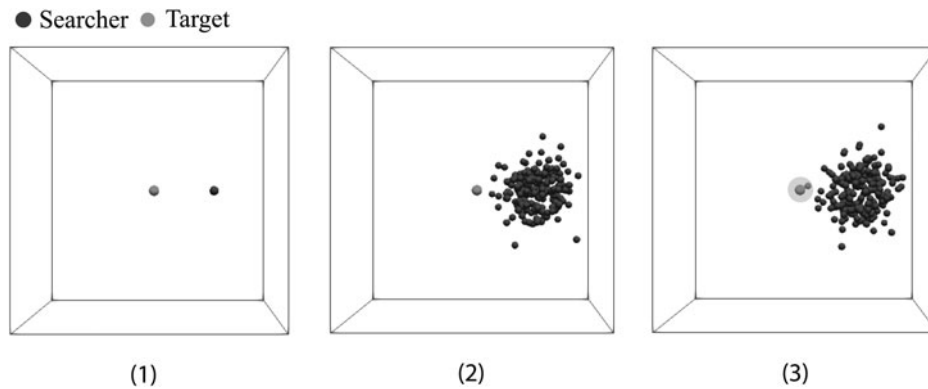


FIG. 2. Schematic showing simulations of Cases 1–3. Multiple searchers (*dark dots*) looking for a target (*light dots*) in a confined cubic region, with the target placed at the center. The figure shows (A) the initial setup where all the searchers are in the same location with a fixed distance $L_{ST} = 0.25L$ from the target at the center of a fixed volume (Case 1). (B) The search progresses at the 250th step, with the searchers moving via Brownian motion and (C) the first contact between a searcher and the target (*very light highlighting*). In Case 1, we fixed volume at 8 mm^3 and systematically increased the number of searchers (N) from 5 to 100,000. In Case 2, we increased V from 8 to $33\,000 \text{ mm}^3$ (L from 2–32) with fixed N . In Case 3, we increased both V and N with the N ranging from 240 to 1,000,000 by scaling with the volume ($N \propto V$). In Case 0, with randomly distributed searchers and multiple targets (not shown); we increased V , T , and N with the N ranging from 20 to 2000 ($N \propto V$), T ranging from 80 to 8000, and V ranging from 8 to 790. Each simulation was replicated 30 times for each combination of parameters.

2.4. Computational experiments

To test the t_{fin} prediction for Case 0 and the competing predictions of t_{fin} and t_{∞} , for Cases 1–3, we conduct a series of experiments using an agent-based model implemented in BioDynaMo (Breitwieser et al., 2022). We consider that a contact happens when the centers of a searcher and a target are within 10 distance. Hence, the physical sizes of searchers and targets do not affect search times. To simulate T cells searching for dendritic cells in the lymph node, we utilized a model of Brownian motion parameterized from empirical data of T cell movement in lymph nodes from Fricke et al. (2016). Figure 2 illustrates the experimental setup of our simulations designed to test the predictions of the t_{fin} and t_{∞} approaches to estimating the initial first contact times in the four cases shown in Figure 1.

The scarcity of empirical data presents a challenge in directly matching the numerical simulations to real biological systems. We acknowledge this limitation and took steps to align the simulation parameters in Case 0 with available data from the immune system. We chose search volumes to match the range of sizes of mammalian lymph nodes and estimated numbers of searchers to match the number of T cells in each clonal line that search for antigen-bearing dendritic cells in each lymph node and also estimated the number of targets for Case 0 to match the number of dendritic cells in lymph node carrying the same type of antigens. The purpose of the computational simulations is to establish a proof of principle showing how the initial first contact times vary with N and various other specifications of the search problem.

Our intent here is not to test our predictions with biological data, but rather to demonstrate through simulations how various mathematical assumptions and predictions hold given simple simulations using finite numbers of searchers, targets, and volumes that approximate different biological scenarios. Our simulations validate cases where different approaches (t_{fin} and t_{∞}) are (and are not) able to predict biologically relevant phenomena. All the data used in our analysis are summarized in an online data set.*

2.5. Fitting exponents using the agent-based model (ABM)

To fit model predictions to simulated data, we estimate the best-fit exponents that relate initial first contact times to the dependent variable in each scenario. We first consider the generic formula in Equation (4) as follows,

*<https://docs.google.com/spreadsheets/d/1AiWJ2onZYN5F-42LNLv-14kUluWyQhGSD3JFk-mNT4Q/edit?usp=sharing>

$$E[t_{\text{fin}}] \propto \frac{V^i}{N^j T^k} \quad (16)$$

For ease of presentation, we convert Equation (16) to base 10:

$$\begin{aligned} \log_{10}(E[t_{\text{fin}}]) &= \log_{10}\left(\frac{aV^i}{N^j T^k}\right) \\ &= \log_{10} a + i \log_{10}(V) - j \log_{10}(N) - k \log_{10}(T) \end{aligned} \quad (17)$$

If we set, $C_1 = \log_{10} a$ then,

$$\log_{10}(E[t_{\text{fin}}]) = C_1 + i \log_{10}(V) - j \log_{10}(N) - k \log_{10}(T) \quad (18)$$

and exponentiating,

$$E[t_{\text{fin}}] = 10^{C_1 + i \log_{10}(V) - j \log_{10}(N) - k \log_{10}(T)} \quad (19)$$

For Case 0, where numbers of searchers and targets maintain a constant density as volume increases (i.e., $V \propto N$ and $T \propto N$), we further simplify Equation (18) because N , V , and T scale isometrically, and thus we can substitute a single scaling exponent to reflect the scaling in terms of N :

$$\log_{10}(E[t_{\text{fin}}]) = C_1 + \mu \log_{10}(N)$$

We use linear regression on log-transformed simulation data to determine the value of the scaling exponent μ . This provides the final scaling relationship between t_{fin} and N for Case 0.

$$E[t_{\text{fin}}] = 10^{C_1 + \mu \log_{10}(N)}$$

We fit the scaling constant C_1 to the simulated data.

Similarly, for Case 1 where there is a single target ($\log_{10}(1)=0$), N increases and V is constant ($i \log_{10}(V)=C_2$), we simplify Equation (18) as,

$$\begin{aligned} \log_{10}(E[t_{\text{fin}}]) &= C_1 + C_2 - j \log_{10}(N) - 0 \\ &= C_{1'} - j \log_{10}(N) \end{aligned}$$

where $C_{1'} = C_1 + C_2$. We again use linear regression on log-transformed simulation data to determine the value of the scaling exponent j . We use the same procedure to estimate scaling exponents for Cases 2 and 3 for t_{fin} ; and for t_{∞} for Cases 1–3.

3. RESULTS: COMPARING MODEL PREDICTIONS TO SIMULATED DATA

The t_{fin} approach to estimating the initial first contact times was developed for Case 0 in which a uniform random distribution of searchers and targets is placed in a volume with a constant density so that N and T increase linearly with V . The key assumption in t_{fin} is that the initial first contact times (the first time any one of the searchers comes into contact with a target) are exponentially distributed within each simulation. This allows us to predict that the first time the *first* searcher in a given simulation will contact a target is $\frac{1}{N}$. We found that in all Case 0 simulations, an exponential distribution was a good fit. We compare with a Gaussian to demonstrate the difference in sum of squares error and Akaike information criteria (Table 2).

In Figure 3, we show example probability distributions from two sets of simulations. These and the other simulations show a close fit to an exponential distribution. Given that the simulation validated the assumption that first contact times are exponentially distributed for Case 0, we then test the prediction that $t_{\text{fin}} \propto \frac{1}{N}$ from Equation (5). In Figures 4–7 we show model predictions compared with simple regressions through the data. We show only the predicted scaling relationship to N or V by fitting a constant to adjust the height of the line to fit the data. For t_{fin} , we chose a scaling constant to fit the height of the line to correspond to the smallest N , and for t_{∞} , we chose the scaling constant to fit the largest simulated N . Thus, the figures should be interpreted as demonstrating how simulations compare with model predictions for how initial first contact times change systematically with N or V . The fitting procedure is detailed in Section 2.5.

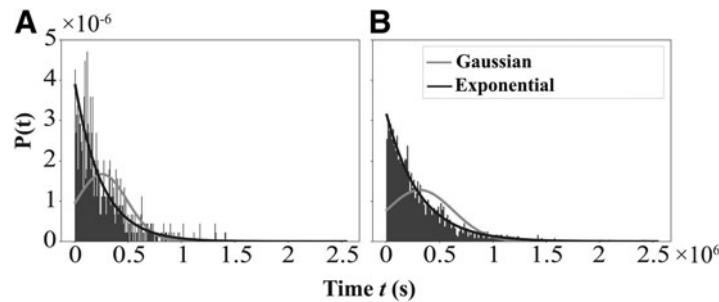


FIG. 3. Distribution of contact times of all searchers within one volume for Case 0. We show the distribution of the first contact times in two (out of 30) simulations of Case 0. (A) Shows the distribution of 629 first contacts for 629 searchers with 2533 targets in a 921 mm^3 search volume. (B) Shows the distribution of 1989 first contact times for 1989 searchers with 8011 targets in a 3640 mm^3 search volume. Data are plotted as a histogram of probabilities of the initial first contact time of each searcher on the y-axis for the contact time given on the x-axis. The theoretical probability density functions of a Gaussian distribution and an exponential distribution for the same mean and standard deviation as the data are shown for comparison. The data are visually consistent with an exponential distribution. The results of the goodness-of-fit tests (sum of squares error and Akaike information criteria) in Table 2 are consistent with our assumption of the exponential distribution of the first contact times.

Figure 4 shows a very close correspondence between the t_{fin} prediction and the simulated data. We also show that initial first contact times are orders of magnitude faster than the mean first contact times. The mean search time is (unsurprisingly) unaffected by N , and it is far less variable. This demonstrates that our model for t_{fin} accurately predicts $\frac{1}{N}$ scaling for its intended use case, and that initial first contact times are substantially faster than mean contact times. We now test whether t_{fin} can be extended to three search

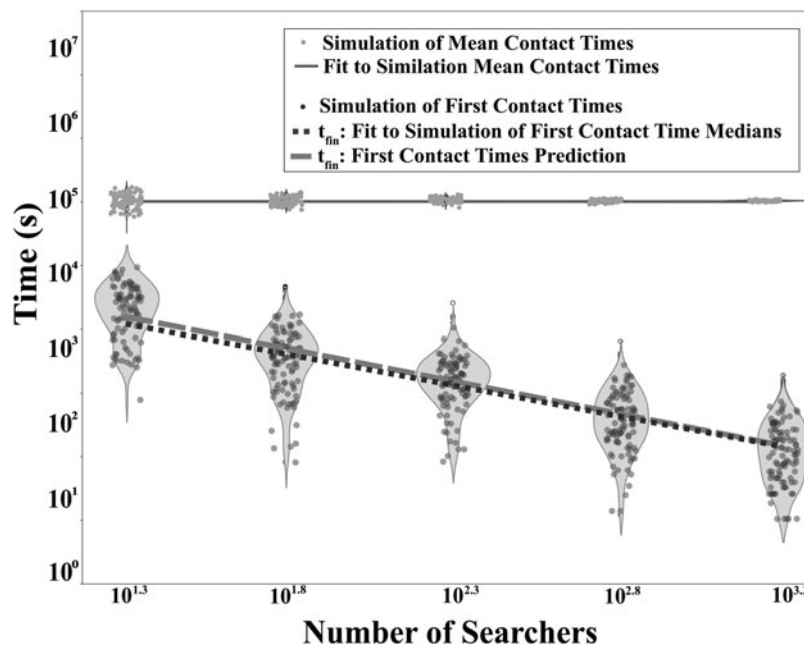


FIG. 4. t_{fin} predictions compared to simulations corresponding to Case 0: constant density, multiple targets. Each *dark circle* represents the time for the first searcher, in a population of N searchers, to contact its target, where N is varied along the x-axis. The initial first contact time for each replicated simulation is shown as *dots*. t_{fin} is predicted to scale as N^{-1} (light dashed line, Eq. (5)). The dark dotted line shows the regression through the simulated data: $E[t_{\text{fin}}] \propto N^{-1.05}$ with a 95% CI $[-1.1 \text{ to } -0.98]$ for the exponent, consistent with the exponent of -1 predicted by t_{fin} . Variance among the initial first contact times for a given N ranges from 0.25 to 0.36. In contrast, the mean search times (light dots), averaged over all N searchers in each simulation, across the same values of N , are much larger and do not vary with N , and have extremely low variance, all less than 0.003. One hundred replicates are used to calculate the mean first contact time and the initial first contact time for each value of N ; thus, there are 100 *light* and 100 *dark* points for each value of N .

scenarios with a single target and all searchers placed at the same initial location. We compare t_{fin} predictions with t_{∞} predictions that assume an infinite number of searchers. For Case 1, we compare t_{fin} and t_{∞} predictions given a single target in a constant search volume while increasing N across simulations.

This was the scenario t_{∞} was designed for, assuming N approaches infinity. Figure 5 shows that for small $N < 10^3$, initial first contact times decrease approximately linearly with N , consistent with our t_{fin} prediction that initial first contact times are proportional to $\frac{1}{N}$; for large $N \geq 10^3$, simulations are approximately consistent with the t_{∞} prediction, proportional to $\frac{1}{\ln(N)}$ [Eq. (13)]. Thus, even in this search scenario that differs from the original scenario envisioned by the t_{fin} approach, t_{fin} approximates initial first contact times for biologically relevant numbers of searchers, while t_{∞} better approximates initial first contact times for larger N .

For Case 2, we compare t_{fin} and t_{∞} predictions to simulations where N is fixed and V increases. Figure 6 shows that the simulations are very close to the t_{∞} prediction that time is proportional to L^2 , and considerably less steep than the t_{fin} prediction, which is proportional to L^3 . Similarly, the simulation tests of Case 3 in which both N and V increase are considerably closer to the t_{∞} prediction ($L^2/\ln(N)$) than the constant t_{fin} prediction.

4. DISCUSSION

4.1. Summary

We developed an analytical model, t_{fin} , designed to predict how quickly the first of many searchers finds its first target, given that searchers and targets are distributed randomly in a volume. Simulations validate the key t_{fin} assumption that search times are exponentially distributed (Fig. 3) and the key prediction that initial first contacts scale as $\frac{1}{N}$ (Fig. 4) for Case 0 with a uniform random distribution of searchers and targets that are kept at constant density as volume increases. To compare t_{fin} with another model of initial first contact times (t_{∞}), we conduct a set of simulations where all searchers start at the same location and

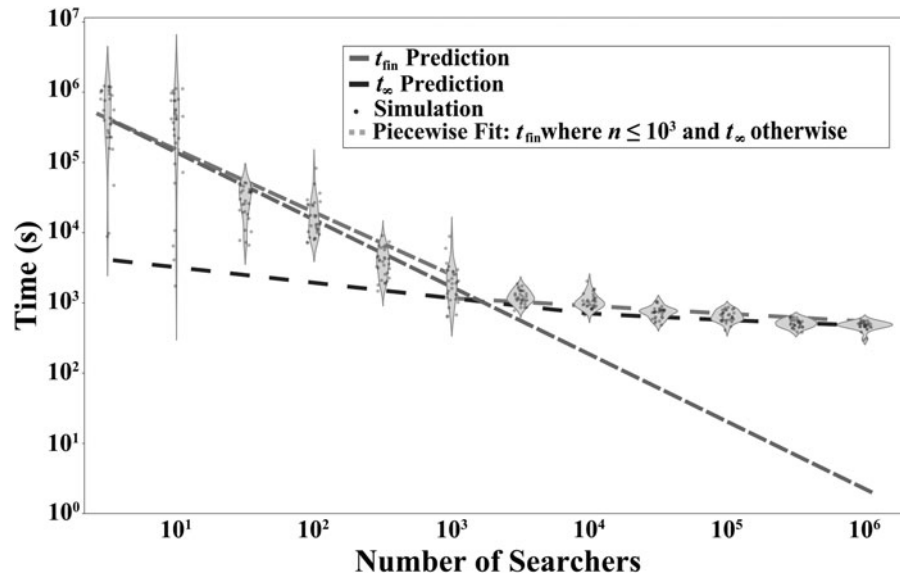


FIG. 5. t_{fin} and t_{∞} predictions compared to simulations corresponding to Case 1: single target, constant volume, increasing N . Each circle represents the time for the first searcher, in a population of N searchers, to contact its target, where N is varied along the x -axis. Thirty simulations were run for each value of N , and the first contact in each simulation is shown. The top dashed line connects the medians of the simulated data. The light dashed line shows the t_{fin} prediction [Eq. (7)] and the horizontal dashed line shows the t_{∞} prediction [Eq. (13)], both fitted to the simulated data. The regression through the simulated first contact points $N < 10^3$ is $N^{-1.1}$ with 95% CI $[-0.86 \text{ to } -1.3]$ consistent with the t_{fin} prediction for small N . The regression through data (top dashed line) where $N \geq 10^3$ gives $t_{\infty} \propto \log(N)^{-1.94}$ with a 95% CI $[-2.1 \text{ to } -1.8]$, which is slightly steeper than the $1/\log(N)$ prediction of t_{∞} . Prediction lines are shown with scaling constants set to fit the median of either the smallest (for t_{fin}) or largest (for t_{∞}) simulated data.

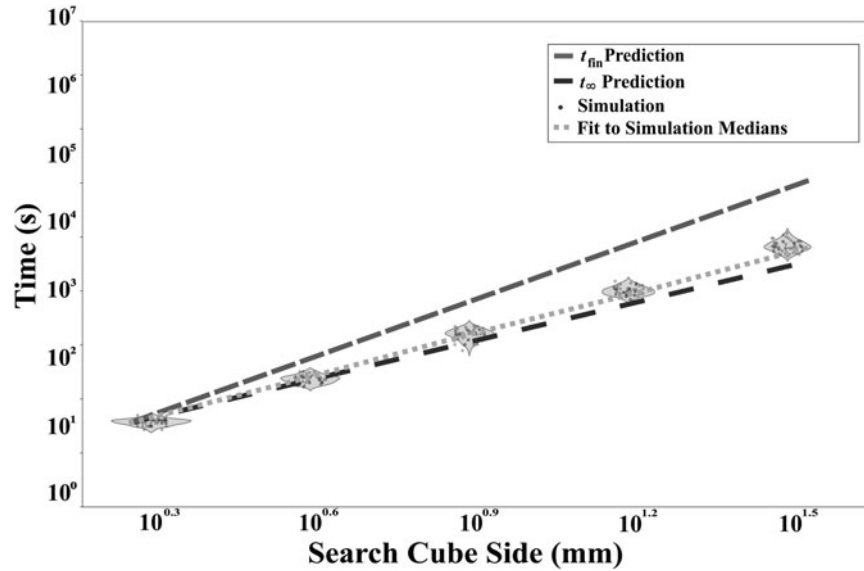


FIG. 6. t_{fin} and t_{∞} predictions compared to simulations corresponding to Case 2: single target, increasing volume, constant N . Symbols are replicated from Figure 5. The *top dashed line* shows the $t_{\text{fin}} \propto L^3$ prediction and the *lower dashed line* shows the $t_{\infty} \propto L^2$ prediction. The regression through the simulated data is $t \propto L^{2.17}$ with 95% CI [2.14–2.2], close to the t_{∞} prediction.

search for a single target. We find that for Case 1 (where the distance between searchers and target is fixed), the $t_{\text{fin}} \propto \frac{1}{N}$ prediction is consistent with simulations for small N , and the $t_{\infty} \propto \frac{1}{\ln(N)}$ prediction is close to simulations for large N (Fig. 5).

In Cases 2 and 3, we systematically increase the distance between the initial placement of searchers and the target. The initial first contact times are close to t_{∞} predictions and quite far from t_{fin} predictions (Figs. 6 and 7). Our analysis of t_{fin} in Case 0 suggests that the fastest searcher is one that is lucky in two respects: it happens to be placed near a target, and it takes a relatively direct path to the target. Because t_{fin} predictions do not hold in Cases 2 and 3 when distance increases, this suggests that $1/N$ scaling only holds when distances between searchers and targets are fixed. We suggest that the existence of lucky first

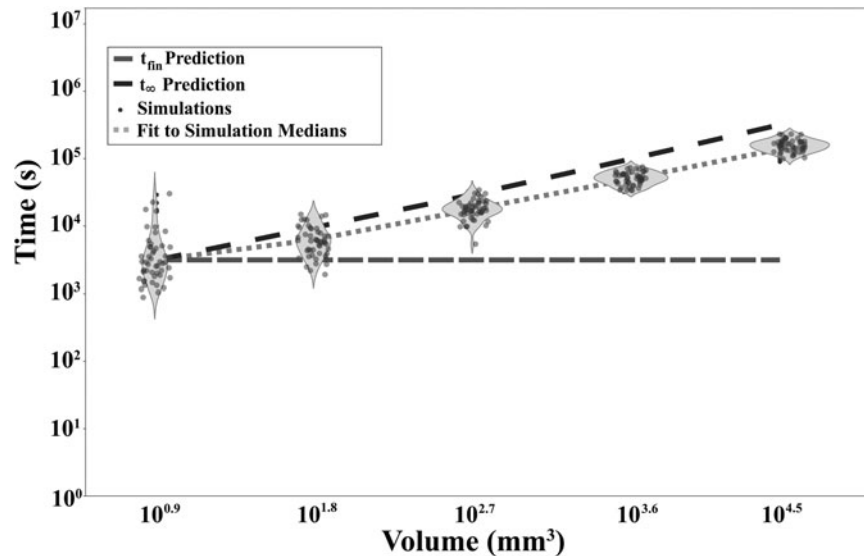


FIG. 7. t_{fin} and t_{∞} predictions compared to simulations corresponding to Case 3: single target, increasing volume, increasing N . Symbols are replicated from Figure 5. The *lower dashed line* shows the $t_{\text{fin}} \propto 1$ prediction and the *top dashed line* shows the $t_{\infty} \propto L^2/\ln(N)$ prediction.

searchers is an important benefit of collective search, particularly when the success of the first searcher causes important downstream events.

4.2. Interpretation of scenarios

In this article, we simulated idealized searches that correspond to the volume of lymph nodes across mammal sizes with estimates of the numbers of T cell searchers and dendritic cell targets within those lymph nodes, where the first T cell that finds a dendritic cell presenting cognate antigen initiates the adaptive immune response motivated by ongoing work (Ferdous et al., 2024). However, we suggest that the initial first contact times are relevant to a broader range of biological phenomena.

To better characterize the circumstances under which t_{fin} makes valid predictions, we compare our analytical predictions with the similar Extreme First Passage Time models developed by physicists to describe the time for a population of searchers to find a single target. That theory (that we call t_{∞}) uses the useful simplifying assumptions of an infinite number of searchers diffusing in one dimension to find a single target. Specifically, we compare the t_{fin} predictions to t_{∞} predictions extended by Lawley (Lawley and Madrid, 2020) to predict the initial first contact times in three dimensions. While t_{fin} predicts that larger populations of searchers find their targets proportionally faster, t_{∞} predicts a much slower logarithmic speedup in search times as N increases.

We compared t_{∞} predictions with those of t_{fin} under three experimental cases. In Case 1, when we hold the size of the search volume constant (and consequently the distance between the initial location of the searchers and the target is held constant), we find that simulated initial first contact times are consistent with the linear decrease with N predicted by t_{fin} when N is less than 1000; for $N > 1000$, simulations are consistent with the t_{∞} logarithmic prediction (Fig. 5). This result suggests that when the only variation across experiments is N , the t_{fin} predictions are reasonable approximations for biologically relevant numbers of searchers. We note that in this case the key assumption holds that search times are exponentially distributed among the searchers within each simulation.

However, there is a crossover, such that t_{∞} applies when N approaches infinity. Once N is sufficiently large there is only a logarithmic benefit to increasing N . We expect that the values of N for which the $\frac{1}{N}$ scaling regime holds depend on the details of how far apart the searchers are placed from the target, the behavior of searchers at the boundaries, and the dimension of the system (i.e., whether searchers are in 2D or 3D). Some of these issues are explored in Ro and Kim (2017), which also found a transition from $\frac{1}{N}$ to $\frac{1}{\ln(N)}$ scaling although they presented the transition in more general dimensionless terms rather than for a particular number of searchers. We leave additional analysis to future work and suggest that interdisciplinary approaches among physicists, biologists, and computer scientists are needed to further understand which factors most influence initial first contact times in different scenarios with biologically relevant numbers of searchers.

In Cases 2 and 3 when we simulate increasing distances between the target and initial searcher placements, t_{fin} predictions are not close to simulated results (Figs. 6 and 7). We predict that t_{fin} does not hold in these cases because the most important factor in these cases is that the distance between searchers and targets increases across simulations. This sheds light on an important feature of Case 0, where searchers and targets are distributed at random and with constant density. In Case 0, the distances between the initial placement of searchers and targets are determined by a random process, and therefore, when there are more searchers and targets in a larger volume, the *closest* distance between a searcher and the nearest target will decrease.

Thus, the lucky first searcher will be one that happens to be very close to its target as well as one that happens to take a relatively direct path to that nearby target. The t_{∞} predictions developed by physicists only consider the directness of the path because that approach assumes a fixed distance.

The $\frac{1}{N}$ scaling of initial first contact times suggests a substantial and underappreciated advantage of collective search. Case 0 suggests that in a bigger lymph node with 100 times more T cell searchers and dendritic cell targets, T cell activation would occur 100 times faster. Case 1 suggests that a bee colony (or in 2D, an ant colony) with 100 more searchers starting in a single nest and foraging independently in a fixed-size territory would find a single resource (e.g., a rare patch of flowers or food) 100 times faster. These advantages would not directly accrue to the average searcher, which would not find a target any faster if it continued to search independently. It is only the lucky first searcher that finds a target faster in a larger population.

However, if there is any signal (e.g., waggle dance or pheromone communication) or structural change (T cell replication or changes in influx and efflux to and from lymph nodes) following the first target discovery, then the subsequent searchers could modify their search to take advantage of that information. This is

consistent with findings from Donaldson-Matasci et al. (2013) that larger bee colonies found resources faster and that the overall foraging rates were greater in larger colonies. Similarly, the first lucky mutation in a population or the lucky first receptor to bind in a cellular interaction means that larger populations of individual agents have a lucky first encounter faster and the whole system can benefit from that lucky first encounter. Thus, the $\frac{1}{N}$ initial first contact time confers a significant advantage to a large population size.

However, we note a caveat that real biological systems may deviate in important ways from the simple simulations we used here. For example, the density of cognate T cells may not be constant in lymph nodes, and forager territory may scale nonlinearly with the forager number. The transition from search times that scale as $\frac{1}{N}$ to search times that scale with $\frac{1}{\ln(N)}$ is also important for analyzing the costs and benefits of increased population size. In many realistic scenarios, more searchers can lead to a proportionally faster search, but as size increases that benefit may diminish. This suggests that, combined with other trade-offs, initial first contact times may contribute to an optimal colony size under different conditions.

Our findings are relevant not only for understanding collective search in biology but also for engineering collective searchers in swarm robotics, a field that takes inspiration from collective search in biology (Dorigo et al., 2020; Fricke et al., 2013; Hecker and Moses, 2015; Lu et al., 2020; Talamali et al., 2021). Understanding the transition from linear to logarithmic dependence on the number of searchers and the shift from cubic to squared initial distance between the searchers and the target dependence with increasing search volume points to important constraints on scalability is a key concern in swarm robotics.

We suggest that swarm robotic systems can both benefit from an understanding of initial first contact times and serve as real-world testbeds that can help to develop and refine biologically relevant theories about how initial first contact times depend on numbers, search areas, or volumes and the distribution of searchers and targets in space. Direct biological data to test the initial first contact time predictions are scarce. In another work (Ferdous et al., 2024), we show that t_{fin} predictions are consistent with the time to initiate adaptive immunity, but the details of how quickly each individual T cell finds its target are poorly understood. We hope that this article will encourage biologists to report initial first contact times in empirical studies.

ACKNOWLEDGMENTS

We thank the University of New Mexico Center for Advanced Research Computing, supported, in part, by NSF, for high-performance computing resources. We thank Sid Redner, Judy Cannon, and the Moses Biological Computation Lab for helpful discussions that informed this article. We thank the anonymous reviewers for particularly helpful and insightful suggestions.

AUTHORS' CONTRIBUTIONS

J.F. wrote the software, performed computational experiments, and created the figures under the supervision of G.M.F. and M.E.M. J.F., G.M.F., and M.E.M. wrote the article. M.E.M. supervised J.F. on the selection of simulation parameters and place our results in context. J.F. derived the mathematical expressions under the supervision of G.M.F. and M.E.M.

AUTHOR DISCLOSURE STATEMENT

The authors declare they have no conflicting financial interests.

FUNDING INFORMATION

The authors thank NSF for funding on Awards #2020247, #2030037, and #2024520.

REFERENCES

- Adler FR, Gordon DM. Information collection and spread by networks of patrolling ants. *Am Nat* 1992;140(3):373–400; doi: 10.1086/285418
- Anderson C, McShea DW. Individual versus social complexity, with particular reference to ant colonies. *Biol Rev* 2001;76(2):211–237.
- Basnayake K, Schuss Z, Holcman D. Asymptotic formulas for extreme statistics of escape times in 1, 2 and 3-dimensions. *J Nonlinear Sci* 2019;29:461–499; doi: 10.1007/s00332-018-9493-7
- Beekman M, Sumpter DJT, Ratnieks FLW. Phase transition between disordered and ordered foraging in pharaoh's ants. *Proc Natl Acad Sci U S A* 2001;98(17):9703–9706; doi: 10.1073/pnas.161285298
- Bénichou O, Coppey M, Moreau M, et al. Optimal search strategies for hidden targets. *Phys Rev Lett* 2005;94:198101; doi: 10.1103/PhysRevLett.94.198101
- Breitwieser L, Hesam A, De Montigny J, et al. Biodynamo: A modular platform for high-performance agent-based simulation. *Bioinformatics* 2022;38(2):453–460.
- Celli S, Day M, Müller AJ, et al. How many dendritic cells are required to initiate a t-cell response? *Blood* 2012;120(19):3945–3948; doi: 10.1182/blood-2012-01-408260
- Donaldson-Matasci MC, DeGrandi-Hoffman G, Dornhaus A. Bigger is better: Honeybee colonies as distributed information-gathering systems. *Anim Behav* 2013;85(3):585–592; doi: 10.1016/j.anbehav.2012.12.020
- Dorigo M, Theraulaz G, Trianni V. Reflections on the future of swarm robotics. *Sci Robot* 2020;5(49):eabe4385; doi: 10.1126/scirobotics.abe4385
- Dornhaus A, Klügl F, Oechslein C, et al. Benefits of recruitment in honey bees: Effects of ecology and colony size in an individual-based model. *Behav Ecol* 2006;17(3):336–344; doi: 10.1093/beheco/arj036
- Dornhaus A, Powell S, Bengston S. Group size and its effects on collective organization. *Annu Rev Entomol* 2012;57(1):123–141; doi: 10.1146/annurev-ento-120710-100604
- Ferdous J, Fricke GM, Cannon JL, et al. Bigger is faster: The scalable adaptive immune response. *arXiv* 2024.
- Ferdous J, Fricke GM, Moses ME. Modeling immune search through the lymphatic network. In: *Swarm Intelligence*. (Dorigo M, et al. eds.). Springer International Publishing: Cham, 2022; pp. 332–340.
- Flanagan TP, Letendre K, Burnside WR, et al. Quantifying the effect of colony size and food distribution on harvester ant foraging. *PLoS One* 2012;7(7):1–9; doi: 10.1371/journal.pone.0039427
- Fricke GM, Asperti-Boursin F, Hecker J, et al. From microbiology to microcontrollers: Robot search patterns inspired by T cell movement. In: *Artificial Life Conference Proceedings, volume ECAL 2013: The Twelfth European Conference on Artificial Life of ALIFE 2023: Ghost in the Machine: Proceedings of the 2023 Artificial Life Conference*. MIT Press: Cambridge, MA, USA, 2013; pp. 1009–1016; doi: 10.1162/978-0-262-31709-2-ch151
- Fricke GM, Letendre KA, Moses ME, et al. Persistence and adaptation in immunity: T cells balance the extent and thoroughness of search. *PLOS Comput Biol* 2016;12(3):1–23; doi: 10.1371/journal.pcbi.1004818
- Hecker JP, Moses ME. Beyond pheromones: Evolving error-tolerant, flexible, and scalable ant-inspired robot swarms. *Swarm Intell* 2015;9:43–70; doi: 10.1007/s11721-015-0104-z
- Krummel MF, Bartumeus F, Gérard A. T cell migration, search strategies and mechanisms. *Nat Rev Immunol* 2016;16(3):193–201; doi: 10.1038/nri.2015.16
- Lawley SD. Distribution of extreme first passage times of diffusion. *J Math Biol* 2020a;80(7):2301–2325; doi: 10.1007/s00285-020-01496-9
- Lawley SD. Extreme first-passage times for random walks on networks. *Phys Rev E* 2020b;102:062118; doi: 10.1103/PhysRevE.102.062118
- Lawley SD. Universal formula for extreme first passage statistics of diffusion. *Phys Rev E* 2020c;101(1):012413; doi: 10.1103/PhysRevE.101.012413
- Lawley SD, Madrid JB. A probabilistic approach to extreme statistics of Brownian escape times in dimensions 1, 2, and 3. *J Nonlinear Sci* 2020;30(3):1207–1227; doi: 10.1007/s00332-019-09605-9
- Lu Q, Fricke GM, Erickson JC, et al. Swarm foraging review: Closing the gap between proof and practice. *Curr Robot Rep* 2020;1:215–225.
- Moses ME, Cannon JL, Gordon DM, et al. Distributed adaptive search in T cells: Lessons from ants. *Front Immunol* 2019;10:1357; doi: 10.3389/fimmu.2019.01357
- Perelson AS, Wiesel FW. Scaling aspects of lymphocyte trafficking. *J Theor Biol* 2009;257(1):9–16; doi: 10.1016/j.jtbi.2008.11.007
- Popp S, Dornhaus A. Ants combine systematic meandering and correlated random walks when searching for unknown resources. *iScience* 2023;26(2):105916; doi: 10.1016/j.isci.2022.105916
- Redner S. *A guide to first-passage processes*. Cambridge University Press, 2001.
- Ro S, Kim YW. Parallel random target searches in a confined space. *Phys Rev E* 2017;96:012143; doi: 10.1103/PhysRevE.96.012143

- Talamali MS, Saha A, Marshall JAR, et al. When less is more: Robot swarms adapt better to changes with constrained communication. *Sci Robot* 2021;6(56):eabf1416; doi: 10.1126/scirobotics.abf1416
- Weiss GH, Shuler KE, Lindenberg K. Order statistics for first passage times in diffusion processes. *J Stat Phys* 1983;31:255–278; doi: 10.1007/BF01011582
- Wiegel FW, Perelson AS. Some scaling principles for the immune system. *Immunol Cell Biol* 2004;82(2):127–131; doi: 10.1046/j.0818-9641.2004.01229.x

Address correspondence to:

Dr. Melanie E. Moses
Department of Computer Science
The University of New Mexico
1 University of New Mexico
Albuquerque, NM 87131
USA

E-mail: melaniem@unm.edu