Flow-mediated olfactory communication in honeybee swarms

Dieu My T. Nguyen[a,b], Michael L. Iuzzolino[a], Aaron Mankel[c], Katarzyna Bozek[d,e], Greg J. Stephens[d,f], and Orit Peleg[a,b,g,1]

[a]Department of Computer Science, University of Colorado Boulder, Boulder, CO 80309; [b]BioFrontiers Institute, University of Colorado Boulder, Boulder, CO 80309; [c]Department of Physics, University of Colorado Boulder, Boulder, CO 80309; [d]Biological Physics Theory Unit, Okinawa Institute of Technology, Okinawa 904-0495, Japan; [e]Center for Molecular Medicine Cologne, University of Cologne, 50931 Cologne, Germany; [f]Department of Physics and Astronomy, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands; and [g]Santa Fe Institute, Santa Fe, NM 87501

Edited by Gene E. Robinson, University of Illinois at Urbana–Champaign, Urbana, IL, and approved February 3, 2021 (received for review June 11, 2020)

Honeybee swarms are a landmark example of collective behavior. To become a coherent swarm, bees locate their queen by tracking her pheromones. But how can distant individuals exploit these chemical signals, which decay rapidly in space and time? Here, we combine a behavioral assay with the machine vision detection of organism location and scenting (pheromone propagation via wing fanning) behavior to track the search and aggregation dynamics of the honeybee *Apis mellifera* L. We find that bees collectively create a scenting-mediated communication network by arranging in a specific spatial distribution where there is a characteristic distance between individuals and directional signaling away from the queen. To better understand such a flow-mediated directional communication strategy, we developed an agent-based model where bee agents obeying simple, local behavioral rules exist in a flow environment in which the chemical signals diffuse and decay. Our model serves as a guide to exploring how physical parameters affect the collective scenting behavior and shows that increased directional bias in scenting leads to a more efficient aggregation process that avoids local equilibrium configurations of isotropic (nondirectional and axisymmetric) communication, such as small bee clusters that persist throughout the simulation. Our results highlight an example of extended classical stigmergy: Rather than depositing static information in the environment, individual bees locally sense and globally manipulate the physical fields of chemical concentration and airflow.

honeybee | olfactory communication | signal propagation | computer vision | agent-based model

Animals routinely navigate unpredictable and unknown environments in order to survive and reproduce. One of the prevalent communication strategies in nature is conducted via volatile signal communication, for example, pheromones (1, 2). As the range and noise tolerance of information exchange is limited by the spatiotemporal decay of these signals (3, 4), animals find creative solutions to overcome this problem by leveraging the diffusivity, decay, and interference with information from other individuals (5–8).

For olfactory communication, honeybees use their antennae to recognize and respond to specific odors. Recent studies have revealed the bees' distinct electrophysiological responses to different chemicals with quantifiable preferences (9). Olfactory communication with pheromones is crucial for many coordinated processes inside a honeybee colony, such as caste recognition, regulating foraging activities, and alarm broadcasting (10–12). Studies have shown that the queen mandibular pheromone regulates gene expression in the brains of workers, inducing changes in downstream behaviors, such as nursing and foraging (13). Among worker bees, adult foragers produce ethyl oleate, a chemical inhibitory factor that plays a role in delaying foraging in younger workers (14). In this work, we study how bees use pheromones in the context of swarm formation. To become a coherent swarm, honeybees must locate their queen by tracking her pheromones that decay rapidly in time and space. How can honeybees that are far away from the queen locate her? The specific mechanisms of the collective signal propagation strategy are still unknown.

The mechanism to locating the queen involves a behavior called "scenting," where bees raise their abdomens to expose the Nasonov pheromone gland and release the chemicals (15–17) (Fig. 1*A* and Movie S1, *Left*). This common behavior is also seen in other scenarios, such as scenting at the hive's entrance to orient foragers coming back to the hive and on food locations (e.g., flowers) (18, 19). So far, scenting has only been characterized at the level of an individual bee. We investigate how this behavior manifests at the level of a group of bees performing the task of localizing the queen for swarming.

In the traditional chemical signaling chemotaxis and quorum sensing scheme, the produced chemical signal by an individual is isotropic, or nondirectional and axisymmetric, as seen in early embryonic development (20, 21) and aggregation of amebae in *Dictyostelium* (22–24). Conversely, scenting bees create a directional bias by fanning their wings, which draws air along the pheromone gland along their anteroposterior axis (Fig. 1*A*). Thus, in addition to diffusion or the transport of the chemicals through random motion, the pheromone signals are also subject to the process of advection, that is, the transport of the chemicals by air motion created by the wing fanning. We show that, when bees perform the scenting behavior, they fan

Significance

We show that bees locate their queen by performing a cascade of "scenting" events, where individual bees direct their pheromone signals by fanning their wings. The bees create a dynamic spatiotemporal network that recruits new broadcasting bees over time, as the pheromones travel a distance that is orders of magnitude the length of an individual. We develop high-throughput machine learning tools to identify the locations and timings of scenting events, and demonstrate that these events integrate into a global "map" that leads to the queen. We use these results to build an agent-based model that illustrates the advantage of the directional signaling in amplifying the pheromones, thus leading to an effective search and aggregation process.

Author contributions: O.P. designed research; D.M.T.N. and A.M. performed research; M.L.I., K.B., and G.J.S. contributed new reagents/analytic tools; D.M.T.N., A.M., and O.P. analyzed data; and D.M.T.N. and O.P. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: orit.peleg@colorado.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2011916118/-/DCSupplemental.

Published March 23, 2021.

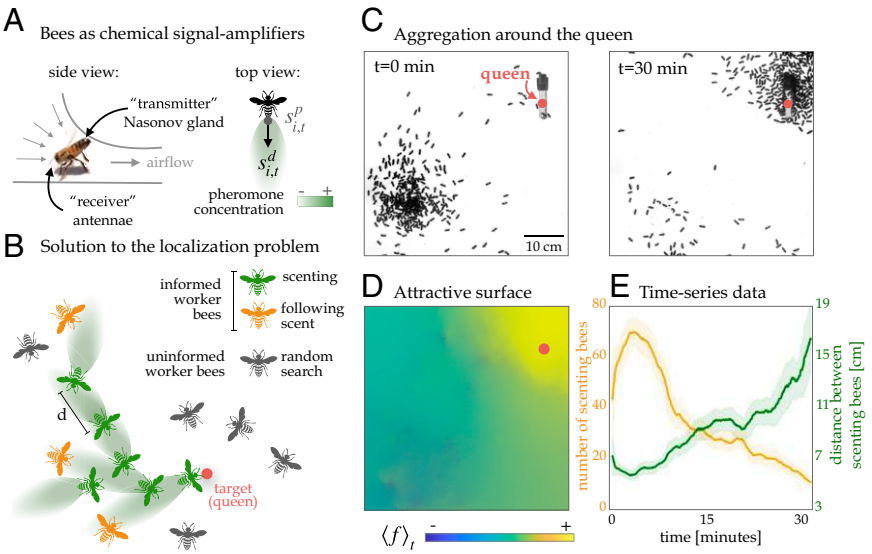

Fig. 1. Honeybees use directed volatile communication to solve a localization problem of finding their queen. (*A*) Pheromone sensing bees amplify the signal by opening the Nasonov gland in their abdomen and fan their wings to transmit their pheromones backward (a behavior called "scenting"). In the top view of the scenting bee, $\mathbf{s}^p_{i,t}$ is the position of a given bee, and $\mathbf{s}^d_{i,t}$ is the unit vector representing the direction of scenting. (*B*) Solution to the queen localization problem: By amplifying the queen's pheromones by emitting their own pheromones directionally, some worker bees (green) transmit the volatile signal over long distances while keeping a certain distance *d* from one another. Other informed worker bees (orange) follow the scent up the pheromone gradient. Uninformed bees (gray) perform a random search. (*C*) Example of a typical experiment where, at $t = 0$, the queen is confined to a cage located at a corner of the arena, and the worker bees are located at the farthest corner. After $\sim$30 min, the bees aggregate around the queen's cage. (*D*) The mean reconstruction of the attractive surface *f* according to *SI Appendix*, Eq. **S2**, averaged over 30 min. The mutual information (*SI Appendix*, Eq. **S4**) between this surface and the density of the bees at the end of the experiment illustrates the correlation between the scenting directions and the aggregation location of the bees. (*E*) The average number of scenting bees and the average distance between scenting bees over time. The aggregation process is accompanied by a sharp increase and gradual fall in the number of scenting bees. In contrast, the distance between scenting bees is lowest at the peak in number of scenting bees at the beginning and gradually increases over time.

their wings to direct the signal away from the queen and toward informed bees following the scent and the rest of the uninformed swarm (Fig. 1*B*). This directional bias increases the probability that distant bees may sense those amplified pheromones, upon which they also stop at a certain distance from the scenting bee and amplify the signal along their own anteroposterior axis. The combination of detection and "rebroadcast" leads to a dynamic network that recruits new broadcasting bees over time, as the pheromones now travel a distance that is orders of magnitude the length of an individual. What are the dynamics in the number of scenting bees during the swarming process? Is there a characteristic distance that defines the pheromone detection range for scenting bees? What are the advantages of a directional communication strategy vs. an isotropic one? And how do the bees harness the physics of directed signals to create an efficient communication network? To address these questions, we combined a behavioral assay, machine learning solutions for organism tracking and scenting detection and computational agent-based modeling of the communication strategy in honeybees and characterize the advantage of collective directional scenting.

Materials and Methods

To quantify the correlation between the scenting behavior of the bees, localization of the queen, and the aggregation process, we established a behavioral assay in which worker bees search for a stationary caged queen in a semi–two-dimensional (2D) arena (see *SI Appendix*, section A for more details). We recorded the search and aggregation behavior of the bees from an aerial view for 1 h to 2 h (see Movie S1 for a closeup example scenting bee with her abdomen pointed upward and wing-fanning behavior in this experimental arena, in contrast with a nonscenting bee standing still). To extract data from the recordings, we then developed a markerless, automatic, and high-throughput analysis method using computer vision methods and convolutional neural networks (CNNs) (25). This pipeline allows us to detect individual bees as well the positions and orientations of scenting bees (see *SI Appendix*, section B and Fig. S2 for more details, and see Movie S2 for a movie of the example in *SI Appendix*, Fig. S2 *B* and *C*).

To identify the unifying behavioral principles that harness the dynamics of volatile signals, we developed a model that captures two important physical dynamics surrounding individual bees, substance advection–diffusion and sensing local concentration gradients, while they perform search and identification (26). Our agent-based model is embedded in a flow environment where individuals can sense local concentrations of pheromones and propagate them backward as well as move up the gradient (see *SI Appendix*, section F and Fig. 3 for full model details).

Behavioral Assay Results

Localization Is Correlated with Scenting. In the 2D experimental arena, worker bees placed at the opposite corner from the queen explore the space and eventually aggregate around the queen's cage after $\sim$30 min (Fig. 1*C*). For each scenting bee i at time t, we collected its position, $\mathbf{s}^p_{i,t}$, and direction of scenting, $\mathbf{s}^d_{i,t}$ (unit vector). We then correlated the scenting events with the spatiotemporal density of bees in the arena by treating $\mathbf{s}^p_{i,t}$ and $\mathbf{s}^d_{i,t}$ as a set of gradients that define a minimal surface of height $f(x, y, t)$. This surface height $f(x, y, t)$ corresponds to the probability that a randomly moving nonscenting bee will end up at position (x, y) by following the local scenting directions of scenting bees (see *SI Appendix*, section E for formal definitions and derivations).

To show that the attractive surface $f(x, y, t)$ is correlated with the final concentration of bees $\rho(x, y, t')$, we compute the normalized mutual information, $\mathrm{NMI}(\langle f \rangle_t ; \langle \rho(t_{end}) \rangle)$, between the attractive surface averaged over the entire experiment (Fig. 1*D*) and the density of the bees averaged over the last 5 min of the experiment. The mutual information measures the information that the two variables, $\langle f \rangle_t$ and $\langle \rho(t_{end}) \rangle$, share, and determines how much knowing one variable reduces uncertainty about the other. The NMI is scaled between zero (no mutual information)

and one (perfect correlation). We averaged the density of bees over the last 5 min of the experiment to capture the density of the entire group and avoid discrete peaks of density resulting from movement of individual bees (see *SI Appendix*, section E for definitions). For the experiment shown in Fig. 1 *C–E*, the NMI is 0.21. The top right region in the arena with maximal $\langle f \rangle_t$ corresponds to the location around the queen, illustrating the correlation between the scenting directions and the queen localization of the bees. We performed $N = 14$ experiments of various group sizes (rounded to the nearest 10), and report an average NMI of 0.11 ($\sigma^2 = 0.002$). See *SI Appendix*, Table S1 for values for individual experiments, and see *SI Appendix*, Fig. S3 for pairs of the average frame and the average attractive surface for all experiments.

A Characteristic Distance between Scenting Bees. The aggregation process is accompanied by a sharp increase in the average number of scenting bees, followed by a slow decay, until there are few scenting bees and the majority of the bees are aggregated around the queen (Fig. 1*E*, orange curve). See Movie S3 for the experiment shown in Fig. 1 *C–E*, with the attractive surface reconstruction and time series data. We also characterize the temporal dynamics of the distance between neighboring scenting bees, which are treated as adjacent points in the Voronoi diagram for each frame (see example diagram in *SI Appendix*, Fig. S2 *E* and *F* and section C for more details). Throughout the growth in the number of scenting bees, the distance between scenting bees decreases to a minimal value, then increases as the number of scenters decreases (Fig. 1*E*, green curve).

To show the reproducibility of the behavioral assay and assess the effect of group size in the aggregation process, we tested 14 group sizes ranging from approximately 180 to 1,000 bees. Three example experiments with $N_{\text{bees}} = 320$, 500, and 790 are shown in Fig. 2*A* and Movies S4–S6. Each example includes three snapshots showing the process of aggregation around the queen's cage over 60 min. The average numbers of scenting bees over time for all experiments of various group sizes are shown in Fig. 2*B*. Generally, there are more scenting bees over time as the number of bees in the arena increases. Across densities, we observe a typical characteristic of a sharp increase in scenting bees at the beginning as the bees initially search for the queen, and a slow decrease as they slowly aggregate around the queen's cage. Based on observations, we assume that bees in tight clusters, such as those at the very beginning, do not usually scent, as they are not sufficiently spread out to fan their wings.

The time evolution of the average distance from all bees in the arena to the queen is shown in Fig. 2*C*. There are no clear distinctions between the temporal dynamics of this distance as a function of density. For most densities, there is a sharp drop in this distance to the queen very early on (~100 s to 200 s), corresponding to the early increase in number of scenting bees. The average distance to the queen then gradually falls as bees make their way to her vicinity.

We also measured the distance between neighboring scenting bees over time for all group sizes (Fig. 2*D*). This distance shows a characteristic gradual increase over time. At the beginning of the aggregation process at the peaks in scenting bees, this distance is fairly constant across all sizes, ranging from 5.00 cm to 7.35 cm, with an outlier from the lowest group size experiment at 9.06 cm (Fig. 2*E*). A linear model is fitted to the data (excluding the outlier) with the slope $m = 0.00031\ \text{cm}/\text{n}_{\text{bees}}$. The presence of this constant distance between nearest-neighbor scenting bees

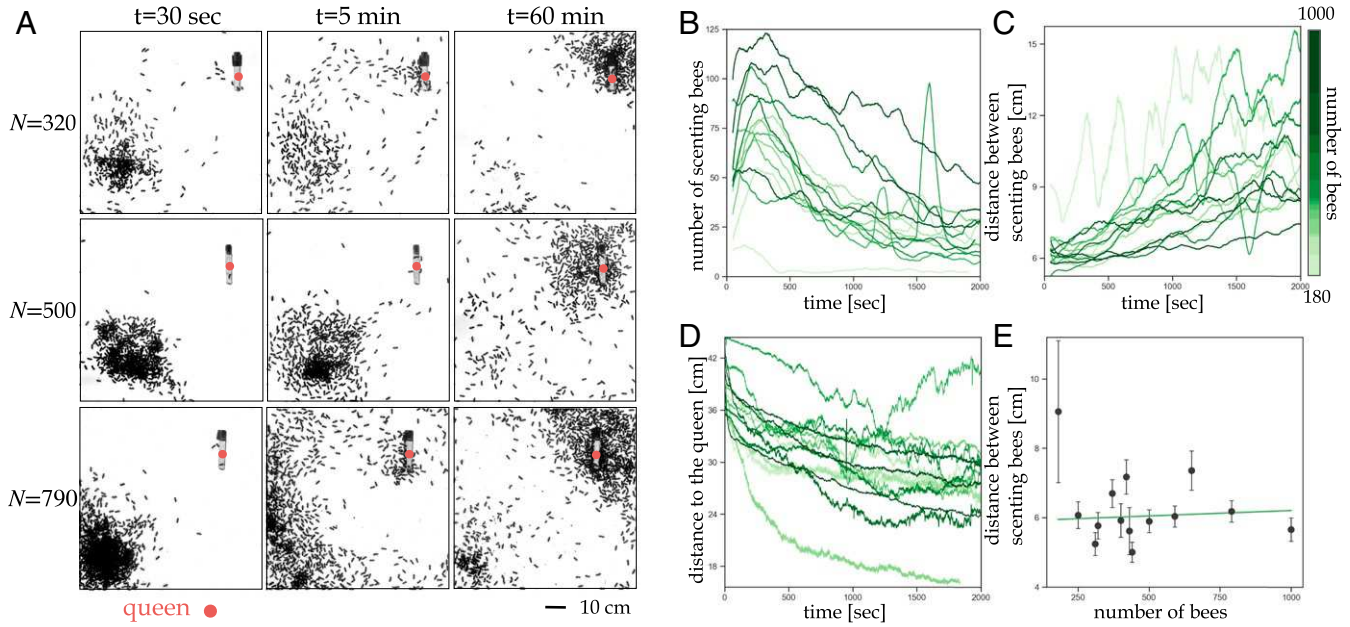

Fig. 2. Distance between scenting bees is invariant to bee density. (*A*) Example experiments of various bee densities, each with three snapshots showing the process of aggregation around the queen's cage over 60 min. (*B*) The average number of scenting bees over time for all experiments of various densities. There are more scenting bees with higher number of bees in the arena. Across densities, we observe a typical characteristic of a sharp increase in scenting bees at the beginning and decrease over time. (*C*) The average distance to the queen over time for all experiments of various densities. This distance gradually decreases as bees aggregate around the queen. The positive correlation between distance to queen and number of scenting bees is evidence for the functional importance of scenting events to the problem of queen finding. (*D*) The average distance between scenting bees over time for all experiments of various densities. Across densities, these distances are fairly constant at the beginning around the peaks in *B*, and increase over time. (*E*) The average distance between scenting bees at the peak in number of scenting bees as a function of density. Error bars represent the standard deviation over all of the distances for each density. The distance between scenting bees is approximately constant across densities, except for a low-density outlier. A linear model (green line) is fit to the data, excluding the outlier, with a slope of $m = 0.00031\ \text{cm}/\text{n}_{\text{bees}}$.

during the initial stage of the aggregation suggests the possibility of a pheromone concentration threshold that turns on scenting for individual bees in this collective communication network. As more bees aggregate around the queen, the bees collectively "turn off" scenting. The higher distance between scenting neighbors later in the recordings suggest that the few remaining scenting events are more stochastic.

These experimental results, scenting behavior with wing fanning to direct pheromones, the threshold-dependent triggering of this behavior, and a characteristic distance between scenting bees, serve as core ingredients for our agent-based model. We will use the model as a proof of concept of our proposal of the mechanistic localization behavior, described in more detail in the next section. The goal of our modeling is to test hypotheses of the mechanisms behind the phenomenon and explore the possible emergent patterns that arise to assess the effect of the directional signaling strategy employed by the bees.

Agent-Based Model Results.

Model Predicts Optimal Signal Propagation within a Range of Behavioral Parameters. We systematically explore a range of values for the directional bias w_b, and the threshold T, for various numbers of bees in the arena N. Parameters w_b and T are potential behavioral parameters that bees could adjust based on input from the environment. The directional bias w_b represents the magnitude of the directional advection–diffusion of pheromone released by a bee (see Fig. 3*B* for a comparison of isotropic [$w_b = 0$] vs. advection–diffusion [$w_b = 10$]). The threshold T is the local concentration of pheromone above which a bee is activated from the random walk (i.e., a bee scents or walks up the concentration gradient when this threshold is met). Each combination of the three parameters (w_b, T, and N) is a simulation repeated 20 times. All other parameters, including C_0 (the initial pheromone concentration), D (diffusion coefficient), and γ (decay constant) from *SI Appendix*, Eq. **6**, remain constant across all simulations.

We quantify aggregation around the queen through the average distance of worker bees to the queen, $\langle d(t)\rangle = 1/N \sum_{i=1}^{N} \sqrt{(x_i(t)-x_q)^2 + (y_i(t)-y_q)^2}$, where $x_i(t)$ and $y_i(t)$ are the x and y positions of worker bee i, respectively, and x_q and y_q are the x and y positions of the stationary queen bee. We also extract various other properties of the aggregation processes in the model: the queen's cluster size and the number of clusters as additional measures of efficiency, the distance to the queen from the farthest active bee (i.e., a bee that is scenting or performing the directed walk up the pheromone gradient) to assess how far signals as a function of w_b and T propagate, and the number of scenting bees as a measure of energy expenditure. Our model predicts four distinct phases for all densities, which are determined by (w_b, T). Note that "phase" does not refer to a period in a time sequence but to the dynamics and outcome observed from different combinations of the parameters.

Phase 1. For low values of both w_b and T, the bees aggregate into small clusters that are homogeneously spread throughout the simulation box (Fig. 4*E*). This is reflected by a sharp initial decrease and then gradual decrease of $\langle d(t)\rangle$ throughout the simulation (Fig. 4*A*), a consistently high number of scenting bees (Fig. 4*B*), a consistently small cluster of bees in the queen's vicinity (Fig. 4*C*), and a consistently high distance of the farthest active bee from the queen (Fig. 4*D*). A simulation is provided in Movie S7.

Phase 2. At higher values of T and w_b, only bees in the vicinity of the queen are activated and join the local cluster around the queen. This is reminiscent of diffusion-limited aggregation that results in a sparse fractal-like cluster around the queen (Fig. 4*F*). This phase is reflected by a gradual decrease in $\langle d(t)\rangle$ and a gradual increase in the number of scenting bees, the queen's cluster size, and the distance of the farthest active bee from the

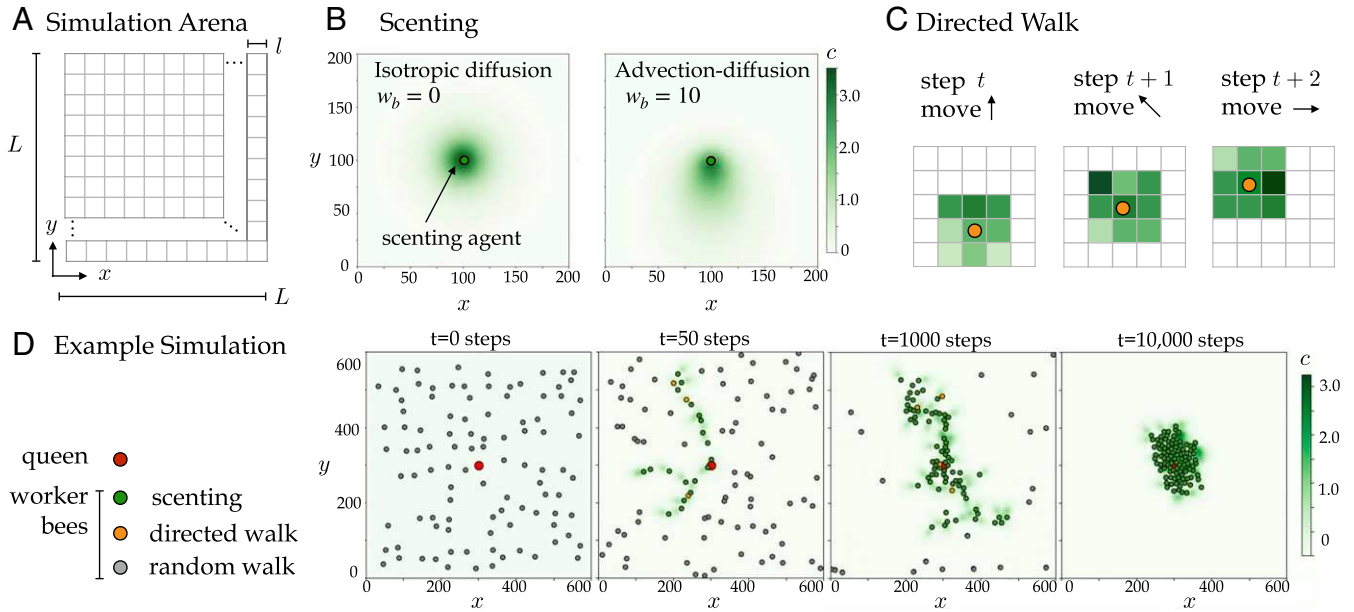

Fig. 3. Scenting model. (*A*) Our $L \times L$ 2D simulation box is discretized into $l \times l$ sized pixels. (*B*) Scenting bees can produce directional bias (example on *Right* with $w_b = 10$), compared to no directional bias on *Left*. (*C*) When a bee detects a local pheromone concentration that is above the activation threshold ($C(x, y, t) > T$), the bee calculates the concentration gradient around it (using the nearest nine pixels, highlighted in different shades of green) and walks up gradient toward higher pheromone concentrations. (*D*) Example of a simulation showing a series of snapshots. The queen is shown as a red filled circle, and worker bees are shown as a red filled circle colored by their internal state: scenting (green), activated (orange), and nonactivated (gray). Activated bees perform a directed walk up the pheromone gradient, and nonactivated bees perform a random walk. The instantaneous pheromone concentration $C(x, y, t)$ corresponds to the green color scale. Simulation parameters are $N = 100$, $w_b = 30$, $T = 0.01$ $C_0 = 0.0575$, $D = 0.6$, and $\gamma = 108$.

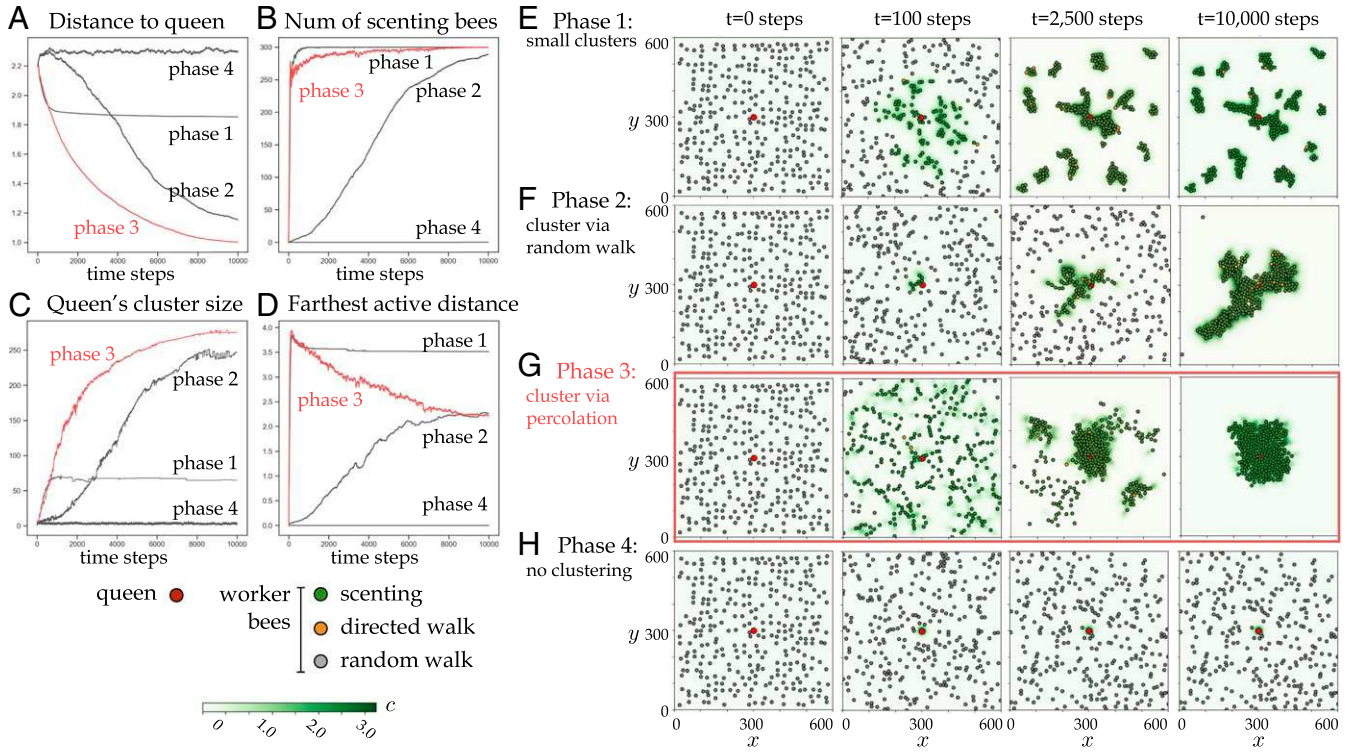

Fig. 4. Directional bias is associated with optimal search and aggregation in the scenting model. (*A*) The average distance of the worker bees to the queen as a function of time steps, for examples of the four different phases. (*B*) The average number of scenting bees as a function of time steps, for examples of the four different phases. (*C*) The average queen's cluster size as a function of time steps, for examples of the four different phases. (*D*) The average distance of the farthest active bee to the queen as a function of time steps, for examples of the four different phases. (*E–H*) Example of a simulation from the four different phases, showing a temporal series of snapshots. The queen is shown as a red filled circle, and worker bees are shown as a red filled circle colored by their internal state: scenting (green), performing a directed walk up the pheromone gradient (orange), and performing a random walk (gray). The instantaneous pheromone concentration C(x,y,t) corresponds to the green color scale. Simulation parameters are $N = 300$, $C_0 = 0.0575$, $D = 0.6$, and $\gamma = 108$. For phase 1, $w_b = 0$, $T = 0.01$ (*E*); for phase 2, $w_b = 10$, $T = 0.5$ (*F*); for phase 3, $w_b = 30$, $T = 0.05$ (*G*); and, for phase 4, $w_b = 0$, $T = 1.0$ (*H*). Phase 3, which is associated with the fastest aggregations around the queen, is highlighted in red.

queen as bees slowly cluster via the random walk (Fig. 4 *A–D*). A simulation is provided in Movie S8.

Phase 3. At low values of T and high values of w_b, the activated bees create a percolating network of senders and receivers of the pheromone signal (Fig. 4*G*). This combination of T and w_b ensures the fastest aggregation process around the queen and the fastest growth of the queen's cluster (Fig. 4 *A* and *C*). This process keeps most bees active with the scenting task throughout the simulation (Fig. 4*B*). Although bees in phases 2 and 3 eventually cluster at the queen's location, the pheromone signals in phase 3 reach a much farther distance at the beginning than in phase 2, where bees slowly cluster only via the random walk (Fig. 4*D*). A simulation is provided in Movie S9.

Phase 4. When the activation threshold T is high enough, no worker bees are activated, and no clusters are formed (Fig. 4*H*). In this phase, the bees simply perform a random walk. A simulation is provided in Movie S10.

The existence of a phase 3 in the computational model suggests that using a directional signal is advantageous, as this phase does not exist in the absence of directional bias ($w_b = 0$). Although, in both phase 2 and phase 3, the bees are able to aggregate around the queen with similar values of the average final distance to the queen (such as for $N = 100$, where the values are 0.56 and 0.53 units, respectively, in *SI Appendix*, Fig. S5*D*), bees in phase 3 are able to reach a plateau distance to the queen earlier (on average at $t = 5{,}117$ compared to $t = 6{,}864.6$ in phase 2, in *SI Appendix*, Fig. S5*E*) while requiring less scenting events to reach that plateau (on average, 5,566.18 events compared to 7,732.50 in phase 2, in *SI Appendix*, Fig. S5*F*).

The Spatial–Temporal Shapes of Clusters in Experiments and Simulations. To connect the results of the model and the experiments, we analyze the clusters that change in shape and size over time in the experiments (for densities ranging from $N = 250$ to 400) and simulations (for density $N = 100$ where the percolation networks are best observed as elongated clusters). See *SI Appendix*, section I and Fig. S9 for how we isolate and define the clusters for this analysis.

To quantitatively compare the experiments to the different phases in the model, we measure farthest active distance, cluster area, cluster circularity, and cluster angle to the queen and show the time series of these attributes in Fig. 5. We exclude phase 4 simulations from these analyses, as the virtual bees are never activated, and no clusters form. First, in Fig. 5*A*, we show the distance to the queen from the farthest active bee (in experiments: scenting bees; in model: scenting or bee walking up the gradient) as a measure that differentiates the phases based on how far the signals propagate. In phase 1, this distance plateaus at a relatively high value when bees remain in small clusters throughout the arena. In phase 2, this distance plateaus at a relatively low value as the bees must be closer to the queen to be activated, and signals do not propagate as far as in phase 3, where we see a high initial distance as the percolation network forms, and a gradual decrease as bees cluster around the queen. The temporal dynamics of the experimental curve (green) are similar to the phase 3 simulation. Note that the oscillations in the experimental

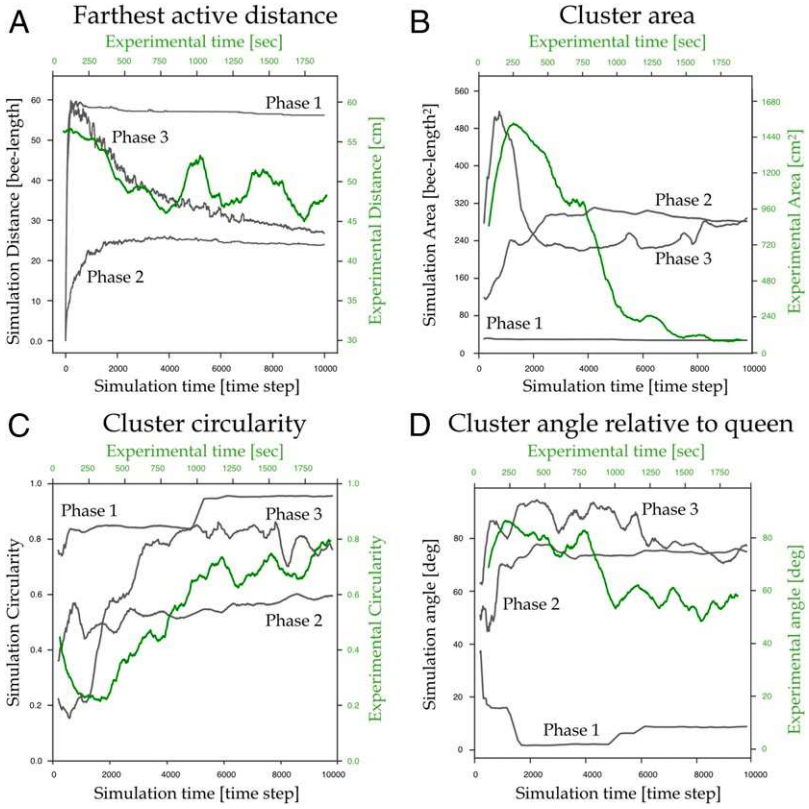

Fig. 5. Cluster properties of the experiments and simulations. For all panels, the green curve shows the average in experiments, and labeled gray curves show the average in different phases of the simulations. (*A*) The distance of the farthest active bee to the queen as a measure of how far the signals propagate. (*B*) The cluster area over time. (*C*) The cluster circularity over time. (*D*) The cluster angle relative to the queen over time.

series are likely due to the spontaneous scenting events that may come from bees far away and not yet in the queen's cluster. For each cluster attribute, we also measure the Pearson correlation between the experiments and the simulations in each phase, presented in *SI Appendix*, Table S4. For the farthest active distance, the coefficients between the experiments and the simulations in phases 1, 2, and 3 are −0.075, −0.662, and 0.759, respectively. For all attributes, the number of samples used to compute the correlation is $N = 1{,}975$ (simulation data are down-sampled to match the number of data points in the experimental data).

Second, we quantify the cluster area for both the experiments and the simulations (Fig. 5*B*). In phase 1, the cluster around the queen is very small, as most bees are stuck in local equilibria. In phase 2, the cluster gradually increases and plateaus at a relatively intermediate size. In phase 3 and in the experiments, the cluster area quickly increases as the percolation network activates, and gradually decreases as bees approach and swarm around the queen. Shown in *SI Appendix*, Table S4, the coefficients between the experiments and the simulations in phases 1, 2, and 3 are 0.527, −0.784, and 0.762, respectively.

Third, we quantify the circularity of the clusters over time, in Fig. 5*C*, defined as $(4 * Area * \pi)/(Perimeter^2)$. A perfect circle has a circularity value of one. In phase 1, the cluster around the queen forms quickly and stays very small and relatively round, with a high circularity. Phase 2 simulations tend to have irregular clusters that form via bees finding the queen via the random walk, resulting in a relatively low to intermediate circularity values. In phase 3, we observe the clusters starting at relatively low circularity as the percolation networks activate, and becoming more circular over time as the compact cluster around the queen forms. The experimental clusters gradually increase in circularity similar to phases 2 and 3. Shown in *SI Appendix*, Table S4, the coefficients between the experiments and the simulations in phases 1, 2, and 3 are 0.684, 0.716, and 0.790, respectively.

Finally, in Fig. 5*D*, we show the angular deviation of the cluster orientations from the orientation leading to the queen. This relative angle plateaus for phases 1 and 2, in which the process of clustering around the queen is less dynamic and the cluster's final shape is determined earlier. In phase 3 and the experiments, the cluster gradually orients toward the queen over time. We compute the angular correlation between the relative angle time series using $\langle \cos(\theta_{experiment} - \theta_{simulation}) \rangle$. Shown in *SI Appendix*, Table S4, the coefficient between the experiments and the simulations in phases 1, 2, and 3 are 0.49, 0.94, and 0.96, respectively.

Overall, we compare the structural phases we observe in the simulations and the spatiotemporal structures formed by the bees in the experiments. By extracting the four properties of the clustering dynamics and comparing the experimental data to the simulation data from phases 1 through 3, we show that the clustering of the bees in our experiments tend to behave inconsistently with virtual bees in phases 1 and 4. Hence, the real-world system likely exists within phase 2 and/or phase 3 of the model.

The Effect of Bee Density on the Phase Boundaries. We use our model to construct phase diagrams for three different densities, low ($L^2/50$), medium ($L^2/100$), and high ($L^2/300$). Across densities, low T and low w_b result in phase 1 with multiple clusters (Fig. 6). In this phase, the range of T increases, while the range of w_b decreases as density increases. On the other hand, phase 4 is the result of high T and high w_b values. With more bees in the arena, other phases expand, while the range of phase 4 gets smaller.

Phase 2 is typically the result of intermediate T values and low to intermediate w_b values. These higher T values prevent

the percolation seen in phase 3, as the pheromone signals are not as far-reaching when bees are less sensitive to sense them to become propagators. Phase 3 with the successful aggregation via percolation is achieved with intermediate to high w_b and low T values. As density increases, the range of w_b and T for phase 3 is greater, as there are more bees to create and sustain the communication network in the arena. Phase 3 in $N = 50$ is $103,051$ pixels, or 15.9% of the total diagram; phase 3 in $N = 100$ is $137,282$ and 21.2%, respectively; and phase 3 in $N = 300$ is $321,807$ and 49.7%, respectively. When there is a higher density of bees, the bees are able to aggregate around the queen while not needing to be as highly sensitive (i.e., lower T value) to the signals. Across all densities, phase 3 never occurs when there is no directional bias (i.e., $w_b = 0$). This illustrates the importance of the directed signaling strategy to create an effective communication network within the model. The existence of a network of transmitters and receivers of pheromones that percolates across the entire computational simulation arena is crucial to achieve fast and successful aggregation around the queen.

Discussion and Conclusion

Combining experiments and high-throughput machine vision tracking of location and scenting behavior, we investigated the communication mechanisms that honeybee swarms employ to collectively locate their queen, a difficult problem given the limited information available from short-lived pheromone signals. We find that individual bees act as receivers and senders of signals by using the Nasonov scenting behavior, releasing pheromones from the glands and fanning their wings to direct the signals backward (Fig. 1 *A* and *B* and Movie S1, *Left*). In an arena with a caged queen, bees were quick to activate a collective communication network, as we see a sharp, early-time increase in the number of scenting bees (Figs. 1*E* and 2*B*). In this network, scenting bees stand at a characteristic distance from their neighbors while dispersing signals (Fig. 2*E*), which suggests a concentration threshold in the activation mechanism of individual bees' scenting behavior. We show that the scenting events are highly correlated with the collective aggregation around the queen (Fig. 1*D* and *SI Appendix*, Fig. S3). Together, these experimental results provide testable hypotheses of the mechanisms of this collective communication strategy and of whether the threshold-dependent directional signaling behavior is advantageous—concepts that we explore with the agent-based model.

The experimental findings guide our design and implementation of an agent-based model of the queen localization phenomenon. We model individual bees as agents with simple rules for movement, detection of pheromone based on a concentration threshold, directed signaling, and chemotaxis to move up the local gradient (Fig. 3). The bee agents are not aware of the queen's location or of the global pheromone profile in the flow environment. We show that, by only local interactions, these agents are able to aggregate around the queen mostly quickly and efficiently when they implement directed signaling within a range of bias values (i.e., $10 \leq w_b < 60$) (Fig. 4). When the density of bees increases in the simulated arena, effective aggregation can also occur with a wider range of T as opposed to being limited to lower T or lower sensitivity to pheromones (Fig. 6). Thus, with more bees active in the communication network, individuals can afford to be less sensitive while still achieving swarm formation. Overall, the four different phases that are present in the model show us the possible emergent properties of this honeybee phenomenon. Importantly, our modeling emphasizes the significance of the wing-fanning behavior, which allows for directed signaling and therefore more efficient aggregation in the swarm. In the absence of the directional bias where the signal is isotropic (i.e., $w_b = 0$), the successful aggregation in phase 3 is never achieved. Finally, we present a comparison of the spatial–temporal structures of the clusters in the experiments and the simulations, which demonstrates that the real-world behavior of the bees is inconsistent with phases 1 and 4 but does exhibit elements of phase 2 and phase 3. Additional experiments and simulations are required to make more precise distinctions. For example, an experiment with a single bee and a stationary queen can decouple the behavior of an individual from the swarm to tune the model parameters.

In addition to presenting mechanistic details about the queen localization phenomenon in honeybees, we have also developed an effective image analysis pipeline for markerless, automatic, and high-throughput honeybee detection and behavior recognition (*SI Appendix*, Fig. S2). Our approach uses state-of-the-art neural network models that are trained on our honeybee data and can be retrained and applied to other systems. Our high-throughput pipeline can be easily improved for future applications. First, the detection of individual bees currently employs a classical computer vision approach of simple Otsu's adaptive thresholding, morphological transformations, and connected components. While these methods are quick and

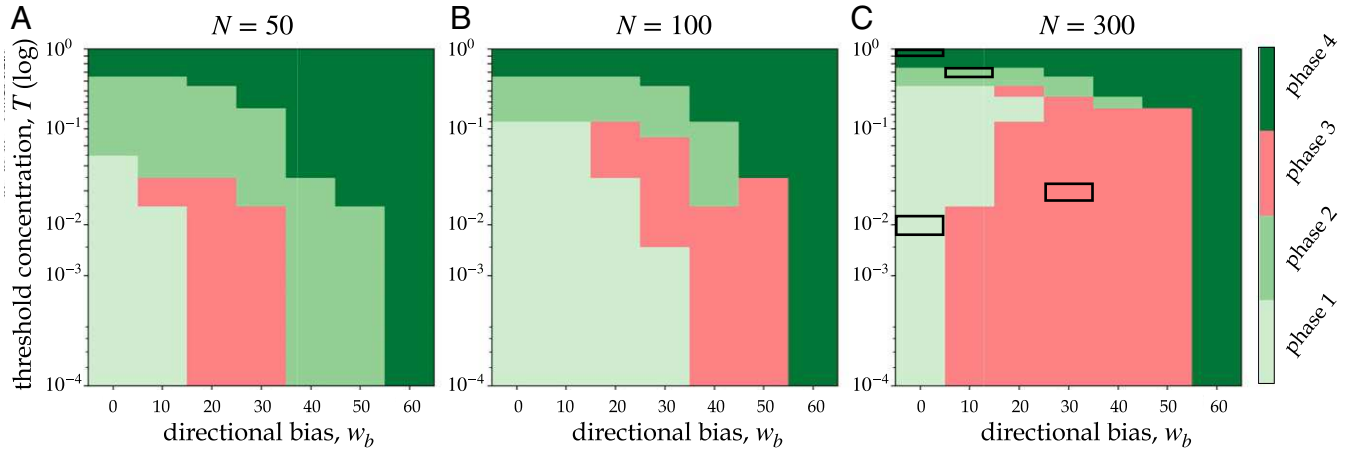

Fig. 6. The effect of bee density on phase boundaries. Phase diagrams constructed from scenting model dynamics using summary heat maps in *SI Appendix*, Fig. S5 as a function of T and w_b for all three densities, $N = 50$ (*A*), $N = 100$ (*B*), and $N = 300$ (*C*). The black rectangles indicate the example simulations we showed in detail in Fig. 4. The area of the optimal phase (phase 3) grows with N. Phase 3 in $N = 50$ is 103, 051 pixels, or 15.9% of the total diagram; phase 3 in $N = 100$ is 137, 282 and 21.2%, respectively; and phase 3 in $N = 300$ is 321, 807 and 49.7%, respectively.

sufficient to identify bees from background, they struggle to separate bees when they significantly touch or overlap in the image, a problem exacerbated by our backlight system. In future designs, we will modify our experimental setup to use different lighting systems where more features on bees are visible, allowing the usage of CNNs for the task of image segmentation to detect more individual bees in dense environment (e.g., ref. 27). Additionally, while using the wing angle as a proxy for scenting in static images is effective, scenting is a time-dependent behavior which could be more accurately identified using information from multiple frames. Temporal information can be incorporated using activity recognition networks (28, 29) on labeled videos of bees. This will require tracking individual bees over time to build a labeled dataset composed of short movies of the scenting behavior. We will explore recent efforts in automatic tracking of bees, such as works by ref. 30. The ability to track the scenting behavior of bees over time will allow us to answer interesting questions regarding the roles worker bees play in this swarming context. For example, does every bee eventually scent or does only a proportion of the same bees scent while others follow the pheromone trails to the queen?

Various future directions arise. The model will allow testable predictions about the resilience of the communication network. This concept includes assessing the effect of node failure via removal of some signaling bees, interference with a secondary signal via introducing artificial pheromone to the network, and the disruption of pheromone flows in the presence of wind. Experiments can then be performed to test the model's phenomenological predictions. Together, the tools will allow better understanding of how the dynamic nature of the network allows the swarm to overcome local obstacles such as solid objects, and deal with a nonstationary search target, as well as turbulent airflow and conflicting chemical signals.

From a physics perspective, our active system functions by coupling flows and forces in the presence of feedback (31–33). The individual building blocks (in this case, a bee) can sense their microenvironment (flow, forces, chemical content) and respond in a way that promotes survival; typically, the response changes the macroenvironment the individual is embedded in, thus creating a perpetual coupling between the individuals, the group, and the environment. From a biological perspective, our approach is an extension of the studies of classical stigmergy (34) wherein organisms deposit and respond to static information in the environment. In contrast, the bees in our system are able to sense local chemicals but also manipulate the global physical fields by actively directing signals with the scenting behavior. Harnessing the bees' natural solutions to communication—honed by eons of evolution, selection, and refinement—we can not only more deeply understand collective animal behavior but leverage that understanding to create bioinspired system designs in the fields of dynamic construction materials, swarm robotics, and distributed communication.

Data Availability. Computational code for the ML-based image analysis and for the agent-based model has been deposited in GitHub at https://github.com/peleg-lab/CollectiveScentingABM and https://github.com/peleg-lab/CollectiveScentingCV.

ACKNOWLEDGMENTS. This work was supported by the NSF Graduate Research Fellowship under Grant DGE 1650115 (D.M.T.N.), and NSF Physics of Living Systems Grant 2014212 (O.P.). Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. We also acknowledge funding from the University of Colorado Boulder, BioFrontiers Institute (internal funds), the Interdisciplinary Research Theme on Autonomous Systems (O.P.), Okinawa Institute of Science and Technology Graduate University Graduate University (K.B. and G.J.S.), and Vrije Universiteit Amsterdam (G.J.S.). We thank Seneca Kristjonsdottir and Christopher Borke for bee management, Gary Nave and Michael Neuder for assistance in developing image analysis pipeline, Aubrey Kroger and Emily Walker for annotating images, and Raphael Sarfati and Chantal Nguyen for reading and commenting on the manuscript. We thank Prof. L. Mahadevan, Prof. Massimo Vergassola, Prof. Gene E. Robinson, Prof. Olav Rueppell, and members of the O.P. laboratory for insightful feedback and discussions.

1. Y. L. Conte, Hefetz, Primer pheromones in social hymenoptera. *Annu. Rev. Entomol.* **53**, 523–542 (2008).
2. M. R. Kant *et al.*, Mechanisms and ecological consequences of plant defence induction and suppression in herbivore communities. *Ann. Bot.* **115**, 1015–1051 (2015).
3. A. Celani, E. Villermaux, M. Vergassola, Odor landscapes in turbulent environments. *Phys. Rev. X* **4**, 956–17 (2014).
4. J. K. Wilson, A. Kessler, H. A. Woods, Noisy communication via airborne infochemicals. *Bioscience* **65**, 667–677 (2015).
5. R. T. Cardé, M. A. Willis, Navigational strategies used by insects to find distant, wind-borne sources of odor. *J. Chem. Ecol.* **34**, 854–866 (2008).
6. O. Bénichou, C. Loverdo, M. Moreau, R. Voituriez, Intermittent search strategies. *Rev. Mod. Phys.* **83**, 81–129 (2011).
7. D. H. Gire, V. Kapoor, A. Arrighi-Allisan, A. Seminara, V. N. Murthy, Mice develop efficient strategies for foraging and navigation using complex natural stimuli. *Curr. Biol.* **26**, 1261–1273 (2016).
8. A. M. Hein, F. Carrara, D. R. Brumley, R. Stocker, S. A. Levin, Natural search algorithms as a bridge between organisms, evolution, and ecology. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9413–9420 (2016).
9. J. Zhao, Z. Li, Z. Zhao, Y. Yang, S. Yan, Electroantennogram reveals a strong correlation between the passion of honeybee and the properties of the volatile. *Brain Behav.* **10**, e01603 (2020).
10. M. Trhlin *et al.*, Chemical communication in the honeybee (*Apis mellifera* L.): A review. *Vet. Med.* **56**, 265–73 (2011).
11. T. Pankiw *et al.*, Queen mandibular gland pheromone influences worker honey bee (*Apis mellifera* L.) foraging ontogeny and juvenile hormone titers. *J. Insect Physiol.* **44**, 685–692 (1998).
12. Y. Lensky, P. Cassier, The alarm pheromones of queen and worker honey bees. *Bee World* **76**, 119–129 (1995).
13. C. M. Grozinger, N. M. Sharabash, C. W. Whitfield, G. E. Robinson, Pheromone-mediated gene expression in the honey bee brain. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 14519–14525 (2003).
14. I. Leoncini *et al.*, Regulation of behavioral maturation by a primer pheromone produced by adult worker honey bees. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 17559–17564 (2004).
15. J. A. Pickett, I. H. Williams, M. C. Smith, A. P. Martin, Nasonov pheromone of the honey bee, *Apis mellifera* L. (Hymenoptera, Apidae). Part III. Regulation of pheromone composition and production. *J. Chem. Ecol.* **7**, 543–554 (1981).
16. N. E. McIndoo, The scent-producing organ of the honey bee. *Proc. Acad. Nat. Sci. Philadelphia* **66**, 542–555 (1914).
17. J. M. Peters, N. Gravish, S. A. Combes, Wings as impellers: Honey bees co-opt flight system to induce nest ventilation and disperse pheromones. *J. Exp. Biol.* **220**, 2203–2209 (2017).
18. L. Bortolotti, C. Costa, "Chemical communication in the honey bee society" in *Neurobiology of Chemical Communication*, C. Mucignat-Caretta, Ed. (CRC Press/Taylor and Francis, 2014), pp. 147–210.
19. I. H. Williams, J. Pickett, A. Martin, The Nasonov pheromone of the honeybee *Apis mellifera* L. (Hymenoptera, Apidae). Part II. Bioassay of the components using foragers. *J. Chem. Ecol.* **7**, 225–237 (1981).
20. M. E. Cates, J. Tailleur, Motility-induced phase separation. *Annu. Rev. Condens. Matter Phys.* **6**, 219–244 (2015).
21. T. Bäuerle, A. Fischer, T. Speck, C. Bechinger, Self-organization of active particles by quorum sensing rules. *Nat. Commun.* **9**, 3232 (2018).
22. M. P. Brenner, L. S. Levitov, E. O. Budrene, Physical mechanisms for chemotactic pattern formation by bacteria. *Biophys. J.* **74**, 1677–1693 (1998).
23. H. Levine, Modeling spatial patterns in Dictyostelium. *Chaos* **4**, 563–568 (1994).
24. D. Kessler, H. Levine, Pattern formation in Dictyostelium via the dynamics of cooperative biological entities. *Phys. Rev. E* **48**, 4801–4804 (1993).
25. D. M. T. Nguyen, M. L. Iuzzolino, CollectiveScentingCV.GitHub. https://github.com/peleg-lab/CollectiveScentingCV. Deposited 15 June 2020.
26. D. M. T. Nguyen, M. L. Iuzzolino, O. Peleg, CollectiveScentingABM.GitHub. https://github.com/peleg-lab/CollectiveScentingABM. Deposited 11 June 2020.
27. K. Bozek, L. Hebert, A. S. Mikheyev, G. J. Stephens, "Towards dense object tracking in a 2D honeybee hive" in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE Computer Society, 2018), pp. 4185–4193.
28. A. Karpathy et al., "Large-scale video classification with convolutional neural networks" in *CVPR '14* (IEEE Computer Society, 2014), pp. 1725–1732.
29. J. Carreira, A. Zisserman, "Quo vadis, action recognition? A new model and the kinetics dataset" in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2017), pp. 6299–6308.
30. K. Bozek, L. Hebert, Y. D. Portugal, G. J. Stephens, Markerless tracking of an entire insect colony. bioRxiv [Preprint] (2020). https://doi.org/10.1101/2020.03.26.007302 (Accessed 15 May 2020).
31. Peleg, Mechanical hive mind. *Phys. Today* **72**, 66–67 (2019).
32. J. M. Peters, Peleg, Mahadevan, Collective ventilation in honeybee nests. *J. R. Soc. Interf.* **16**, 20180561 (2019).
33. Peleg, J. M. Peters, M. K. Salcedo, Mahadevan, Collective mechanical adaptation of honeybee swarms. *Nat. Phys.* **14**, 1193–1198 (2018).
34. G. Theraulaz, E. Bonabeau, A brief history of stigmergy. *Artif. Life* **5**, 97–116 (1999).