Context-dependent interaction leads to emergent search behavior in social aggregates

Colin Torney[a,b,1], Zoltan Neufeld[a], and Iain D. Couzin[b]

[a]School of Mathematical Sciences and Complex and Adaptive Systems Laboratory, University College Dublin, Belfield, Dublin 4, Ireland; and [b]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved November 3, 2009 (received for review July 15, 2009)

Locating the source of an advected chemical signal is a common challenge facing many living organisms. When the advecting medium is characterized by either high Reynolds number or high Peclet number, the task becomes highly nontrivial due to the generation of heterogeneous, dynamically changing filamental concentrations that do not decrease monotonically with distance to the source. Defining search strategies that are effective in these environments has important implications for the understanding of animal behavior and for the design of biologically inspired technology. Here we present a strategy that is able to solve this task without the higher intelligence required to assess spatial gradient direction, measure the diffusive properties of the flow field, or perform complex calculations. Instead, our method is based on the collective behavior of autonomous individuals following simple social interaction rules which are modified according to the local conditions they are experiencing. Through these context-dependent interactions, the group is able to locate the source of a chemical signal and in doing so displays an awareness of the environment not present at the individual level. This behavior illustrates an alternative pathway to the evolution of higher cognitive capacity via the emergent, group-level intelligence that can result from local interactions.

collective intelligence | olfactory search | cooperation

Throughout the natural world, organisms are constantly faced with the challenge of locating the resources required for their survival. Often this means navigating their environment based on spatiotemporally varying information such as advected chemical cues, thermal gradients, or magnetic fields. It has been noted that collective behavior can greatly assist animal navigation. One explanation for this, known as the "many wrongs" principle (1), is that inherent noise in the environment is dampened due to multiple sampling by individuals within a group. A quantitative study of an effect of this type was made by Grünbaum (2) and the benefits of sociality clearly illustrated. However, this effect does not fully capture the potential emergent properties of social aggregations, which often display complex behaviors entirely absent at the individual level (3–6). In this context, complex systems, such as fish schools or animal herds, can be viewed as information-processing entities with a collective awareness of their environment (7). Understanding their capacity for performing search tasks will have important consequences for the development of distributed technologies, such as olfactory robot swarms, with applications in the detection of explosives, landmines, or locating people in search-and-rescue operations (8–10).

The use of chemical signals by organisms in order to gain information about their environment is a ubiquitous behavior commonly seen in aquatic animals or flying insects and observed in a diverse range of species over a range of scales. For low Reynolds number, viscous environments chemotaxis in organisms such as *Escherichia coli* has been studied extensively in the context of chemical signalling pathways (11, 12), the fluid mechanics of locomotion (13), and optimal strategies (14). However, when chemoattractants are advected by stochastically fluctuating, chaotic flows, the mechanisms that allow organisms to respond effectively to these signals are poorly understood. This issue has been addressed previously, and algorithms based on a statistical (15) or information-theoretic (16) approach have been developed. Here we consider an approach that requires minimal cognitive or sensing capacity on the part of an individual and instead relies on interactions with conspecifics to generate a search response that is effective in tracking an advected chemical filament. The key mechanism, which may be generalized to other situations, lies in the continual adjustment of an individual's behavior, from being more or less independent to entirely following its neighbors, as a function of its level of confidence in its own current strategy. Consequently, the leadership structure of the group changes, continuously adapting to the quality of local information available to the members. For the search problem, the individual strategy is the presumed direction toward the target along the concentration filament, and confidence is assessed based on the changes in the olfactory signal sensed along the trajectory in the recent past.

It is assumed that the transport properties of the flow considered exhibit characteristics observed in chaotic advection that lead to thin filaments of chemical concentration in which the steepest gradients are transverse to the direction of the source and density is nonmonotonically decreasing with distance to the source location (i.e. patches of high concentration are advected downstream). The dynamics of the chemical field are represented by

$$\frac{\partial C}{\partial t} + \mathbf{v_f} \cdot \nabla C = S(\mathbf{r}_0) - bC, \qquad [1]$$

where $S(\mathbf{r}_0)$ is a constant source at a point in space, b is a decay rate, and $\mathbf{v_f}$ is a stochastic velocity field with an imposed mean flow along the y-axis. The velocity field is generated by a stochastic process in Fourier space with a prescribed energy spectrum (17, 18) that decays exponentially for large wavenumbers (see *SI Appendix* for details).

Interaction Rules

Individuals in our model are advected by the flow and move at a constant speed in a direction defined by their orientation

$$\dot{\mathbf{r}}_i = \mathbf{v_f}(\mathbf{r}_i, t) + v_s \mathbf{p}_i. \qquad [2]$$

The flow is simulated in a domain of length L, and a time scale unit for the model is selected by relating this length to a characteristic velocity of the model; the absolute velocity of an individual if it is perfectly aligned against the imposed mean flow. Therefore, $T = L/(v_s - \bar{v}_{fy})$; i.e., an individual moving against the mean flow will cross the domain in unit time. All further parameters are then defined in terms of these units.

The time evolution of the orientation vector results from interaction rules based on an abstraction of aggregation tendencies

Author Contributions: C.T., Z.N., and I.D.C. designed research; C.T. performed research; C.T., Z.N., and I.D.C. analyzed data; and C.T., Z.N., and I.D.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: ctorney@princeton.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0907929106/DCSupplemental.

observed in biological systems. A range of theoretical models have been introduced to describe the collective motion of animal groups (see e.g., refs. 19–21); here we assume that the preferred direction results from local alignment, repulsion, and attraction (5, 6, 22, 23). However, we assume the interaction zone over which an individual responds to neighbors varies as a function of current local conditions. This approach follows a similar concept outlined in ref. 24, which demonstrated that simulated fish schools were able to track regions of improved abiotic conditions by modifying reactions to conspecifics and reducing speed when located in preferred regions.

The direction of motion $\mathbf{p}_i$ is modified on the basis of the position and velocity of neighboring individuals. A desired direction $\mathbf{d}_i$ is defined by three social interaction rules. In order to maintain a minimum distance between neighbors, individuals move away from those within a repulsion zone Z_R (this precludes any other behavioral response)

$$\mathbf{d}_i = -\sum_{j \in Z_R} \frac{\mathbf{r}_j - \mathbf{r}_i}{|\mathbf{r}_j - \mathbf{r}_i|}. \quad [3]$$

If no others are present within the repulsion zone, individuals move toward those within an attraction zone Z_A and seek to align themselves with their neighbors within an orientation zone Z_O

$$\mathbf{d}_i = \sum_{j \in Z_A} \frac{\mathbf{r}_j - \mathbf{r}_i}{|\mathbf{r}_j - \mathbf{r}_i|} + \sum_{j \in Z_O} \mathbf{p}_j, \quad [4]$$

where the attraction zone is typically assumed to be larger than the alignment zone. Each individual turns toward $\mathbf{d}_i$ at a rate proportional to the difference between the current and desired direction with a maximum angular velocity defined by a parameter γ. If no neighbors are present within the interaction zone, the direction vector $\mathbf{p}_i$ remains unchanged.

Context Dependence

In order to couple environment to behavior and create an effective search algorithm, each individual varies the size of the interaction zones of local alignment and attraction depending on the current concentration value experienced. It should be noted that this is not a directional measure, no gradient is detected, and flow velocity is ignored; therefore at the level of a single individual, a search strategy does not exist.

The kernel of the model search behavior lies in the expansion and contraction of the attraction and orientation zones depending on the local concentration experienced by an individual. The radius of each zone is a function of a normalized individual concentration parameter $C_i(t) \in [0, 1]$ that measures the confidence in the actual direction of motion based on the recently encountered signal, defined as

$$C_i(t) = \frac{C(\mathbf{r}_i, t)}{\max_{0<\tau<t}[C(\mathbf{r}_i(t-\tau), t-\tau)e^{-\tau/\alpha}]}. \quad [5]$$

The denominator is a record of the history of the individual's trajectory, where α controls the decay time and acts as a short-term memory that allows individuals to assess the local concentration in the context of recent experience only. Note that this type of temporal comparison of sensed concentration values is also an essential component of the well-known run-and-tumble model of bacterial chemotaxis (25).

In our 2D model, an individual controls the radius of the attraction and orientation zone it responds to by applying a scaling factor to the maximum size of these values (R_A and R_O, respectively) dependent on the concentration parameter. The radius of interaction for each response is then defined as (see Fig. 2*B Inset*)

$$R_A(t) = (1 - C_i(t))^2 R_{A,max}$$
$$R_O(t) = \sin^2(\pi C_i(t)) R_{O,max}. \quad [6]$$

Although the exact functional form of this relation is reasonably arbitrary, the key behavioral response they represent can be summarized as

- a sharp decrease in signal results in attraction only. An individual then moves toward the center of the group within its maximum interaction zone;
- a weakening signal results in moderate attraction and strong alignment with neighbors. It is this intermediate response that allows net movement along the filament and prevents a low polarity swarm from forming;
- if an improving signal is being experienced, this is the optimal direction, interaction zones are reduced to zero, all neighbors are ignored, and current direction is maintained.

Qualitatively similar results to those presented here were obtained by using different functions for each scaling factor provided the dynamical responses listed were approximated.

By following the type of rules outlined, an individual acting within a group is able to effectively track a filament and locate the source of the signal, be it a nutrient source, a potential mate, or the location of a suitable habitat. Sequential snapshots of the search behavior being performed are shown in Fig. 1 for 60 individuals, and animations (Movies S1 and S2) are included in the *SI Appendix*.

This model is not aimed to be an accurate representation of the behavior of a particular species, and a range of other qualitatively similar interactions could be employed by different types of animal groups and in various environments. For example, varying the turning rates or some weighting factors of the alignment, attraction and actual direction as a function of some measure of confidence can lead to qualitatively similar results so long as there exists a local context dependence of the type we describe above.

Evidence that animals can, and do, employ rules of this type comes from experiments involving schooling fish where observed group-level behavior has been accurately modeled by implementing variable, context-dependent interaction ranges. In ref. 26, it was shown that fish appear to dynamically alter the range of interactions with other individuals according to locally perceived stimuli, such as increasing this range when alarm odor [indicative of predation risk (27)] is detected. Experimental studies of fish have also demonstrated that individuals will tend to restrict their schooling tendency when they can gather reliable direct information from the environment, such as the location of resources, but increase their tendency to group with others when this information is perceived to be unreliable or is scarce (28). Potential examples are not limited to aquatic animals, and a similar mechanism has been postulated as an explanation for the active recruitment of others to food patches by the cliff swallow *Hirundo pyrrhonota* (29). This hypothesis is supported by the increased frequency of recruitment signals when food patches are localized and advected by moderate wind speeds.

Our model captures behavioral interactions without specifying the sensory modalities used and in some systems individuals may alter their range of interaction by changing dynamically the dominant modality used or by changing their response within one or multiple modalities. The memory parameter that affects this change can be either a physiological constraint (e.g., the biochemical mechanism used by *E. coli*) and therefore optimized through evolutionary processes or a variable decision process about how to discount recent concentration values based on learned behavior. Although the sensory modalities used have not yet been explored in the examples concerning schooling fish, it is likely that vision plays a key role and both lateral-line (water-pressure detection)

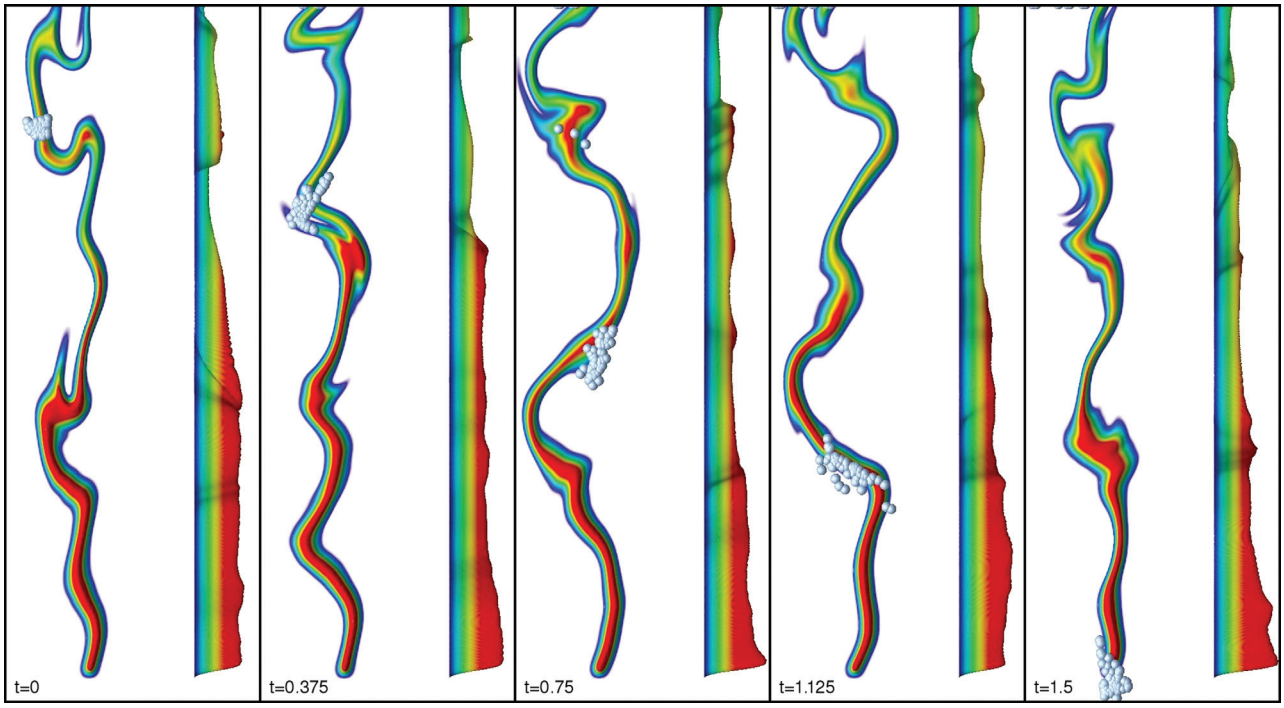


Fig. 1. Snapshots of 60 individuals performing filament tracking at intervals of 0.375. Fifty-three successfully locate the source, repulsion zone size 2×10^{-3}, $\alpha = 12.5 \times 10^{-3}$. Top view of filament and individuals is shown on left of each sequence, and on the right is the rotated concentration profile.

sensitivity and vision could be employed at different stages to respond to conspecifics at different distances (30).

Several testable predictions can be made in order to demonstrate a context-dependent mechanism in these example systems. If experimentalists can modify the modalities used by organisms (through genetic, pharmacological, or other means) such that they cannot alter their range dynamically, we expect that there will be characteristic and predictable deficiencies in response to environmental stimuli of the type we discuss here. Furthermore, existing experimental evidence demonstrates that the chemoattractive capability of schooling fish does not linearly improve with group size (31) as would be expected by the "many wrongs" principle, and it can therefore be hypothesized that this is a result of some nontrivial interactions between individuals. Although experiments that reproduce the predictions of the model are no guarantee that the exact nature of the interactions are fully understood, a success rate which is both a nonmonotonic function of group size and has a minimum threshold below which success rate drops would indicate that a higher-level functionality exists beyond the individual. Examining this optimum group size for different signal properties would also provide evidence for a group-level search mechanism.

Group-Size Effects

To further quantify this behavior and understand the role of the parameters in the model, we investigate first the effects of the group size on the probability of successfully locating the source. We refer here to group size as the number of individuals present at the beginning of each simulation and ignore the potential for populations to fracture as the search is performed. Numerical simulations were performed with a given number of individuals positioned on the chemical filament with a random initial orientation distributed around the half-circle pointing towards the source. The number of individuals able to locate the source was recorded and the probabilities of success calculated for varying parameter values. It should be noted that due to the lack of significant persistent gradient parallel to the desired direction, navigating a filament upstream or downstream is almost equivalent; hence the anisotropy of the initial conditions. However, it can be assumed any organism wishing to track a filament will have rudimentary knowledge of the flow direction, and this may even be passively obtained through reorientation by the rheology of the flow.

Fig. 2*A* shows how success rate is affected for different group sizes when the radius of the repulsion zone is varied. From this, we can see that this distance is potentially more than an avoidance mechanism, as it allows the group to act as an efficient network of sensors independently sampling the spatiotemporally evolving environment. To examine this further, we rescale the rates of success in terms of the effective area occupied by the group. The effective area is defined as the ideal exclusion area that an individual wishes to maintain multiplied by the number of individuals in the population. The result of this rescaling is included in Fig. 2*B*. This figure shows the existence of an optimum area with a sharp increase in success as this optimum is reached then a slow exponential decay as the area is exceeded. These results suggest that although the optimal area is invariant, maximum success is somewhat improved with larger numbers as large populations, tightly packed together, are more robust and better able to track these filaments.

However, this comes at a cost; our simulations show the percentage of the group that will arrive at the source on a successful trial will reduce as total group size is increased (see *SI Appendix*). For larger groups, there is a fission effect, and groups will split with some losing the signal and moving out of range of the successful group. To maintain group cohesion, the number of individuals present cannot be too large. This may not be a disadvantage in some robotic search applications where the proportion of the group reaching the target is unimportant.

The Role of Memory

The decay rate α of the stored concentration value (with which current values are compared) represents the length of an individual's memory. This memory allows comparison of experienced concentration levels and the assessment of the current trajectory. The value of α defines a time scale for the memory decay, and this effectively controls the sensitivity to environmental conditions and

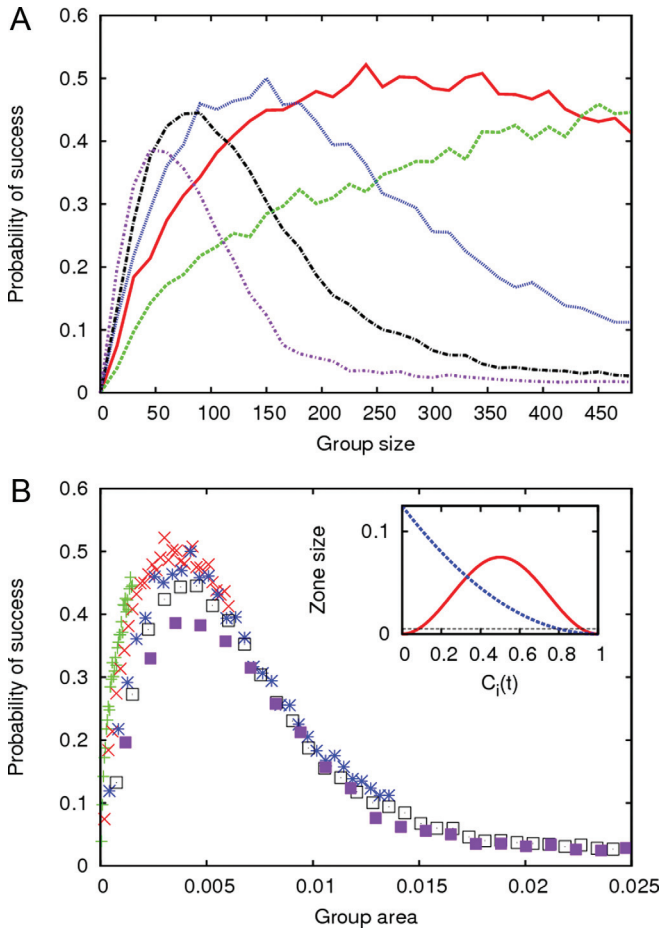


Fig. 2. Effect of group size and group area on probability of success. (*A*) Probability of success as a function of group size. Lines represent varying repulsion zone size. Green dash, 1×10^{-3}; red solid, 2×10^{-3}; blue dot, 3×10^{-3}; black dot-dash, 4×10^{-3}; purple dot-small dash, 5×10^{-3}. (*B*) Probability of success as a function of group area. Green plus sign, 1×10^{-3}; red x, 2×10^{-3}; blue asterisk, 3×10^{-3}; black square (outline), 4×10^{-3}; purple square (solid), 5×10^{-3}. (*Inset*) Interaction zone sizes as a function of $C_i(t)$. Blue dash, attraction zone; red solid, orientation zone; black dash, repulsion zone.

represents an individual level parameter that can be tuned through evolution or experience to affect the success of the group.

In the flow regime we are considering, the filamental structure of the advected chemical signal means that the largest gradient is transverse to the desired direction. The memory parameter therefore controls how responsive an individual is to a reduction in signal strength as it moves towards the edges of the filament. To place the time scale of the memory in the context of the properties of the signal, we approximate the memory parameter required to effectively trace the edges of the filament. By measuring the average width of the filament and assuming its profile follows a Gaussian distribution, we calculate the time taken to traverse one standard deviation travelling at 45° to the parallel direction as ≈0.0125. As the role of memory is to allow an individual to determine if conditions are improving or deteriorating, this value defines a memory length for which it can be assumed that an individual is able to detect that it is moving toward the edges of the filament.

Decay rates which are greater than ≈0.0125 lead to less responsiveness and more binary behavior; individuals ignore others when experiencing any concentration value and then attract when the filament is lost. In the limiting case of no memory, this corresponds to a low-polarization swarm as seen in groups of mosquitoes or midges. In this region of parameter space there is no net motion, but the group is able to maintain its position centered on the plume for long periods.

In the other extreme, high values of α lead to fluctuations in concentration having a larger impact on success rates due to a weakening responsiveness to the decay of the filament perpendicular to the desired direction. Effectively, individuals that have experienced a high-concentration parcel are subsequently less able to accurately respond to the horizontal profile of the filament within the time frame defined by the decay rate. The negative impact of longer memory is a result of intermittent fluctuations in the signal and is therefore a weak effect, meaning that once the required memory is reached, the success rate declines slowly as it is increased.

To examine the intermediate regime, the value of α was varied and success rates recorded for different group sizes. These results are shown in Fig. 3*A*, from which it is clear that optimum memory length is dependent on group size. Although the global optimum value corresponds to the value required to detect the edges of the filament while also being able to forget high-concentration patches, different group sizes do not all share this same optimum. The reason for this can be found in the tradeoff between exploration and exploitation of the signal where exploration means the effective sampling of the signal and exploitation corresponds to

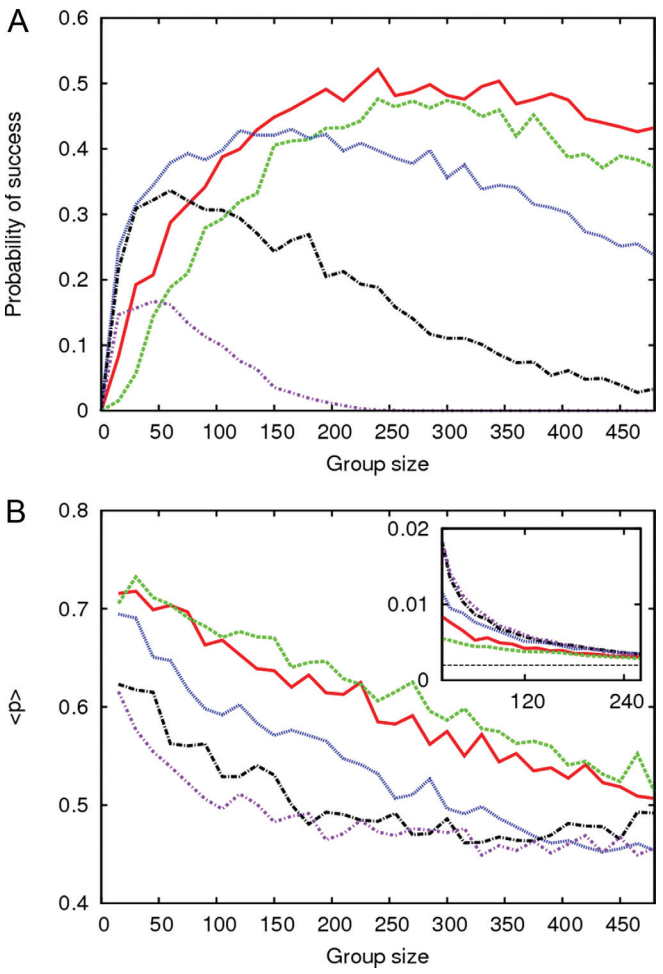


Fig. 3. Effects of group size and memory length. (*A*) Probability of success as a function of group size for various memory lengths. Green dash, $\alpha = 25 \times 10^{-3}$; red solid, $\alpha = 12.5 \times 10^{-3}$; blue dot, $\alpha = 2.5 \times 10^{-3}$; black dot-dash, $\alpha = 1.25 \times 10^{-3}$; purple dash-dot, $\alpha = 0.5 \times 10^{-3}$. (*B*) Group polarity as a function of group size. Legend as in *A*. (*Inset*) Mean nearest-neighbor distance. Legend as in *A*. The black thin-dashed line represents repulsion radius. These values represent ensemble average values, time averaged within and across multiple simulation instances.

net motion toward its source. We illustrate this by measuring two properties of the group dynamics, the average nearest neighbor distance

$$\langle \Delta r \rangle = \frac{1}{N} \sum_{i=1}^{N} \min_{j}(|\mathbf{r}_j(t) - \mathbf{r}_i(t)|), \quad [7]$$

and the average group polarity

$$\langle p \rangle = \frac{1}{N} \left| \sum_{i=1}^{N} \mathbf{p}_i(t) \right|, \quad [8]$$

which are measured numerically and shown in Fig. 3*B*. It can be seen the fast decay of memory (which causes more asocial behavior until concentration is greatly reduced) leads to consistently lower polarity and also for smaller groups to be less-densely packed. For smaller populations, becoming too sensitive to their environment leads to a more compact and highly polarized grouping which is less stable and is more likely to lose track of the filament when required to respond quickly to curves generated by vortices in the flow. This effect may be compensated for by decreasing α, which results in groups covering a larger spatial area, lower alignment and a reduced tendency to collectively lose the filament. The polarity of the group defines the speed at which it will travel toward the source, but the optimum value is a tradeoff between speed and exploration of the surrounding area, which enables the group as a whole to move in different directions when necessary.

Discussion

Our model illustrates how simple adaptive social interactions can lead to cooperative behavior that, in this case, produces an emergent group-level search. By modifying their behavior based on local conditions, autonomous individuals enable the group to collectively act as a spatial nonlocal gradient sensor that is able to track a chemical signal and locate its source. It has previously been shown how small numbers of individuals can control decision making (6) and lead groups in a given direction. Here the context-dependent interactions mean leaders are dynamically changing and automatically selected. As these leaders are those that are experiencing increases in local concentration, group direction is toward the source.

Although we have chosen to study an advection-dominated regime where chaotic dynamics lead to patchy, heterogeneous information, the mechanism described here is equally applicable to diffusion-dominated environments where concentration values increase smoothly and steadily as their source is approached. In this case, group dynamics can be easily seen to follow a trajectory toward the source; however, asocial strategies such as a biased random walk (32) are equally effective in these conditions, and social behavior is redundant (even maladaptive due to effects of competition). The interesting case from both an evolutionary and technological perspective is when the nature of the environment renders individual strategies ineffective.

In the case of the model presented here, the center of the filament can be considered as an unstable position for any organism to maintain. That the flow also advects the individuals in our model is both beneficial, as it aids in following the signal, and disadvantageous, as a particle is rapidly moved away if the signal is lost. A strategy based on temporal sampling and biased random turns in these environments is ineffective, as once the signal is lost, it becomes highly difficult to relocate. Responding to the edges of the filament through a counter-turning mechanism may be effective and is observed in nature (33) but requires spatially separate sensors to accurately determine the turning speed required (34).

The advantage of sociality therefore lies in two factors. Firstly the stability it provides in maintaining position on the signal; although lone individuals are easily dislodged and moved away from the signal as a result of exponentially diverging trajectories, collective interactions anchor the group to the filament. Secondly, the group is able to sample over a much larger area simultaneously and, due to the context-dependent interactions, makes a consensus decision based on these samples. The result of these factors is the smoothing out of fluctuations in the chemical concentration, a comparison of different trajectories without losing track of the signal and the automatic selection of the optimal path. The parameter values and the exact form of the response functions could be tuned further by some evolutionary selection or learning process to produce optimal behavior adapted to the specific characteristics of the environment.

The recent study of animal groups as complex adaptive systems has shown that they are able to perform sophisticated tasks through simple local interactions. These distributed systems are topics of considerable interest in the robotics community (35–37) because they have many advantages over the traditional paradigm of centralized control, notably the absence of communication time (in autonomous systems), inherent robustness (any given component is expendable), and cost effectiveness as large numbers of simple robotic components can be manufactured efficiently. In the traditional sense, the results outlined in this paper may be applied to robotic search strategies but may also be applied to more abstract search problems in the domain of genetic algorithms, e.g., particle swarm optimization, as our model illustrates mechanisms of this type can be effective even in the absence of global information.

Materials and Methods

The isotropic flow field was generated in Fourier space by a stochastic Ornstein–Uhlenbeck process with an exponentially decaying energy spectrum and the length scale of the highest energy modes set to 0.31. Root mean squared velocity was 0.25 and a mean flow of 0.6 was imposed in a constant direction. Advection of the chemical signal was performed by using a semi-Lagrangian method with a gridsize of 512. Particle positions were updated by using a second-order scheme with positions and velocities of neighbors updated once per time step. Particle speed $v_s = 1.6$, turning rate $\gamma = 140$ (radians), $dt = 0.00025$, maximum orientation radius $R_{O,max} = 0.075$, maximum attraction radius $R_{A,max} = 0.125$. Success rates were obtained by using 1,000 trials and recording number of individuals reaching within a distance of 0.025 of the source location. Individuals were initially located in a ball centered on the filament at a distance of 0.8 from the source. For further details of numerical methods, see the *SI Appendix*.

ACKNOWLEDGMENTS. The authors wish to thank Istvan Scheuring, Sepideh Bazazi, and Christos Ioannou for useful discussions. C. T. and Z. N. were supported by Science Foundation Ireland; I. D. C. was supported by Searle Scholar Award 08-SPP-201. This work was supported by National Science Foundation Award PHY-0848755 and Office of Naval Research Award N00014-09-1-1074. High-performance computational facilities and support were provided by the Irish Centre for High-End Computing.

1. Simons A (2004) Many wrongs: The advantage of group navigation. *Trends Ecol Evol* 19:453–454.
2. Grünbaum D (1998) Schooling as a strategy for taxis in a noisy environment. *Evol Ecol* 12:503–522.
3. Parrish JK, Edelstein–Keshet L (1999) Complexity, pattern, and evolutionary trade-offs in animal aggregation. *Science* 284:99–101.
4. Camazine S, et al. (2003) *Self-Organization in Biological Systems* (Princeton Univ Press, Princeton).
5. Couzin I, Krause J, James R, Ruxton G, Franks N (2002) Collective memory and spatial sorting in animal groups. *J Theor Biol* 218:1–11.
6. Couzin ID, Krause J, Franks NR, Levin SA (2005) Effective leadership and decision-making in animal groups on the move. *Nature* 433:513–516.
7. Couzin ID (2007) Collective minds. *Nature* 445:715.
8. Marques L, de Almeia A (2006) Mobile robot olfaction. *Auton Robot* 20:183–184.
9. Marques L, Nunes U, de Almeia A (2006) Particle swarm-based olfactory guided search. *Auton Robot* 20:277–287.
10. Hayes A, Martinoli A, Goodman R (2003) Swarm robotic odor localization: Off-line optimization and validation with real robots. *Robotica* 21:427–441.
11. Shimizu TS, Aksenov SV, Bray D (2003) A spatially extended stochastic model of the bacterial chemotaxis signalling pathway. *J Mol Bio* 329:291–309.
12. Keymer JE, Endres RG, Skoge M, Meir Y, Wingreen NS (2006) Chemosensing in escherichia coli: Two regimes of two-state receptors. *Proc Natl Acad Sci USA* 103:1786–1791.
13. Lauga E, Powers TR (2009) The hydrodynamics of swimming microorganisms. *Rep Prog Phys*, 72:096601.
14. Nicolau DV, Armitage JP, Maini PK (2009) Directional persistence and the optimality of run-and-tumble chemotaxis. *Comput Biol Chem* 33:269–274.
15. Balkovsky E, Shraiman B (2002) Olfactory search at high Reynolds number. *Proc Natl Acad Sci USA* 99:12589–12593.

16. Vergassola M, Villermaux E, Shraiman BI (2007) Infotaxis as a strategy for searching without gradients. *Nature* 445:406–409.
17. Torney C, Neufeld Z (2008) Phototactic clustering of swimming microorganisms in a turbulent velocity field. *Phys Rev Lett* 101:078105.
18. Martí AC, Sancho JM, Sagués F, Careta A (1997) Langevin approach to generate synthetic turbulent flows. *Phys Fluids* 9:1078–1084.
19. Vicsek T, Czirok A, Ben–Jacob E, Cohen I, Shochet O (1995) Novel type of phase transition in a system of self-driven particles. *Phys Rev Lett* 75:1226–1229.
20. Gueron S, Levin S, Rubenstein D (1996) The dynamics of herds: From individuals to aggregations. *J Theor Biol* 182:85–98.
21. Mogilner A, Edelstein–Keshet L (1999) A non-local model for a swarm. *J Math Biol* 38:534–570.
22. Aoki I (1982) A simulation study on the schooling mechanism in fish. *Bull Jpn Soc Sci Fish* 48:1081–1088.
23. Huth A, Wissel C (1992) The simulation of the movement of fish schools. *J Theor Biol* 156:365–385.
24. Huth A, Wissel C (1994) The analysis of behaviour and the structure of fish schools by means of computer simulations. *Comments Theor Biol* 3:169–201.
25. Segall J, Block S, Berg H (1986) Temporal comparisons in bacterial chemotaxis. *Proc Natl Acad Sci USA* 83:8987–8991.
26. Hoare D, Couzin I, Godin JG, Krause J (2004) Context-dependent group size choice in fish. *Anim Behav* 67:155–164.
27. Kats LB, Dill L (1998) The scent of death: Chemosensory assessment of predation risk by prey animals. *Ecoscience* 5:361–394.
28. Van Bergen Y, Coolen I, Laland K (2004) Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. *Proc R Soc London Ser B* 271:957–962.
29. Brown C, Brown M, Shaffer M (1991) Food-sharing signals among socially foraging cliff swallows. *Anim Behav* 42:551–564.
30. Partridge B, Pitcher T (1980) The sensory basis of fish schools: Relative roles of lateral line and vision. *J Comp Physiol A* 135:315–325.
31. Steele C, Scarfe A, Owens D (1991) Effects of group size on the responsiveness of zebrafish, *Brachydanio rerio* (hamilton buchanan), to alanine, a chemical attractant. *J Fish Biol* 38:553–564.
32. Berg H (1993) *Random Walks in Biology* (Princeton Univ Press, Princeton).
33. Grasso FW, Basil JA (2002) How lobsters, crayfishes, and crabs locate sources of odor: current perspectives and future directions. *Curr Opin Neurobiol* 12:721–727.
34. Grasso FW, Consi TR, Mountain DC, Atema J (2000) Biomimetic robot lobster performs chemo-orientation in turbulence using a pair of spatially separated sensors: Progress and challenges. *Robot Auton Syst* 30:115–131.
35. Bonabeau E, Dorigo M, Theraulaz G (1999) Swarm intelligence: From natural to artificial systems (Oxford Univ Press, New York).
36. Desai J, Ostrowski J, Kumar V (2001) Modeling and control of formations of nonholonomic mobile robots. *IEEE Trans Robotic Autom* 17:905–908.
37. Ogren P, Fiorelli E, Leonard N (2004) Cooperative control of mobile sensor networks: Adaptive gradient climbing in a distributed environment. *IEEE Trans Automat Contr* 49:1292–1302.

Supporting Information

Numerical methods

The following details the methodology behind the simulation of a velocity field which displays the characteristics of turbulent flow following an approach termed *synthetic turbulence* [1, 2, 3, 4], the coupling of this flow model to a semi-Lagrangian scheme [5] for the advection of the chemical signal, and the numerical methods which simulate the behavioral rules of the self-propelled particles.

The method used for the stochastic flow is based on a representation of the statistical properties of the fluid dynamics. The aim of this method is to generate a statistically stationary, homogeneous, and isotropic two-dimensional stochastic flow with zero mean velocity and a well-defined energy spectrum.

If a two-dimensional flow is divergence free, $\nabla \cdot \mathbf{v} = 0$, it can be represented by a stream function, ψ where

$$\mathbf{v} = \left(\frac{\partial \psi}{\partial y}, -\frac{\partial \psi}{\partial x} \right). \tag{1}$$

The real-valued, scalar field ψ completely defines the flow field, with the streamlines of the flow corresponding to the contour lines of ψ. To ensure the flow is homogeneous and stationary the stream function is taken to be a solution to the stochastic partial differential equation

$$\frac{\partial \psi}{\partial t} = \nu \nabla^2 \psi + \sqrt{\xi} \frac{\partial W}{\partial t} \tag{2}$$

where W is a coloured noise [6] with a prescribed spectrum which defines the energy spectrum of the flow. The parameter ν acts as viscosity and dampens the fluctuations in the stream function, while ξ defines the magnitude of the noise term.

As W is coloured noise it can be defined as a Fourier series by

$$W(x,t) = \sum_{k \in K} \sqrt{\lambda_k} \hat{W}_k(t) e^{ikx}, \qquad K = 2\pi\mathbb{Z}^2 \tag{3}$$

where λ_k is the energy for the given Fourier mode of the noise (we take the square root for convenience later when this term acts as a variance), and $\hat{W}_k$ is a sequence of Brownian motions, complex valued and with the constraint that $\hat{W}_k = \hat{W}^*_{-k}$ to ensure the stream function remains real valued. The use of x in this equation and those that follow refers to x/L where L is the domain size. Taking the series expansion of ψ and incorporating Eqn. 3 into Eqn. 2 leaves

$$\sum_{k \in K} \frac{\partial \hat{\psi}_k}{\partial t} e^{ikx} = -\nu \sum_{k \in K} |k|^2 \hat{\psi}_k e^{ikx} + \sum_{k \in K} \sqrt{\xi \lambda_k} \frac{\partial \hat{W}_k}{\partial t} e^{ikx} \tag{4}$$

As this is a linear equation each Fourier mode can be solved independently and this leads directly to a system of Ornstein-Uhlenbeck equations of the form

$$d\hat{\psi}_k = -\nu |k|^2 \hat{\psi}_k dt + \sqrt{\xi \lambda_k} d\hat{W}_k. \tag{5}$$

The solution to this can be obtained exactly as

$$\hat{\psi}_k(t) = \hat{\psi}_k(0) e^{-\nu |k|^2 t} + \int_0^t \sqrt{\xi \lambda_k} e^{\nu |k|^2 (s-t)} d\hat{W}_k(s) \tag{6}$$

From the theory of stochastic integration [7] the integral in Eqn. 6 is a Gaussian process with mean zero and variance given by

$$\xi \lambda_k \int_0^t e^{2\nu |k|^2 (s-t)} ds = \frac{\xi \lambda_k}{2\nu |k|^2} \left(1 - e^{-2\nu |k|^2 t} \right). \tag{7}$$

Each Fourier mode of ψ can now be evolved in discrete time steps from a given initial condition by the formula

$$\hat{\psi}_k(t + \Delta t) = \hat{\psi}_k(t) e^{-\nu |k|^2 \Delta t} + \sqrt{\frac{\xi \lambda_k}{2\nu |k|^2} \left(1 - e^{-2\nu |k|^2 \Delta t} \right)} Z_k \tag{8}$$

where Z_k are random numbers sampled from a $N(0,1)$ distribution and subject to the constraint $Z_k = Z^*_{-k}$. We take as an initial condition the steady state by letting $t \to \infty$ and sampling according to the uncorrelated variance.

The resulting flow field is modified by introducing a constant velocity along the vertical direction and then coupled to a semi-Lagrangian advection scheme (see [8] for detailed methodology) which advects a decaying chemical field entered into the system at a constant rate from a point in the domain. A grid size of 512x256 was used and simulations were performed using a parallel MPI algorithm.

The final stage for the numerics involves the simulation of self-propelled particles which are advected by the flow and orient themselves depending on local conditions. Individuals firstly assess the concentration parameter $C_i(t)$ which is a function of the instantaneous local concentration and a stored memory value which decays with time. At each timestep the desired vector $\mathbf{d}_i$ is calculated with a naive order N^2 method using behavioral rules based on the interaction zones defined by $C_i(t)$.

In two-dimensions both the current orientation vector $\mathbf{p}_i$ and the desired direction $\mathbf{d}_i$ can be represented by angles denoted by θ and ϕ respectively, where $-\pi < \theta, \phi < \pi$. The equations of motion for individual i are then defined by

$$\begin{aligned} \dot{x} &= V_{fx} + V_s \cos(\theta) \\ \dot{y} &= V_{fy} + V_s \sin(\theta) \\ \dot{\theta}_i &= \gamma (\phi - \theta) \end{aligned} \tag{9}$$

where $(\phi - \theta)$ is rotated clockwise or counterclockwise by 2π to ensure it lies within the region $-\pi \leq (\phi - \theta) \leq \pi$, V_{fx} and V_{fy} are the x and y components of the carrier flow, V_s is the swimming speed and γ controls the maximum turning rate.

The equations of motion are integrated using the second order modified Euler's method.

1. **Juneja A, Lathrop DP, Sreenivasan KR, Stolovitzky G (1994) Synthetic turbulence. *Physical Review E* 49:5179–5194.**
2. **Martí AC, Sancho JM, Sagués F, Careta A (1997) Langevin approach to generate synthetic turbulent flows. *Physics of Fluids* 9:1078–1084.**
3. **Sigurgeirsson H, Stuart AM (2002) A model for preferential concentration. *Physics of Fluids* 14:4352–4361.**
4. **Sigurgeirsson H (2001) *Particle-field models: Algorithms and applications*. Ph.D. thesis, Program in scientific computing and computational mathematics, Stanford University.**
5. **Staniforth A, Côté J (1991) Semi-Lagrangian Integration Schemes for Atmospheric Models - A Review. *Monthly Weather Review* 119:2206–2223.**
6. **García-Ojalvo J, Sancho JM, Ramírez-Piscina L (1992) Generation of spatiotemporal colored noise. *Physical Review A* 46:4670–4675.**
7. **Gardiner C (1985) *Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences* (Springer-Verlag, Berlin).**
8. **Durran D (1999) *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics* (Springer-Verlag Inc., New York).**

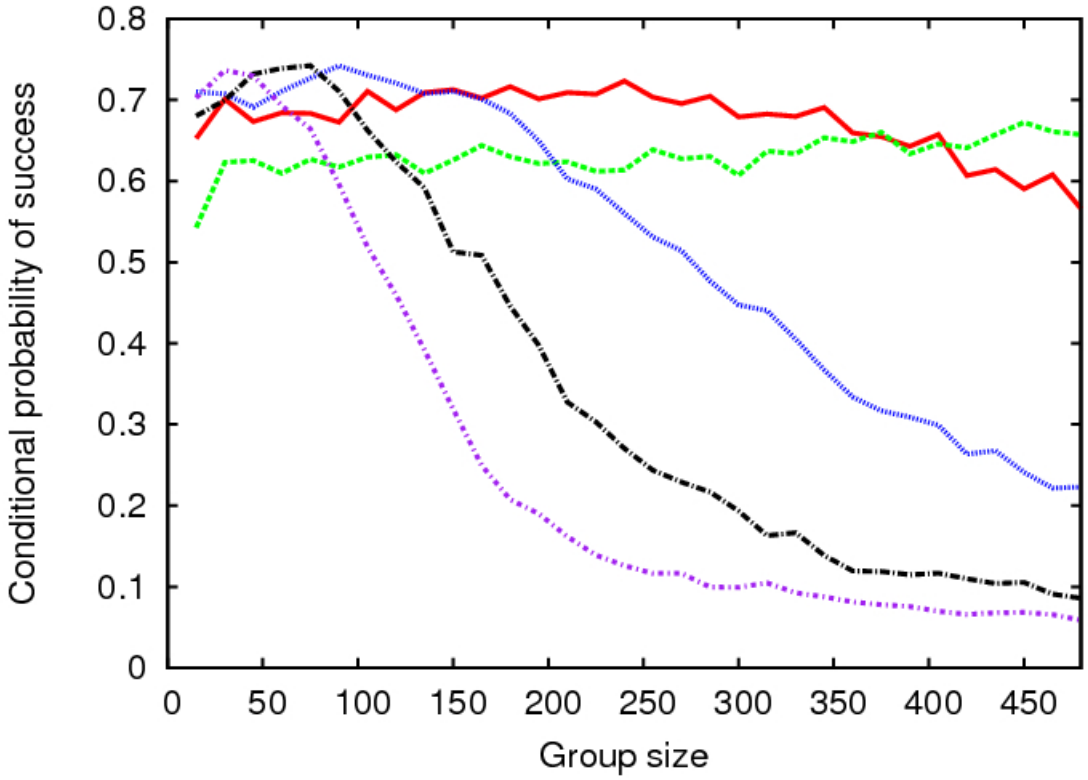
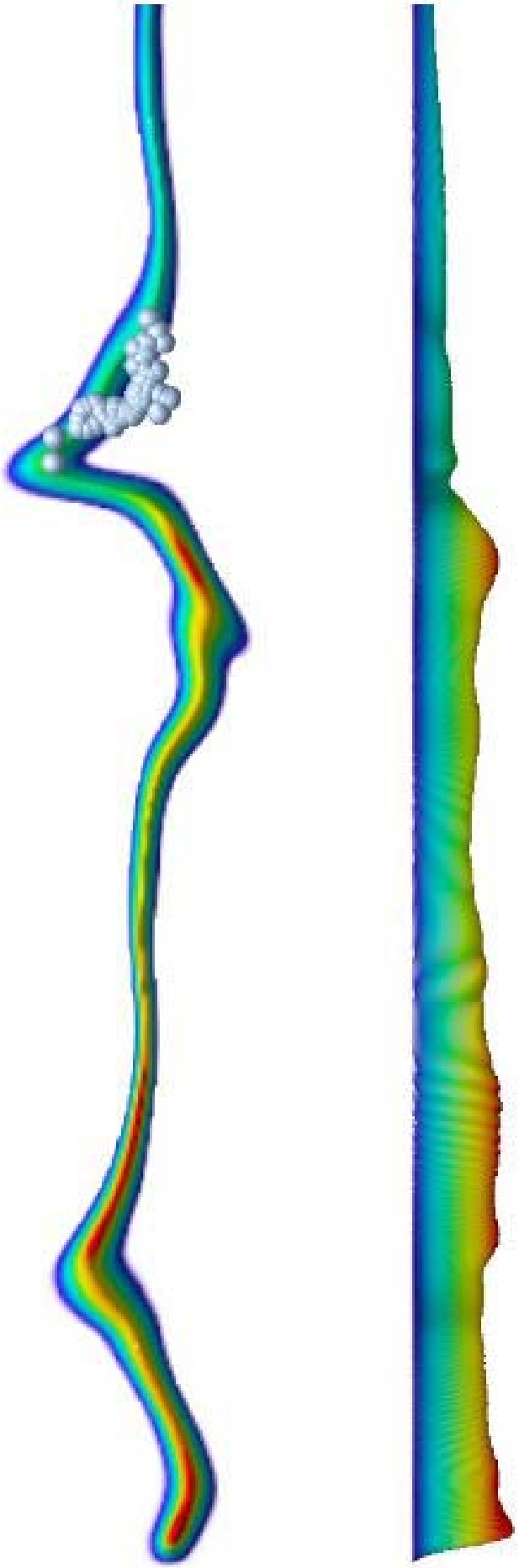
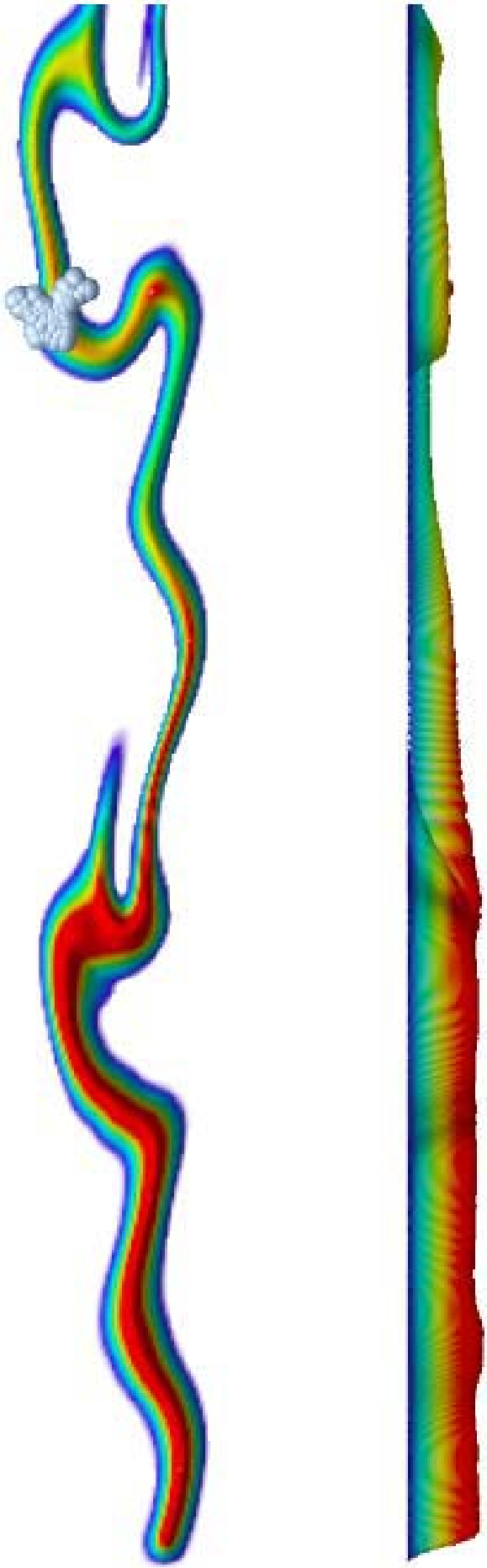


Fig. 1. Conditional success rate. Probability of successfully locating the source when at least one other member has been successful as a function of group size. Lines represent varying repulsion zone size, green dash 1×10^{-3}, red solid 2×10^{-3}, blue dot 3×10^{-3}, black dot-dash 4×10^{-3}, purple dot-small dash 5×10^{-3}.



Movie S1 (AVI).



Movie S2 (AVI).