Sector search strategies for odor trail tracking

Gautam Reddy

NSF-Simons Center for Mathematical & Statistical Analysis of Biology,
Harvard University, Cambridge, MA 02138 and
Department of Physics, University of California, San Diego, La Jolla, CA, USA

Boris I. Shraiman[*]

Kavli Institute of Theoretical Physics, University of California Santa Barbara, Santa Barbara, CA 93106

Massimo Vergassola[†]

Laboratoire de physique de l'École Normale Supérieure, CNRS,
PSL Research University, Sorbonne Université, Paris, France and
Department of Physics, University of California, San Diego, La Jolla, CA, USA

Terrestrial animals such as ants, mice and dogs often use surface-bound scent trails to establish navigation routes or to find food and mates, yet their tracking strategies are poorly understood. Tracking behavior features zig-zagging paths with animals often staying in close contact with the trail. Upon sustained loss of contact, animals execute a characteristic sequence of sweeping "casts" – wide oscillations with increasing amplitude. Here, we provide a unified description of trail-tracking behavior by introducing an optimization framework where animals search in the angular sector defined by their estimate of the trail's heading and its uncertainty. *In silico* experiments using reinforcement learning based on this hypothesis recapitulate experimentally observed tracking patterns. We show that search geometry imposes limits on the tracking speed, and quantify its dependence on trail statistics and memory of past contacts. By formulating trail-tracking as a Bellman-type sequential optimization problem, we quantify the basic geometric elements of optimal sector search strategy, effectively explaining why and when casting is necessary. We propose a set of experiments to infer how tracking animals acquire, integrate and respond to past information on the tracked trail. More generally, we define navigational strategies relevant for animals and bio-mimetic robots, and formulate trail-tracking as a novel behavioral paradigm for learning, memory and planning.

Experimental studies demonstrate the ability of ants, dogs, humans, and rodents to track odor trails[1–6]. Rodents accurately track trails in the dark, remaining close to the trail and casting when contact is lost (Figure 1a)[5]. Carpenter ants closely follow a trail while sampling it using a "criss-cross" pattern with their two antennae (Figure 1b)[1]. Current models of this behavior rely on variants of chemotaxis[7] based on continuous estimates of the rising and falling odor gradients as the trail is crossed. One such strategy compares simultaneous odor concentrations detected by two spatially separated sensors[8]. Yet, rats with a blocked nostril[5] and ants with a single antenna[1] are still able to track trails, although less accurately. An alternative chemotaxis strategy has the animal measuring odor gradients along its trajectory and turning when a significant decrease is perceived[5].

While chemotaxis-based strategies can allow for trail tracking when trails are continuous, they fail when trails are broken and gradients are absent, which is certainly relevant for animals tracking trails in the wild. In experiments with broken trails[1,5], the absence of signal triggers casting, which is a fundamental feature shared with olfactory searches in a turbulent medium[9–11]. Even though turbulent searches also feature sporadic cues, airborne odor signals tend to be localized in a cone and, even within the cone, the signal is highly fluctuating[12,13]. Therefore, beyond qualitative similarities between terrestrial trail tracking and airborne olfactory searches, the specific statistics of detections, geometric constraints and behavioral patterns are distinct.

In contrast with chemotaxis-based algorithms, we propose an alternative framework built on the searcher exploiting past contacts with the trail to maintain an estimate of the trail's local heading and its uncertainty, which defines an angular sector of probable trail headings that radiates from the most recent detection point. The resulting "sector search" provides a quantitative description of trail-tracking behavior that unifies its various phases and yields specific experimental predictions.

We first show that reinforcement learning (RL) based on the sector search idea can recapitulate natural behavior. An RL agent in this scheme learns to traverse the trail as quickly as possible while minimizing the probability of losing it (see Methods for details). Our *in silico* RL experiments show that general aspects of

* shraiman@ucsb.edu
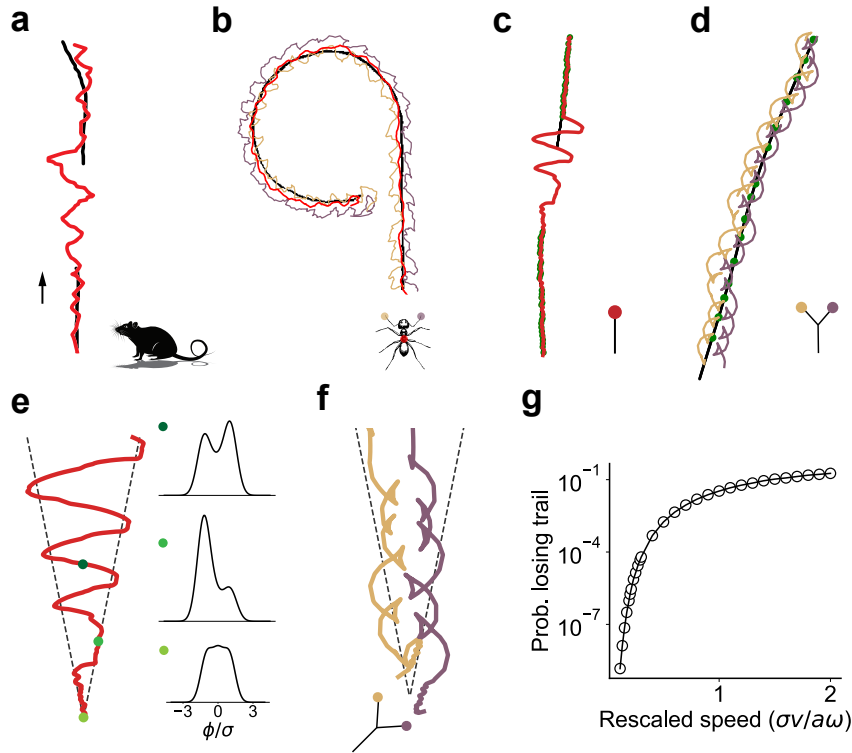† massimo.vergassola@phys.ens.fr

2

Figure 1. Sample trail tracking trajectories from previous experiments and our reinforcement learning (RL) simulations. (a) A rat (head position in red) tracking a trail (in black). Adapted from ref. 5. Note the wide casts on extended loss of contact with the trail. (b) A carpenter ant tracking an odor trail (black) using a stereotyped "criss-crossing" strategy[1]. (c,d) Sample trajectories obtained from RL for agents with one sensor (panel c) and two sensors (panel d) recapitulate experimentally observed tracking patterns in panels a,b. (e) Left: Search paths executed by RL agents with a single sensor upon loss of contact with the trail. Right: The initial prior distribution (bottom plot) over trail headings transforms into a bimodal posterior distribution (top two plots), which drives the oscillatory pattern of casting. (f) RL agents with two sensors show a characteristic criss-crossing pattern close to the last detection point. The search path is similar to the single-sensor agent at large distances (Figure M1a). (g) RL agents show a trade-off between tracking speed (re-scaled by the sector angle σ, sensor size a and sampling frequency ω) and the probability of losing the trail entirely.

animal tracking behavior naturally emerge (see Figure 1c,d). Specifically, casts are observed around the most likely heading of the trail, and their amplitude is within the angular sector defined by the initial uncertainty σ of the trail's heading ϕ. The reason for the oscillatory pattern of casting is intuitive. Indeed, while moving along a path $\mathcal{C}$ without detecting the trail, the estimated heading's probability distribution $P(\phi)$ (see Figure 1e) is updated into $P_{\mathcal{C}}(\phi)$ as

$$P(\phi) \rightarrow P_{\mathcal{C}}(\phi) \propto \Gamma_{\mathcal{C}}(\phi)P(\phi), \qquad (1)$$

where $\Gamma_{\mathcal{C}}(\phi)$ is the probability of not detecting the trail headed along ϕ. Irrespective of the explicit form of $\Gamma_{\mathcal{C}}(\phi)$, the depletion of headings already explored generally leads to a bimodal posterior distribution, with the two modes at the edges of the angular sector (see Figure 1e). Oscillations are then understood in terms of marginal value theory[14,15]: we show using a minimal model of casting (Methods) that the turning point of a cast occurs when the marginal value of continuing on one side of the sector is outweighed by the probability of finding the trail on the opposite side.

We proceed now by establishing geometric limits on tracking speed. A typical RL curve for the probability of losing the trail versus speed is shown in Figure 1g. Its monotonicity epitomizes universal limits that "staying on the trail" imposes on tracking speed. Intuitively, searching slowly reduces the distance between detections (the inter-detection interval (IDI)), decreases the uncertainty in the estimate of the trail's heading, and thus the probability of losing the trail. However, these benefits come at the cost of slow forward progression along the trail. In contrast, moving quickly reduces the detection rate, leading to longer IDIs, increased uncertainty and loss probability.

We quantify the above trade-off using simple scaling arguments. Suppose the tracking agent has a sensor of size a, samples at a frequency ω and moves with a

fixed forward speed v. As shown in Figure 2a, the angle subtended by the detector at distance r from the last contact is a/r. The agent searching over an angular sector scans then at a rate $d\phi/dt = \omega a/r$. Integrating the above expression for the angular rate, we obtain the typical time for re-establishing contact with the trail: $t_c \sim \omega^{-1}e^{\sigma v/a\omega}$. The corresponding distance $L \sim vt_c$ is obtained using $r \sim vt$. The heading of the trail is known with uncertainty σ, which is the opening angle of the conical sector shown in Figure 2a. Uncertainty is expected to depend on the distance L' from the *previous* detection as $\sigma(L') = (L'/\ell)^\gamma$ where ℓ and γ characterize the statistics of trails (see below and Figure 2d). Importantly, a stable strategy for long-term tracking requires that successive IDIs should *on average* be equal, i.e., $L = L'$. Combining $L = vt_c$ with $L' = \sigma^{1/\gamma}\ell$ and the expression for t_c, we finally obtain an upper bound on the tracking speed v:

$$\frac{v^{1+\gamma}}{a\ell^\gamma\omega^{1+\gamma}} = (\omega t_c)^{-\gamma}\log(\omega t_c) \le \gamma^{-1}e^{-1}\,. \qquad (2)$$

Its maximum $v_{\max} \sim \omega(a\ell^\gamma)^{\frac{1}{1+\gamma}}$ defines the optimal stable tracking speed in terms of the tracker's sensory parameters and trail statistics. The basic element that leads to this bound is the geometric factor $1/r$ that underlies searching over an angular sector. The result from (2) that ωt_c is of order one ($e^{1/\gamma}$) explains experimental observations (Figure 1) that tracking animals typically take only a few samples to re-establish contact with the trail.

Ideas from polymer physics suggest a statistical description of trails. We ask how detecting the trail at a set of points $\boldsymbol{r}_0, \boldsymbol{r}_1, \boldsymbol{r}_2, \dots$ (Figure 2b) constrains its future heading? We consider the case when the searcher keeps track of the two most recent points of contact; a more extended memory is discussed further below. Intuition for the two-point case is provided by the familiar "curve" tool in graphical design software, which draws a cubic spline through a set of prescribed points (Figure 2c). The tool captures the simple intuition that tangents to a curve are continuous, i.e., the trail's heading has local persistence, which is a plausible, minimal assumption about trails. We show in the SI that cubic spline interpolation corresponds to the most likely path (through a fixed set of points) in the so-called worm-like chain (WLC) ensemble (originally introduced for polymers)[16,17]. In this ensemble, the tangent direction undergoes diffusion with rate κ, and the uncertainty is then $\sigma = (\kappa L/3)^{\frac{1}{2}}$, which determines the two parameters $\gamma = 1/2$ and $\ell = 3\kappa^{-1}$ in Eq. (2). Actual trails could be smoother and have a well-defined curvature (the rate of change of heading) that persists on a characteristic length scale λ. We capture this ensemble of curves by an extended WLC model (EWLC) with two parameters: persistence length λ and typical radius of curvature ξ (Methods). Uncertainty is then given by $\sigma \approx L/2\xi$ (hence $\gamma = 1$ and $\ell = 2\xi$) at distances $L < \lambda$ while diffusive scaling is recovered at larger distances with an effective diffusivity $\kappa = 2\lambda\xi^{-2}$. Finally, the two models are combined in the general WLC (GWLC) ensemble with crossovers across the various regimes (Methods). In summary, each model leads to a "propagator" which encodes how information about past contacts is integrated to form an estimate of the trail's heading while taking into account geometric aspects of trails. A feature shared across models is that the headings at two consecutive contacts are anti-correlated (Figure 2e), which reflects the bending of the spline relative to the chord seen in Figure 2c.

Why and when do searchers need to cast? The question stems from our previous result that a few samples are typically sufficient to re-establish contact with the trail. To address it quantitatively, we consider again the setup of Eqs. (1), (2). The non-detection probability averaged over the ensemble of trails that pass through past contact points is

$$\Gamma_{\mathcal{C}} = \langle e^{-\omega \int_{\mathcal{C}} \frac{ds}{v} I_a(\boldsymbol{r}(s), \boldsymbol{y}(s))}\rangle_{\boldsymbol{y}}, \qquad (3)$$

where s parametrizes the searcher's path $\mathcal{C}$ and the Boolean indicator function I_a measures if the agent at $\boldsymbol{r}(s)$ is within sensing range a of the trail at $\boldsymbol{y}$, i.e, the integral is the time spent in contact with the trail. Numerical simulations of the search show a power-law scaling regime for $\Gamma_{\mathcal{C}}$, which is cut-off at short distances by the initial surge and at large distances by trails escaping out of the casting envelope (Figure 3a-c). We proceed to explain these three regimes shown in Figure 3b. Intuitively, at short radial distances $r < a/\sigma \sim v/\omega$ (the latter from Eq. (2)), the sensor covers the entire sector of likely headings, the searcher can just move forward and $\Gamma_{\mathcal{C}} \propto e^{-\omega r/v}$ (Figure 3b). Casting sets in if the searcher reaches, without detection, a distance $r \gtrsim a/\sigma$, i.e., when the sector is not fully covered any longer by the sensor of size a. The sector geometry in Figure 3a implies that the length of a single casting sweep is proportional to r. A fixed forward speed then implies that the distance between successive encounters with the trail also scales with r. Hence, the number of times the searcher crosses the trail (and thus the time spent on the trail) per unit radial distance decreases as $1/r$. This $1/r$ scaling in the overlap then leads to a logarithmic integral in the exponent of Eq. (3) and thus a power-law regime $\Gamma_{\mathcal{C}} \propto r^{-\beta}$ during casting. The optimal exponent β^* depends on the statistics of the trails, yet it generally satisfies $\beta^* > 1$ ($\beta^* = 1.63$ for the EWLC model, see Eq. (M15)). Since $\Gamma_{\mathcal{C}}$ is a cumulative distribution, $\beta > 1$ implies that the mean distance is indeed determined by the lower cutoff, i.e., the trail is typically found in a few samples, as estimated in Eq. (2). However, the power-law decay implies that
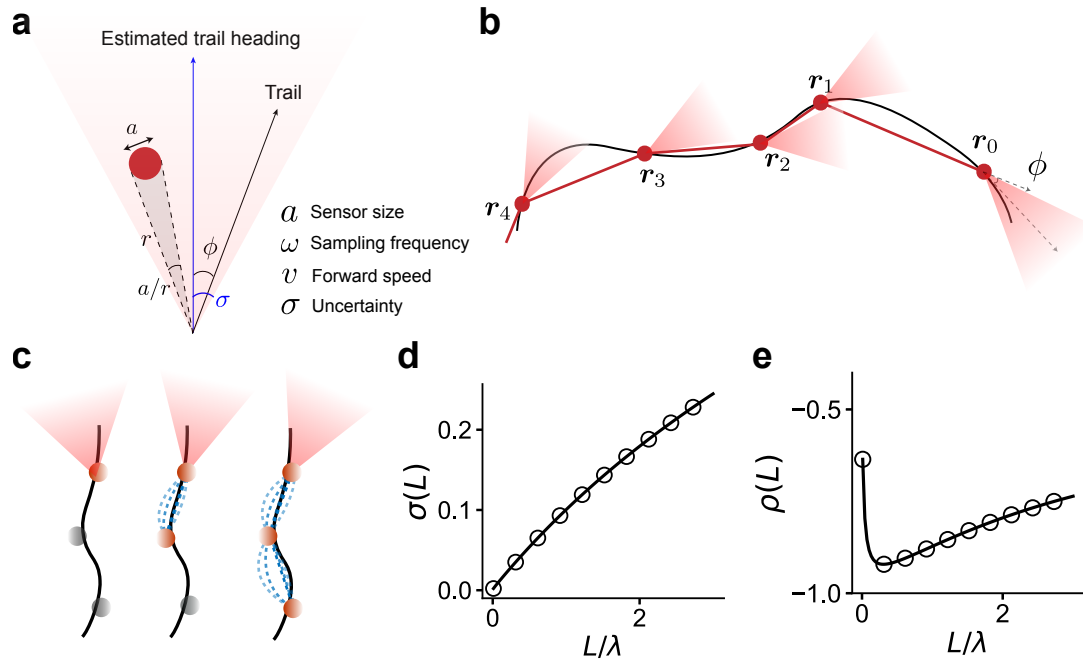
4

Figure 2. History-dependence and trail models. (a,b) Trail tracking naturally splits into distinct episodes punctuated by trail detections by the searcher. In each episode, we propose that the tracker searches for the trail using an estimate of the trail's heading updated based on the past points of contact with the trail and a model of trail statistics. We affix a polar coordinate system with the origin at the most recent contact point and the azimuthal angle defined relative to the estimated trail heading. The uncertainty σ fixes the angular width of the search. The searcher moves forward with a speed v while sampling at a frequency ω. A sensor of size a spans a/r radians at distance r, which determines the rate at which the angular space is searched. (c) To estimate where the trail is headed and its uncertainty from past contacts, the tracker can either use local anisotropy estimated from a single contact (left) or extrapolate from previous points of contact using a model of trail statistics (middle, right). In the latter case, the most likely trail paths (dashed blue lines) are similar to interpolated splines, which capture basic geometric notions of persistence in heading and curvature. (d) The uncertainty in trail heading (in radians) as a function of the distance between points of contact for the GWLC model of trails discussed in the text. (e) The correlation between trail heading at the most recent and second-most recent point of contact for the GWLC model changes with the distance between these points yet it is generally expected to be negative.

casting phases are frequent and can span up to the upper cutoff where all aspects of trail statistics and search geometry come into play, as discussed next.

How should an agent perform sustained casts so as to minimize the probability of losing the trail while maximizing tracking speed? To go beyond the above scaling arguments, we now consider the geometry of sector search in detail. The searcher's path is parameterized by the sequence of turning points of its casting trajectory, $\{r_k, \theta_k\}$ (in polar coordinates w.r.t the most recent point of contact), which are to be optimized. We maximize the average tracking speed L/T, where L and T are the distance and duration between the most recent and the subsequent point of contact with the trail. As discussed previously, the uncertainty estimated for a bout of sector search depends on the inter-detection interval (IDI). To constrain the uncertainty, we therefore constrain the average, $\langle L \rangle$. Hence, we consider the following optimization problem:

$$V_{v,\Lambda} = \max_{\{r_k, \theta_k\}} \left[\left\langle \frac{L}{T} \right\rangle - \Lambda \langle L \rangle \right], \qquad (4)$$

where Λ is the Lagrange multiplier enforcing the $\langle L \rangle$ constraint. The turning points and the searcher's speed v affect the probability of detecting the trail in a single cast, which is implicit in the expectation in Eq. (4). We solve the Bellman equation corresponding to the above optimization problem using dynamic programming (see Methods), which sequentially optimizes for the turning points by considering at each step the two possibilities that either the trail is detected during the cast or the agent advances to the next cast. The probabilities for these two events are controlled by the non-detection probability given by Eq. (3). In the event of no detection, the estimate of the trail's heading is updated according to Eq. (1). The resulting optimization yields a search strategy with an increasing sequence of cast-
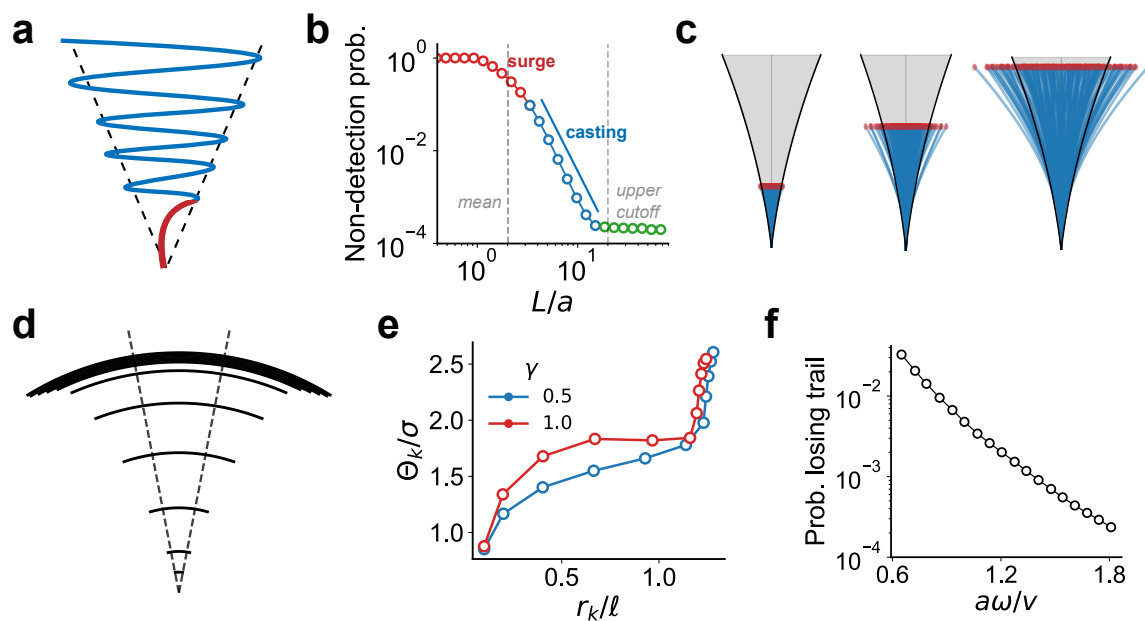
Figure 3. The role of casting in trail tracking. (a,b,c) The probability of not detecting the trail at distance L decays exponentially during the initial forward surge (red) and as a power-law during casting (blue) over a conical envelope. The searcher often finds the trail within the initial surge (mean in panel b) whereas casting determines the probability of losing the trail (determined by the upper cutoff) if the trail is not found during the surge. Beyond a characteristic lengthscale ℓ, the detection-rate becomes negligible (green region in panel b) as trails initially well-inside the casting envelope "escape" out of the envelope (panel c). (d,e) The casting policy obtained from Bellman optimization. The specific casting strategy depends on trail statistics. Note the increasing casting angle and the slowing down of the agent at $r \approx \ell$ (panel e). (f) The trade-off between the probability of missing the trail for fixed $\langle L \rangle$ and searcher speed v.

ing angles (Figure 3d). The specific casting strategy depends on how trails meander and curve (Figure 3e). The choice of v controls the trade-off between the tracking speed and the probability of losing the trail entirely (Figure 3f). Independent of the choice of v and $\langle L \rangle$, azimuthal excursions are by and large conical but extend dramatically (Figure 3d,e) when trails that were initially inside the cone escape from it with high probability (Figure 3c). This happens at a distance that scales with ℓ but also depends on the sector envelope (which in turn depends on σ). Intuitively, at lengthscales $\sim \ell$ the trail's heading decorrelates from its initial value, the relevance of information on past detections has expired, the trail is effectively lost and it is optimal to stop progressing forward.

A number of transformative experimental assays are suggested by our theoretical framework. The broad theme is whether and how animals adapt their behavior to the statistics of trails. For field experiments, it would be informative to measure the statistics of natural trails, analogous to the statistics of natural images that has brought insight into the adaptation of visual responses[19–22]. Specifically, one can measure the auto-correlation of local heading and curvature of natural trails, which would test the validity and fix the param-

eters of our WLC-type models. In laboratory settings, the statistics of trails can be controlled by varying persistence, curvature or using broken trails (Figure 4a). The general issue of adaptation is articulated in the following four specific questions that stem from our work.

First: how long after the loss of contact do animals "give up" tracking? Our prediction is that they should when they get beyond the characteristic correlation length of the trails. At this point, the value of past information has expired and it is best to turn back or start a new search. This prediction can be tested by varying trail statistics, interrupting the trails and measuring when animals give up.

Second: does the amplitude of casting depend on the statistics of trails and the inter-detection interval? We predict that it should, and the specific quantitative relationship is a signature of the underlying predictive model employed by the animal (Figure 4b). Our prediction should be contrasted with the non-adaptive casting envelope assumed in ref. 5. The IDI can be experimentally manipulated by generating dashed trails as shown in Figure 4b, which forces the animal to detect the trail sporadically yet at controllable intervals. It would be particularly informative to verify whether or not animals include curvature in their estimates of the trails'
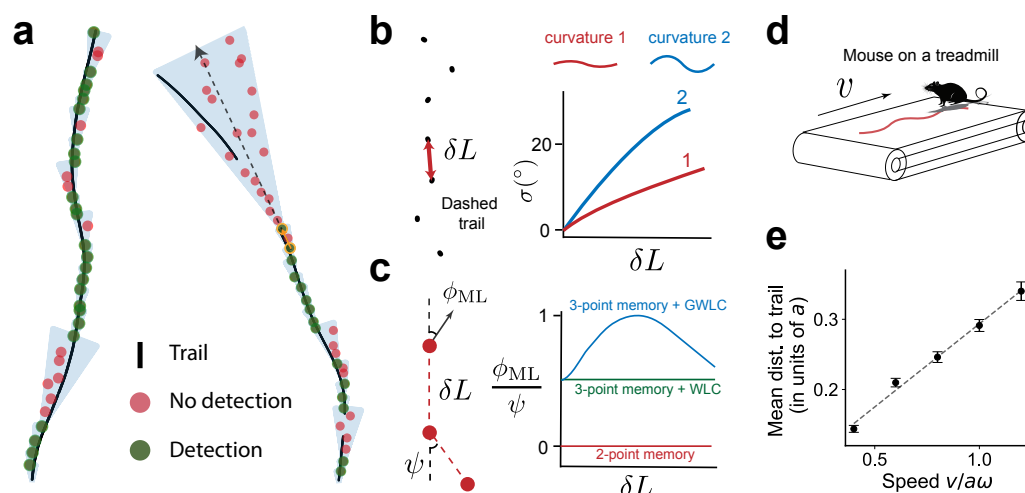
6

Figure 4. Behavioral assays to dissect trail tracking strategies. (a) A tracker executing a sector search often maintains continuous contact with the trail (left). Extended sector searches can instead be systematically elicited using broken trails (right). The subsequent search envelope's orientation and width relative to past contact points informs the tracker's internal estimate and uncertainty σ of the trail's heading. (b) We propose dashed trails as an assay to infer how a tracking animal integrates past information. The distance between dashes forces contact points to be separated by at least δL and the subsequent search sector yields an estimate of σ. (c) Our theory predicts that the most likely trail heading ϕ_{ML} is proportional to the angle ψ between the line segments joining points of contact (red dots), with a pre-factor that depends on the distance δL between the two most recent contact points, the trail model and memory. (d) Automated behavioral tracking of rodents on a treadmill allows control of tracking speed and trail statistics[5,18]. (e) The mean distance to trail with speed in simulations of sector search (as in panel a (left)), which recapitulates the linear relationship (dashed line) observed in experiments with rats[5]. We use a sector search strategy where the longitudinal speed v is fixed and the tracker rapidly casts within a conical envelope (see Figure M4a for a sample trajectory in a single bout). For each v, we simulate 100 trials, where each trial consists of successive 10 successful contacts with the trail. Trails are generated from the GWLC ensemble with $\kappa = 0, \xi/\lambda = 3, a/\lambda = 0.1$. Error bars are 1 s.e.m.

future heading or limit to persistence.

Third: what is the memory of past trail contacts? Experiments with forked trails[5] show that rats exhibit a predictive component, suggesting a memory that extends over the recent past. Our theory posits that the tracker remembers (at least) the two most recent detection points. For the case of two-point memory, we predict the sector search is oriented along the line connecting those two points. If more than two points are remembered, the expected heading deviates from this line (by an angle that we calculate explicitly in the Supplement) as illustrated in Figure 4c. Note that the heading is not an average of the past headings, as assumed in ref. 5, and actually depends on the IDI between recent contacts: this prediction could again be tested by using curved, dashed trails as in Figure 4b.

Fourth: does the tracking speed vary with the typical IDI, reducing with increased uncertainty as predicted by the speed-accuracy trade-off Eq. (2)? This can be tested by varying the speed, for instance of the treadmill in ref. 5 (see Figure 4d), and measuring tracking

accuracy. Available data for three speeds in ref. 5 are captured by our theory (see Figure 4e), which highlights the importance of revisiting those pioneering experiments and measuring additional quantities, namely the explicit prediction in Figure 3f.

In conclusion, we show that an optimized sector search strategy based on the memory of two or more recent detection events yields an oscillatory search path with increasing amplitude that naturally unifies the observed low-amplitude "zig-zagging" and larger-amplitude "casting" behaviors into the same quantitative framework. This framework elucidates the geometric and computational constraints faced by tracking animals, and identifies general features of the algorithms that efficiently solve the task, which can also be implemented for robotic applications. Insights and predictions developed here impact and should inform the design and analysis of future animal behavior experiments.

[1] Draft RW, McGill MR, Kapoor V, Murthy VN. Carpenter ants use diverse antennae sampling strategies to track odor trails. Journal of Experimental Biology. 2018;221(22).

[2] Thesen A, Steen JB, Doving K. Behaviour of dogs during olfactory tracking. Journal of Experimental Biology. 1993;180(1):247–251.

[3] Hepper PG, Wells DL. How many footsteps do dogs need to determine the direction of an odour trail? Chemical Senses. 2005;30(4):291–298.

[4] Porter J, Craven B, Khan RM, Chang SJ, Kang I, Judkewitz B, et al. Mechanisms of scent-tracking in humans. Nature neuroscience. 2007;10(1):27–29.

[5] Khan AG, Sarangi M, Bhalla US. Rats track odour trails accurately using a multi-layered strategy with near-optimal sampling. Nature Communications. 2012;3(1):1–10.

[6] Bhattacharyya U, Bhalla US. Robust and rapid airborne odor tracking without casting. Eneuro. 2015;2(6).

[7] Berg HC. E. coli in Motion. Springer Science & Business Media; 2008.

[8] Riman N, Victor JD, Boie SD, Ermentrout B. The Dynamics of Bilateral Olfactory Search and Navigation. arXiv preprint arXiv:200803358. 2020.

[9] David C, Kennedy J, Ludlow A. Finding of a sex pheromone source by gypsy moths released in the field. Nature. 1983;303(5920):804–806.

[10] Murlis J, Elkinton JS, Carde RT. Odor plumes and how insects use them. Annual review of entomology. 1992;37(1):505–532.

[11] Mafra-Neto A, Cardé RT. Fine-scale structure of pheromone plumes modulates upwind orientation of flying moths. Nature. 1994;369(6476):142–144.

[12] Vergassola M, Villermaux E, Shraiman BI. 'Infotaxis' as a strategy for searching without gradients. Nature. 2007;445(7126):406–409.

[13] Celani A, Villermaux E, Vergassola M. Odor landscapes in turbulent environments. Physical Review X. 2014;4(4):041015.

[14] Charnov EL, et al. Optimal foraging, the marginal value theorem. Theoretical Population Biology. 1976;9(2):129–136.

[15] Stephens DW, Krebs JR. Foraging theory. vol. 1. Princeton University Press; 1986.

[16] Doi M, Edwards SF, Edwards SF. The theory of polymer dynamics. vol. 73. oxford university press; 1988.

[17] Marko JF, Siggia ED. Stretching DNA. Macromolecules. 1995;28(26):8759–8770.

[18] Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nature neuroscience. 2018;21(9):1281–1289.

[19] Field DJ. Relations between the statistics of natural images and the response properties of cortical cells. Josa a. 1987;4(12):2379–2394.

[20] Ruderman DL, Bialek W. Statistics of natural images: Scaling in the woods. Physical review letters. 1994;73(6):814.

[21] Simoncelli EP, Olshausen BA. Natural image statistics and neural representation. Annual review of neuroscience. 2001;24(1):1193–1216.

[22] Wark B, Lundstrom BN, Fairhall A. Sensory adaptation. Current opinion in neurobiology. 2007;17(4):423–429.

[23] Sutton RS, Barto AG. Reinforcement learning: An introduction. MIT press; 2018.

[24] Reddy G, Wong-Ng J, Celani A, Sejnowski TJ, Vergassola M. Glider soaring via reinforcement learning in the field. Nature. 2018;562(7726):236–239.

METHODS

A reinforcement learning framework for trail tracking

Trail tracking naturally splits into discrete episodes where, after each loss of contact, the tracker searches and attempts to re-establish contact with the trail. We use reinforcement learning (RL) to identify optimal search strategies for each episode and explore how factors such as sensory configuration or movement constraints influence the strategy. The task is a one-dimensional search over angular space θ, the geometry of which is illustrated in Figure 2a. The tracker controls its tangential speed $u(t) \equiv r\dot{\theta}$ while its radial speed v is kept constant. For simplicity, we focus here on the configuration featuring a single sensor of size a sampling with a Poisson frequency ω. The generalization to two sensors is found in SI.

In each episode, the (Bayesian) agent maintains a posterior probability distribution function (PDF) $P(\phi)$ over possible trail headings, which is continuously updated based on the locations already visited, until the trail is recontacted. The agent's strategy of decisions about its future trajectory a $priori$ depends on the full high-dimensional distribution $P(\phi)$, which is difficult to learn. To circumvent this issue, we formulate the search task using a tractable parametrization of the posterior as a mixture of K Gaussian basis functions:

$$P(\phi) = \sum_{i=1}^{K} q_i B_i(\phi), \qquad B_i(\phi) = s^{-1}\varphi\left(\frac{|\phi - \mu_i|}{s}\right),$$
(M1)

where φ is the standard normal PDF and the q_i's are normalized weights. The posterior is encoded by the weights $\boldsymbol{q}$, the posterior probabilities of the latent states given the agent's history. The corresponding vector is lower-dimensional and yields to standard RL methods. For simulations in the main text, we used $K = 3$, equal initial weights, $\mu_1, \mu_2, \mu_3 = -\sigma, 0, \sigma$, and $s = 0.5\sigma$. These values were chosen so that the prior has mean
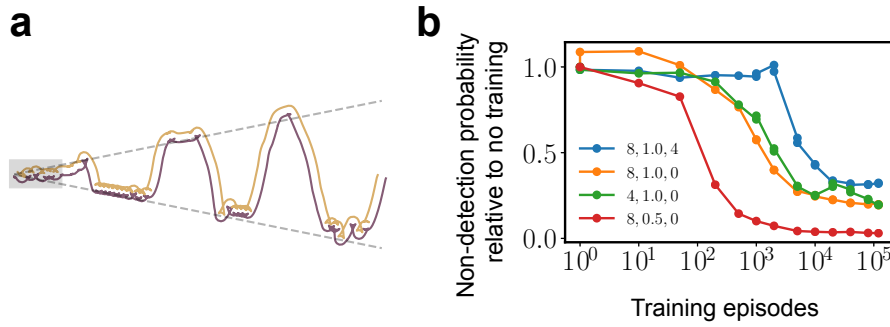
8

Figure M1. Two-sensor RL sector search path and learning curves. (a) The complete search path for an RL agent with two sensors. The gray shaded region close to the last detection point is shown in Figure 1f to highlight the criss-crossing pattern. At large distances, the search path shows casting similar to the one-sensor case (Figure 1e). (b) The probability of not finding the trail until time T (set to 100) as a function of the training time normalized by the probability of not finding the trail without any training. We show results for various values of the parameters $(\alpha, 1/\beta, d)$, where $\alpha \equiv u/a\omega$, $\beta = a\omega/2v\sigma\Theta_0$, and $2d$ is the distance between sensors for the two sensor case (see SI). The curves are meant to demonstrate non-trivial learning for different conditions.

zero and standard deviation $\approx \sigma$. Using $K > 3$ led to similar strategies of search but training was slower. We define the detection probability given latent state i for an agent at r as

$$\bar{\gamma}_i(\boldsymbol{r}) = \int \gamma(\boldsymbol{r}, \phi) B_i(\phi)\, d\phi, \qquad (\text{M2})$$

where $\gamma(\boldsymbol{r}, \phi)$ is the detection probability of finding a trail headed along ϕ if the searcher is at $\boldsymbol{r}$. We assume a Gaussian detection kernel of size a, with distance measured to the closest point on the trail: $\gamma(\boldsymbol{r}, \phi) = e^{-r^2 \sin^2(\theta-\phi)/2a^2} \approx e^{-r^2(\theta-\phi)^2/2a^2}$, where we have used the small-angle approximation. Conditional on no detection at $\boldsymbol{r}$, Bayes' rule yields $\dot{q}_i = -\omega q_i(\bar{\gamma}_i(\boldsymbol{r}) - \bar{\gamma}(\boldsymbol{r}))$, where $\bar{\gamma} = \sum_i q_i \bar{\gamma}_i$ is the total probability of detection. From (M2), we have

$$\bar{\gamma}_i(\boldsymbol{r}) = \sqrt{\frac{\pi}{2}} \frac{\sigma_r}{\sigma_r'} \varphi\left(\frac{\theta - \mu_i}{\sigma_r'}\right), \qquad (\text{M3})$$

where $\sigma_r = a/r$, and $\sigma_r' = \sqrt{\sigma_r^2 + s^2}$.

We use a discount rate λ and provide a reward as discussed below, after (M4). Training is performed in an episodic fashion with each episode lasting time T. The kinematic variables are updated with timestep dt and actions are taken with timestep $dt_{\text{act}} \geq dt$. Movement constraints are imposed by restricting the set of actions to three values, $u/a\omega \in \{-\alpha, 0, \alpha\}$. The state-space has four dimensions, $r, \theta, -\ln q_1/q_2$ and $-\ln q_3/q_2$. We discretize our state space using a non-overlapping tile coding scheme[23]. We refer to the SI for hyperparameter values, details about the state space architecture and the case of two sensors.

We use a SARSA Q-learning algorithm[23], which learns the so-called Q function, that is the value func-

tion for each action in a given state:

$$Q_\pi(\boldsymbol{r}, \boldsymbol{q}, u) = \bar{\gamma}(\boldsymbol{r})dt + \qquad (\text{M4})$$
$$[1 - (\lambda + \bar{\gamma}(\boldsymbol{r}))dt]\, V_\pi\left(\boldsymbol{r} + \boldsymbol{v}dt, \boldsymbol{q} + \dot{\boldsymbol{q}}dt\right),$$

where $V_\pi(\boldsymbol{r}, \boldsymbol{q}) = \sum_u Q_\pi(\boldsymbol{r}, \boldsymbol{q}, u)\pi(u|\boldsymbol{r}, \boldsymbol{q})$, and the index π highlights the dependence on the probabilistic policy $\pi(u|\boldsymbol{r}, \boldsymbol{q})$. The above equation differs from the standard SARSA update by the addition of $\bar{\gamma}dt$ in the discount term, which is due to our formulation of the search as a continuing process conditional on no contact with the trail. Alternatively, one may provide a unit reward when the trail is found, stop the episode and start over. However, the credit assignment problem in goal-oriented tasks makes the training (see (M5) hereafter) problematic even though the final optimal policy is equivalent (see, e.g.,[24]). Our formulation circumvents both issues by (i) giving a local reward $\bar{\gamma}dt$ rather than a final one, which addresses the credit assignment, and (ii) including the detection probability $\bar{\gamma}dt$ into the discount rate to account for the condition of no contact.

To learn the Q-function as defined in (M4), we use a "softmax" training policy, i.e., $\ln \pi(u|\boldsymbol{r}, \boldsymbol{q}) \propto \hat{Q}(\boldsymbol{r}, \boldsymbol{q}, u)/T_{\text{explo}}$, where $\hat{Q}$ is the current estimate of the Q function, and T_{explo} is a "temperature" parameter that is annealed as training progresses to allow for sufficient exploration of actions. Given an action u at state $(\boldsymbol{r}, \boldsymbol{q})$ and a subsequent action u' at state $(\boldsymbol{r}', \boldsymbol{q}')$, $\hat{Q}$ is updated during training as

$$\hat{Q}_\pi(\boldsymbol{r}, \boldsymbol{q}, u) \to \hat{Q}_\pi(\boldsymbol{r}, \boldsymbol{q}, u)(1 - \eta) + \qquad (\text{M5})$$
$$\eta\left(\bar{\gamma}(\boldsymbol{r})dt + (1 - (\lambda + \bar{\gamma}(\boldsymbol{r}))dt)\hat{Q}_\pi(\boldsymbol{r}', \boldsymbol{q}', u')\right),$$

where η is the learning rate. The function $\hat{Q}$ obtained

at the end of the training period yields a search strategy as : $\pi^*(\boldsymbol{r}, \boldsymbol{q}) = \arg\max_u \hat{Q}(\boldsymbol{r}, \boldsymbol{q}, u)$.

We applied the algorithm just described for a range of values of v, α and d (half the distance between sensors for the two-sensor case). For each parameter set, we obtain a search strategy, the corresponding probability $\Gamma_{\mathcal{C}}(T)$ of missing the trail in time T, and the expected number of samples to find the trail. Comparing $\Gamma_{\mathcal{C}}(T)$ on a test set for different number of training episodes (Figure M1b) shows that non-trivial learning takes place, typically saturating at $\sim 10^4$ training episodes.

Casting in a minimal model of sector search

In order to establish the relation between casting and marginal value theory, here we propose a minimal model of sector search. The model lends to an analytical solution which allows us to quantify how the frequency of casting and the efficiency of search depend on the movement and computational constraints imposed upon the tracker.

We consider the same setting as the above episodic RL framework, where the tracker is searching for the trail over a sector after losing contact with it. To focus on casting, we analyze the behavior of the searcher after an initial forward excursion along the most likely heading. This surge decreases the probability weight q of the mode at $\phi = 0$ and yields a symmetric bimodal posterior distribution concentrated at the two modes, $\pm\phi_0$ (typically $\phi_0 \sim \sigma$). The resulting model is equivalent to (M1) with $K = 2$, $\mu_1 = -\phi_0, \mu_2 = \phi_0$ and s small. The searcher moves radially as $r(t)$ (fixed), controls its tangential speed $u = r\dot{\theta}$, receives a unit reward when it finds the trail and incurs a movement cost per unit time, $\mu u^2/2$, where μ sets the movement constraint. The reward and cost are discounted at a rate λ. The two-dimensional state space of the agent consists of θ and the probability q of finding the target at ϕ_0 (the probability is $1-q$ at $-\phi_0$). For full details, we refer to the SI.

The above model is exactly solvable. An optimal searcher exhibits oscillations between $-\phi_0$ and ϕ_0 ("casting") until it finds the trail (Figure M2a). After an initial transient, the searcher traverses a loop in state space (θ, q), alternating between sampling at $-\phi_0$ or ϕ_0, and casting to the other side:

$$(\phi_0, q_s^*) \xrightarrow{\text{cast}} (-\phi_0, q_s^*) \xrightarrow{\text{sample}} (-\phi_0, 1 - q_s^*) \xrightarrow{\text{cast}}$$
(M6)
$$(\phi_0, 1 - q_s^*) \xrightarrow{\text{sample}} (\phi_0, q_s^*),$$

where q_s^* is the optimal switching probability at which the searcher stops sampling at ϕ_0 and traverses to $-\phi_0$.

The speed of traversal from ϕ_0 to $-\phi_0$ is determined by balancing the cost of traversing quickly and the potential value at $-\phi_0$ discounted due to the limited time horizon.

Intuitively, the searcher casts when the marginal value of continuing to sample at ϕ_0 is *just* outweighed by the marginal value that the searcher receives if it stops sampling, traverses from ϕ_0 to $-\phi_0$ and samples at $-\phi_0$. Balancing the marginal value of these two possibilities then yields the optimal switching probability q_s^*. For small q_s^*, we derive (see SI) :

$$q_s^* \approx \frac{\lambda}{1 + \lambda} \frac{\zeta^2}{1 + \zeta + (1 + \lambda)\zeta^2 - \zeta^3}, \qquad (M7)$$

where $\zeta \equiv 1 - \sqrt{2\phi_0^2 \mu \lambda (1 + \lambda)} > 0$. When $\zeta < 0$, the movement cost outweighs the value the agent may receive, and the optimal strategy is to simply not move. q_s^* get smaller and thus the searcher casts less frequently with increasing time horizon ($\lambda \ll 1$) or movement costs ($0 < \zeta \ll 1$). The probability of not detecting the trail after time t is given by

$$\Gamma_{\mathcal{C}}(t) \simeq e^{-I \int_{t_0}^{t} \frac{ds}{r(s)}}, \qquad (M8)$$

where I is interpreted as the rate of information acquisition. Its dependence on μ, λ and ϕ_0 is shown in Figure M2b. As expected, increasing movement cost (decreasing ζ) decreases how quickly information is acquired. Similarly, a large time horizon (λ small) makes the agent sample at $\pm\phi_0$ longer and cast slower, decreasing the rate of information acquisition. For a constant radial speed v, the probability $\Gamma_{\mathcal{C}}(t) \simeq t^{-I/v}$ decreases as a power-law, which arises quite generally from the sector search geometry as discussed in the main text.

Trail statistics

If an agent makes contact with the trail at two points separated by distance L, statistical and geometric information is encoded in the "propagator" $P(\phi_L, \phi_0)$, where ϕ_L, ϕ_0 are the trail headings at the two points (measured relative to the line joining them). If H points of contact are remembered, then the propagator can be used to compute the posterior distribution of the heading: $P(\phi) = \int \mathcal{D}\phi_i P(\phi|\phi_1) P(\phi_1|\phi_2) \ldots P(\phi_{H-2}, \phi_{H-1})$. We introduce an ensemble of trails (which we call the Generalized Worm-Like Chain (GWLC) ensemble) that have persistence in heading and curvature, quantified respectively by the parameters κ and λ (distinct from the discount rate used for RL) and a typical radius of curvature, ξ. Typical samples of trails from this ensemble are presented in Figure M3a. The WLC ensemble[16,17] previously introduced for polymers is a special case with $\xi = \infty$. The propagator $P(\phi_L, \phi_0)$
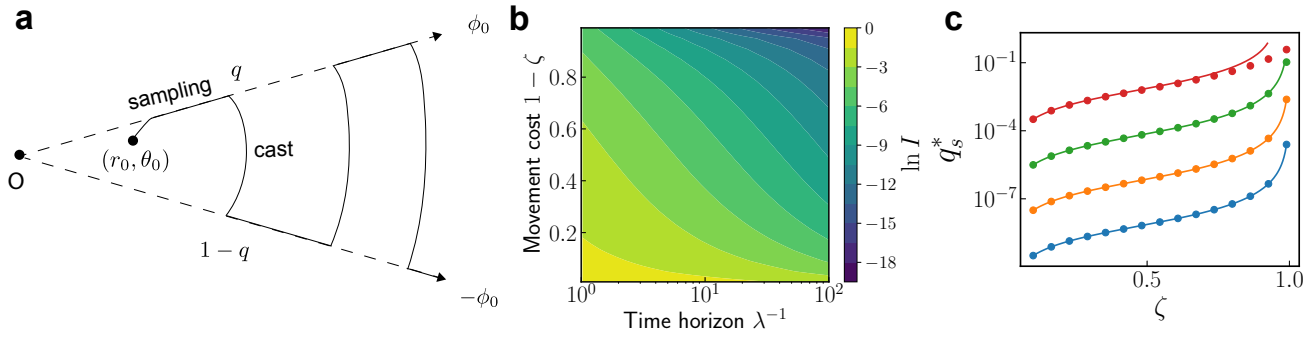
10

Figure M2. (a) The behavior of the agent searching for the trail which is either along $-\phi_0$ or ϕ_0. The agent alternates between sampling at $\pm\phi_0$ and traversing to the other side. q is the agent's estimate of the probability that the trail is along ϕ_0. (b) The rate of information acquisition I as a function of the movement cost $(1 - \zeta)$ and the effective time horizon (λ^{-1}). (c) The optimal switching probability, q_s^*, as a function of ζ for $\lambda = 10^{-8}, 10^{-6}, 10^{-4}, 10^{-2}$ plotted with blue, orange, green and red respectively. The dots are the exact solution obtained numerically (see SI), and the solid lines are from using the approximation (M7).

takes into account all the trails constrained to pass through two contact points, weighted by their probability. In particular, we define

$$P(\phi_L, \phi_0) = Z^{-1} \int \mathcal{D}\phi_x \mathcal{D}\chi_x e^{-\mathcal{E}(\{\phi_x, \chi_x\})}, \quad \text{(M9)}$$

where Z is a normalization constant and the integral is over all possible headings and curvatures ϕ_x, χ_x at the various positions x. The trails satisfy the constraints $y_0 = y_L = 0$ and have end-point headings ϕ_0, ϕ_L. The action $\mathcal{E}$ of a path is given by

$$\mathcal{E}(\{\phi_x, \chi_x\}) = \frac{1}{2\kappa} \int_0^L dx (\dot{\phi}_x - \chi_x)^2 \quad \text{(M10)}$$
$$+ \frac{\lambda\xi^2}{4} \int_0^L dx \left(\dot{\chi}_x + \frac{\chi_x}{\lambda} \right)^2 + \frac{\xi^2 \chi_0^2}{2}.$$

Here, we have oriented the x axis along the line joining the two points of contact and used the small-angle approximation so that $\dot{y}_x = \phi_x$. The model (M10) is a Gaussian process. Since symmetry dictates that $\langle \phi_L \rangle = \langle \phi_0 \rangle = 0$, it follows that $P(\phi_L, \phi_0)$ is defined entirely by the variance $\sigma^2(L) \equiv \langle \phi_L^2 \rangle = \langle \phi_0^2 \rangle$, and by the correlation $\rho(L) \equiv \langle \phi_0 \phi_L \rangle / \sigma_L^2$.

We present the full calculation of $\sigma(L)$ and $\rho(L)$ in the SI. In summary, the first integral over χ_x can be performed using the Gaussian integral formula and leads to an effective action in ϕ_x. The Euler-Lagrange equation of this effective action then yields extremal paths ("splines") that minimize the effective action (see SI for details). The splines have the form:

$$y_x = a_0(L - x) + a_L x + c_0 \frac{(L - x)^3}{6} + c_L \frac{x^3}{6} \quad \text{(M11)}$$
$$+ \lambda^2 d_0 e^{-(L-x)/\lambda} + \lambda^2 d_L e^{-x/\lambda},$$

where the constants in the above equation are set so that $y_0 = y_L = 0$ and can be expressed in terms of ϕ_0, ϕ_L. Figure M3b shows the splines between contact points spaced at increasing intervals. Plugging (M11) in the effective action, we obtain

$$\sigma^2(L) = \frac{\kappa\Omega^2\lambda^2}{L} \frac{b_1 + b_2}{2}, \quad \rho(L) = \frac{b_2 - b_1}{b_2 + b_1}, \quad \text{(M12)}$$

where the two functions b_1 and b_2 are

$$b_1 = \frac{L^2}{2} - \frac{\lambda L}{2\Omega^2 V^2} \left(1 - e^{-\frac{L}{\lambda}} \right) \quad \text{(M13)}$$
$$b_2 = \frac{L^2}{6} - \frac{\lambda(L + 2\lambda)}{2\Omega^2 V^2 L} \left[(L - 2\lambda) + e^{-\frac{L}{\lambda}} (L + 2\lambda) \right],$$

and $V = \sqrt{\kappa\xi^2\lambda/2}, \Omega^2 = \lambda^{-2} + V^{-2}$. The variance $\sigma(L)$ and the correlation $\rho(L)$ are plotted in Figure M3c,d. Three distinct regimes are apparent. For $L/\lambda \ll 1$, we can approximate $\sigma^2(L) \approx \kappa L/3 + L^2/4\xi^2$, which reflects diffusive $\propto L$ and curvature-dominated $\propto L^2$ scalings. When diffusion dominates, $\rho(L) = -1/2$, whereas $\rho(L) \approx -1$ when curvature dominates. The perfect anti-correlation in the curvature-dominated regime is intuitive as the line joining the two points of contact can be viewed as the chord of a circle with radius ξ, whereas $\sigma(L) = L/2\xi$ is the angular deviation of the trail around this chord. When $L/\lambda \gg 1$, the heading is randomized over many correlation lengths and the diffusive scaling $\sigma^2(L) \approx 2\lambda L/3\xi^2$ is recovered.

While $P(\phi_L|\phi_0)$ (the interpolation model) is required to integrate past information, search strategies also require the *forward* propagator $P(\phi_L, y_L|\phi_0)$, which keeps track of trail headings and locations while the agent searches for the trail. The methods described above can be used again (see SI for details) to yield extremal paths as (M11), and the three quantities
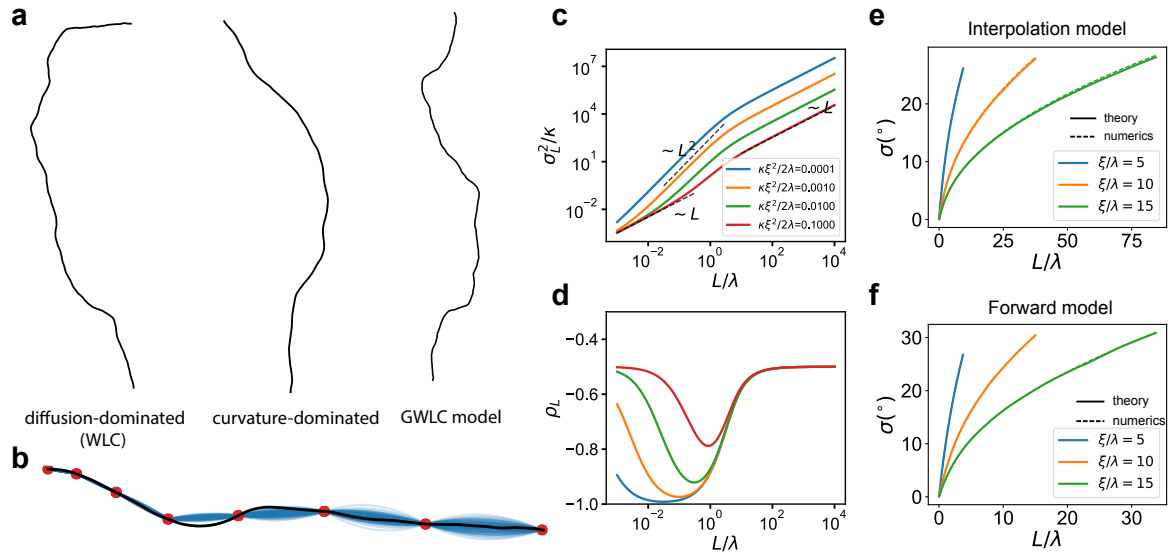
Figure M3. (a) From left to right: samples of trails diffusion-dominated ($\kappa\lambda = 0.05, \xi/\lambda = \infty$), curvature-dominated ($\kappa\lambda = 0, \xi/\lambda = 7$), and trails which have both diffusive and curvature components ($\kappa\lambda = 0.05, \xi/\lambda = 7$). (b) Extremal paths (blue) between detection points (red) on a trail (black) are plotted using (M11) for $\kappa\lambda = 0, \xi/\lambda = 7$. (c,d) The uncertainty σ_L^2 (plotted in log-log scale to highlight the three scaling regimes) and the correlation ρ_L between end-point angles versus the distance between detections, L, for the generalized worm-like chain (GWLC) model. (e,f) Numerical validation (dashed lines) of the theoretical predictions (solid lines) obtained by generating trail samples from the GWLC ensemble.

$\langle\phi_L^2\rangle_{\text{fwd}}, \langle y_L^2\rangle_{\text{fwd}}, \langle\phi_L y_L\rangle_{\text{fwd}}$. These quantities fully describe $P(\phi_L, y_L|\phi_0)$. We validated our interpolation model (M12) and the forward model using numerical simulations (Figure M3e,f).

The non-detection probability during surge and cast

We introduce a sector search strategy that allows us to quantify the non-detection probability taking into account the full dynamics of the trails and yields intuition on the factors that contribute towards losing the trail entirely. We suppose the radial speed v is fixed and the tangential speed $u \gg a\omega$. In other words, the agent casts rapidly within a conical envelope of semi-aperture angle $\sigma\Theta_0$, where σ is the prior uncertainty of the trail's heading and $\Theta_0 \gtrsim 1$ (Figure M4a). To simplify the presentation, we assume curvature-dominated trails, i.e., $\sigma(L) = L/2\xi$, though the arguments below are general.

The rapid casting limit $u/a\omega \gg 1$ allows us to compute the non-detection probability at distance r from the most-recent contact point defined by Eq. (3):

$$\Gamma_{\mathcal{C}}(r) = \langle e^{-\frac{\omega}{v}\int_0^r dx \gamma(x,y(x))}\rangle_{y(x)}, \qquad (M14)$$

where the expectation is over the full ensemble of trails $\{y(x)\}$ that pass through past detection points. Here, we provide intuition for the detection rate $\omega\gamma(x, y(x))$

and refer to the SI for full details. For small x, i.e., $x \lesssim a/2\sigma\Theta_0$, $\gamma(x, y(x)) = 1$ for most trails as the sensor size a spans the entire casting envelope $2\sigma x\Theta_0$. In the casting regime ($x \gtrsim a/2\sigma\Theta_0$), since the time spent on the trail in a single cast is a/u, the probability of non-detection per crossing is $e^{-a\omega/u}$. After n crossings, the non-detection probability is then $e^{-na\omega/u}$. As the agent moves a distance dx in the radial direction, it crosses the trail $n = udx/2v\sigma x\Theta_0$ times, i.e., $\gamma(x, y(x)) = a/2\sigma x\Theta_0$, which is independent of u. From (M14), these relations yield an exponential $\Gamma_{\mathcal{C}}(r)$ during the initial surge followed by a power-law, $r^{-\beta}$ with exponent $\beta \equiv a\omega/2\sigma v\Theta_0$. This heuristic argument aligns well with the results of numerical simulations in Figure M4b.

The spline formulation of the GWLC yields a geometric picture of the dynamics. Intuitively, the non-detection probability (and thus the posterior probability from Eq. (1)) is large if the trail has a large probability of escaping the casting envelope before it is found by the tracker. In order to "escape" the casting envelope, the trails that are initially well within the casting envelope have to bend significantly. This bending incurs a cost in the action (M10), which reduces with increasing r. The non-detection probability conditional on initial trail heading, $\Gamma_{\mathcal{C}}(r|\phi)$, flattens out (i.e., the detection rate vanishes) when a significant fraction of trails escape the casting envelope, including trails with initial
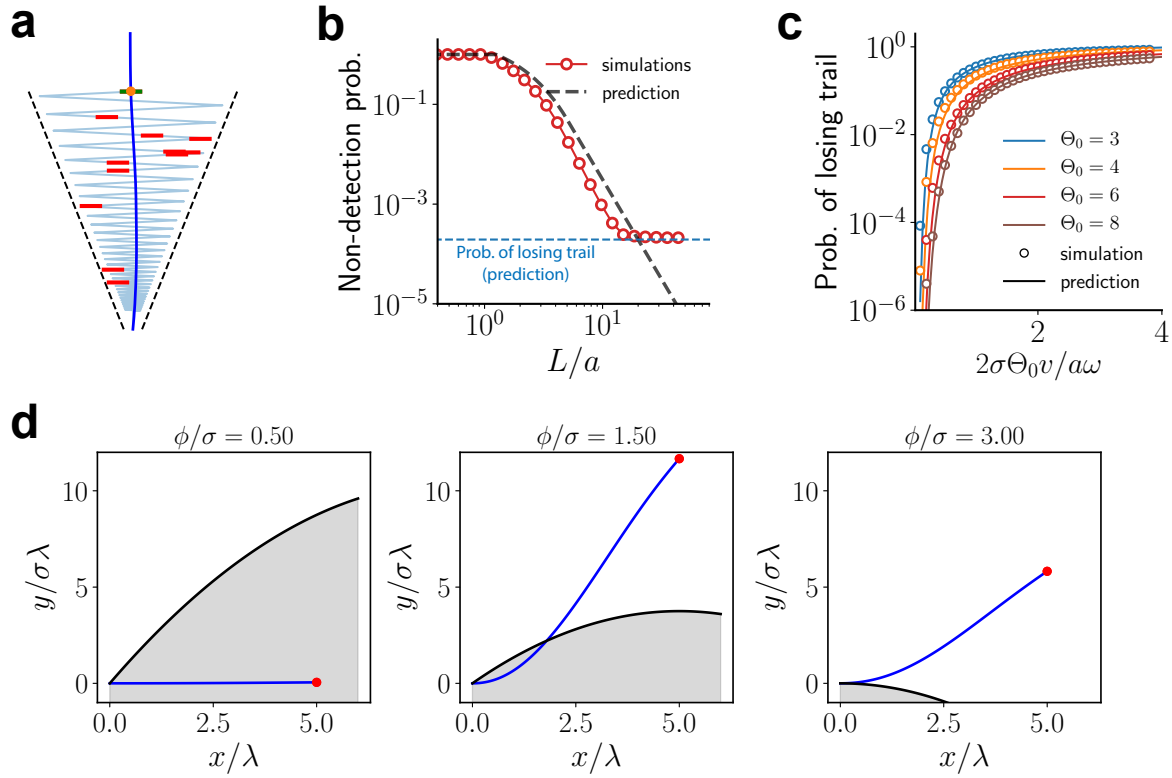
12

Figure M4. Sector search over a conical envelope: simulations. (a) A sample episode of a sector search starting from the most recent point of contact with the trail in dark blue. The tracker samples at discrete locations (red) with Poisson statistics while casting within the conical envelope until it re-establishes contact. (b) The prediction (black dashed line) of an exponential non-detection probability during the surge followed by a power-law decay during casting aligns well with simulations in panel (a) (red circles). Parameters are $1/\beta \equiv 2\sigma v \Theta_0/a\omega = 0.25, \Theta_0 = 5, \sigma = 0.02$. Trail parameters are $\kappa = 0, \xi/\lambda = 5, a/\lambda = 0.1$ for all panels. (c) The probability of losing the trail in simulations (open circles). The solid lines correspond to the theory prediction $\epsilon_{\text{in}} = \Gamma_{\mathcal{C}}(L_{\text{in}})$, where $L_{\text{in}} = 2\sigma\xi\Theta_0$ (see Methods). (d) The path of the trail (blue) that is most likely to go undetected for three values of ϕ. The black line is the rotated casting envelope $R(x) = \sigma x \Theta_{\text{env}}(x) - x\phi, \sigma = 0.17(10°)$.

heading $\phi = 0$ (Figure 3c).

This geometric picture yields a length scale L_{in}, which is the distance at which the trails initially along the most-likely heading $\phi = 0$ escape the casting envelope. For curvature-dominated trails, escaping trails should deviate by an angle $\sigma\Theta_0 \simeq L_{\text{in}}/2\xi$, which gives $L_{\text{in}} \simeq 2\sigma\xi\Theta_0$. The probability of losing the trail can then be estimated as $\epsilon_{\text{in}} = \Gamma_{\mathcal{C}}(L_{\text{in}})$. An additional contribution to the flattening of $\Gamma_{\mathcal{C}}(r)$ comes from the trails which are headed outside of the cone $|\phi| > \Theta_0$ even before reaching L_{in}. The probability $\Gamma_{\mathcal{C}}(r)$ would then flatten out at a different length scale L_{out}, where $\Gamma_{\mathcal{C}}(L_{\text{out}}) = \epsilon_{\text{out}}$. The relative contributions of the two mechanisms are discussed in detail in the SI and they are validated by using numerical simulations as shown in Figure M4c and Figure M5b.

Due to the power-law tail for $\Gamma_{\mathcal{C}}(r)$ with exponent β, the mean distance to find the trail, $\bar{r}$, is determined by the upper or lower cutoff for $\beta < 1$ or $\beta > 1$, respectively. Since the upper cutoff is much larger than the lower cutoff, it is more convenient for the tracker to use $\beta > 1$. By using the self-consistency condition $\sigma = \bar{r}/2\xi$ in the curvature-dominated regime, and the relation $\bar{r} \simeq v\omega^{-1}\left(1 + e^{-\beta}/(\beta - 1)\right)$, we obtain for the maximum speed

$$v^* = \omega\sqrt{a\xi/\Theta_0} \max_{\beta}\left(\beta + \frac{\beta e^{-\beta}}{(\beta - 1)}\right)^{-1} \simeq 0.47\omega\sqrt{a\xi/\Theta_0},$$
(M15)

where the maximum occurs at $\beta^* = 1.63$ (see Figure M5c). The mean number of samples to find the trail at this optimum is ≈ 1.3, indicating that the optimal strategy is for the agent to move slow enough that it typically finds the trail within one or two samples.

Intuition for the above results holds quite generally, including for other statistical ensembles of trails and
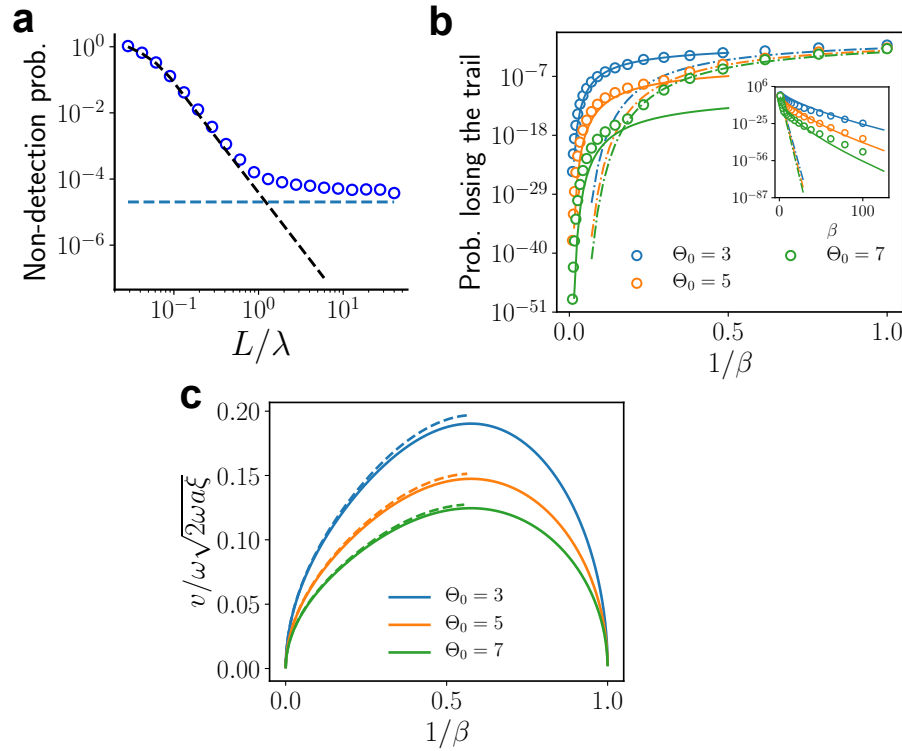
Figure M5. Sector search over a conical envelope: numerics. (a) The non-detection probability obtained by numerically evaluating (M14) (blue circles) instead of simulating the process (Figure M4a). The black and blue dashed lines are the theory predictions as in Figure M4b. (b) The probability of losing the trail computed numerically (circles). The dash-dot lines correspond to predictions (to a constant prefactor of order one) from the regime where the missed trails escape the cone, ϵ_{in}, and the solid lines corresponds to the prediction when the trails outside the cone are the ones most likely to be missed, ϵ_{out}. Inset shows the same data plotted against $\beta \equiv a\omega/2\sigma v\Theta_0$ to highlight the small v behavior. (c) The speed obtained from imposing the self-consistency relationship between IDIs and uncertainty (see Eq. (2)). Solid lines are from theory in the curvature-dominated limit of the GWLC ensemble and dashed lines are from the values that satisfy the self-consistency relationship computed numerically.

for non-conical casting envelopes. The latter is relevant because we expect that a slow down in the radial direction and a widening casting envelope may increase the likelihood of finding the trail. Relaxing the constraints of fixed radial speed and conical envelopes constitutes the aim of the next Section.

Bellman optimization of the casting strategy

Our final step is to formulate the problem of optimizing the casting strategy. The geometry of the search is shown in Figure M6a. To simplify, we parametrize the path by the set of turning points $\{r_i, \theta_i\}$. We assume a fixed speed v (note that the average *radial* speed depends on the strategy), which is set later based on the probability of losing the trail. As discussed in the main text, we optimize the average tracking speed, $\langle L/T \rangle$, after constraining the inter-detection in-

terval, $\langle L \rangle$, using a Lagrange multiplier Λ (see Eq. (4)). This optimization can be recast as a Bellman-type dynamic programming problem by breaking the problem up into discrete steps, each corresponding to a cast from $\{r_i, \theta_i\} \to \{r_{i+1}, \theta_{i+1}\}$ (denoted $\{r', \theta'\} \to \{r, \theta\}$ below for conciseness):

$$\mathcal{V}(r', t', P'(\phi)) = \hspace{3cm} (\text{M16})$$
$$\max_{r,\theta} \left[\left(\frac{r}{t} - \Lambda r \right)(1 - \bar{\Gamma}_{r,\theta}) + \bar{\Gamma}_{r,\theta} \, \mathcal{V}(r, t, P_{r,\theta}(\phi)) \right],$$

where the required reward function in Eq.(4) is $V_{v,\Lambda} = \mathcal{V}(0, 0, P_0(\phi))$ with the Gaussian normal prior $P_0(\phi) = \mathcal{N}(0, \sigma^2)$. The first term in the square bracket returns the reward if the trail is found (weighted by the detection probability $1 - \bar{\Gamma}_{r,\theta}$), while the second term advances to the next cast if detection failed. The Bellman equation thus relates $\mathcal{V}$ before and after a cast by updating the current state variables, $r', t', P'(\phi)$. The time elapsed is updated as $t \to t' + (r - r' + 2r|\theta|)/v$,
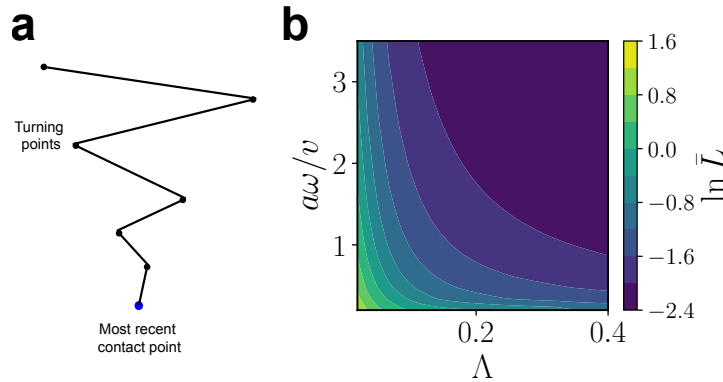
Figure M6. (a) We approximate a zigzagging casting trajectory as a sequence of casts parameterized by turning points $\{r_i, \theta_i\}$, which are defined using polar coordinates with the origin at the last contact point. The turning points are optimized using Bellman optimization so as to maximize the average radial speed $\langle L/T \rangle$ after constraining the average radial distance $\langle L \rangle$. (b) Contours of constant $\bar{L}$ after optimization.

where we approximate $|\theta' - \theta| \sim 2|\theta|$ for simplicity. The prior $P'(\phi)$ is updated using Bayes' rule $P_{r,\theta}(\phi) = \bar{\Gamma}_{r,\theta}(\phi)P'(\phi)/\bar{\Gamma}_{r,\theta}$, where $\bar{\Gamma}_{r,\theta}(\phi)$ is the non-detection probability given the initial trail heading ϕ, and $\bar{\Gamma}_{r,\theta} = \langle \bar{\Gamma}_{r,\theta}(\phi) \rangle_\phi$ is the normalization.

It remains to calculate $\bar{\Gamma}_{r,\theta}(\phi)$, which depends on the *forward* model of the trail. We consider generally that the trail's azimuthal position expands as $\sigma_{\text{fwd}}(r)$ i.e., a trail initially headed along ϕ is located at the azimuthal position, $\phi + \Delta\phi$, where $\Delta\phi \sim \mathcal{N}(0, \sigma_{\text{fwd}}^2(r))$. For the GWLC model, we show in the SI that $\sigma_{\text{fwd}}(r) = \sigma(r)$, where $\sigma(r)$ is given in (M12). Since the time spent on the trail within the casting envelope in a single cast is a/v, we have $\bar{\Gamma}_{r,\theta}(\phi) = e^{-a\omega/v}$ if $|\phi + \Delta\phi| < \theta$ and equal to unity otherwise. Since $\Delta\phi$ is normal-distributed, we have $\langle \mathbb{1}(|\phi + \Delta\phi| < |\theta|) \rangle_{\Delta\phi} = \Phi\left(\frac{\theta - \phi}{\sigma_{\text{fwd}}(r)}\right) - \Phi\left(\frac{-\theta - \phi}{\sigma_{\text{fwd}}(r)}\right)$, where Φ is the normal cumulative distribution function. We finally approximate $\bar{\Gamma}_{r,\theta}(\phi) \approx e^{-\frac{a\omega}{v}\langle \mathbb{1}(|\phi+\Delta\phi|<|\theta|)\rangle_{\Delta\phi}}$, and compute the expectation over ϕ in $\bar{\Gamma}_{r,\theta}$ numerically. The previous approximation greatly simplifies the optimization as $P(\phi)$ is then a sufficient statistic for past measurements (thus allowing the decomposition in Eq. (M16)) and yet captures the effect of the trail's widening trajectory on the optimized casting strategy.

At each casting step, we optimize (using standard black-box optimization methods in SciPy) for $\Delta r = r - r' > 0$ and θ by expanding Eq. (M16) two steps further into the future. The optimization then involves six variables, the immediate pair $\Delta r, \theta$ and the subsequent two pairs. The optimized immediate pair is then used for updating as detailed above, and the process is repeated. Optimizing more than two steps did not yield different results as the $\bar{\Gamma}_{r,\theta}$ factors introduce an effective planning horizon, effectively suppressing contributions from future rewards beyond two steps. For less than two steps, optimization is "greedy" and the corresponding landscape was found to have a qualitatively different, shallow landscape around the optimum. After completing the optimization procedure, the constraint of fixed $\langle L \rangle$ is imposed by varying v and the Lagrange multiplier Λ (see Figure M6b).

Supplementary Information for 'Sector search strategies for odor trail tracking'

Gautam Reddy, Boris I. Shraiman, Massimo Vergassola

1 Additional details of reinforcement learning implementation

Here, we expand on the details of the reinforcement learning implementation presented in the main text and Methods, specifically the case of two sensors and the hyperparameters used for training. We set $a = 1, \omega = 1$ to fix the units.

The agent's position vector $\boldsymbol{r}$ is updated with a time step $dt = 0.05$ and the agent's actions are updated with a time step $dt_{\mathrm{act}} = 0.1$. The agent chooses its tangential speed from three values, $u \in \{-\alpha, 0, \alpha\}$, where α is a parameter. The tangential speed, v_{tang}, is smoothed over a time scale t_{smooth}:

$$\frac{dv_{\mathrm{tang}}}{dt} = \frac{u - v_{\mathrm{tang}}}{t_{\mathrm{smooth}}}, \tag{1}$$

where we use $t_{\mathrm{smooth}} = 1$. The agent's heading is given by $\theta_{\mathrm{agent}} = \tan^{-1}(v_{\mathrm{tang}}/v_{\mathrm{rad}})$, where the radial speed $v_{\mathrm{rad}} = v$ is kept fixed. The two sensors are located at $\boldsymbol{r}_1, \boldsymbol{r}_2$ which are at a distance d from $\boldsymbol{r}$ in a direction perpendicular to the agent's heading. The case of two sensors is similar to the case of one sensor described in the Methods except for the detection probability given latent state i. For one sensor, we have

$$\bar{\gamma}_i(\boldsymbol{r}) = \sqrt{\frac{\pi}{2}} \frac{\sigma_r}{\sigma_r'} \varphi \left(\frac{\theta - \mu_i}{\sigma_r'} \right), \tag{2}$$

where $\sigma_r = 1/r$, and $\sigma_r' = \sqrt{\sigma_r^2 + s^2}$. For two sensors, we assume that each sensor has a Gaussian detection kernel of size a and samples at a frequency $\omega/2$, where the factor of half ensures that the two sensor case reduces to the one sensor case for $d = 0$. The detection probability is then $\bar{\gamma}_i(\boldsymbol{r})|_{\mathrm{two\ sensors}} = (\bar{\gamma}_i(\boldsymbol{r}_1) + \bar{\gamma}_i(\boldsymbol{r}_2))/2$. The rest of the details of the RL implementation remain the same as the one sensor case.

For the figures in the paper, we fix $\sigma = 0.2$ as the strategy depends largely on $\sigma v/u$, which is varied by changing σv for a given α. The state-space has four dimensions, r, θ, $-\ln q_1/q_2$ and $-\ln q_3/q_2$. We discretize our state space using a non-overlapping tile coding scheme: the variable r is discretized using $\sigma v T/5$ (rounded) states between 0 to $\sigma v T$, θ into 15 states between -1.5σ and 1.5σ, and, finally, $-\ln q_1/q_2$ and $-\ln q_3/q_2$ into 5 states each between -4 and 4. We choose a number of samples $\omega T = 100$, which implies that if $v/a\omega = 1$, for instance, we have 20 states for r and a total of $20 \times 15 \times 5 \times 5 = 7500$ states that tile the state space.

As noted in the Methods, we use a softmax policy with an annealing scheme for training. For each state s, we first normalize the Q-values: $Q(s,a) \to \frac{Q(s,a)-\min_{a'} Q(s,a')}{\max_{a'} Q(s,a')-\min_{a'} Q(s,a')}$ so that the normalized Q values lie between 0 and 1. The "temperature" parameter, T_{explo} is first set to 3 and annealed linearly down to 0.2 from the first training episode to the last training episode. For generating trajectories after learning, we set $T_{\text{explo}} = 0.2$. We use a learning rate $\eta = 0.02$ and discount rate λ such that $e^{-\lambda dt_{\text{act}}} = 0.98$ which corresponds to a time horizon $\lambda^{-1} \approx 5$ samples.

2   Optimal stochastic control and searching

We first describe a general formulation of searching in the framework of optimal control and then proceed to analyze the case of sector search in more detail further below.

2.1   The Hamilton-Jacobi-Bellman equation for search problems

Consider an optimal control problem where the agent searches for a target at an unknown location $\boldsymbol{y}$. The agent has prior knowledge encoded in the vector $\boldsymbol{\psi}$ which is updated during the search. The agent receives a reward of one and stops searching upon finding the target. The probability per unit time of finding the target for an agent at $\boldsymbol{x}$ is given by $\omega\gamma(\boldsymbol{x},\boldsymbol{y})$ with γ being the detection kernel of size a. Additionally, there is a cost per unit time of controlling its velocity, $\dot{\boldsymbol{x}} \equiv \boldsymbol{u}$, along various directions, $\mu u^2/2$ (u is speed and μ is some constant), and the rewards and costs are discounted per unit time using a discount rate λ. In this section, we set $\omega = 1, a = 1$.

The agent has an internal model of how $\boldsymbol{\psi}$ is updated when it doesn't detect the target at $\boldsymbol{x}$. In general, $\boldsymbol{\psi}$ represents all the information that the agent maintains about the location of the target. In particular, if we take $\boldsymbol{\psi}$ as a posterior distribution, $P(\boldsymbol{y})$, updated using Bayes' rule starting from some prior, we have

$$\frac{d \ln P(\boldsymbol{y})}{dt} = -\gamma(\boldsymbol{x}(t),\boldsymbol{y}) + \bar{\gamma}(\boldsymbol{x}(t)), \quad \bar{\gamma}(\boldsymbol{x}) = \int \gamma(\boldsymbol{x},\boldsymbol{y}')P(\boldsymbol{y}')d\boldsymbol{y}', \tag{3}$$

where $\boldsymbol{x}(t)$ is the position of the agent at time t. Note that $\bar{\gamma}(\boldsymbol{x})$ is the probability per unit time of detecting the target given current information. We introduce the value function $V(\boldsymbol{x},\boldsymbol{\psi})$, which is the optimal expected sum of discounted rewards given the current position of the agent and available information about the target. The evolution of the value function in a time interval dt satisfies the dynamic programming equation

$$V(\boldsymbol{x},\boldsymbol{\psi}) = \max_{\boldsymbol{u}} \left\{ \bar{\gamma}dt - \frac{\mu u^2}{2}dt + (1-\lambda dt)(1-\bar{\gamma}dt)V(\boldsymbol{x}+\boldsymbol{u}dt, \boldsymbol{\psi}+\dot{\boldsymbol{\psi}}dt) \right\} \tag{4}$$

$$= \max_{\boldsymbol{u}} \left\{ (\bar{\gamma} - \frac{\mu u^2}{2})dt + (1-dt(\lambda+\bar{\gamma}))V(\boldsymbol{x},\boldsymbol{\psi}) + dt\left(\boldsymbol{u}.\frac{\partial V}{\partial \boldsymbol{x}} + \dot{\boldsymbol{\psi}}.\frac{\partial V}{\partial \boldsymbol{\psi}} \right) \right\}. \tag{5}$$

The first two terms on the right hand side of (4) are the probability of detecting the target and the cost of control in the interval dt respectively. The latter term is the discounted value function if the process continues, i.e., target is not detected in dt. The optimal policy is obtained by optimizing the terms in the curly brackets with respect to the control variables

$\boldsymbol{u}$. Taking the gradient w.r.t $\boldsymbol{u}$ and equating to zero, we get

$$\boldsymbol{u}^* = \mu^{-1} \frac{\partial V}{\partial \boldsymbol{x}}. \tag{6}$$

Plugging this back into (5) and canceling $V(\boldsymbol{x}, \boldsymbol{\psi})$ on both sides, we get the Hamilton-Jacobi-Bellman (HJB) equation

$$\frac{1}{2\mu} \left(\frac{\partial V}{\partial \boldsymbol{x}} \right)^2 + \dot{\boldsymbol{\psi}} \cdot \frac{\partial V}{\partial \boldsymbol{\psi}} - (\bar{\gamma} + \lambda)V + \bar{\gamma} = 0. \tag{7}$$

The solution of the above HJB equation yields the optimal policy via (6). In most cases, $\boldsymbol{\psi}$ is high-dimensional and an exact solution is infeasible. We consider a minimal model of sector search which enables an exact solution, as described below.

2.2 A minimal model of casting

We consider a minimal model of trail tracking discussed in the Methods where the search is over a one-dimensional angular space. The model naturally gives rise to casting dynamics and quantifies how the frequency of casts and the efficiency of search depends on various movement and computational constraints.

Consider two point targets located at $\pm\phi_0$. As noted in the Methods, these correspond to the modes of a bimodal posterior distribution of the heading of the trail after an initial forward excursion. Suppose q is the probability of the target at ϕ_0. When the agent is at $\theta = \pm\phi_0$, we assume

$$\dot{q} = \mp q(1-q)/r \tag{8}$$

respectively, where the factor $1/r$ is the angle subtended by the detector at distance r from the last contact (note that we have set $a = 1, \omega = 1$). For intermediate values of θ, $\dot{q} = 0$. The radial path $r(t)$ is assumed fixed a *priori* with $r(0) = r_0 > 0$.

We would like to find the optimal control variable $\dot{\theta}$ that maximizes the reward while minimizing a energy cost per unit time $E = \mu r^2 \dot{\theta}^2/2$. However, the dependence on r makes the problem non-stationary and challenging to analyze. To circumvent this issue, we introduce the coordinate $\tau = \int_0^t \frac{dt'}{r(t')}$. We have $\frac{dq}{dt} = \frac{dq}{d\tau}\frac{1}{r}$ and $r\frac{d\theta}{dt} = \frac{d\theta}{d\tau}$, so that the update rule on q and the energy cost are both independent of r in the new coordinates. We use $\dot{q}$ and $\dot{\theta}$ to denote the derivative w.r.t τ henceforth.

We re-write equation (4) for a one-dimensional search (along θ) with the above assumptions:

$$V(\theta, q) = \max_{\dot{\theta}} \left\{ \left(\bar{\gamma}(\theta) - \frac{\mu\dot{\theta}^2}{2} \right) d\tau \right.$$

$$\left. + (1 - (\lambda + \bar{\gamma}(\theta))d\tau)V(\theta + \dot{\theta}d\tau, q + \dot{q}d\tau) \right\}, \tag{9}$$

where $\bar{\gamma}(\theta)$ is the detection rate and λ is the effective discount rate in the new coordinates; if the time horizon in the t-scale is T, we have an effective time horizon $\lambda^{-1} \sim \int_0^T \frac{dt'}{r(t')}$ in the

3

τ-scale. Taylor expanding the V in the curly brackets above to first order and optimizing for $\dot{\theta}$ we have

$$\dot{\theta}^* = \mu^{-1}\frac{\partial V}{\partial \theta}. \tag{10}$$

Plugging this back into (9) and canceling $V(\theta, q)$ on both sides, we get the HJB equation

$$\frac{1}{2\mu}\left(\frac{\partial V}{\partial \theta}\right)^2 + \dot{q}\frac{\partial V}{\partial q} - (\bar{\gamma}(\theta) + \lambda)V + \bar{\gamma}(\theta) = 0. \tag{11}$$

2.3 Solution to the HJB equation

The dynamics of the agent is as follows: 1) If the agent is initially between $-\phi_0$ and ϕ_0, it chooses to travel to either $-\phi_0$ or ϕ_0 depending on its current θ and q. In other words, there is a location $\tilde{\theta}(q)$, such that the agent travels to the right if the current position $\theta > \tilde{\theta}(q)$ and vice-versa. 2) If the agent chooses to travel to ϕ_0, say, then it travels at a speed that balances out the cost of traveling with the discounted value it expects to obtain at ϕ_0. 3) Once at ϕ_0, the agent begins sampling and if it doesn't find the target, it remains stationary until q decreases to a 'switching' probability q_s. At that point, a decision is made to traverse to $-\phi_0$. The choice of q_s depends on the balance between the cost of traversing (at an optimized speed) and the potential future reward at $-\phi_0$ versus the expected reward from remaining at ϕ_0. Once at $-\phi_0$, by symmetry, the agent samples until q increases to $1 - q_s$ or until it finds the target. Eventually, the agent cycles between $-\phi_0$ and ϕ_0 at an optimized speed, sampling for a fixed amount of time at each end, until it finds the target.

There are three choices made by the agent corresponding to the above three steps of the dynamics: 1) Given the initial θ and q, which direction should the agent move i.e., what is $\tilde{\theta}(q)$? 2) At what speed should the agent move? 3) When should the agent switch from searching at one end to the other i.e., what is the optimal value of q_s, q_s^*? Below, we calculate q_s^* and the optimal speed. We do not calculate $\tilde{\theta}(q)$, since it does not affect the agent's long-term behavior of cycling between $-\phi_0$ and ϕ_0.

We have $\dot{q} = \mp q(1 - q)$ depending on whether $\theta = \pm\phi_0$, and $\dot{q} = 0$ otherwise. Likewise, the detection rate is $\bar{\gamma} = q, 1 - q$ at $\phi_0, -\phi_0$ respectively, and 0 otherwise. Since the agent remains stationary while sampling at $\theta = \phi_0, -\phi_0$, we have $\partial V/\partial \theta = 0$ from (10) at these two points. Putting these together, (11) can be re-written as the set of equations

$$\frac{1}{2\mu}\left(\frac{\partial V}{\partial \theta}\right)^2 - \lambda V = 0, \quad -\phi_0 < \theta < \phi_0, \tag{12}$$

$$q(1 - q)\frac{\partial V}{\partial q} + (q + \lambda)V - q = 0, \quad \theta = \phi_0, q \geq q_s, \tag{13}$$

$$q(1 - q)\frac{\partial V}{\partial q} - (1 - q + \lambda)V + 1 - q = 0, \quad \theta = -\phi_0, q \leq 1 - q_s. \tag{14}$$

The first equation is the HJB equation of a "free particle" (with an additional discount factor) which describes the motion of the agent while it traverses from one end to the other. The latter two equations capture the evolution of the value function as the agent samples at ϕ_0 and $-\phi_0$ respectively conditional on not finding the target.

4

The agent traverses a loop in state space:

$$(\phi_0, q_s) \xrightarrow{\text{cast}} (-\phi_0, q_s) \xrightarrow{\text{sample}} (-\phi_0, 1 - q_s) \xrightarrow{\text{text}} (\phi_0, 1 - q_s) \xrightarrow{\text{sample}} (\phi_0, q_s),$$

The solution is found by following the value function around this loop, obtained by integrating the above set of equations and matching the value function at (ϕ_0, q_s) at the beginning and the end of the loop. We exploit the symmetry $V(\phi_0, q_s) = V(-\phi_0, 1 - q_s)$, which reduces the problem to matching the first and third points in the above loop. For an arbitrary choice of q_s, the solution to the first equation yields the value function during a traversal between $-\phi_0$ and ϕ_0, which is then used to match the value functions at the two boundaries obtained from the latter two equations. Optimizing over q_s then yields the optimal policy.

$-\phi_0 < \theta < \phi_0$: Consider the first equation for $-\phi_0 < \theta < \phi_0$. Equation (12) has two possible solutions

$$V_\pm(\theta, q) = \frac{\mu\lambda}{2} \left(\theta - h_\pm(q)\right)^2, \tag{15}$$

where $+/-$ corresponds to the solution where the agent moves right/left. At $\theta = \tilde{\theta}(q)$, the two solutions meet. Since V is continuous, we have $(\tilde{\theta}(q) - h_+(q))^2 = (\tilde{\theta}(q) - h_-(q))^2$ giving $h_+(q) + h_-(q) = 2\tilde{\theta}(q)$. We re-write $h_\pm(q) = \tilde{\theta}(q) \mp g(q)\sqrt{\frac{2}{\mu\lambda}}$ (with $g(q) > 0$) to get

$$V(\theta, q) = \left(\sqrt{\frac{\mu\lambda}{2}} \left(\theta - \tilde{\theta}(q)\right) \pm g(q)\right)^2. \tag{16}$$

The above expression represents the cost of travel and connects the value at the boundaries; the two unknown functions $\tilde{\theta}(q)$ and $g(q)$ can be found by matching the value functions at $\pm\phi_0$. Note that V is positive as expected; if not, the agent would simply choose to stop and incur no control cost. From the equation for the optimal speed (10), for $\theta > \tilde{\theta}(q)$, the agent moves to the right with speed $\lambda(\theta - \tilde{\theta}(q)) + \sqrt{\frac{2\lambda}{\mu}}g(q)$ until it reaches ϕ_0. The trajectory is therefore an exponential with exponent λ. By symmetry, we expect $\tilde{\theta}(q) = -\tilde{\theta}(1 - q)$ and $V(\theta, q) = V(-\theta, 1 - q)$. In particular, $\tilde{\theta}(1/2) = 0$, i.e., moving to the right or left are equally favorable if the agent is at $\theta = 0$ and $q = 1/2$.

$\theta = \phi_0$: Consider now the HJB equation at $\theta = \phi_0$. By inspection, equation (13) has the particular solution $q(1 + \lambda)^{-1}$. The general solution is then the particular solution plus a constant C times the homogeneous solution $(1 - q)^{1+\lambda}/q^\lambda$:

$$V(\phi_0, q) = C \frac{(1 - q)^{1+\lambda}}{q^\lambda} + \frac{q}{1 + \lambda}. \tag{17}$$

Note that the above expression satisfies the boundary condition $V(\phi_0, 1) = (1 + \lambda)^{-1}$, consistent with the expectation that if the target is known to be at ϕ_0 and the agent is at ϕ_0, the expected discounted reward is $\int_0^\infty e^{-\lambda\tau}e^{-\tau}d\tau = (1 + \lambda)^{-1}$. Equation (17) has an alternative, perhaps more intuitive, derivation. Recall the switching probability q_s at which point the agent decides to leave ϕ_0. The value at $V(\phi_0, q)$ is the expected discounted reward of

5

finding the target before the posterior q gets to q_s plus the discounted value at $V(\phi_0, q_s)$ if the agent doesn't find the target. Since $\dot{q} = -q(1-q)$, the sampling time for q to get to q_s is $\tau_s = \int_{q_s}^{q} dq/q(1-q) = \ln \frac{q}{1-q} - \ln \frac{q_s}{1-q_s}$. The expected discounted reward is therefore $q \int_0^{\tau_s} e^{-\lambda\tau} e^{-\tau} d\tau = \frac{q}{1+\lambda}(1 - e^{-(1+\lambda)\tau_s})$. The probability of not finding the target in τ_s is $1 - q + qe^{-\tau_s}$. Combining these expressions we get

$$V(\phi_0, q) = \frac{q}{1+\lambda}(1 - e^{-(1+\lambda)\tau_s}) \tag{18}$$
$$+ (1 - q + qe^{-\tau_s})e^{-\lambda\tau_s}V(\phi_0, q_s)$$

Plugging in the expression for τ_s yields (17). The equation holds for $q \geq q_s$, for any arbitrary q_s. From the second derivation, it is clearer that the constant C in (17) contains all the dependencies on q_s. Since the optimal switching point, q_s^*, corresponds to the point at which the value is maximized, it is therefore the point at which $C(q_s)$ attains its maximum.

2.3.1 Relationship with marginal value theory

Before proceeding to find q_s^*, it is worth making a few clarifying points regarding the calculation of the optimal switching probability q_s^* as the maximum of $C(q_s)$ and its relationship with marginal value theory.

In marginal value theory as applied to foraging, one computes the optimal switching point as the point at which the value of continuing to sample at a certain location equals the value of switching to another location. If the agent decides to continue sampling at q_s^*, it would imply that (13) applies just below q_s^* as well as above it, where $\frac{\partial V}{\partial q}$ is instead the left derivative at q_s^*. Continuity of V then implies that the left and right derivatives match at the optimal switching probability

$$\frac{\partial V}{\partial q}\Big|_{q_s^*,+} = \frac{\partial V}{\partial q}\Big|_{q_s^*,-}. \tag{19}$$

We now show that this condition is equivalent to maximizing $C(q_s)$ w.r.t q_s. Call $f_1(q) = (1-q)^{1+\lambda}/q^\lambda$ and $f_2(q) = q/(1+\lambda)$. From (17) we have for the right derivative $\frac{\partial V}{\partial q}\Big|_{q_s,+} = C(q_s)f_1'(q_s) + f_2'(q_s)$. For $q < q_s$, the agent immediately traverses from ϕ_0 to $-\phi_0$. Using (16) at ϕ_0 and $-\phi_0$ we get

$$V(\phi_0, q)^{1/2} = -\sqrt{2\mu\lambda}\phi_0 + V(-\phi_0, q)^{1/2} \tag{20}$$
$$= -\sqrt{2\mu\lambda}\phi_0 + V(\phi_0, 1-q)^{1/2} \tag{21}$$

Equation (17) applies to $V(\phi_0, 1-q)$ since $1 - q > 1 - q_s \geq q_s$. Taking the derivative of (21) w.r.t q on both sides

$$\frac{1}{\sqrt{V(\phi_0, q)}}\frac{\partial V}{\partial q} = \frac{1}{\sqrt{V(\phi_0, 1-q)}} \times$$
$$\left(C(q_s)f_1'(1-q) + f_2'(1-q)\right), \tag{22}$$

which yields the left derivative of V at q_s. At the optimal switching point, equating the left

and right derivatives at q_s^*, we get the condition for the optimal switching probability

$$\frac{1}{\sqrt{V(\phi_0, q_s^*)}} \left(C(q_s^*) f_1'(q_s^*) + f_2'(q_s^*) \right)$$
$$= \frac{1}{\sqrt{V(\phi_0, 1 - q_s^*)}} \left(C(q_s^*) f_1'(1 - q_s^*) + f_2'(1 - q_s^*) \right) \quad (23)$$

For an arbitrary switching point q_s, using (16) at q_s, we have

$$V(\phi_0, 1 - q_s)^{1/2} = \sqrt{2\mu\lambda}\phi_0 + V(\phi_0, q_s)^{1/2}. \quad (24)$$

Plugging in (17) for the two value functions, taking the derivative w.r.t q_s on both sides and imposing (23) at q_s^*, it is cumbersome but straightforward to show that $dC(q_s)/dq_s = 0$ at q_s^*.

2.3.2 Optimizing for q_s

Finally, to calculate the maximum of $C(q_s)$, we use (24) to first calculate $C(q_s)$ for any q_s. Using (17) for $V(\phi_0, q_s)$ and $V(\phi_0, 1 - q_s)$ in (24), and solving for $C(q_s)$, we get

$$C(q_s) = -\frac{c_2 + \sqrt{\Delta}}{c_1^2}, \quad \text{where} \quad (25)$$

$$c_1 = \frac{q_s^{1+\lambda}}{(1 - q_s)^\lambda} - \frac{(1 - q_s)^{1+\lambda}}{q_s^\lambda}, \quad (26)$$

$$c_2 = c_1 \left(\frac{1 - 2q_s}{1 + \lambda} - 2\mu\lambda\phi_0^2 \right) - 4\mu\lambda\phi_0^2 \frac{(1 - q_s)^{1+\lambda}}{q_s^\lambda}, \quad (27)$$

$$\Delta = c_2^2 - c_1^2 \left(\left(\frac{1 - 2q_s}{1 + \lambda} - 2\mu\lambda\phi_0^2 \right)^2 - 8\mu\lambda\phi_0^2 \frac{q_s}{1 + \lambda} \right). \quad (28)$$

The value of q_s^* is then obtained using $dC/dq_s = 0$ if there is a maximum for $q_s \leq 1/2$ or otherwise $q_s = 0$. To get a sense of what the solution looks like, for $\lambda \ll 1$ or when the cost of control is large (but below the point where it is no longer optimal to traverse), we expect $q_s^* \ll 1$. In this limit, we use (24) and match terms up to $O(q_s^{2+\lambda})$ to get

$$C(q_s) = \frac{q_s^\lambda}{1 + \lambda} \left(\zeta^2 - q_s \left(1 + \zeta - \zeta^2(1 + \lambda) - \zeta^3 q_s^{2\lambda} \right) \right)$$
$$+ O(q_s^{2+\lambda}) \quad (29)$$

where $\zeta \equiv 1 - \sqrt{2\phi_0^2\mu\lambda(1 + \lambda)}$. The above expression holds for $\zeta > 0$. Equating dC/dq_s to zero, we get

$$q_s^* \approx \frac{\lambda}{1 + \lambda} \frac{\zeta^2}{1 + \zeta + (1 + \lambda)\zeta^2 - \zeta^3}, \quad (30)$$

where we have used $q_s^{2\lambda} \approx 1$ for $q_s \sim \lambda$ and $\lambda \ll 1$. For $\zeta \ll 1$, that term can be ignored as its pre-factor is ζ^3. For $\zeta \leq 0$, (24) cannot be satisfied, and we have $q_s^* = 0$. In Figure M2c, we verify the above approximation for q_s^*.

7

2.4 Rate of information acquisition

An agent has a constant probability of finding the target per cycle. We compute the rate at which the probability of *not* detecting the target decays with time, which quantifies the rate of information acquired by the agent about the target. After $\tau = n\tau_c$ i.e., after n cycles of τ_c each, the probability of not detecting the target $\Gamma_\mathcal{C}(\tau)$ is

$$\Gamma_\mathcal{C}(\tau) \simeq \nu^\tau, \tag{31}$$

where $\nu^{\tau_c} < 1$ is the probability of not finding the target per cycle computed below. Denote τ_s as the time spent sampling at ϕ_0 (or $-\phi_0$) per cycle respectively. Since at ϕ_0, the agent samples such that q goes from $1 - q_s^*$ to q_s^*, we have $\tau_s = -\int_{1-q_s^*}^{q_s^*} dq/q(1-q) = 2\ln\frac{1-q_s^*}{q_s^*}$. The probability of finding the target in this period is $q_s^* \int_0^{\tau_s} e^{-\tau} d\tau = q_s^*(1 - e^{-\tau_s})$. Therefore, we have

$$\nu^{\tau_c} = \left(1 - q_s^*(1 - e^{-\tau_s})\right)^2. \tag{32}$$

It remains to calculate τ_c. If τ_t is the time required for traversing from ϕ_0 to $-\phi_0$ (or vice-versa), we have $\tau_c = 2(\tau_t + \tau_s)$. In (16) note that $\tilde{\theta}(q_s^*) = \phi_0$ so that $g(q_s^*) = \sqrt{V(\phi_0, q_s^*)}$ (which can be evaluated using (17)). The leftward speed when $q = q_s^*$ is thus

$$\frac{d\theta}{d\tau} = \lambda(\theta - \phi_0) - \sqrt{\frac{2\lambda V(\phi_0, q_s^*)}{\mu}}. \tag{33}$$

This is easily integrated to give

$$\tau_t = \lambda^{-1} \ln\left(1 + \sqrt{\frac{2\lambda\mu\phi_0^2}{V(\phi_0, q_s^*)}}\right). \tag{34}$$

To summarize,

$$\Gamma_\mathcal{C}(\tau) \simeq e^{-I\tau}, \quad \text{where} \tag{35}$$

$$I = -\frac{1}{\tau_t + \tau_s} \ln\left(1 - q_s^*(1 - e^{-\tau_s})\right) \tag{36}$$

is interpreted as the rate of information acquisition. Reverting back to the original t-scale, we get

$$\Gamma_\mathcal{C}(t) \simeq e^{-I \int_0^t \frac{dt'}{r(t')}}. \tag{37}$$

For the specific case of constant radial speed v, $r(t) = vt + r_0$, the probability of not detecting decays as a power law,

$$\Gamma_\mathcal{C}(t) \simeq t^{-\frac{I}{v}}. \tag{38}$$

The mean time to find the target therefore diverges above a speed $v^*/a\omega = I$, which is set by a combination of the uncertainty ϕ_0, the movement cost μ and the discount rate λ.

8

3 The Generalized Worm-like Chain (GWLC) ensemble of trails

Suppose ϕ_0, ϕ_L are the headings at two points separated by distance L. We first compute the joint distribution $P(\phi_0, \phi_L)$ for an Gaussian ensemble of trails with persistence in heading (WLC) and then introduce a new model with persistence in curvature, which we call the Generalized Worm-like Chain (GWLC) ensemble. While $P(\phi_0, \phi_L)$ is the "interpolation" model, we will subsequently also calculate the forward model of how the trail's heading and transverse position changes with distance from a fixed end.

Since ϕ_0, ϕ_L are jointly normal, it is sufficient to compute their means and the covariance matrix. We denote b_1, b_2 for the diagonal and off-diagonal elements respectively of the inverse covariance matrix. In the interpolation models, the agent makes contact with the trail at $x = 0, y = 0$ and $x = L, y = 0$. Small-angle approximation is used throughout; its validity is examined further below. $\phi_x = \dot{y}_x$ is the local heading (i.e., the tangent) and χ_x is the curvature.

3.1 Wormlike-chain

We begin by considering a simple wormlike-chain (WLC) model with no force, which corresponds to the model in the main text without curvature. The probability of a given path is

$$P(\{\phi_x\}) \sim e^{-\mathcal{E}(\{\phi_x\})}, \quad \text{where} \quad \mathcal{E} = \frac{1}{2\kappa} \int_0^L dx \dot{\phi}_x^2 = \frac{1}{2\kappa} \int_0^L dx \ddot{y}^2 \tag{39}$$

The Euler-Lagrange equation is $d^4y/dx^4 = 0$, so that the extremal path is a cubic polynomial, $y_x = \sum_{k=0}^3 a_k x^3$. There are four parameters. Setting $y_0 = 0$ and $y_L = 0$, writing a_k's in terms of ϕ_L, ϕ_0 and plugging this in (39), we get the joint distribution

$$P(\phi_0, \phi_L) \sim \exp\left(-\frac{2}{\kappa L}\left(\phi_0^2 + \phi_0\phi_L + \phi_L^2\right)\right) \tag{40}$$

The inverse covariance matrix therefore has diagonal elements, $b_1 = 4/\kappa L$ and off-diagonal elements, $b_2 = 2/\kappa L$. From inverting the matrix, the correlation is $\rho_L = -b_2/b_1$ and $\sigma_L^2 \equiv \langle\phi_L^2\rangle = \langle\phi_0^2\rangle = b_1/(b_1^2 - b_2^2)$, which gives

$$\rho(L) = -1/2, \quad \sigma^2(L) = \kappa L/3. \tag{41}$$

3.2 Fixed but unknown curvature

We now consider the case when the trail has a fixed but unknown curvature $\chi \sim \mathcal{N}(0, \xi^{-2})$, where ξ is the typical radius of curvature. Given χ, the WLC action is modified as

$$\mathcal{E}(\chi) = \frac{1}{2\kappa} \int_0^L dx(\dot{\phi}_x - \chi)^2. \tag{42}$$

When the hidden variable χ is integrated out, we get

$$\mathcal{E} = \frac{1}{2\kappa}\left(\int_0^L dx \dot{\phi}_x^2 - \frac{(\phi_L - \phi_0)^2}{L + \kappa\xi^2}\right). \tag{43}$$

9

This looks much like the WLC case from before except for the extra boundary terms. These don't affect the extremal path which is still a cubic polynomial. Plugging this in as before after setting $y_0, y_L = 0$, we get

$$P(\phi_0, \phi_L) \sim \exp\left(-\frac{2}{\kappa L}\left((\phi_0^2 + \phi_0\phi_L + \phi_L^2) - \frac{(\phi_L - \phi_0)^2}{4\left(1 + \kappa\xi^2/L\right)}\right)\right). \tag{44}$$

We can immediately see that when $\kappa\xi^2/L \gg 1$, this reduces to the WLC case. This is simply the case of $(\kappa L)^{1/2} \gg L/\xi$ i.e., the curve diffuses much more than it curves. Call $\mu_L = \kappa\xi^2/L$. Expanding the above equation, we have

$$b_1 = \frac{4}{\kappa L}\left(1 - \frac{1}{4(1 + \mu_L)}\right), \tag{45}$$

$$b_2 = \frac{4}{\kappa L}\left(\frac{1}{2} + \frac{1}{4(1 + \mu_L)}\right), \tag{46}$$

so that

$$\rho(L) = -\frac{1 + 2\mu_L/3}{1 + 4\mu_L/3}, \quad \sigma^2(L) = \frac{\kappa L}{3}\left(1 + \frac{3}{4\mu_L}\right). \tag{47}$$

When $\mu_L \gg 1$, we recover $\rho(L) \simeq -1/2$ and $\sigma^2(L) \simeq \kappa L/3$. On the other hand, when $\mu_L \ll 1$, we have

$$\rho(L) \simeq -1 + 2\mu_L/3, \quad \sigma^2(L) \simeq L^2/4\xi^2. \tag{48}$$

3.3 Varying curvature with a finite correlation length (GWLC)

In the previous section, we calculated the case of a fixed curvature, i.e., the curvature has infinite correlation length. Now, we assume the curvature is an Ornstein-Uhlenbeck process with correlation length λ. This corresponds to the full model considered in the main text. In this case, we have to integrate out the uncertainty over the initial curvature χ_0 as well as the stochastic curvature field.

The full action including the χ_x curvature field and the stationary prior is

$$\mathcal{E}(\chi) = \frac{1}{2\kappa}\int_0^L dx(\dot{\phi}_x - \chi_x)^2 + \frac{\lambda\xi^2}{4}\int_0^L dx\left(\dot{\chi}_x + \frac{\chi_x}{\lambda}\right)^2 + \frac{\xi^2\chi_0^2}{2} \tag{49}$$

The prefactor $\lambda\xi^2/4$ ensures the stationary distribution of χ is $\mathcal{N}(0, \xi^{-2})$. We integrate χ_x out using the Gaussian integral formula and obtain the effective action in ϕ_x. To do this, we require the correlation matrix of the curvature process χ_x, which we calculate below.

Let's first calculate the correlation function for the curvature OU process without coupling it to the $\dot{\phi}$ field, i.e., we ignore the first term in the action above. Let $\chi_0, \chi_1, \ldots, \chi_N$ be the curvature at lengths $0, dx, 2dx, \ldots, Ndx$, where $dx = L/N \ll \lambda$. The dynamics is defined by

$$\chi_{i+1} - \chi_i = -\chi_i dx/\lambda + \sqrt{\frac{2}{\lambda\xi^2}}dW(x), \tag{50}$$

10

where $dW(x) \sim \mathcal{N}(0, \sqrt{dx})$. Then, the joint distribution of this set including the prior is

$$P(\chi_0, \chi_1, \ldots, \chi_N) \sim e^{-\mathcal{E}} \sim e^{-\frac{\lambda\xi^2}{4}dx \sum_{i=0}^{N-1}\left(\frac{\chi_{i+1}-\chi_i}{dx}+\frac{\chi_i}{\lambda}\right)^2 - \frac{\xi^2\chi_0^2}{2}} \tag{51}$$

The cross term when the parenthesis is expanded is $\propto \int_0^L dx(2\dot{\chi}_x\chi_x) = \chi_N^2 - \chi_0^2$. The action is

$$\mathcal{E} = \frac{\lambda\xi^2}{4}dx \left(\sum_{i=0}^{N-1}\left(\frac{\chi_{i+1}-\chi_i}{dx}\right)^2 + \frac{\chi_i^2}{\lambda^2}\right) + \frac{\xi^2}{4}\left(\chi_N^2 - \chi_0^2\right) + \frac{\xi^2\chi_0^2}{2}, \tag{52}$$

$$= \frac{\lambda\xi^2}{4dx}\left\{\left(\sum_{i=0}^{N-1}(\chi_{i+1}-\chi_i)^2 + \chi_i^2\frac{dx^2}{\lambda^2}\right) + \left(\chi_N^2 + \chi_0^2\right)\frac{dx}{\lambda}\right\} \tag{53}$$

$$= \frac{\lambda\xi^2}{4dx}\left(\sum_{i,j=0}^{N}\chi_i G_{ij}^{-1}\chi_j\right), \tag{54}$$

where G is the correlation matrix and G^{-1} is a tridiagonal matrix,

$$G^{-1} = \begin{pmatrix} 1+dx/\lambda+dx^2/\lambda^2 & -1 & 0 & \ldots & 0 \\ -1 & 2+dx^2/\lambda^2 & -1 & \ldots & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & \ldots & 2+dx^2/\lambda^2 & -1 \\ 0 & 0 & \ldots & -1 & 1+dx/\lambda \end{pmatrix}. \tag{55}$$

The inverse of a tridiagonal matrix can be calculated using a pair of recurrence equations [1]:

$$G_{ij} = \begin{cases} R_{i-1}S_{j+1}/R_N, & \text{if } i \leq j \\ R_{j-1}S_{i+1}/R_N, & \text{if } i > j, \end{cases} \tag{56}$$

where R and S satisfy the recurrence relations and boundary conditions $(d_i = G_{ii}^{-1})$

$$R_i = d_i R_{i-1} - R_{i-2}, \quad i = 1, 2, \ldots, N, \quad R_{-1} = 1, R_0 = d_0, \tag{57}$$
$$S_i = d_i S_{i+1} - S_{i+2}, \quad i = N-1, N-2, \ldots, 0, \quad S_{N+1} = 1, S_N = d_N, \tag{58}$$
$$\tag{59}$$

Since $d_i = 2 + dx^2/\lambda^2$ for $i \neq 0, N$, we can immediately see that in the continuous limit away from the boundaries

$$\ddot{R} - R/\lambda^2 = 0, \tag{60}$$
$$\ddot{S} - S/\lambda^2 = 0, \tag{61}$$

The difference between R and S is in the boundary conditions. We have $R_{-1} = 1, S_{N+1} = 1$, $R_0 - R_{-1} = dx/\lambda + dx^2/\lambda^2$ and $S_{N+1} - S_N = -dx/\lambda$, which implies in the continuum

$$R(0) = 1, \quad \dot{R}(0) = \frac{1}{\lambda}, \quad S(L) = 1, \quad \dot{S}(L) = -\frac{1}{\lambda}. \tag{62}$$

11

Solving for R and S, we get

$$R(x) = e^{x/\lambda}, \quad S(x) = e^{(L-x)/\lambda}. \tag{63}$$

Finally, we should be careful about R_N which appears at the denominator of the correlation matrix (56). We have $R_N = (1+dx/\lambda)R_{N-1} - R_{N-2}$. $R(x)$ is not continuous at the rightmost point. Call $R_N = Z(L)$, $R_{N-1} = R(L - dx)$ and $R_{N-2} = R(L - 2dx)$, we have as $dx/\lambda \to 0$,

$$Z(L) = (1 + dx/\lambda) R(L - dx) - R(L - 2dx), \tag{64}$$

$$= dx \left(\frac{R(L)}{\lambda} + \dot{R}(L)|_- \right), \tag{65}$$

where $\dot{R}(L)|_-$ is the left derivative of R at L. We get $Z(L) = (2dx/\lambda)e^{L/\lambda}$

Plugging into (56), the correlation matrix in the continuous limit is

$$G(x, x') = \frac{\lambda}{2dx} e^{-|x-x'|/\lambda}, \tag{66}$$

as expected for the correlation function of an OU process (note that we multiply by $2dx/\lambda$ in $\mathcal{E}$).

Now, let's get back to the full problem. The full action is

$$\mathcal{E}(\chi) = \frac{1}{2\kappa} \int_0^L dx(\dot{\phi}_x - \chi_x)^2 + \frac{\lambda\xi^2}{4} \int_0^L dx \left(\dot{\chi}_x + \frac{\chi_x}{\lambda} \right)^2 + \frac{\xi^2\chi_0^2}{2} \tag{67}$$

$$= \frac{1}{2\kappa} \int_0^L dx(\dot{\phi}_x)^2 - \frac{1}{\kappa} \int_0^L dx\dot{\phi}_x\chi_x + \frac{\lambda\xi^2}{4} \int_0^L dx \left((\dot{\chi}_x)^2 + \Omega^2\chi_x^2 \right) + \frac{\xi^2}{4} \left(\chi_L^2 + \chi_0^2 \right), \tag{68}$$

where the boundary terms are the same as above and we have defined $\Omega^2 = \lambda^{-2} + V^{-2}$ and $V^2 = \kappa\xi^2\lambda/2$. We need the correlation matrix G of χ in this case.

However, notice now that even though the effective dynamics contain Ω^2 instead of λ^{-2}, the boundary conditions remain the same. That is, we should still get (62), but the solutions of (60) are now of the form $R(x) = Ae^{\Omega x} + Be^{-\Omega x}$ (and similarly for S). We instead get

$$R(x) = \cosh(\Omega x) + \frac{1}{\lambda\Omega} \sinh(\Omega x) \tag{69}$$

$$S(x) = \cosh(\Omega(L - x)) + \frac{1}{\lambda\Omega} \sinh(\Omega(L - x)), \tag{70}$$

i.e., $A = (1 + 1/\lambda\Omega)/2, B = (1 - 1/\lambda\Omega)/2$. Since the boundary conditions are the same as above, we have for the normalization

$$Z(L) = dx \left(\frac{R(L)}{\lambda} + \dot{R}(L)|_- \right) \tag{71}$$

$$= dx \left(A \left(\frac{1}{\lambda} + \Omega \right) e^{\Omega L} + B \left(\frac{1}{\lambda} - \Omega \right) e^{-\Omega L} \right) \tag{72}$$

$$= \frac{dx}{2\lambda^2\Omega} \left((1 + \Omega\lambda)^2 e^{\Omega L} - (1 - \Omega\lambda)^2 e^{-\Omega L} \right). \tag{73}$$

12

Once we multiply by the pre-factor, $\lambda/2dx$, the re-scaled correlation matrix, $G(x, x')$ is

$$G(x, x') = \frac{K(L)}{\xi^2} \begin{cases} R(x)S(x'), & \text{if } x \leq x' \\ R(x')S(x), & \text{if } x > x', \end{cases}, \tag{74}$$

$$\text{where} \quad K(L) = \frac{4\lambda\Omega}{(1 + \Omega\lambda)^2 \, e^{\Omega L} - (1 - \Omega\lambda)^2 \, e^{-\Omega L}} \tag{75}$$

Using the Gaussian integral formula in (67), we get for the effective action

$$\mathcal{E}_{\text{eff}} = \frac{1}{2\kappa} \int_0^L dx (\dot{\phi}_x)^2 - \frac{1}{2\kappa^2} \int_0^L \int_0^L dx dx' \dot{\phi}_x G(x, x') \dot{\phi}_{x'} \tag{76}$$

3.3.1 Fixed curvature limit

We check if the case derived for fixed curvature is recovered when $\lambda \to \infty$. Since $\Omega = \lambda^{-1}\sqrt{1 + 2\lambda/\kappa\xi^2}$, when λ is taken to be large, we get $(\lambda\Omega)^2 \approx 2\lambda/\kappa\xi^2$. Also, $\Omega L \ll 1$. From (75), in this limit,

$$K(L) \approx \frac{4\lambda\Omega}{4\lambda\Omega + 2(1 + (\lambda\Omega)^2)\Omega L} \tag{77}$$

$$= \frac{1}{1 + \frac{1+(\lambda\Omega)^2}{2\lambda}L} \approx \frac{1}{1 + L/\kappa\xi^2} \tag{78}$$

When $\Omega L \ll 1$, both $R(x)$ and $S(x)$ are 1. The correlation matrix is therefore

$$G(x, x') \approx \frac{1}{\xi^2 + L/\kappa} \tag{79}$$

Plugging in (76), we get

$$\mathcal{E}_{\text{eff}} \approx \frac{1}{2\kappa} \left(\int_0^L dx (\dot{\phi}_x)^2 - \frac{(\phi_L - \phi_0)^2}{\kappa\xi^2 + L} \right), \tag{80}$$

which is the one derived assuming fixed curvature.

3.3.2 Extremal equation

With the correlation matrix in hand, we proceed to compute the extremal equation in y_x satisfied by the effective action. From (76), the Euler-Lagrange equation is

$$\frac{d^2}{dx^2} \left(\ddot{y}_x - \frac{1}{\kappa} \int_0^L G(x, x') \ddot{y}_{x'} dx' \right) = 0, \tag{81}$$

which gives

$$\ddot{y}_x - \frac{1}{\kappa} \int_0^L G(x, x') \ddot{y}_{x'} dx' = (\Omega\lambda)^{-2} \left(c_0(L - x) + c_L x \right), \tag{82}$$

where the form of the r.h.s has been chosen for notational convenience which will be apparent later. c_0 and c_L are constants. The integral over G can be eliminated by differentiating w.r.t

13

x twice. Define $f(x) = \lambda R(x)$. Then the correlation matrix is $G(x,x') = K(L)f(x)f(L-x')/\lambda^2\xi^2$ for $x \le x'$ and $G(x,x') = K(L)f(L-x)f(x')/\lambda^2\xi^2$ for $x \ge x'$.

Differentiating once w.r.t x gives

$$y_x^{(3)} = \frac{K(L)}{\kappa\xi^2\lambda^2}\left(\int_x^L f'(x)f(L-x')\ddot{y}_{x'}dx' - \int_0^x f'(L-x)f(x')\ddot{y}_{x'}dx'\right) + (\Omega\lambda)^2(c_L - c_0) \quad (83)$$

Differentiating once more, we can use

$$f'(x)f(L-x) + f(x)f'(L-x) = \Omega^{-1}\left(\left(1+(\lambda\Omega)^2\right)\cosh(\Omega x) + 2\lambda\Omega\sinh(\Omega x)\right) \quad (84)$$
$$= 2\lambda/K(L) \quad (85)$$

and $f''(x) = \Omega^2 f(x)$ and $f''(L-x) = \Omega^2 f(L-x)$ to get

$$y_x^{(4)} = \frac{K(L)}{\kappa\xi^2\lambda^2}\left(-\frac{2\lambda}{K(L)}\right)\ddot{y}_x + \Omega^2\ddot{y}_x - \frac{c_0(L-x) + c_Lx}{\lambda^2}, \quad (86)$$

$$= -\frac{1}{V^2}\ddot{y}_x + \Omega^2\ddot{y}_x - \frac{c_0(L-x)+c_Lx}{\lambda^2}, \quad (87)$$

$$= \frac{1}{\lambda^2}\ddot{y}_x - \frac{c_0(L-x)+c_Lx}{\lambda^2}, \quad (88)$$

where we have used $\Omega^2 = \lambda^{-2} + V^{-2}$ and (82) to get rid of the integral over G.

3.3.3 The joint distribution of end-point angles

The general solution of the extremal equation (88) can be written as

$$y_x = a_0(L-x) + a_Lx + c_0\frac{(L-x)^3}{6} + c_L\frac{x^3}{6} + \lambda^2 d_0 e^{-(L-x)/\lambda} + \lambda^2 d_L e^{-x/\lambda} \quad (89)$$

We'd like to get the coefficients in terms of ϕ_0 and ϕ_L given $y_0, y_L = 0$. The coefficients d_0 and d_L are not independent of c_0 and c_L and are to be obtained by plugging in (89) into (82). For convenience, we work in terms of $\psi(x) \equiv \ddot{y}_x$. We have

$$\psi(x) \equiv \ddot{y}_x = c_0(L-x) + c_Lx + d_0 e^{-(L-x)/\lambda} + d_L e^{-x/\lambda}. \quad (90)$$

We can use integration by parts to compute the integral in (82). We have

$$\frac{1}{\kappa}\int_0^L G(x,x')\psi(x')dx' = \frac{K(L)}{\kappa\lambda^2\xi^2}\left(f(L-x)\int_0^x f(x')\psi(x')dx' + f(x)\int_x^L f(L-x')\psi(x')dx'\right). \quad (91)$$

Write $\psi(x) = h(x) + g(x)$, where $h(x) = c_0(L-x) + c_Lx$ and $g(x)$ has the exponential terms. Then $h''(x) = 0$ and $g''(x) = g(x)/\lambda^2$. Let $I(g) = \int_0^x f(x')g(x')dx'$. We have

$$I(g) = \Omega^{-2}\int_0^x f''(x')g(x')dx' = \Omega^{-2}\left(f'(x)g(x) - f'(0)g(0) - f(x)g'(x) + f(0)g'(0)\right) \quad (92)$$

$$+ (\lambda\Omega)^{-2}I(g). \quad (93)$$

14

Sending the $I(g)$ to the left side and using $1 - (\lambda\Omega)^{-2} = (V\Omega)^{-2}$, we get

$$I(g) = V^2 \left(f'(x)g(x) - f'(0)g(0) - f(x)g'(x) + f(0)g'(0) \right) \tag{94}$$

Similarly, define $I'(g) = \int_x^L f(L - x')g(x')dx'$.

$$I'(g) = \Omega^{-2} \int_x^L f''(L - x')g(x')dx' = \Omega^{-2} \left(f'(0)g(L) + f'(L - x)g(x) - f(L - x)g'(x) - f(0)g'(L) \right) \tag{95}$$

$$+ (\lambda\Omega)^{-2} I'(g), \tag{96}$$

which gives

$$I'(g) = V^2 \left(f'(0)g(L) + f'(L - x)g(x) - f(L - x)g'(x) - f(0)g'(L) \right) \tag{97}$$

Eventually, after some simplifications we get

$$\frac{K(L)}{\kappa\lambda^2\xi^2} \left(f(L - x)I(g) + f(x)I'(g) \right) = \tag{98}$$

$$g(x) - \frac{K(L)}{2\lambda} \left(f(x) \left(f'(0)g(L) + f(0)g'(L) \right) + f(L - x) \left(f'(0)g(0) - f(0)g'(0) \right) \right). \tag{99}$$

When we compute the same integrals for $h(x)$ instead of $g(x)$, since $h''(x) = 0$, when we integrate by parts we don't get the $I(g)$ as we did on the r.h.s of (93). We can replace V^2 by Ω^{-2} and repeat the same calculation for $I(h) \equiv \int_0^x f(x')h(x')dx'$ and $I'(h)$ to get

$$\frac{K(L)}{\kappa\lambda^2\xi^2} \left(f(L - x)I(h) + f(x)I'(h) \right) = \tag{100}$$

$$(\Omega V)^{-2} \left(h(x) - \frac{K(L)}{2\lambda} \left(f(x) \left(f'(0)h(L) + f(0)h'(L) \right) + f(L - x) \left(f'(0)h(0) - f(0)h'(0) \right) \right) \right). \tag{101}$$

We subtract (99) and (101) from $\psi(x)$ to evaluate the l.h.s of (82). The $g(x)$ cancels out, whereas $h(x)(1 - (\Omega V)^{-2}) = h(x)(\Omega\lambda)^{-2}$ cancels out with the r.h.s of (82). The leftover terms read

$$f(x) \left(f'(0)g(L) + f(0)g'(L) \right) + f(L - x) \left(f'(0)g(0) - f(0)g'(0) \right) + \tag{102}$$

$$(\Omega V)^{-2} \left(f(x) \left(f'(0)h(L) + f(0)h'(L) \right) + f(L - x) \left(f'(0)h(0) - f(0)h'(0) \right) \right) = 0 \tag{103}$$

Define $m(x) = g(x) + (\Omega V)^{-2}h(x)$. Using $f(0) = \lambda$ and $f'(0) = 1$, we have

$$f(x) \left(m(L) + \lambda m'(L) \right) + f(L - x) \left(m(0) - \lambda m'(0) \right) = 0. \tag{104}$$

Since $f(x)$ and $f(L - x)$ are independent, we have $m(L) + \lambda m'(L) = 0$ and $m(0) - \lambda m'(0) = 0$, which we re-write as $m(L) + m(0) = \lambda(m'(0) - m'(L))$ and $m(0) - m(L) = \lambda(m'(0) + m'(L))$. From the definition of $g(x)$ and $h(x)$, we have

$$m(0) = (\Omega V)^{-2}c_0 L + d_0 e^{-L/\lambda} + d_L, m(L) = (\Omega V)^{-2}c_L L + d_0 + d_L e^{-L/\lambda}, \tag{105}$$

$$m'(0) = (\Omega V)^{-2}(c_L - c_0) + \lambda^{-1} \left(d_0 e^{-L/\lambda} - d_L \right), m'(L) = (\Omega V)^{-2}(c_L - c_0) + \lambda^{-1} \left(d_0 - d_L e^{-L/\lambda} \right). \tag{106}$$

15

From the above two conditions $m(L) + \lambda m'(L) = 0$ and $m(0) - \lambda m'(0) = 0$, it is easy to show that

$$-\frac{L}{2\Omega^2 V^2}(c_0 + c_L) = d_0 + d_L, \tag{107}$$

$$\frac{L + 2\lambda}{2\Omega^2 V^2}(c_0 - c_L) = d_0 - d_L. \tag{108}$$

We want to write the effective action as a quadratic form in ϕ_0 and ϕ_L. Notice that the effective action can be written as

$$\mathcal{E}_{\mathrm{eff}} = \frac{1}{2\kappa}\left(\int_0^L dx\, \ddot{y}_x\left(\ddot{y}_x - \frac{1}{\kappa}\int_0^L G(x,x')\ddot{y}_{x'}dx'\right)\right). \tag{109}$$

When we integrate by parts twice, two of the terms are zero, one because $y_0, y_L = 0$ and the other from the extremal equation (82). The leftover term is a boundary term

$$\mathcal{E}_{\mathrm{eff}} = \frac{1}{2\kappa}\left(y_L'\left(\ddot{y}_L - \frac{1}{\kappa}\int_0^L G(L,x')\ddot{y}_{x'}dx'\right) - y_0'\left(\ddot{y}_0 - \frac{1}{\kappa}\int_0^L G(0,x')\ddot{y}_{x'}dx'\right)\right) \tag{110}$$

$$= \frac{L}{2\kappa\Omega^2\lambda^2}\left(\phi_L c_L - \phi_0 c_0\right), \tag{111}$$

where we have used (82). It remains to express c_0, c_L in terms of ϕ_0, ϕ_L. The inverse covariance matrix should have equal diagonal elements and equal off-diagonal elements, and so we should have the form $b_1(c_L + c_0) = \phi_L - \phi_0$ and $b_2(c_L - c_0) = \phi_L + \phi_0$ for some b_1, b_2 which are to be found. From (89), we have

$$0 = a_0 L + c_0 L^3/6 + \lambda^2 d_0 e^{-L/\lambda} + \lambda^2 d_L, \tag{112}$$

$$0 = a_L L + c_L L^3/6 + \lambda^2 d_0 + \lambda^2 d_L e^{-L/\lambda}, \tag{113}$$

$$\phi_0 L = a_L L - a_0 L - c_0 L^3/2 + \lambda L d_0 e^{-L/\lambda} - \lambda L d_L, \tag{114}$$

$$\phi_L L = a_L L - a_0 L + c_L L^3/2 + \lambda L d_0 - \lambda d_L L e^{-L/\lambda}. \tag{115}$$

After some straightforward manipulations, we get

$$b_1 = \frac{L^2}{2} - \frac{\lambda L}{2\Omega^2 V^2}\left(1 - e^{-L/\lambda}\right) \tag{116}$$

$$b_2 = \frac{L^2}{6} - \frac{\lambda(L + 2\lambda)}{2\Omega^2 V^2 L}\left(L\left(1 + e^{-L/\lambda}\right) - 2\lambda\left(1 - e^{-L/\lambda}\right)\right). \tag{117}$$

It can be checked that

$$\rho(L) = \frac{b_2 - b_1}{b_2 + b_1}, \quad \sigma^2(L) = \frac{\kappa\Omega^2\lambda^2}{L}\frac{b_1 + b_2}{2} \tag{118}$$

To recover the WLC and fixed curvature limits, we can expand b_1 and b_2 for small $L/\lambda \ll 1$ to $O(L^3)$. We get

$$b_1 \approx \frac{L^2}{2\Omega^2\lambda^2} + \frac{L^3}{4\lambda\Omega^2 V^2} \tag{119}$$

$$b_2 \approx \frac{L^2}{6\Omega^2\lambda^2} \tag{120}$$

Plugging in (118), we recover $\sigma^2(L) \approx \kappa L/3 + L^2/4\xi^2$.

16

3.3.4 Numerical validation

Before comparing to numerics, we would like to examine the limits of validity of our original small-angle approximation, i.e., the conditions for $\langle \phi_L^2 \rangle \ll 1$ to hold. For simplicity, we consider the case when κ is small w.r.t the other length scales in the problem. Then, we can show that

$$\langle \phi_L^2 \rangle \approx \frac{\lambda}{2L\xi^2} \left(\frac{4L^2}{3} - L\lambda \left(1 - e^{-L/\lambda} \right) - \frac{\lambda(L+2\lambda)}{L} \left(L - 2\lambda + (L+2\lambda) e^{-L/\lambda} \right) \right) \quad (121)$$

We assume $\xi \gg \lambda$ since otherwise the trail forms circles. When $L \gg \lambda$, we can show diffusive scaling (also apparent in Figure M3c), $\langle \phi_L^2 \rangle \approx 2\lambda L/3\xi^2$. This expression is valid when $1 \ll L/\lambda \ll 3\xi^2/2\lambda^2$. When $L \ll \lambda$, (121) reduces to $\langle \phi_L^2 \rangle \approx L^2/4\xi^2$ derived previously, which is always much less than one in that limit if $\xi \gg \lambda$. Thus, the small-angle approximation is valid for $L/\lambda \ll 3\xi^2/2\lambda^2$.

For numerical validation, we integrate over s, $dx/ds = \cos\theta$, $dy/ds = \sin\theta$, $d\theta/ds = \chi$ and $d\chi/ds = -\chi/\lambda + \sqrt{2/\lambda\xi^2}\eta$, where η is white noise. The initial conditions are $x(0) = y_0 = \theta(0) = 0$ and $\chi(0) \sim \mathcal{N}(0, \xi^{-2})$. Rotating the θ axis so that initial and final y are zero, we have $\phi_0 = -y_L/L$ and $\phi_L = \theta(L) - y_L/L$. We plot σ vs L, where $\sigma^2 = (\langle \phi_0^2 \rangle + \langle \phi_L^2 \rangle)/2$. Results are shown in Figure M3e.

3.3.5 Forward propagator

We exploit our previous calculation of the extremal path to compute the forward propagator, i.e., $P(\phi_L, y_L | \phi_0 = 0, y_0 = 0)$ (if $\phi_0 \neq 0$, we can simply re-orient axes). In this case, the extremal path remains the same, but when we integrate the effective action (109) by parts we get a different expression:

$$\mathcal{E}_{\text{eff}} = \frac{1}{2\kappa} \left(y_L' \left(\ddot{y}_L - \frac{1}{\kappa} \int_0^L G(L, x') \ddot{y}_{x'} dx' \right) - y_L \frac{d}{dx} \left(\ddot{y}_x - \frac{1}{\kappa} \int_0^L G(x, x') \ddot{y}_{x'} dx' \right) \Big|_L \right) \quad (122)$$

$$= \frac{1}{2\kappa\Omega^2\lambda^2} \left(\phi_L L c_L - y_L(c_L - c_0) \right). \quad (123)$$

We need to express c_L and c_0 in terms of ϕ_L and y_L. We have from (89)

$$0 = a_0 L + c_0 L^3/6 + \lambda^2 d_0 e^{-L/\lambda} + \lambda^2 d_L, \quad (124)$$

$$y_L = a_L L + c_L L^3/6 + \lambda^2 d_0 + \lambda^2 d_L e^{-L/\lambda}, \quad (125)$$

$$0 = a_L L - a_0 L - c_0 L^3/2 + \lambda L d_0 e^{-L/\lambda} - \lambda L d_L, \quad (126)$$

$$\phi_L L = a_L L - a_0 L + c_L L^3/2 + \lambda L d_0 - \lambda d_L L e^{-L/\lambda}. \quad (127)$$

The relationship between d_0, d_L and c_0, c_L in (107) remains the same as before. From these equations, it is lengthy but straightforward to show that

$$\phi_L = (c_L + c_0) b_1, \quad (128)$$

$$y_L = \frac{(c_0 + c_L) b_1 L}{2} + \frac{(c_0 - c_L) b_2 L}{2}. \quad (129)$$

17

Solving for c_L and c_0 and plugging in to (123), we eventually get

$$\langle \phi_L^2 \rangle_{\text{fwd}} = 2\kappa(\Omega\lambda)^2 b_1/L, \tag{130}$$

$$\langle y_L^2 \rangle_{\text{fwd}} = \kappa(\Omega\lambda)^2 L \frac{b_1 + b_2}{2} = L^2 \langle \phi_L^2 \rangle,$$

$$\langle \phi_L y_L \rangle_{\text{fwd}} = \kappa(\Omega\lambda)^2 b_1$$

These expressions are numerically validated as shown in Figure M3f. Note also that the widening of the azimuthal position y_L/L is given by $\sigma_{\text{fwd}}^2(L) \equiv \langle y_L^2 \rangle_{\text{fwd}}/L^2 = \langle \phi_L^2 \rangle$, which is the variance of end-point angles for the interpolation model.

4 The non-detection probability during surge and cast

In this section, we analyze a sector search strategy where the agent casts rapidly and moves radially with constant speed. In contrast to the previous section, the full dynamics of trails is considered. The key quantity here is the non-detection probability of not detecting the trail at radial distance L after loss of contact, $\Gamma_{\mathcal{C}}(L)$, using which we can calculate mean times to detection and the probability of missing the trail. We compute the non-detection probability conditioned on initial path heading, $\Gamma_{\mathcal{C}}(L|\phi)$. We have

$$\Gamma_{\mathcal{C}}(L) = \int_{-\infty}^{\infty} \frac{d\phi}{\sqrt{2\pi}} e^{-\frac{\phi^2}{2\sigma^2}} \Gamma_{\mathcal{C}}(L|\phi), \tag{131}$$

where σ is the uncertainty. We assume constant longitudinal speed v, transverse speed u, sampling rate ω and kernel size a. We also assume there is a duration, t_r, after the most recent detection where the agent does not detect which represents the minimum delay after a sample (for e.g., the exhalation phase of a sniff). We expect $t_r \approx \omega^{-1}$. We set $t_r = 0$ for the results presented in the main text and Methods. Throughout, we use the GWLC ensemble of trails (Section 3.3).

4.1 Non-detection probabilities in the cast phase

The probability that a particular trail path, $\{y(x)\}$, is not detected is the probability that the agent does not sample during the period when it is within detection range of the trail. The non-detection probability is thus $e^{-\omega T(\{y(x)\})}$ where $T(\{y(x)\})$ is the total time of overlap between the agent's detection range and the trail's path. We assume the tangential speed $u/a\omega \gg 1$ and the detection kernel has full-width a. Let the turning points be at the azimuthal envelope $\sigma x \Theta_{\text{env}}(x)$, where x is the distance along the most likely heading. During casting, the agent repeatedly traverses from $-\sigma x \Theta_{\text{env}}(x)$ to $\sigma x \Theta_{\text{env}}(x)$ (and vice versa) in time $2\sigma x \Theta_{\text{env}}(x)/u$. It is easy to show that if $|y(x)| < a/2 + \sigma x \Theta_{\text{env}}(x)$, then the time spent within detection range of the trail during a single traversal is given by

$$T_{\text{traversal}} = u^{-1} \left(\min(2\sigma x \Theta_{\text{env}}(x), y(x) + a/2 + \sigma x \Theta_{\text{env}}(x)) - \max(0, y(x) - a/2 + \sigma x \Theta_{\text{env}}(x)) \right). \tag{132}$$

In the time it takes to travel dx in the longitudinal direction, the total number of traversals is $\frac{u}{2\sigma x \Theta_{\text{env}}(x)} \frac{dx}{v}$. Therefore, the total time spent within detection range while the agent travels

18

dx is

$$\gamma(x, y(x))dx/v \equiv \min(2\sigma x\Theta_{\text{env}}(x), y(x) + a/2 + \sigma x\Theta_{\text{env}}(x)) - \max(0, y(x) - a/2 + \sigma x\Theta_{\text{env}}(x)))$$
$$\times \frac{dx}{2\sigma vx\Theta_{\text{env}}(x)} \tag{133}$$

In this $u/a\omega \gg 1$ limit, the detection probability is thus independent of u. To compute $\Gamma_{\mathcal{C}}(L|\phi)$, we need to integrate over all trail paths with $y(0) = 0, y'(0) = \phi$ up to length L. That is

$$\Gamma_{\mathcal{C}}(L|\phi) = \int \mathcal{D}y(x)e^{-\mathcal{E}(\{y(x)\})_{y(0)=0,y'(0)=\phi} - \frac{\omega}{v}\int_{vt_r}^{L} dx\gamma(x,y(x))\mathbb{1}(|y(x)|<a/2+\sigma x\Theta_{\text{env}}(x))}, \tag{134}$$

where $\mathcal{E}$ is the action defined by the specific model of the trail and the integral is over trails with $y(0) = 0, y'(0) = \phi$. It is useful to rotate axes such that the paths $y(x)$ have initial heading 0, i.e., equivalently,

$$\Gamma_{\mathcal{C}}(L|\phi) = \int \mathcal{D}y(x)e^{-\mathcal{E}(\{y(x)\})_{y(0)=0,y'(0)=0} - \frac{\omega}{v}\int_{vt_r}^{L} dx\gamma(x,y(x)+\phi x)\mathbb{1}(|y(x)+\phi x|<a/2+\sigma x\Theta_{\text{env}}(x))}, \tag{135}$$

We now exploit the spline formulation from the previous sections. The simplification is that the extremal paths of the action have relatively simple forms like the one in eq. (89). Since the extremal paths are constrained only by the initial and final positions and headings, they have four parameters, two of which are fixed by $y(0) = 0, y'(0) = 0$. The action for the extremal case thus has two free parameters, viz., the final transverse position, y_L, and final heading, ϕ_L. From our previous results, the action is a quadratic form in y_L and ϕ_L, which have mean 0 and covariance $\langle\phi_L^2\rangle_{\text{fwd}}, \langle y_L^2\rangle_{\text{fwd}}, \langle\phi_L y_L\rangle_{\text{fwd}}$ in (130). We will assume the diffusivity κ is negligible in the analysis hereafter. The results can be easily generalized to diffusive trails. Note that $\xi \gg \lambda$, otherwise the trail forms circles. The path $y(x)$ has the form (89), where the coefficients $a_0, a_L, c_0, c_L, d_0, d_L$ can be written in terms of y_L, ϕ_L using the relations derived previously. The expectation over paths in (135) is replaced with an expectation over ϕ_L, y_L:

$$\Gamma_{\mathcal{C}}(L|\phi) = \langle e^{-\frac{\omega}{v}\int_{vt_r}^{L} dx\gamma(x,y(x)+\phi x)\mathbb{1}(|y(x)+\phi x|<a/2+\sigma x\Theta_{\text{env}}(x))}\rangle_{y_L,\phi_L}, \tag{136}$$

which is efficiently computed numerically by sampling (sampled ϕ_L, y_L can be re-used for all ϕ).

4.2 Geometric interpretation

When v is small, the first order contributions to the non-detection probability come from the path most likely to survive the casting search. That is, to first order we can approximate the integral (135) as

$$-\ln\Gamma_{\mathcal{C}}(L|\phi) \simeq \min_{y_L,\phi_L}\left\{\mathcal{E}(y_L,\phi_L) + \frac{\omega}{v}\int_{vt_r}^{L} dx\gamma(x,y(x)+\phi x)\mathbb{1}(|y(x)+\phi x| < a/2+\sigma x\Theta_{\text{env}}(x))\right\}, \tag{137}$$

where $\mathcal{E}(y_L,\phi_L)$ is the action. Note that the minimum is rather shallow, so the first order approximation is not a very good one (since the trails are flexible at the length scales of interest), but leads to some geometric intuition.

In Figure M4d we show examples of the path that minimizes the quantity in the parenthesis above. Intuitively, when ϕ is small, the minimizer path is 'immersed' in the rotated casting profile and should expend significant bending cost in order to escape the casting region. The minimizer is simply a straight line. In this case, $\mathcal{E}(y_L, \phi_L) = 0$ and all the ϕ within this range have approximately equal probability of being absorbed. As ϕ increases, the casting profile gets lower and there is a transition point, ϕ^* (which depends on L), where the minimizer path can escape the casting region without incurring too much bending cost. This is shown in the middle panel of Figure M4d. For $\phi > \phi^*$, the non-detection probability rapidly increases. At large angle, the casting profile no longer has a significant effect on the non-detection probability and the minimizer path begins to straighten out again.

The posterior distribution of the initial heading $P(\phi, L)$ is $P(\phi, L) \propto \Gamma_{\mathcal{C}}(L|\phi)e^{-\phi^2/2\sigma^2}$, which can be numerically computed from (136). The prior is unimodal, which transitions to a trimodal shape at intermediate values of L. At large L, the posterior becomes bimodal. The locations of the modes at large L are computed below for small v and conical envelopes.

4.3 Conical casts

We analyze the case when $\Theta_{\mathrm{env}}(x) = \Theta_0$ is a constant. In this case, $\Gamma_{\mathcal{C}}(L)$ can be approximated and the optimal speed derived. The contribution to $\Gamma_{\mathcal{C}}(L)$ from trails outside the cone, $|\phi| > \sigma\Theta_0$ is calculated further below. For trails directed inside the cone, their non-detection probability decreases independent of their initial heading until they escape the cone. The point at which they escape the cone is easily calculated. We give a heuristic argument (partly reproduced in the Methods) which captures the basic intuition and leads to a good approximation for $\Gamma_{\mathcal{C}}(L)$.

The idea is that to first order, for $L/\lambda \ll 1$, the trail paths have a single curvature scale $\chi \sim \xi^{-1}$. If $|\phi| < \sigma\Theta_0$, the trails need to curve $\sigma\Theta_0 - \phi$ radians to leave the cone, which is at $L^* \sim 2\xi(\sigma\Theta_0 - |\phi|)$. The non-detection probability can be calculated from (136), but γ has a complicated dependence on x and $y(x)$. To simplify, we assume for trails inside the cone $\gamma(x, y(x)) \approx \gamma(x, 0)$, which corresponds to the trail directed along the cone's midline. Then, $\gamma(x, y(x)) = 1$ if $a/2 > \sigma\Theta_{\mathrm{env}}(x)$ and $\gamma(x, y(x)) = a/2\sigma\Theta_{\mathrm{env}}(x)$ otherwise. Define $\beta \equiv a\omega/2\sigma v\Theta_0$. From (136), we get for $|\phi| < \sigma\Theta_0$

$$\Gamma_{\mathcal{C}}(L|\phi) \sim e^{\omega t_r - \beta}\left(\frac{a/2\sigma\Theta_0}{L}\right)^{\beta} \quad \text{if } a/2\sigma\Theta_0 > vt_r, \tag{138}$$

$$\sim \left(\frac{vt_r}{L}\right)^{\beta} \quad \text{if } a/2\sigma\Theta_0 < vt_r,$$

until it flattens out at some L_{in} calculated below. For small L and $\Theta_0 \gg 1$, $\Gamma_{\mathcal{C}}(L)$ is determined by the trails directed within the cone and thus $\Gamma_{\mathcal{C}}(L) \approx \Gamma_{\mathcal{C}}(L|0)$ also decays as (138) until it flattens out. There are two possibilities for $\Gamma_{\mathcal{C}}(L)$ to flatten out, whichever occurs earlier: 1) a significant fraction of trails leave the cone, or 2) $\Gamma_{\mathcal{C}}(L)$ decays until it matches the probability of missing the trails that began outside the cone.

We estimate the length scale at which a significant fraction of trails leave as $L_{\mathrm{in}} \sim 2\xi\sigma\Theta_0$, when even the trails initially directed towards the center of the cone escape. This underestimates L_{in} since the curvature of trails with large initial curvature (the ones that escape) may

20

revert back to the mean. At $L = L_{\text{in}}$, $\Gamma_{\mathcal{C}}(L)$ flattens out at

$$\epsilon_{\text{in}} \sim \Gamma_{\mathcal{C}}(L_{\text{in}}) \approx \Gamma_{\mathcal{C}}(L_{\text{in}}|0), \quad \text{where} \tag{139}$$

$$L_{\text{in}} \sim 2\xi\sigma\Theta_0 \tag{140}$$

This fixed curvature estimate is valid when $2\xi\sigma\Theta_0 \ll \lambda$. When $L \gg \lambda$, the trails' azimuth diffuses as $\Delta\phi^2 \sim \lambda L/\xi^2$, and thus it takes $L_{\text{in}} \sim \xi^2\sigma^2\Theta_0^2/\lambda$ for most trails to diffuse $\sigma\Theta_0$ and out of the cone.

We now calculate the probability of missing the trails that began outside the cone, ϵ_{out}, i.e., the trails with $|\phi| > \sigma\Theta_0$. These trails leave the cone early before the casting phase $2\sigma\Theta_0 L \gg a$ begins. Since L is small, we can assume these are stiff, straight paths with angle ϕ. The distance at which they escape is $L^* \sim a/2(|\phi| - \sigma\Theta_0)$. Define $\beta' \equiv \sigma v/a\omega$. The conditional non-detection probability is

$$\Gamma_{\mathcal{C}}(L|\phi) \simeq e^{-\omega \int_{vt_r}^{L^*} dx/v} = e^{\omega t_r - \frac{1}{2\beta'(|\phi|/\sigma - \Theta_0)}}. \tag{141}$$

Then,

$$\epsilon_{\text{out}} \sim \frac{2e^{\omega t_r}}{\sqrt{2\pi\sigma^2}} \int_{\sigma\Theta_0}^{\infty} e^{-\frac{1}{2\beta'(\phi/\sigma - \Theta_0)} - \phi^2/2\sigma^2} d\phi. \tag{142}$$

Transforming variables $\phi \to \phi/\sigma - \Theta_0$ and expanding the prior term, we get

$$\epsilon_{\text{out}} \sim \frac{2e^{\omega t_r - \Theta_0^2/2}}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{1}{2\beta'\phi} - \phi^2/2 - \phi\Theta_0} d\phi. \tag{143}$$

We calculate the above integral in the small β' limit. This is because, as shown below, when $\beta' \ll 1$, the detection rate of trails towards the center of the cone scales as β'^{-1}, whereas for trails that begin outside the cone the detection rate scales as $\beta'^{-2/3}$ or $\beta'^{-1/2}$. Thus the small β' limit is the relevant one for ϵ_{out} i.e., when $\epsilon_{\text{out}} > \epsilon_{\text{in}}$.

When $\beta' \ll 1$, (143) can be approximated using Laplace's method, but the maximum is hard to calculate analytically since its a cubic root. It is easy to show that depending on whether $2\beta'\Theta_0^3 \ll 1$ or $2\beta'\Theta_0^3 \gg 1$, either $\phi^2/2$ or $\phi\Theta_0$ respectively dominate at the maximum. For the two cases, we substitute $\phi = \beta'^{-1/3}\phi'$ or $\phi = \beta'^{-1/2}\phi'$ respectively to transform the moving maxima and use Laplace's method to get up to constant prefactors

$$\epsilon_{\text{out}} \sim e^{\omega t_r - \frac{\Theta_0^2}{2} - \frac{3}{2}(2\beta')^{-2/3}}, \quad \text{if } 2\beta'\Theta_0^3 \ll 1, \tag{144}$$

$$\epsilon_{\text{out}} \sim (\beta'\Theta_0^3)^{-1/4} e^{\omega t_r - \frac{\Theta_0^2}{2} - \frac{(2\Theta_0\beta')^{1/2}}{\beta'}}, \quad \text{if } 2\beta'\Theta_0^3 \gg 1. \tag{145}$$

When $\epsilon_{\text{out}} > \epsilon_{\text{in}}$, we can define a length scale, L_{out} when $\Gamma_{\mathcal{C}}(L)$ flattens out, which satisfies

$$\epsilon_{\text{out}} \sim \Gamma_{\mathcal{C}}(L_{\text{out}}). \tag{146}$$

We can also estimate the initial heading of the trails that are most likely missed. When $\epsilon_{\text{out}} < \epsilon_{\text{in}}$, this is the trail directed towards the center of the cone, $\phi \approx 0$, since the trails missed are the ones inside the cone and $\phi = 0$ has the highest prior probability. When

$\epsilon_{\text{out}} > \epsilon_{\text{in}}$, this is located at the extrema in the integrals above, i.e., $|\phi|/\sigma = \Theta_0 + (2\beta')^{-1/3}$ and $|\phi|/\sigma = \Theta_0 + (2\beta'\Theta_0)^{-1/2}$ for $2\beta'\Theta_0^3 \ll 1$ and $2\beta'\Theta_0^3 \gg 1$ respectively.

In summary, we estimate if $a/2\sigma\Theta_0 > vt_r$,

$$\Gamma_{\mathcal{C}}(L) \simeq \begin{cases} 1, & \text{for } L \leq vt_r \\ e^{\omega t_r - \omega L/v}, & \text{for } vt_r \leq L \leq a/2\sigma\Theta_0, \\ e^{\omega t_r - \beta}\left(\frac{a/2\sigma\Theta_0}{L}\right)^\beta, & \text{for } a/2\sigma\Theta_0 \leq L \lesssim \min(L_{\text{in}}, L_{\text{out}}) \\ \max(\epsilon_{\text{in}}, \epsilon_{\text{out}}), & \text{for } L \gtrsim \min(L_{\text{in}}, L_{\text{out}}) \end{cases}, \tag{147}$$

If $a/2\sigma\Theta_0 < vt_r$, we have

$$\Gamma_{\mathcal{C}}(L) \simeq \begin{cases} 1, & \text{for } L \leq vt_r \\ \left(\frac{vt_r}{L}\right)^\beta, & \text{for } vt_r \leq L \lesssim \min(L_{\text{in}}, L_{\text{out}}) \\ \max(\epsilon_{\text{in}}, \epsilon_{\text{out}}), & \text{for } L \gtrsim \min(L_{\text{in}}, L_{\text{out}}) \end{cases}, \tag{148}$$

In Figure M5a we compare the prediction for $\Gamma_{\mathcal{C}}(L)$ to the numerically obtained result using (136). Figure M5b shows the comparison between numerics and the prediction for $\Gamma_{\mathcal{C}}(\infty)$ for a range of values of β and Θ_0.

References

[1] Mallik RK. The inverse of a tridiagonal matrix. Linear Algebra and its Applications. 2001;325(1-3):109–139.