Supplemental material: "Optimal switching strategies for navigation in stochastic settings"

xx
yy

DERIVATION OF THE OPTIMAL CONTROL EQUATION

The dynamics of the angle $\theta(t)$ in a small time interval dt can be written as

$$\theta(t+dt) = \begin{cases} \theta(t) + \eta(t)dt & \text{if } s(\theta,t) = 0\,, \\ 0 & \text{if } s(\theta,t) = 1\,, \end{cases} \tag{1}$$

where $s(\theta,t)$ is a binary variable, describing a reorientation policy. We recall that $\eta(t)$ is Gaussian white noise with zero mean and correlator $\langle \eta(t)\eta(t')\rangle = 2D\delta(t-t')$.

Let us clarify that we have written the dynamics in Eq. (1) in discrete time to provide a more intuitive explanation. However, the stochastic process $\theta(t)$ remains well-defined in the continuous-time limit as $dt \to 0$, which is the regime considered throughout the paper for deriving explicit results. For any sufficiently regular policy (as is the case for all policies analyzed in this work), the trajectory $\theta(t)$ consists of segments of free diffusion, which converge to a well-defined continuous limit described by Brownian motion, punctuated by resetting events to the origin. These resetting events occur whenever the process first enters a region where $s(\theta,t)=1$. It is important to emphasize that the discrete time increment dt does not affect our theoretical results, as they are derived in the $dt \to 0$ limit. Nevertheless, all numerical simulations are conducted with finite but small dt, and they show excellent agreement with the theoretical predictions.

Our goal is to identify the optimal policy $s^*(\theta,t)$ that minimizes the cost function. We define the cost to-go function

$$\mathcal{F}_{\theta,t}[s] = \left\langle -\int_t^{t_f} v_0 \cos(\theta(\tau))dt + x_c[N(t_f) - N(t)] \right\rangle_\theta\,, \tag{2}$$

which describes the average cost incurred between time t and time t_f, starting at θ at time t, and following a given policy $s(\theta,t)$. The symbol $\langle\cdot\rangle_\theta$ indicates the average with respect to all stochastic trajectories starting at θ at time t. We recall that v_0 is the speed of the agent and that $N(t)$ indicates the number of reorientations up to time t. The cost function $\mathcal{F}_{\theta,t}[s]$ is a functional, depending on the specific form of the function $s = s(\tilde{\theta},\tilde{t})$. The cost function defined in the main text corresponds to $\mathcal{F} = \mathcal{F}_{0,0}[s]$. The optimal cost-to-go $J(\theta,t)$ can be defined as

$$J(\theta,t) = \min_s \mathcal{F}_{\theta,t}[s]\,, \tag{3}$$

where the minimization is performed over all piecewise-continuous binary functions $s(\tilde{\theta},\tilde{t}) \in \{0,1\}$, for $\tilde{\theta} \in [-\pi/2,\pi/2]$ and $\tilde{t} \in [t,t_f]$.

To perform the minimization, we adopt a dynamic programming approach [1]. We consider the evolution of the system in a small time interval dt, as given in Eq. (1). Then, Eq. (2) can be rewritten by splitting the cost incurred in the short time interval $[t,t+dt]$ and in the remaining time interval $[t+dt,t_f]$ as

$$\mathcal{F}_{\theta,t}[s] = -v_0\cos(\theta)dt + \left(x_c + \mathcal{F}_{0,t+dt}[s]\right)s(\theta,t) + \left\langle \mathcal{F}_{\theta+\eta(t)dt,t+dt}[s]\right\rangle (1 - s(\theta,t))\,. \tag{4}$$

The first term on the right-hand side comes from the integral term in Eq. (2). In addition, if the system is reset, corresponding to $s(\theta,t)=1$, a unit cost x_c is paid and the remaining cost up to time t_f is $\mathcal{F}_{0,t+dt}$, since $\theta=0$ after a reorientation. On the other hand, if $s(\theta,t)=0$, the system evolves freely and the new position is $\theta+\eta(t)dt$, leading to the last term in Eq. (4). Note that $\langle\cdot\rangle$ now indicates averaging over the one-time noise $\eta(t)$, which is a Gaussian random variable with zero mean $\langle\eta(t)\rangle = 0$ and variance $\langle \eta(t)^2\rangle = 2D/dt$.

We perform the minimization in Eq. (3) in two steps: first we minimize over the binary function $s(\tilde{\theta},\tilde{t})$ for $\tilde{t} \in [t+dt,t_f]$ and then over the binary variable $s(\theta,t)$, where θ and t are fixed. We obtain

$$J(\theta,t) = \min_{s(\theta,t)} \left[-v_0\cos(\theta)dt + \left(x_c + J(0,t+dt)\right)s(\theta,t) + \langle J(\theta+\eta(t)dt,t+dt)\rangle (1-s(\theta,t))\right]\,. \tag{5}$$

Expanding to first order in dt and using the fact that $s(\theta,t)$ is a bianry variable, we get

$$J(\theta,t) = -v_0\cos(\theta)dt + \min\left[x_c + J(0,t); J(\theta,t) + \partial_t J(\theta,t)dt + D\partial_\theta^2 J(\theta,t)dt\right] . \tag{6}$$

Therefore, if $x_c + J(0,t) \leq J(\theta,t)$ the optimal policy is $s^*(\theta,t) = 1$ and hence $x_c + J(0,t) = J(\theta,t)$. In the opposite case $x_c + J(0,t) > J(\theta,t)$, we obtain $s^*(\theta,t) = 0$ and

$$-\partial_t J(\theta,t) = -v_0\cos(\theta) + D\partial_\theta^2 J(\theta,t) . \tag{7}$$

The evolution equation can be rewritten in a more compact form using the domain $\Omega(\theta,t)$ as

$$-\partial_t J(\theta,t) = D\partial_\theta^2 J(\theta,t) - v_0\cos(\theta) , \quad \theta\in\Omega(t) , \tag{8}$$

where

$$\Omega(t) = \{\theta : J(\theta,t) < J(0,t) + x_c\} . \tag{9}$$

Note that the domain $\Omega(t)$ directly determines the optimal reorientation policy. Indeed,

$$s^*(\theta,t) = \begin{cases} 0 , & \text{if } \theta\in\Omega(t) , \\ 1 , & \text{otherwise} . \end{cases} \tag{10}$$

This differential equation has to be solved with the final condition $J(\theta,t_f) = 0$, which can be obtained by setting $t = t_f$ in Eq. (2). The boundary condition $\partial_\theta J(\theta,t) = 0$ for $\theta\in\Omega(\theta,t)$ is derived in Ref. [2]. The boundary domain must be updated dynamically at each integration step according to Eq. (9). Finally, once $\Omega(t)$ is determined at all times, the optimal strategy can be obtained from (10).

We also consider the region where $(t_f - t) \ll 1/D$. Indeed, when the final time t_f is approached, the optimal policy is to let the system evolve freely, without reorientations. Hence, in this region, the differential equation is defined over the full interval $[-\pi,\pi]$ (with periodic boundary conditions) and can be solved analytically. The most general solution reads

$$J(\theta,t) = -\frac{v_0}{D}(1 - e^{-D(t_f - t)})\cos(\theta) . \tag{11}$$

Note that when $D(t - t_f) \ll 1$, the condition $J(\theta,t) < J(0,t) + x_c$ is verified everywhere in $[-\pi,\pi]$. The solution in Eq. (11) is therefore valid until, increasing $(t_f - t)$, the condition $J(\theta,t) = J(0,t) + x_c$ is verified for the first time. This occurs at the critical time $t = t^c$, where

$$t^c = t_f - \frac{1}{D}\log\left(\frac{2}{2-z}\right) . \tag{12}$$

Therefore, if $z < z_c = 2$, the system is only actively controlled for $t < t^c$ and is left free to evolve without reorientations for $t > t^c$. For $z \to z_c$ the critical time diverges. Hence, for $z > z_c$ the cost is too high and it is never convenient to reorient the direction.

STEADY-STATE PROPERTIES

In this section, we investigate the steady-state distribution of the angle θ in the regime where $tD \gg 1$ and $(t_f - t)D \gg 1$. Assuming that the angle θ is reset to 0 as soon as $|\theta| > \theta_a$, the steady state probability density function $P_{\rm ss}(\theta)$ satisfies the stationary Fokker-Planck equation [3]

$$0 = D\partial_\theta^2 P(\theta) - 2\delta(\theta)\left[D\partial_\theta P(\theta_a)\right] , \tag{13}$$

where the amplitude of the δ function indicates that the particles flowing at the absorbing barrier at $\theta = \pm\theta_a$ are reset at $\theta = 0$. The boundary conditions are absorbing, since particles are reset as soon as they reach the bounday at $\theta = \pm\theta_a$, corresponding to $P(\pm\theta_a) = 0$. The most general normalized solution that satisfies the boundary conditions reads

$$P_{ss}(\theta) = \frac{\theta_a - |\theta|}{\theta_a^2} . \tag{14}$$

TIME BETWEEN TWO REORIENTATION EVENTS

In this section, we compute the distribution of the time T between two subsequent reorientation events. We denote by $S(\theta,t)$ the survival probability, i.e., probability that an angle starting from θ is not reset up to time t.

The survival probability satisfies the backward Fokker-Planck equation [4]

$$\partial_t S = D\partial_\theta^2 S\,, \tag{15}$$

with initial condition $S(\theta,0)=1$ and boundary conditions $S(\pm\theta_a,t)=0$. Taking a Laplace transform with respect to t, we find

$$s\tilde{S}(\theta,s)-1 = D\partial_\theta^2\tilde{S}(\theta,s)\,, \tag{16}$$

where

$$\tilde{S}(\theta,s)=\int_0^\infty dt\; S(\theta,t)e^{-st}\,. \tag{17}$$

Solving the equation and imposing the boundary conditons, we get

$$\tilde{S}(\theta,s)=\frac{1}{s}\left[1-\frac{\cosh\left(\sqrt{\frac{s}{D}}\theta\right)}{\cosh\left(\sqrt{\frac{s}{D}}\theta_a\right)}\right]\,. \tag{18}$$

We denote by $F(\theta,T)$ the probability density function (PDF) of the time T of the first reorientation event, assuming that the initial angle is θ. Note that, since after a reorientation the angle is set to $\theta=0$, the PDF of the time T between two reorientation events is $F(T)\equiv F(0,T)$.

Using the relation

$$S(\theta,t)=\int_t^\infty F(\theta,t')dt'\,, \tag{19}$$

we find

$$\tilde{F}(\theta,s)=\int_0^\infty dt\; e^{-st}F(\theta,t)=\frac{\cosh\left(\sqrt{\frac{s}{D}}\theta\right)}{\cosh\left(\sqrt{\frac{s}{D}}\theta_a\right)}\,. \tag{20}$$

Setting $\theta=0$ we find that the distribution of the time T between resetting events reads

$$F(0,T)=\frac{1}{2\pi i}\int_\Gamma ds\; e^{sT}\frac{1}{\cosh\left(\sqrt{\frac{s}{D}}\theta_a\right)}\,, \tag{21}$$

where Γ is a Bromwich contour parallel to the imaginary axis in the complex-s plane. Using Cauchy's residue theorem, we find

$$F(0,T)=\frac{D}{\theta_a^2}f\left(\frac{DT}{\theta_a^2}\right)\,, \tag{22}$$

where

$$f(w)=\pi\sum_{n=0}^{\infty}(-1)^n(2n+1)\exp\left[-\frac{\pi^2(2n+1)^2w}{4}\right]\,. \tag{23}$$

This function $f(w)$ is non-monotonic, with a maximum at $w\approx 0.17$.

The large-w regime can be immediately obtained from Eq. (23) by taking the $n=0$ term

$$f(w)\approx \pi e^{-\pi^2 w/4}\,. \tag{24}$$

To investigate the small-w limit, we use Poisson summation formula to find the alternative expression

$$f(w)=\frac{1}{2\sqrt{\pi}}\sum_{m=-\infty}^{\infty}(-1)^{m+1}(2m-1)w^{-3/2}e^{-(2m-1)^2/(4w)}\,. \tag{25}$$

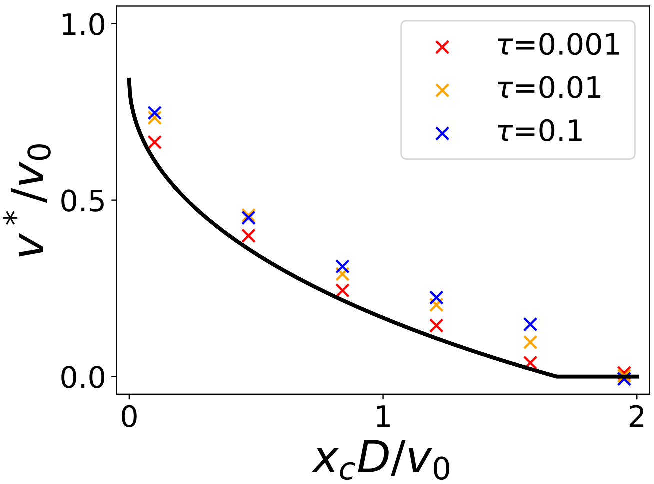

FIG. 1. Optimal speed v^*/v_0 as a function of $z = Dx_c/v_0$ in the presence of execution error with uniform distribution in $[-1, 1]$ and correlation timescale τ. The symbols indicate numerical simulations with $dt = 0.001$, $v_0 = D = 1$, and different values of τ. We observe that as $\tau \to 0$ the simulations approach the theoretical curve. The execution error $\epsilon(t)$ is obtained as a Brownian motion with periodic boundary conditions in $[-1, 1]$ and diffusion constant $D_\epsilon = 1/\tau$.

The small-w behavior is given by the $m = 0$ and $m = 1$ terms, yielding

$$f(w) \approx \frac{1}{\sqrt{\pi} w^{3/2}} e^{-1/(4w)} . \tag{26}$$

The mean time T between two resetting events can be evaluated as

$$\langle T \rangle = \int_0^\infty dt\; S(0, t) = \tilde{S}(\theta, 0) = \frac{\theta_a^2}{2D} . \tag{27}$$

EFFECT OF EXECUTION ERRORS

Here we introduce execution errors in the model. The dynamics of θ is modified as

$$\begin{cases} \dot{x}(t) = v_0 \cos(\theta(t) + \epsilon(t)) , \\ \dot{y}(t) = v_0 \sin(\theta(t) + \epsilon(t)) , \\ \dot{\theta}(t) = \eta(t) . \end{cases} \tag{28}$$

We assume that the execution error $\epsilon(t)$ is uniformly distributed in $[-\delta, \delta]$ and with time correlations $\langle \epsilon(t)\epsilon(t') \rangle \sim e^{-|t-t'|/\tau}$ with some timescale τ which we assume to be fast, i.e., $\tau \ll \frac{1}{D}$, where we recall that $1/D$ is the correlation timescale of $\theta(t)$. Then, the cost function becomes

$$\mathcal{F}_{0,0} = \langle x_c n_r \rangle_{\vec{\theta}} - \langle \int_0^{t_f} v_0 \cos(\theta(t) + \epsilon(t)) dt \rangle_{\vec{\epsilon}, \vec{\theta}} , \tag{29}$$

where we have made explicit that the average is computed over the full trajectories $\vec{\epsilon} = \{\epsilon(t')\}_{0<t'<T}$ and $\vec{\theta} = \{\theta(t')\}_{0<t'<T}$. Since the additional noise $\epsilon(t)$ is a fast ergodic process, we can approximate the right-hand side with an ensemble average over the distribution of ϵ. To show this, let us first choose a time increment Δ such that $\tau \ll \Delta \ll 1/D$ and rewrite the integral over t as

$$\mathcal{F}_{0,0} = \langle x_c n_r \rangle_{\vec{\theta}} - \langle v_0 \sum_{i=0}^{T/\Delta} \int_{i\Delta}^{(i+1)\Delta} \cos(\theta(t) + \epsilon(t)) dt \rangle_{\vec{\epsilon}, \vec{\theta}} . \tag{30}$$

Since $\Delta \gg \tau$, we can use the ergodicity of $\epsilon(t)$ and replace the time average with an ensemble average

$$\mathcal{F}_{0,0} \approx \langle x_c n_r \rangle_{\vec{\theta}} - \langle v_0 \sum_{i=0}^{T/\Delta} \int_{i\Delta}^{(i+1)\Delta} \langle \cos(\theta(i\Delta) + \epsilon) \rangle_\epsilon dt \rangle_{\vec{\theta}} \,. \tag{31}$$

where $\langle \cdot \rangle_\epsilon$ indicates the ensemble average over the random variable ϵ. Note that we have replaced $\theta(t)$ with $\theta(i\Delta)$ since $\Delta \ll 1/D$. Finally, by explicitly computing the average over ϵ, we find

$$\mathcal{F}_{0,0} \approx \langle -\frac{\sin(\delta)}{\delta} v_0 \int_0^{t_f} \cos(\theta(t)) dt + x_c n_r \rangle_{\vec{\theta}} \,. \tag{32}$$

In Fig. 1 we observe that as expected this approximations works well for small τ. In the limit $\delta \to 0$, we recover the standard case above. For finite δ, the optimal policy will be modified. By rewriting the cost function as

$$\mathcal{F}_{0,0} \approx \frac{\sin(\delta)}{\delta} \langle -v_0 \int_0^{t_f} \cos(\theta(t)) dt + \frac{\delta}{\sin(\delta)} x_c n_r \rangle_{\vec{\theta}} \,. \tag{33}$$

Thus, execution noise increases the effective cost by a factor $\delta/\sin(\delta) \geq 1$, while the overall displacement is scaled by a factor $\sin(\delta)/\delta \leq 1$.

EFFECT OF PARTIAL OBSERVATIONS

In this section, we consider the case where the agent has only partial knowledge of its current direction. We assume that the real direction θ can be affected by various sources of environmental noise. The agent can only detect a subset of these sources. By integrating the available information, the agent derives an estimate, denoted as θ_1, of its current direction. The remaining unmeasurable factors give rise to an angle mismatch, denoted as θ_2, which grows independently of θ_1 over time. We consider the following dynamics

$$\begin{cases} \dot{x}(t) = v \cos(\theta(t)) \,, \\ \dot{y}(t) = v \sin(\theta(t)) \,, \\ \theta(t) = \theta_1(t) + \theta_2(t) \,, \\ \dot{\theta}_1(t) = \eta_1(t) \,, \\ \dot{\theta}_2(t) = \eta_2(t) \,, , \end{cases} \tag{34}$$

where η_1 and η_2 are zero-mean uncorrelated Gaussian white noise with $\langle \eta_i(t) \eta_j(t') \rangle = 2D_i \delta_{ij} \delta(t - t')$. The angle θ_1 represents the estimate of the agent of the real direction θ. After a reorientation event, the agent has perfect information on its direction and hence both θ_1 and θ_2 are set to zero. We consider the average speed

$$v = \lim_{t_f \to \infty} \frac{1}{t_f} \langle v_0 \int_0^{t_f} \cos(\theta(t)) dt - x_c N(t_f) \rangle \,, \tag{35}$$

where $N(t_f)$ is the total number of resetting events. We are interested in the restarting policy that minimizes the total cost. Note that in principle this policy could depend on the total time τ elapsed since the last resetting event, since the amplitude of the noise θ_2 grows as $\sqrt{\tau}$. However, here we only focus on time-independent strategies. In other words, we assume that the agent cannot measure θ_2 and has to take decisions only knowing θ_1. Hence, the optimal strategy consists once again in resetting the system if $|\theta| > \theta_a$.

We rewrite the cost average speed as

$$v = v_0 \langle \cos(\theta) \rangle_{\mathrm{ss}} - x_c / \langle T \rangle \,, \tag{36}$$

where $\langle \cdot \rangle_{\mathrm{ss}}$ indicates the steady state average and T is the time between two reorientation. From Eq. (27), we find $\langle T \rangle = \theta_a^2/(2D_1)$. We next compute the steady-state distribution of $\theta = \theta_1 + \theta_2$.

We define $G(\theta, t)$ as the probability density of reaching θ at time t starting from $\theta_0 = 0$ with no resetting events, i.e., always remaining in the interval $[-\theta_a, \theta_a]$. Similarly, we define $P(\theta, t)$ as the probability density of reaching θ at time t starting from $\theta_0 = 0$ with any number of resetting events. We also denote by $F(T)$ the distribution of the time T between two successive resetting events. The distribution $P(\theta, t)$ satisfies the relation

$$P(\theta,t) = G(\theta,t) + \int_0^t dt'\ F(t')P(\theta, t-t')\,. \tag{37}$$

The first term corresponds to the case where θ is reached without any resetting events, while the second term describes the case where the first resetting occurs at time t'. Performing a Laplace transform on both sides, we find

$$\tilde{P}(\theta,s) = \frac{\tilde{G}(\theta,s)}{1-\tilde{F}(s)}\,, \tag{38}$$

where we use the notation

$$\tilde{f}(s) = \int_0^\infty dt\ e^{-st} f(t)\,. \tag{39}$$

From Eq. (20), we have

$$\tilde{F}(s) = \frac{1}{\cosh\left(\sqrt{\frac{s}{D_1}}\theta_a\right)}\,. \tag{40}$$

The steady-state distribution

$$P_{ss}(\theta) = \lim_{t\to\infty} P(\theta,t)\,, \tag{41}$$

can then be obtained taking the small-s limit

$$P_{ss}(\theta) = \lim_{s\to 0} s\frac{\tilde{G}(\theta,s)}{1-\tilde{F}(s)}\,. \tag{42}$$

Using Eq. (40), we obtain

$$P_{ss}(\theta) = \frac{2D_1}{\theta_a^2}\tilde{G}(\theta,0) = \frac{2D_1}{\theta_a^2}\int_0^\infty dt\ G(\theta,t)\,. \tag{43}$$

Therefore, we need to compute the constrained propagator $G(\theta,t)$.

To proceed, we define the propagator $G(\theta_1,\theta_2,t)$ as the probability of reaching the angles θ_1 and θ_2 at time t without resetting, i.e., with the constraint that $|\theta_1| < \theta_a$ up to time t. The propagator satisfies the Fokker-Plank equation

$$\partial_t G(\theta_1,\theta_2,t) = D_1\partial^2_{\theta_1} G(\theta_1,\theta_2,t) + D_2\partial^2_{\theta_2} G(\theta_1,\theta_2,t)\,, \tag{44}$$

with initial condition $G(\theta_1,\theta_2,t=0) = \delta(\theta_1)\delta(\theta_2)$ and boundary conditions $G(\theta_1,\theta_2,t) = 0$ if $|\theta_1| = \theta_a$. Integrating both sides over t and integrating by parts, we find

$$-\delta(\theta_1)\delta(\theta_2) = D_1\partial^2_{\theta_1}\tilde{G}(\theta_1,\theta_2,0) + D_2\partial^2_{\theta_2}\tilde{G}(\theta_1,\theta_2,0)\,. \tag{45}$$

Performing a Fourier transform with respect to θ_2, we find

$$-\delta(\theta_1) = D_1\partial^2_{\theta_1} g(\theta_1,k) - D_2k^2 g(\theta_1,k)\,, \tag{46}$$

where

$$g(\theta_1,k) = \int_{-\infty}^{\infty} d\theta_2\ e^{ik\theta_2}\tilde{G}(\theta_1,\theta_2,0)\,. \tag{47}$$

Solving the differential equation and imposing the boundary conditions, we get

$$g(\theta_1,k) = \frac{1}{2\sqrt{a}D_1}\frac{\sinh(\sqrt{a}(\theta_a - |\theta|))}{\cosh(\sqrt{a}\theta_a)}H(\theta_a - |\theta|)\,, \tag{48}$$

where $a = D_2 k^2/D_1$ and $H(z)$ is the Heaviside step function. Inverting the Fourier transform, we get

$$\tilde{G}(\theta_1, \theta_2, 0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\ e^{-ik\theta_2} g(\theta_1, k)\,. \tag{49}$$

Integrating over θ_2 keeping $\theta = \theta_1 + \theta_2$ fixed, we find

$$\tilde{G}(\theta, 0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\ \int_{-\infty}^{\infty} d\theta_2\ e^{-ik\theta_2} g(\theta - \theta_2, k)\,. \tag{50}$$

Evaluating the integral over θ_2 with Mathematica, we find

$$\tilde{G}(\theta, 0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\ e^{-ik\theta}\frac{1}{(D_1+D_2)k^2}\left[1 - \frac{\cos(k\theta_a)}{\cosh\left(k\theta_a\sqrt{\frac{D_2}{D_1}}\right)}\right]\,. \tag{51}$$

Finally, using Eq. (43), we obtain the steady state distribution

$$P_{ss}(\theta) = \frac{2D_1}{\theta_a^2}\frac{1}{2\pi}\int_{-\infty}^{\infty} dk\ e^{-ik\theta}\frac{1}{(D_1+D_2)k^2}\left[1 - \frac{\cos(k\theta_a)}{\cosh\left(k\theta_a\sqrt{\frac{D_2}{D_1}}\right)}\right]\,. \tag{52}$$

As a check, we can verify that this distribution is correctly normalized to one

$$\int_{-\infty}^{\infty} d\theta\ P_{ss}(\theta) = \frac{2D_1}{\theta_a^2}\int_{-\infty}^{\infty} dk\ \delta(k)\frac{1}{(D_1+D_2)k^2}\left[1 - \frac{\cos(k\theta_a)}{\cosh\left(k\theta_a\sqrt{\frac{D_2}{D_1}}\right)}\right] = 1\,. \tag{53}$$

We can now evaluate the average in Eq. (36)

$$\begin{aligned}
\langle\cos(\theta)\rangle_{\rm ss} &= \frac{D_1}{\pi\theta_a^2}\int_{-\infty}^{\infty} d\theta\ \cos(\theta)\int_{-\infty}^{\infty} dk\ e^{-ik\theta}\frac{1}{(D_1+D_2)k^2}\left[1 - \frac{\cos(k\theta_a)}{\cosh\left(k\theta_a\sqrt{\frac{D_2}{D_1}}\right)}\right] \\
&= \frac{D_1}{\theta_a^2}\int_{-\infty}^{\infty} d\theta\ \frac{e^{i\theta}+e^{-i\theta}}{2\pi}\int_{-\infty}^{\infty} dk\ e^{-ik\theta}\frac{1}{(D_1+D_2)k^2}\left[1 - \frac{\cos(k\theta_a)}{\cosh\left(k\theta_a\sqrt{\frac{D_2}{D_1}}\right)}\right] \\
&= \frac{D_1}{\theta_a^2}\int_{-\infty}^{\infty} dk\ \left[\delta(k+1)+\delta(k-1)\right]\frac{1}{(D_1+D_2)k^2}\left[1 - \frac{\cos(k\theta_a)}{\cosh\left(k\theta_a\sqrt{\frac{D_2}{D_1}}\right)}\right] \\
&= \frac{2D_1}{\theta_a^2}\frac{1}{(D_1+D_2)}\left[1 - \frac{\cos(\theta_a)}{\cosh\left(\theta_a\sqrt{\frac{D_2}{D_1}}\right)}\right]
\end{aligned} \tag{54}$$

Therefore, the speed in Eq. (36) can be written as

$$v = v_0\frac{2D_1}{\theta_a^2}\frac{1}{(D_1+D_2)}\left[1 - \frac{\cos(\theta_a)}{\cosh\left(\theta_a\sqrt{\frac{D_2}{D_1}}\right)}\right] - \frac{2D_1 x_c}{\theta_a^2}\,, \tag{55}$$

which can be rewritten as

$$\frac{v}{v_0} = \frac{2}{\theta_a^2(1+1/\ell)}\left[1 - \frac{\cos(\theta_a)}{\cosh\left(\theta_a/\sqrt{\ell}\right)}\right] - \frac{2z}{\theta_a^2}\,, \tag{56}$$

where we have defined $\ell = D_1/D_2$ and $z = D_1 x_c/v_0$. In the limit of low cost $z \to 0$, we find the following asymptotic behaviors

$$\theta_a \approx \left(\frac{24\ell}{5+\ell}\right)^{1/4} z^{1/4}\,, \tag{57}$$

and

$$\frac{v^*}{v_0} \approx 1 - \sqrt{\frac{2(5+\ell)z}{3\ell}}\,. \tag{58}$$

EFFECT OF MEASUREMENT ERRORS

In this section, we consider the effect of measurement errors on the navigation strategies. Instead of assuming that the agent has access to only a subset of the factors affecting its direction, we assume that the agent has access to all relevant factors, albeit through noisy sensory channels. As a consequence, the internal representation angle θ_1, which the agent employs to approximate the real direction θ, will inevitably deviate from the true θ over time. To keep things simple, we assume that this deviation, denoted as θ_n, grows independently of the actual angle θ. The resulting dynamics read

$$\begin{cases} \dot{x}(t) = v_0 \cos(\theta(t))\,, \\ \dot{y}(t) = v_0 \sin(\theta(t))\,, \\ \theta_1(t) = \theta(t) + \theta_n(t)\,, \\ \dot{\theta}(t) = \eta(t)\,, \\ \dot{\theta}_n(t) = \eta_n(t)\,, \end{cases} \tag{59}$$

where $\eta(t)$ and $\eta_n(t)$ are independent Gaussian white noises with zero mean and correlators $\langle \eta(t)\eta(t')\rangle = 2D\delta(t-t')$ and $\langle \eta_n(t)\eta_n(t')\rangle = 2D_n\delta(t-t')$. We denote by $r = D/D_n$ the signal-to-noise ratio. Note that, despite the similarity with Eq. (34), the two models are actually quite different. Note for instance that the variance of the internal angle θ_1 is larger than that of the real angle θ for the model in Eq. (59), while it is the other way around for Eq. (34).

Once more, we examine the strategy in which the agent adjusts its direction when $|\theta_1| > \theta_a$ and we focus on the steady-state properties. As done in the previous section, we first derive the steady-state distribution $P_{\rm ss}(\theta)$ of the real direction θ for a fixed value of θ_a. Eq. (43) can be easily adapted to the model in Eq. (59), yielding

$$P_{ss}(\theta) = \frac{2(D+D_n)}{\theta_a^2}\int_0^\infty dt\, \int_{-\theta-\theta_a}^{-\theta+\theta_a} d\theta_n\; G(\theta,\theta_n,t)\,, \tag{60}$$

where $G(\theta,\theta_n,t)$ is the probability that θ and θ_n are reached at time t without any reorientations, assuming that initially $\theta = \theta_n = 0$.

This probability density satisfies the Fokker-Planck equation

$$\partial_t G(\theta,\theta_n,t) = D\partial_\theta^2 G(\theta,\theta_n,t) + D_n \partial_{\theta_n}^2 G(\theta,\theta_n,t)\,, \tag{61}$$

with boundary conditions $G(\theta,\theta_n,t) = 0$ if $|\theta+\theta_n| = \theta_a$ and initial condition $G(\theta,\theta_n,0) = \delta(\theta)\delta(\theta_n)$. Integrating over $t > 0$, we find

$$-\delta(\theta)\delta(\theta_n) = D\partial_\theta^2 \int_0^\infty dt\; G(\theta,\theta_n,t) + D_n\partial_{\theta_n}^2 \int_0^\infty dt\; G(\theta,\theta_n,t)\,. \tag{62}$$

Making the change of variables $\theta_1 = \theta + \theta_n$ and $w = \theta - w$, we find

$$-2\delta(\theta_1)\delta(w) = D(\partial_{\theta_1} + \partial_w)^2 f(\theta_1,w) + D_n(\partial_{\theta_1} - \partial_w)^2 f(\theta_1,w)\,, \tag{63}$$

where we have defined

$$f(\theta_1,w) = \int_0^\infty dt\; G\left(\theta = \frac{\theta_1+w}{2}, \theta_n = \frac{\theta_1 - w}{2}, t\right)\,. \tag{64}$$

Performing a Fourier transform with respect to w, we find

$$-2\delta(\theta_1) = (D+D_n)\partial_{\theta_1}^2 \hat{f}(\theta_1,k) + 2ik(D_n - D)\partial_{\theta_1}\hat{f}(\theta_1,k) - (D+D_n)k^2\hat{f}(\theta_1,k)\,, \tag{65}$$

where

$$f(\theta_1,w) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\; e^{-ikw}\hat{f}(\theta_1,k)\,. \tag{66}$$

Solving the equation and imposing the boundary conditions $\hat{f}(\pm\theta_a,k) = 0$, we find

$$\hat{f}(\theta_1,k) = \frac{1}{2|k|\sqrt{DD_n}} \frac{\exp\left(4|k|(\theta_a - |\theta_1|)\sqrt{DD_n}/(D+D_n)\right) - 1}{\exp\left(4|k|\theta_a\sqrt{DD_n}/(D+D_n)\right) + 1} \exp\left(i\frac{D-D_n}{D+D_n}k\theta_1 + 2\frac{\sqrt{DD_n}}{D+D_n}|k\theta_1|\right)\,. \tag{67}$$

Combining Eqs. (60) and (64), we get

$$P_{ss}(\theta) = \frac{2(D+D_n)}{\theta_a^2} \int_{-\theta-\theta_a}^{-\theta+\theta_a} d\theta_n \; f(\theta+\theta_n, \theta-\theta_n)\,, \tag{68}$$

where $f(\theta_1, w)$ is given in Eq. (66). We have checked that $P_{ss}(\theta)$ is correctly normalized to one.

Finally, the average velocity can be written as (see Eq. (36))

$$v = v_0 \langle \cos(\theta) \rangle_{\rm ss} - x_c/\langle T \rangle\,, \tag{69}$$

where $\langle \cdot \rangle_{\rm ss}$ denotes the average with respect to $P_{ss}(\theta)$ in Eq. (68). Performing the average and using $\langle T \rangle = \theta_a^2/(2(D+D_n))$ (see Eq. (27)), we find

$$\frac{v}{v_0} = \frac{2}{\theta_a^2}(1+1/r)\left[1 - \frac{\cos\left(\frac{\theta_a}{1+1/r}\right)}{\cosh\left(\frac{\theta_a\sqrt{1/r}}{1+1/r}\right)}\right] - \frac{2z}{\theta_a^2}(1+1/r)\,, \tag{70}$$

where $r = D/D_n$. In the limit of small cost $z \to 0$, we find

$$\theta_a \approx \frac{2^{3/4}3^{1/4}(1+r)^{3/4}}{r^{1/2}(5+r)^{1/4}} z^{1/4}\,, \tag{71}$$

and

$$\frac{v^*}{v_0} \approx 1 - \sqrt{\frac{2r^2(r+5)z}{3r^2(r+1)}}\,. \tag{72}$$

COMPARISON WITH POISSONIAN REORIENTATIONS

In this section, we compare the optimal reorientation strategy obtained in the main text for the simplest model (without measurement or observation errors) with the standard strategy of performing correction at random Poissonian times with a fixed rate $r_{\rm corr}$. In other words, in a small time interval dt, the agent performs a reorientation with probability $r_{\rm corr}dt$, irrespectively of its position.

The late-time ($t \gg 1/D$) average displacement in the case of Poissonian resetting was computed in [5] and reads

$$\langle x \rangle \approx \frac{v_0 r_{\rm corr} t}{r_{\rm corr} + D}\,. \tag{73}$$

The average number of correction events after time t is simply $r_{\rm corr}t$, therefore the effective average velocity in the x direction reads

$$v \approx \frac{v_0 r_{\rm corr}}{r_{\rm corr}+D} - x_c r_{\rm corr}\,, \tag{74}$$

where x_c is the cost per correction. By minimizing this expression, we find that optimal reorientation rate therefore reads

$$r_{\rm corr} = \begin{cases} D\frac{1-\sqrt{z}}{\sqrt{z}} & \text{if } z<1 \\ 0 & \text{if } z>1\,, \end{cases} \tag{75}$$

where we recall that $z = x_c D/v_0$. The corresponding optimal velocity reads

$$v^* = \begin{cases} v_0(1-\sqrt{z})^2 & \text{if } z<1 \\ 0 & \text{if } z>1\,. \end{cases} \tag{76}$$

The optimal speed is shown in Fig. 2, where we compare it with the optimal reorientation strategy derived in this paper. We observe the optimal reorientation strategy outperforms the Poissonian one for all values of $z = Dx_c/v_0$. Moreover, for $1 < z < 2$, while the Poissonian strategy has zero velocity, the optimal stategy manages to achieve finite speed.

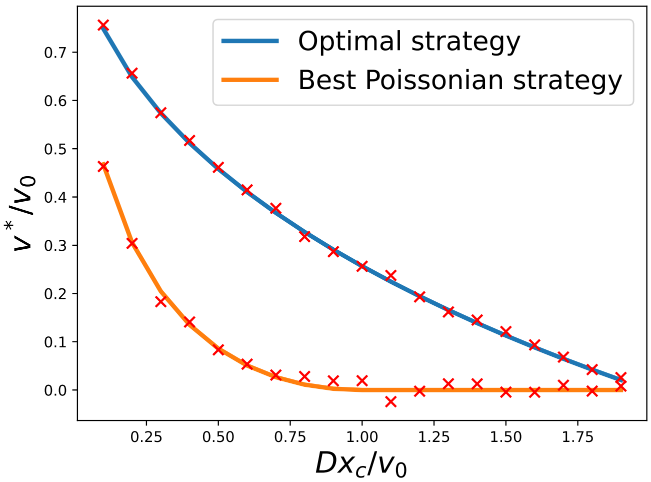

FIG. 2. Comparison between the optimal strategy derived in this paper and the best Poissonian strategy. The continuous lines show the exact results derived in the text, while the symbols correspon to numerical simulations.

NUMERICAL SIMULATIONS

In this section, we describe the details of the numerical simulations presented in the main text. To verify our analytical predictions we implement a standard Langevin algorithm. In the stationary regime $D(t_f - t) \gg 1$ the optimal strategy is time-independent. Hence, in a small time interval dt the optimal dynamics of the direction θ reads

$$\theta(t+dt) = \begin{cases} \theta(t) + \sqrt{2Ddt}\eta(t) & \text{if } |\theta(t)| < \theta_a\,, \\ 0 & \text{if } |\theta(t)| > \theta_a\,, \end{cases} \tag{77}$$

where $\eta(t)$ is a zero-mean unit-variance Gaussian variable and θ_a is the optimal activation angle for a given noise level $z = Dx_c/v_0$. For each value of z, we generate numerically a single trajectory of duration $t_f = 10^4$. We choose $v_0 = 1$, $x_c = z$, $D = 1$ and $dt = 0.01$. Finally, the average speed is evaluated as the time average

$$v^* \approx \frac{1}{t_f}\left[\int_0^{t_f} v_0 \cos(\theta(t))dt - x_c N(t_f)\right]\,, \tag{78}$$

where $N(t)$ is the total number of reorientation events up to time. This algorithm can be easily adapted to the models of sensory or motor noise.

To evaluate the average time T between two reorientation events, we integrate Eq. (77) starting from $\theta(t=0) = 0$ and stopping the dynamics when a reorientation occurs for the first time. We repeat this procedure to obtain 10^7 samples.

The code used to numerically integrate the optimal control equations, perform numerical simulations and generate figures is available at `https://anonymous.4open.science/r/optimal_reorientations-48DE/README.md`

[1] R. Bellman and Kalaba, *Dynamic programming and modern control theory*, Vol. 81 (Academic Press, New York, 1965).
[2] B. De Bruyne and F. Mori, Phys. Rev. Research **5**, 013122 (2023).
[3] B. De Bruyne, J. Randon-Furling, and S. Redner, Phys. Rev. Lett. **125**, 050602 (2020).
[4] S. Redner, *A guide to first-passage processes* (Cambridge university press, 2001).
[5] V. Kumar, O. Sadekar, and U. Basu, Phys. Rev. E **102**, 052129 (2020).