A LETTERS JOURNAL EXPLORING THE FRONTIERS OF PHYSICS

EPL, xx (2022) xxxxx
doi: 10.1209/0295-5075/ac6620

www.epljournal.org

Gradient sensing in Bayesian chemotaxis

ANDREA AUCONI[1(a)], MAJA NOVAK[1,2] and BENJAMIN M. FRIEDRICH[1,3]

[1] *cfaed, Technische Universität Dresden - 01069 Dresden, Germany*
[2] *Department of Physics, Faculty of Science, University of Zagreb - Bijenička cesta 32, 10000 Zagreb, Croatia*
[3] *Cluster of Excellence "Physics of Life" - 01307 Dresden, Germany*

received 23 December 2021; accepted in final form 11 April 2022
published online xx Xxxx xxxx

Abstract – Bayesian chemotaxis is an information-based target search problem inspired by biological chemotaxis. It is defined by a decision strategy coupled to the dynamic estimation of target position from detections of signaling molecules. We extend the case of a point-like agent previously introduced (Vergassola *et al.*, *Nature* (2007)), which establishes concentration sensing as the dominant contribution to information processing, to the case of a circle-shaped agent of small finite size. We identify gradient sensing and a Laplacian correction to concentration sensing as the two leading-order expansion terms in the expected entropy variation. Numerically, we find that the impact of gradient sensing is most relevant because it provides direct directional information to break symmetry in likelihood distributions, which are generally circle shaped by concentration sensing.

Introduction. – Biological cells and organisms navigate their environments, *e.g.*, in search for nutrient sources, guided by the detection of signaling molecules [1–3]. Knowledge of the time and location of such detections, typically receptor-ligand binding events, provides information about the target position if the target is a source of ligands. More precisely, an agent can reduce its uncertainty about target position if the binding events time series is interpreted with respect to a *model* of the environment. In other words, a chemotactic agent should know something about the relation between binding statistics and target position in order to process the information from binding events.

The estimation of a time-varying hidden state from noisy measurements is the subject of stochastic filtering theory [4–6]. The celebrated Kalman filter provides the analytical form of the estimation update scheme for Gaussian systems, and it was recently applied to biological examples such as concentration and direction sensing in the linear regime [7–11]. Chemotaxis can be formalized as stochastic filtering coupled to a decision-making problem, namely the decision of where to move next.

Just knowing that the binding rate increases towards the target is sufficient for a simple strategy of gradient ascent, where the agent only needs to repeatedly estimate the local gradient from the binding asymmetry across its diameter [12], or across short-distance runs like in bacterial run-and-tumble chemotaxis [13], referred to as spatial or temporal comparison in the biological literature, respectively.

However, if concentrations are dilute and therefore binding events are sparse, then the estimation of gradients becomes very slow and inefficient. Nevertheless, if an agent is equipped with a model of the binding rate as a function of the relative target position, then it can gain information also by estimating the local binding rate. Indeed, the signal-to-noise ratio in the estimation of a concentration from binding events is typically much larger than that of a gradient.

The chemo-sensory navigation of a point-like agent performing stochastic filtering on a likelihood map was discussed in the pioneering work of Vergassola *et al.* [14,15], where the authors also introduced a decision strategy that consists in maximizing the expected information gain locally in time, there named *infotaxis*. Analytical expressions for the information gain and decision making were subsequently given in [16]. We would like to refer to this and similar search problems as Bayesian chemotaxis.

Intuitively, a point-like agent can estimate the local gradient, but only indirectly by movement and temporal comparison of estimated concentrations. On the contrary, an agent with small finite size can also obtain direct information on the local gradient, as the location of a single binding event on its surface is itself a directional cue [9,17].

(a) E-mail: andrea.auconi@gmail.com (corresponding author)

Added by us as per EPL standard

Therefore, we interpret the case of point-like agent as pure concentration sensing, while the small finite-size agent performs simultaneously concentration and gradient sensing.

In this letter, we introduce gradient sensing from an information gain expansion with respect to agent size. The zeroth-order terms are the motility noise, and the concentration sensing discussed in [14–16]. The second order of the expansion gives the gradient sensing term, as well as a Laplacian correction to concentration sensing which, to the best of our knowledge, was not yet discussed in the literature. We evaluate the relative importance of these terms in a steady-state chemotactic search model in two space dimensions. We discuss the increase on target search efficiency deriving from a finite agent size. Finally, we briefly discuss the decision process comparing infotaxis to the simpler maximum-likelihood and minimum-distance direction strategies.

The model. – A disk-shaped agent of radius a moving in two-dimensional space is immersed in a concentration profile of signaling molecules released by a target. These molecules are detected on the agent's circumference through binding events happening at a rate proportional to the local concentration. Let us use a coordinate system centered at and co-rotating with the agent (called "material frame" [9]), and let $r(\mathbf{x})$ denote the binding rate at the agent position $\mathbf{0}$ when the target is at position $\mathbf{x}$. We denote by $p(\mathbf{x}_t)$ the smooth likelihood distribution formalizing the agent's knowledge about target position at time t, which is based on previous measurements and dynamics up to time t. The agent moves with a velocity $\mathbf{v}_t$ having a constant speed v, and a direction angle θ_t, which is continuously decided based on the current likelihood. The target further undergoes translational diffusion with constant diffusion coefficient D.

Stochastic filtering. – The likelihood $p(\mathbf{x}_t)$ is continuously updated in response to the binding events time series and to account for the target stochastic dynamics in a Bayesian inference scheme called stochastic filtering. Please note that in this section all probabilities are implicitly conditional to measurement information up to $t = 0$ only, unless specified differently. Consider a time interval $[0, \tau]$ with $\tau > 0$ sufficiently small so that variations of the binding rate due to movement within $[0, \tau]$ are negligible. Let us further assume that the agent's velocity in real space follows a smooth protocol decided at $t = 0$. The evolution of the likelihood is then decomposed into a prediction step $p(\mathbf{x}_\tau) = \int \mathrm{d}\mathbf{x}_0\, p(\mathbf{x}_0)p(\mathbf{x}_\tau|\mathbf{x}_0)$, which is driven by the transition density of the target position $p(\mathbf{x}_\tau|\mathbf{x}_0)$, and of an update step $p(\mathbf{x}_\tau|\mathbf{m}_\tau) = p(\mathbf{x}_\tau)p(\mathbf{m}_\tau|\mathbf{x}_\tau)/p(\mathbf{m}_\tau)$, which is driven by the measurement $\mathbf{m}_\tau$, that here is a record of the number and locations on the agent's circumference of the binding events happened within the interval $[0, \tau)$. The measurement model $p(\mathbf{m}_\tau|\mathbf{x}_\tau)$ formalizes the relation between binding rate and target position, and such relation is known by the agent in Bayesian chemotaxis. The current measurement probability $p(\mathbf{m}_\tau) = \int \mathrm{d}\mathbf{x}_\tau\, p(\mathbf{x}_\tau)p(\mathbf{m}_\tau|\mathbf{x}_\tau)$ formalizes the agent's uncertainty at time 0 about the outcome of the measurement $\mathbf{m}_\tau$, which is delivered to it at time τ. In the limit $\tau \to 0$, the Bayesian updating is integrated in the non-anticipating Ito stochastic calculus scheme to obtain a time-continuous inference and decision process, see Supplementary Material `Supplementarymaterial.pdf` (SM).

Expansion in the agent's size. – If we neglect the effect of the agent body and movement on the sensing field, the inference problem is equivalent to that made by a density of point-like agents distributed on circumference $a\{\mathbf{e}^{\mathrm{i}\theta}\}_{\theta=[0,2\pi)}$, where $\mathbf{e}^{\mathrm{i}\theta}$ is a unit vector in 2D with direction defined by the angle θ from a fixed axis in the agent's reference frame. Accordingly, the event probability density per unit angle in a small time interval τ is $p(\mathbf{m}_\tau = \mathbf{e}^{\mathrm{i}\theta}|\mathbf{x}) = \frac{\tau}{2\pi} r(\mathbf{x} - a\mathbf{e}^{\mathrm{i}\theta}) + \mathcal{O}(\tau^2)$, where the normalization is chosen so that in the limit $a \to 0$ one recovers the case of a single point-like agent. Note that the measurement $\mathbf{m}_\tau$ is vector valued, assuming $\mathbf{0}$ (no events) almost surely for $\tau \to 0$, and $\mathbf{e}^{\mathrm{i}\theta}$ in the case of a binding event at angle θ. If the binding rate differences along the agent's body are small compared to the absolute binding rate, then we can study the effect of the agent size on stochastic filtering through a second-order Taylor expansion of the binding rate. For an event at angle θ, we write

$$r(\mathbf{x} - a\mathbf{e}^{\mathrm{i}\theta}) = r(\mathbf{x}) - a\,\mathbf{e}^{\mathrm{i}\theta} \cdot \nabla r(\mathbf{x}) + \frac{1}{2}a^2 \mathbf{e}^{\mathrm{i}\theta} \cdot (\mathbf{H}(\mathbf{x})\,\mathbf{e}^{\mathrm{i}\theta}) + \mathcal{O}(a^3), \quad (1)$$

where $\mathbf{H}(\mathbf{x}) \equiv \nabla \otimes \nabla r(\mathbf{x})$ is the Hessian matrix of the binding rate evaluated at $\mathbf{x}$. Using the expansion eq. (1) in the stochastic filtering problem, we obtain the corresponding expansion of the likelihood stochastic evolution equation, whose expression is given in the SM.

Entropy dynamics. – We quantify the agent's current uncertainty about target position by the Shannon entropy $S[p(\mathbf{x})] = -\int \mathrm{d}\mathbf{x}\, p(\mathbf{x}) \ln p(\mathbf{x})$. The expected entropy variation up to second order in a reads

$$\frac{\langle \mathrm{d}S \rangle}{\mathrm{d}t} = -D\langle \nabla^2 \ln p \rangle + \left\langle \left(r + \frac{a^2}{4}\nabla^2 r \right) \ln\left(\frac{\langle r \rangle}{r} \right) \right\rangle + \frac{a^2}{4}\left(\frac{\|\langle \nabla r \rangle\|^2}{\langle r \rangle} - \left\langle \frac{\|\nabla r\|^2}{r} \right\rangle \right) + \mathcal{O}(a^3), \quad (2)$$

where $\|\cdot\|$ denotes the Euclidean norm, the averaging brackets are meant with respect to the current likelihood p, and we dropped the explicit dependence on $\mathbf{x}$. Equation (2), whose derivation is given in the SM text, is the main result of this letter. Let us now discuss its physical meaning.

The first zeroth-order term, $-D\langle \nabla^2 \ln p \rangle \geq 0$, is the information erasure due to translational diffusion, and it is always non-negative having the form of a Fisher information [18]. The second zeroth-order term, $\langle r \ln(\langle r \rangle / r) \rangle \leq 0$,

is the concentration sensing already discussed in [14,16], and it is always non-positive by Jensen's inequality.

We identify the second-order term $\|\langle\nabla r\rangle\|^2/\langle r\rangle - \langle\|\nabla r\|^2/r\rangle \leq 0$ as *gradient sensing*, and it is always non-positive having the form of the expectation of a negative square. Indeed, once a binding event has happened, the expected information gain from further knowing its location on the circumference can only be positive, being a type of conditional mutual information. Note that for this term, if we consider an agent who periodically erases all of its current directional knowledge so that $\langle\nabla r\rangle = \mathbf{0}$, then we obtain simply $-\langle\|\nabla r\|^2/r\rangle$. Depending on whether additionally distance information becomes erased, this expectation value is either taken with respect to a fixed prior or and adaptive one [19].

The other second-order term $\langle(\nabla^2 r)\ln(\langle r\rangle/r)\rangle$ is interpreted as a *Laplacian correction* to concentration sensing. This is due to the convexity of the binding rate field, which is the leading-order correction in the total binding rate on the agent's circumference, $\int_0^{2\pi}(\mathrm{d}\theta/2\pi)\, r(\mathbf{x}+a\mathbf{e}^{\mathrm{i}\theta}) = r(\mathbf{x}) + (a^2/4)\nabla^2 r(\mathbf{x}) + \mathcal{O}(a^4)$. The impact of the Laplacian correction on the entropy variation can be positive or negative depending on the particular likelihood configuration and binding rate field shape.

From eq. (2), we can derive as limiting cases two known examples of stochastic filtering in biology, which are mathematically analogous to the Kalman filter [6,7].

Limiting cases. – We start with pure gradient direction sensing of [9]. Consider a target constrained at a fixed position at distance R, which generates a conic field $r(\mathbf{x}) = r_0 - \|\nabla r\|\,\|\mathbf{x}\|$, with $\|\nabla r\| \leq \frac{r_0}{R+a}$. The agent does not move but undergoes rotational diffusion with coefficient D_{rot}, defined as Brownian motion in the polar variable θ (denoted ψ in [9]). In the small noise limit $D_{\mathrm{rot}} \to 0$, the angular likelihood distribution is Gaussian with mean μ_θ and variance σ_θ^2. Concentration sensing is irrelevant here because we can just substitute $r = r(R) > 0$, so that only gradient sensing and rotational diffusion contribute to the entropy variation. The entropy variation due to rotational diffusion is a type of Fisher information [18], and in the linear regime the corresponding Cramer-Rao bound $-\langle\partial_\theta^2 \ln p\rangle \geq 1/\sigma_\theta^2$ is saturated, as can be immediately seen by taking the x-axis along the direction of the current estimate μ_θ, and linearizing $y = R\theta + \mathcal{O}(\theta^2)$. Similarly, the gradient sensing is found by linearizing $\|\langle\nabla r\rangle\|^2 = \|\nabla r\|^2(1-\sigma_\theta^2) + \mathcal{O}(\sigma_\theta^3)$. Then the angular variance which makes the expected entropy variation vanish satisfies

$$\sigma_\theta^2\big|_{\langle dS\rangle=0} = \sqrt{D_{\mathrm{rot}}}\,\frac{2\sqrt{r(R)}}{a\|\nabla r\|}. \tag{3}$$

We see that the average binding rate $r(R)$ acts as Poissonian noise for the signal $2a\|\nabla r\|$, which is the difference of binding rate across the agent's diameter. The angular variance scales as the square root of the rotational noise D_{rot}, reproducing the result of [9].

Similarly, we derive the pure concentration sensing case of [8] by considering eq. (2) in one-dimensional space, for a point-like agent ($a = 0$) with translational diffusion whose likelihood at $t = 0$ is non-zero only on one side, say for $x > 0$. We take again the linear binding rate $r(x) = r_0 - |\nabla r||x|$ for simplicity, and expand the concentration sensing to obtain $\sigma_x^2\big|_{\langle dS\rangle=0} = \sqrt{D}\,\frac{2\sqrt{r(R)}}{|\nabla r|}$, which also scales as the square root of motility noise.

Gradient sensing in steady-state Bayesian chemotaxis. – To evaluate the relative importance of the expansion terms of eq. (2) in the full Bayesian chemotaxis model, we plot the histogram of their realized values over the steady-state dynamics obtained with a binding rate field $r(\mathbf{x}) = \lambda\|\mathbf{x}\|^{-1}$, with $\lambda > 0$, whose particular form is taken for numerical convenience. Here, we already employ a decision strategy (infotaxis) detailed in the next section. To ensure the existence of an ergodic steady-state dynamics for the chemotaxis process, we impose reflecting boundary conditions for the relative target distance $R_{\min} < \|\mathbf{x}\| < R_{\max}$. The lower bound $R_{\min}$, which can be interpreted as the target size, is taken for numerical stability and to ensure that the second-order expansion holds in the allowed region.

We find that gradient sensing is of larger magnitude compared to the Laplacian correction, see fig. 1. This can be understood by looking at typical likelihood shapes, see fig. 2 and the SM text. The main drivers of the likelihood dynamics are the velocity-induced translation, the diffusion-induced smoothing, and the concentration sensing. In particular, concentration sensing gives rise to circle-shaped distributions, because we consider a concentration field, which is symmetric around the target and the agent knows the relation between binding rate and distance. The likelihood is never an exact circle but rather an annulus whose width is increased by motility noise (effective diffusion of the target). Such annuli are then transformed into semicircles, bimodal, and single peak distributions as a result of the dynamics and measurements with efficient decision strategies.

In particular, bimodal distributions can arise from an annulus when the agent moves to a new position, where a new measurement effectively amounts to an intersection of the original annulus with a second annulus corresponding to the new measurement, especially when the agent moves perpendicular to the direction pointing towards the target. The change of annular distributions into bimodal and then unimodal distributions can be rationalized in terms of simple triangulation [20], see SM. More complex likelihood shapes are possible, especially when the speed of the agent is large and the binding rate low.

In this framework, we find that angular information from gradient sensing helps in discriminating the correct target direction from circle-shaped or bimodal likelihoods. In terms of entropy variations, this effect results to be greater than the Laplacian correction, which here helps

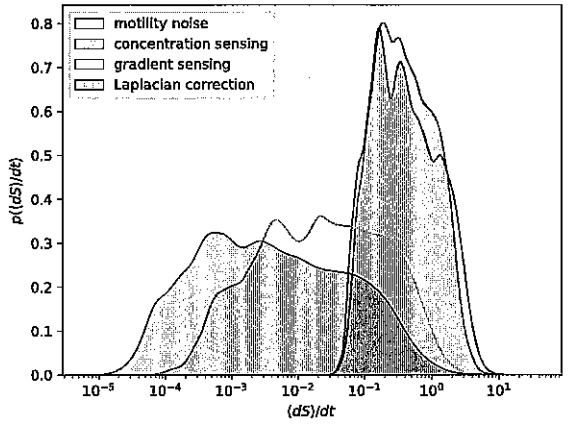

Fig. 1: Distribution of the absolute values of the four terms in the expected entropy variation equation, eq. (2), evaluated in a steady-state dynamics with parameters $a = 0.01$, $\lambda = 2$, $v = 0.01$, $D = 2.5 \cdot 10^{-4}$, $R_{\min} = 0.03$, $R_{\max} = 0.87$. Note that the contribution due to motility noise is positive, while the three contributions from the measurement components are all negative.

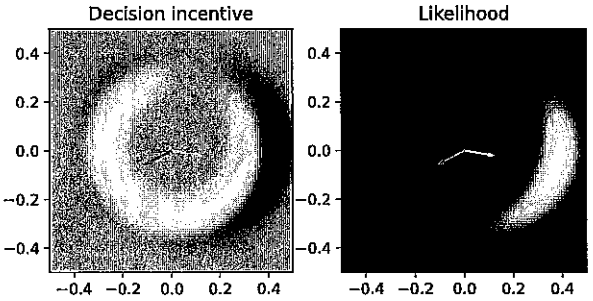

Fig. 2: Typical likelihood configuration p, and corresponding decision incentive q, see eq. (4) and following discussion. Dark blue regions in the decision incentive plot correspond to $q < 0$, *i.e.*, are repulsive. The orange arrow corresponds to the maximum likelihood direction, while the green arrow to the infotaxis direction. The particular configuration was selected from a long steady-state realization with parameters $a = 0.01$, $\lambda = 2$, $v = 0.01$, $D = 2.5 \cdot 10^{-4}$, $R_{\min} = 0.03$, $R_{\max} = 0.87$.

only slightly by increasing the effective variation of the binding rate over spatial distances because of the convexity $\nabla^2 r = \lambda \|\mathbf{x}\|^{-3} > 0$.

Decision making. – We use the average distance from the target at steady state, $\langle\langle\|\mathbf{x}\|\rangle\rangle$, as our metric to evaluate the efficiency of chemotactic search. Then it is natural to introduce an *exploitation* strategy, which minimizes the expected distance variation locally in time, see eq. (18) in SM, that means to move in the direction of $\langle \mathbf{x}/\|\mathbf{x}\| \rangle$. However, for $a = 0$, this strategy forces the likelihood shape to be always symmetric around the centre, and diffusion leads effectively away from the target. Having a finite $a > 0$ breaks the symmetry and improves search performance, see the Supplementary movie `animation1.mp4`.

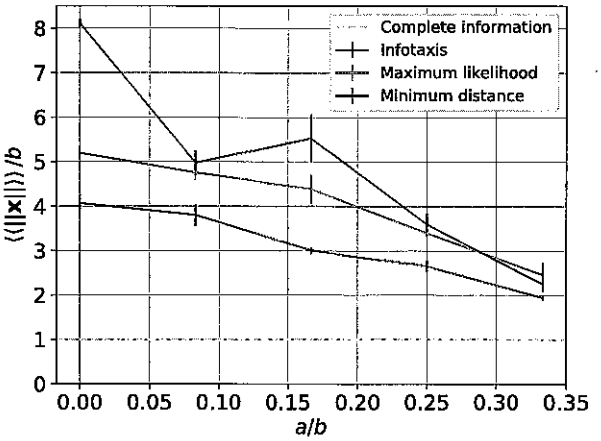

Fig. 3: Search performance increases with agent's size a, in a steady-state dynamics with parameters $\lambda = 2$, $v = 0.01$, $D = 2.5 \cdot 10^{-4}$, $R_{\min} = 0.03$, $R_{\max} = 0.87$. Green is infotaxis, orange is maximum likelihood decision, gray is minimum distance decision, and in gold is the lower bound corresponding to the benchmark case of complete information, which here gives $b = 0.06$. Error bars denote standard error of the mean based on 5 replicas.

A second exploitative strategy is to go for the maximum-likelihood direction, which results to have a slightly larger directional persistence, therefore enabling a more efficient comparison of concentration measurements, see the Supplementary movie `animation2.mp4`. Also here, having a finite $a > 0$ improves the search performance because gradient sensing helps in removing the angular degeneracy of likelihood distributions.

The most efficient local strategy that we consider here, see fig. 3, is the infotaxis strategy [14–16], which prescribes to choose the movement direction by a maximization of the information gain locally in time. Being based on expected entropy variations only, we regard it as a purely *explorative* strategy. We see that the expected entropy variation (eq. (2)) is independent of the velocity, therefore we need to compute and minimize the second entropy variation with respect to the velocity direction. This gives an optimal direction

$$\mathbf{q} = \left\langle (\nabla r) \ln\left(\frac{\langle r \rangle}{r}\right) \right\rangle + a^2 \mathbf{f}[p, r], \tag{4}$$

whose precise expression of the decision correction functional $\mathbf{f}[p, r]$ is given in the SM text. A realization of this dynamics is shown in the Supplementary movie `animation3.mp4`. The impact of corrections on the infotaxis decision making is locally small, but it may be relevant and it will be investigated in future numerical simulations.

Since we are dealing with a binding rate field, which is symmetric around the target, the integrand in the decision integral of eq. (4) points always in the radial direction. We can therefore interpret the magnitude of such integrand,

namely $q = -p\,\|\nabla r\|\,\ln(\langle r\rangle/r) + \mathcal{O}(a^2)$, as a density of *decision incentive*, and simply plot it as a heatmap, see fig. 2. In other words, we rewrite the decision integral as $\mathbf{q} = \int \mathrm{d}\mathbf{x}\, q\, \mathrm{e}^{\mathrm{i}\theta(\mathbf{x})}$, where $\mathrm{e}^{\mathrm{i}\theta(\mathbf{x})}$ is the unit vector in the direction of $\mathbf{x}$. Generally, $q \lesseqgtr 0$ for $r \lesseqgtr \langle r\rangle$. We see that the infotaxis decision integral drives the agent towards regions of high likelihood, high binding rate relative to the expected $\langle r\rangle$, and high binding rate gradient.

As benchmark to compare the performance of the three strategies considered, we study the optimal case of complete target position information. The corresponding dynamics is simply $\mathrm{d}\mathbf{x} = -v(\mathbf{x}/\|\mathbf{x}\|)\mathrm{d}t + \sqrt{2D}\,\mathrm{d}\mathbf{W}$, where $\mathbf{W}$ is a standard 2D Brownian motion. Having circular symmetry, we write the stochastic differential equation for the distance $\|\mathbf{x}\|$ using Ito's Lemma [21], namely $\mathrm{d}\|\mathbf{x}\| = (D/\|\mathbf{x}\| - v)\mathrm{d}t + \sqrt{2D}\,\mathrm{d}W$, and from stationarity $\langle\langle \mathrm{d}\|\mathbf{x}\|\rangle\rangle = 0$ and Jensen's inequality, we obtain $\langle\|\mathbf{x}\|\rangle \geq D/v$. A more accurate estimate of the benchmark lower bound b (which also considers the $R_{\min}, R_{\max}$ boundaries) is found numerically.

For all three decision strategies tested, the mean distance from the target at steady state decreases with agent size, as expected, see fig. 3. For the infotaxis strategy, which generally performs best, this distance $\langle\langle\|\mathbf{x}\|\rangle\rangle$ was about two fold smaller for the largest agent size tested ($a = b/3$) compared to the case of a point-like agent ($a = 0$). In fact, $\langle\langle\|\mathbf{x}\|\rangle\rangle \approx 2b$ for the parameters used, *i.e.*, comparable to the benchmark b of complete information.

two times ~ (ok?)

We did not test larger agent sizes because numerical accuracy decreases with agent size, as the second-order expansion will not be as accurate, and the space discretization becomes too coarse.

Discussion. – In this letter, we derived the information-processing properties of an agent performing chemotaxis based on a likelihood map evolving with a Bayesian updating scheme, which we like to call Bayesian chemotaxis. In particular, we extend the point-like agent model of [14] to a finite-size agent, where the knowledge of the angle of a binding event directly provides directional information.

Like in [14], we made the strong assumption that the measurement model is exact, meaning that the agent knows the binding rate dependence on distance. While this may be unrealistic for chemotaxis of bacteria, where information-processing capabilities are limited, it is certainly feasible to implement in olfactory robots [22].

The leading-order expansion terms in the expected entropy variation, which derive from having a circle of radius a instead of a point-like agent, are interpreted as gradient sensing and a Laplacian correction to concentration sensing. With this expansion, we provide a theoretical framework to predict a typical size of the agent above which the contribution from gradient sensing becomes macroscopically relevant.

We do not attempt at a realistic model of biological chemotaxis; in particular, single cells do not possess the information-processing capabilities to perform stochastic filtering based on likelihood maps. Instead, we establish a theoretical baseline of optimal stochastic filtering in chemotaxis, and we analytically characterize the contribution of gradient sensing therein. Future work will explore approximations of this optimal model, which will allow a more direct comparison with modes of cellular chemotaxis.

In the original ref. [14] it was assumed that concentration sensing is the dominant contribution and effects due to the finite size of the agent can be ignored; this previous assumption is consistent with the findings reported here. Indeed, even a Laplacian correction to concentration sensing is of the same expansion order as that of gradient sensing.

However, we numerically establish that the impact of gradient sensing is larger compared to that of the Laplacian correction if averages are taken over a steady-state chemotaxis dynamics. This is understood as gradient sensing, which is in some way orthogonal to concentration sensing, can remove the degeneracy in the angular marginal likelihood, even without movement.

Finally, we observe that also in our framework infotaxis outperforms other greedy strategies.

* * *

AA was supported by the DFG through FR3429/3-1 to BMF; BMF is supported by the DFG through FR3429/4-1 (Heisenberg grant); AA, MN, and BMF were supported through the Excellence Initiative by the German Federal and State Government (Cluster of Excellence PoL EXC-2068), and the cfaed - Center for Advancing Electronics Dresden.

Data availability statement: The data that support the findings of this study are included within the article (and any supplementary files).

REFERENCES

[1] Bialek W., *Biophysics: Searching for Principles* (Princeton University Press) 2012.
[2] Alon U., Surette M. G., Barkai N. and Leibler S., *Nature*, **397** (1999) 168.
[3] Friedrich B. M. and Jülicher F., *Proc. Natl. Acad. Sci. U.S.A.*, **104** (2007) 13256.
[4] Fujisaki M., Kallianpur G. and Kunita H., *Osaka J. Math.*, **9** (1972) 19.
[5] Rogers L. C. G. and Williams D., *Diffusions, Markov Processes and Martingales: Volume 2, Itô calculus*, Vol. **2** (Cambridge University Press) 2000.
[6] Petris G., Petrone S. and Campagnoli P., *Dynamic linear models* in *Dynamic Linear Models with R* (Springer) 2009, pp. 31–84.
[7] Husain K., Pittayakanchit W., Pattanayak G., Rust M. J. and Murugan A., *Cell Syst.*, **9** (2019) 459.
[8] Mora T. and Nemenman I., *Phys. Rev. Lett.*, **123** (2019) 198101.

Added by us as per EPL standard

[9] Novak M. and Friedrich B. M., *New J. Phys.*, **23** (2021) 043026.
[10] Malaguti G. and Ten Wolde P. R., *eLife*, **10** (2021) e62574.
[11] Siggia E. D. and Vergassola M., *Proc. Natl. Acad. Sci. U.S.A.*, **110** (2013) E3704.
[12] Devreotes P. N. and Zigmond S. H., *Ann. Rev. Cell Biol.*, **4** (1988) 649.
[13] Segall J. E., Block S. M. and Berg H. C., *Proc. Natl. Acad. Sci. U.S.A.*, **83** (1986) 8987.
[14] Vergassola M., Villermaux E. and Shraiman B. I., *Nature*, **445** (2007) 406.
[15] Masson J. B., Bechet M. B. and Vergassola M., *J. Phys. A: Math. Theor.*, **42** (2009) 434009.
[16] Barbieri C., Cocco S. and Monasson R., *EPL*, **94** (2011) 20005.
[17] Hu B., Chen W., Rappel W.-J. and Levine H., *Phys. Rev. Lett.*, **105** (2010) 048104.
[18] Amari S. I., *Information Geometry and its Applications*, Vol. **194** (Springer) 2016.
[19] Ito S. and Sagawa T., *Nat. Commun.*, **6** (2015) 1.
[20] Dobramysl U. and Holcman D., *Phys. Rev. Lett.*, **125** (2020) 148102.
[21] Shreve S. E., *Stochastic Calculus for Finance II: Continuous-Time Models*, Vol. **11** (Springer Science & Business Media) 2004.
[22] Chen X.-X. and Huang J., *Robot. Auton. Syst.*, **112** (2019) 123.