doi: 10.1038/nature05464

Supplementary Materials for "Infotaxis: searching without gradients"

by M. Vergassola, E. Villermaux, B. I. Shraiman

1 Constraints on concentration gradient measurements

Constraints on the chemoreception process due to finite concentration and finite time of measurement were originally investigated by Berg and Purcell [1]. Here, we simply recall their conclusions for the sake of completeness. The rate of encounters between a spherical particle of radius a and molecules diffusing with effective diffusivity D is given by the classical Smoluchowski's [2] expression:

$$J(\boldsymbol{r}) = 4\pi\, D\, a\, c(\boldsymbol{r})\,. \tag{1}$$

For a reliable assessment of the local concentration $c(\boldsymbol{r})$, one collects detection events over time T_{int}. The average number of detection events will then be $J(\boldsymbol{r})\, T_{int}$. Typical fluctuations are of the order of the square root of the mean (the precise calculation of order-unity constants is provided in [1]). The condition for the signal to emerge out of the noise reads then:

$$\sqrt{T_{int} D a c(\boldsymbol{r})} \gg 1\,. \tag{2}$$

Reliable measurements of concentration gradients require the *difference* in counts across the interval of measurement to be above the noise level. The corresponding conditions read:

$$\left(vT_{int}\frac{dc}{dr}\right) T_{int} D a \gg \sqrt{D c T_{int} a}\,; \qquad vT_{int}\frac{d\log c}{dr} \ll 1\,. \tag{3}$$

Here, v is the velocity of the searcher and dc/dr is the concentration gradient. The first inequality in (3) gives the condition that the signal-to-noise ratio for the difference in the number of hits experienced by the searcher across the integration time T_{int} be larger than unity. The second inequality is the requirement of locality, i.e. that the change in concentration across the distance spanned during T_{int} be small compared to the concentration itself. Assuming an exponentially decaying concentration $\exp(-r/\lambda)$, reliable integration time T_{int} scales as $\exp(r/3\lambda)$.

2 Odor encounter rates in a model of turbulent medium

The simplest reasonable statistical description of odor encounters in a turbulent flow is provided by a model where detectable "particles" are emitted by the source at a rate R, have a finite lifetime τ, propagate with isotropic effective diffusivity D and are advected by a mean current or wind $\boldsymbol{V}$ [3, 4]. Other propagation models can be treated by the same tools discussed in the body of the paper and there are no particular restrictions on their nature (see Section 6).

The *mean* stationary concencentration field $c\,(\boldsymbol{r}|\boldsymbol{r}_0)$ generated by a source located at position $\boldsymbol{r}_0$ will satisfy the following advection-diffusion equation

$$0 = V\,\nabla_y\, c\,(\boldsymbol{r}|\boldsymbol{r}_0) + D\Delta c\,(\boldsymbol{r}|\boldsymbol{r}_0) - \frac{1}{\tau} c\,(\boldsymbol{r}|\boldsymbol{r}_0) + R\,\delta\,(\boldsymbol{r} - \boldsymbol{r}_0)\ , \tag{4}$$

where the wind has been taken to blow in the negative y-direction. Note that $\boldsymbol{V}$ is the *mean* velocity and that the instantaneous velocity fluctuates both in direction and amplitude due to the noise term described by the effective diffusivity term D, the sum of both the turbulent diffusivity and the (usually much smaller) molecular diffusivity. Both in two and three dimensions, equation (4) admits an analytical solution. In 2D, the solution reads:

$$c(\boldsymbol{r}|\boldsymbol{r}_0) = \frac{R}{2\pi D} e^{\frac{-(y-y_0)V}{2D}} K_0\left(\frac{|\boldsymbol{r}-\boldsymbol{r}_0|}{\lambda}\right)\ ; \qquad \lambda = \sqrt{\frac{D\tau}{1+\frac{V^2\tau}{4D}}}\ , \tag{5}$$

where K_0 is the modified Bessel function of order zero [5]. A similar expression is derived in three dimensions as:

$$c(\boldsymbol{r}|\boldsymbol{r}_0) = \frac{R}{4\pi D|\boldsymbol{r}-\boldsymbol{r}_0|}\, e^{\frac{-(y-y_0)V}{2D}}\, e^{-\frac{|\boldsymbol{r}-\boldsymbol{r}_0|}{\lambda}}\ , \tag{6}$$

where λ is given by the same expression as in (5).

A spherical object of small linear size a moving into such media will experience a series of encounters at rates $R(\boldsymbol{r}|\boldsymbol{r}_0)$ given by the Smoluchowski's arguments discussed in the previous Section. In three dimensions, the expression for the hit rate given in the Methods of the paper immediately follows from (1) and (6).

In two dimensions, the time of return to a given location for a diffusive particle is known to have a logarithmic divergence. That induces a divergence in the usual arguments employed to derive (1). The logarithmic divergence is regularized by the presence of a finite lifetime of detected particles and (1) takes the form

$$R(\boldsymbol{r}|\boldsymbol{r}_0) = \frac{2\pi D c\,(\boldsymbol{r}|\boldsymbol{r}_0)}{\ln\left(\frac{\lambda}{a}\right)} = \frac{R}{\ln\left(\frac{\lambda}{a}\right)}\, e^{\frac{(y_0-y)V}{2D}} K_0\left(\frac{|\boldsymbol{r}-\boldsymbol{r}_0|}{\lambda}\right)\ , \tag{7}$$

for the concentration profile given by (5).

3 Search time and the entropy of the probability distribution of source location

The aim of this Section is to derive a lower bound on the expected search time T as a function of the entropy S of the probability distribution of source location.

The definition of Shannon's entropy S for a (discrete) random variable X reads [6, 7]:

$$S = -\sum_{j=1}^{N} p_j \ln p_j \,, \tag{8}$$

where X takes N possible values x_j ($j = 1, \ldots, N$) with probabilities p_j normalized to unity: $\sum_{j=1}^{N} p_j = 1$. It is easy to show [6, 7] that $S \geq 0$ and that $S = 0$ if and only if there is no uncertainty on X, i.e. one of the events, say x_{j^*}, occurs with certainty $p_{j^*} = 1$ and all others are impossible $p_{j \neq j^*} = 0$. The maximum value of S is $\ln N$, corresponding to the situation where all possible outcomes of X have equal probability, i.e. the situation of maximum uncertainty. The definition (8) was in fact originally derived by C. Shannon as the only expression satisfying certain properties stemming from intuitive notions of uncertainty [6]. Considering random variables defined on the same ensemble, those having a larger entropy are more unpredictable while those with a smaller entropy have a probability distribution more concentrated and localized in a few points. Reducing the entropy of the probability distribution $P_t(\boldsymbol{r}_0)$ for the source location $\boldsymbol{r}_0$ thus amounts to having less uncertainty, i.e. more information, on its location. This is the rationale for quantifying the rate of acquisition of information on the location of the odor source by the rate of reduction of the entropy of its estimated probability distribution, as we have done in the body of the paper.

To derive the lower bound on the search time, let us now consider an ensemble of probability distributions for the location of the source of odors, all having fixed entropy S, and compute the expected minimal time to locate the source. In obtaining a lower bound we can relax continuity constraints and allow the searcher to jump, like a grasshopper, from one site to any other. For simplicity, we consider a spatial lattice with unit mesh size, set by the linear dimension of the searcher which one can choose as a unit of length. Using a lattice is justified by the fact that short-distance source recognition is often influenced by multiple sensory modalities, including tactile and visual.

Denoting by p_j the probability to find the source at the j-th (in time) point visited, the expected search time reads $T = \sum_j j\, p_j$. This expression stems from the following simple calculation. Upon visiting the first point, either the source is found (probability p_1) and the search process stops or we know that the source is elsewhere and we rescale all other probabilities, that is $p_{j \neq 1} \mapsto p_{j \neq 1}/(1 - p_1)$. Similar rescalings take place if successive visits happen to be unsuccessful. It follows that the average search time reads:

$T = 1\,p_1 + (1-p_1)\left\{2\frac{p_2}{(1-p_1)} + \left(1 - \frac{p_2}{1-p_1}\right)\left[3\frac{p_3}{(1-p_1-p_2)} + \ldots\right]\right\}$. Simplifying the various terms, one can check that the aforementioned expression for T is recovered. The best possible option to minimize T is to visit points in the order of decreasing probability. The desired lower bound is obtained by minimization with respect to all possible distributions p_j with fixed entropy S. Thus, we minimize

$$T' = \sum_j j\,p_j + \alpha\left(\sum_j p_j - 1\right) + \beta^{-1}\left(\sum_j p_j \log p_j + S\right), \tag{9}$$

where α and β are the Lagrange multipliers enforcing the normalization of probability and the constraint on the entropy. The probability distribution corresponding to the minimum of (9) has the Gibbs form: $p_j \propto \exp(-\beta j)$. If boundaries can be neglected, i.e. $S \ll \log N$ where N is the total number of points, the inverse temperature $\beta \gg 1/N$. The relation between the entropy and the search time then reads: $S = T\log T - (T-1)\log(T-1)$. In the interesting cases where S and T are both large compared to unity, this reduces to:

$$T = \sum_j jp_j \simeq e^{S-1} \equiv \overline{n}/e\,. \tag{10}$$

In other words, the search time is bounded in terms of the effective number of points $\overline{n} = e^S$, which is determined by the entropy of the distribution. Note incidentally that $T = \sum j\,p_j$ is the expected code length for a message composed of words having probabilities p_j [7]. The alphabet here is quite special as it is composed of a single letter and different words can be discriminated by their length only. That is the reason for the linear (rather than logarithmic) dependence on $\overline{n}$.

The bound (10) is a lower bound. Its purpose is to relate the "complexity" of the search (as measured by the average time that it takes) to the entropy of the distribution. The reduction of entropy of the estimated distribution along the search trajectory then becomes a *necessary, although not sufficient*, condition for efficient searching. The actual average time of the search depends on how well the estimated distribution approximates the true one. (Note that the "true distribution" here means a posterior distribution, i.e. a distribution of sources that could have generated the recorded set of odor encounters.) The ultimate proof of efficiency of the algorithm would require an upper bound on the actual search time. Upper bounds are generally dominated by "worst" case scenarios, that do not represent more typical cases. We are not presently able to provide a rigorous bound on the efficiency of the infotaxis algorithm and rely on numerical simulations for demonstration.

4 Exploitation *vs* exploration and comparison among search strategies

The tradeoff between exploitation and exploration discussed in the body of the paper is a classical issue that has emerged in reinforcement learning [8]. The latter is a wide class of methods and techniques devoted to the general problem of developing an "effective policy", that is choosing the probability of actions to be taken by an agent as a function of the state of its environment. Effectiveness is measured by a numerical reward and the agents' goal is to maximize the global value of the reward, cumulated over a series of actions. Situations where these ideas have found fruitful applications range from strategies in games such as backgammon, chess, etc., to processes of allocation of resources and system control.

Emphasis on the interactions with an *uncertain* environment is one of the essential features of reinforcement learning: the agent is not offered an instructive training set of good choices to be made and has to discover by himself the value of different possible actions by a trial-and-error evaluative process [8]. Various methods have been developed to reach this goal, *viz.* dynamic programming, Monte Carlo methods, temporal-difference updating, eligibility traces, and we refer to [8] for a detailed presentation. For Markov problems, where future states and rewards depend only on the action and the state of the environment at the actual time, exact recursive equations (the so-called Bellman equations) are available and one can get a hold on the convergence process to the optimal policy [9].

All methods of reinforcement learning feature two competing tendencies. On one hand, to maximize their reward, agents might choose actions already tried in the past and greedily favor the action currently estimated to be most rewarding in that particular situation. This is the exploitive attitude which tends to capitalize on available information. The name "reinforcement learning" stems precisely from the intuitive notion that, throughout the evaluation process, the policy should evolve so as give higher chance to actions yielding large rewards and penalize those appearing unfavourable. On the other hand, different actions must be tried and risks taken to discover valuable actions, possibly superior to the action estimated to be the best on the basis of current information. Pure exploration would correspond to trying all possible actions without any particular preference. The general lesson learnt in reinforcement learning [8] is that neither pure exploitation nor pure exploration are effective: they should be blended and, for effective learning, an optimal tradeoff must be found. This has been proven with rigor in some simple situations, such as the n-bandit problem, and has been demonstrated numerically in a range of other cases [9, 8].

As we have already discussed in the article, the infotaxis policy indeed ensures a tradeoff between exploitation and exploration. Considering the expression of the expected

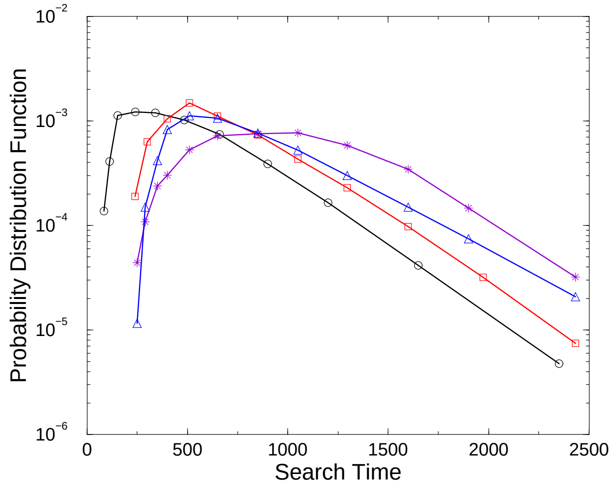

Figure 1: A comparison among the probability distribution function of search times for four different strategies. (i) black: infotaxis; (ii) blue: the greedy strategy where the searcher chooses local moves towards locations of highest estimated probability of source location; (iii) purple: the strategy locally maximizing the probability of detection; (iv) red: the strategy complementary to (ii), where the second term in the right-hand side of (11) is locally maximized. Simulations were performed for a diffusivity $D = 1$, emission rate $R = 1$, without any mean wind, lifetime $T = 400$, on a grid 256^2. The typical length traveled by particles is 20 and the starting position is at distance 40 from the source, along one of the axes.

variation of entropy

$$\overline{\Delta S}\left(\boldsymbol{r} \mapsto \boldsymbol{r}_j\right) = P_t(\boldsymbol{r}_j)\ [-S] + [1 - P_t\left(\boldsymbol{r}_j\right)]\left[\rho_0\left(\boldsymbol{r}_j\right)\Delta S_0 + \rho_1\left(\boldsymbol{r}_j\right)\Delta S_1 + \ldots\right], \quad (11)$$

it is indeed clear that the searcher (agent) will not generally choose the move (action) that gives the maximum estimated reward (probability to find the source). This would correspond to locally maximizing only the first term on the right hand side in (11). Our purpose here will be to give a concrete demonstration of the efficiency of the tradeoff quantified in (11).

Specifically, in Fig. S1 we compare search time statistics for different search strategies: (i) Infotaxis, locally maximizing the expected reduction of entropy as quantified by (11); (ii) The greedy exploitive strategy locally maximizing the currently estimated probability to find the source $P_t(\boldsymbol{r}_j)$; (iii) The strategy choosing the local move which maximizes the expected rate of detection $h(\boldsymbol{r}_j) \equiv \int P_t(\boldsymbol{r}_0)\, R(\boldsymbol{r}_j|\boldsymbol{r}_0)\, d\boldsymbol{r}_0$, with $R(\boldsymbol{r}|\boldsymbol{r}_0)$ denoting the mean rate of hits at position $\boldsymbol{r}$ if the source is located in $\boldsymbol{r}_0$, as defined in the body of the paper; (iv) The strategy complementary to (ii), where only the second-term on the right hand side of (11) is locally maximized. Note that (iii) and (iv) are both biased toward exploration in the sense that none of them features explicitly a greedy term in the probability $P_t(\boldsymbol{r}_j)$ to detect the source. Both will of course still tend to favor approaching the source (this is a means to detect more particles) yet the process occurs in a more distributed way, being driven by the hit rate which involves a space-integral.

Fig. S1 clearly shows that infotaxis is more efficient than both (iii) and (iv). Another relevant remark is the importance of the information brought by detections and not just their sheer number. This is illustrated by the faster searches ensured by (iv) with respect to (iii), where only the occurrence of detections matters. Note also that (iii) has a tendency to feature long periods where the searcher remains still. We have found that the process of search is more rapid if a stochastic policy is employed, where the move towards the point $\boldsymbol{r}_j$ is selected with a probability $\propto exp\left(\beta h(\boldsymbol{r}_j)\right)$ and the parameter β is linearly decreased as the waiting time increases. The histogram for strategy (iii) in Fig. S1 has been obtained using this stochastic policy.

Finally, let us comment on the greedy strategy (ii), where one chooses moves toward the neighboring point(s) of maximum estimated probability $P_t(\boldsymbol{r}_j)$ to detect the source. Simulations show that this strategy is extremely unstable and easily leads to the loss of the searcher, by which we mean searcher reaching the boundary of the domain. Indeed, the searcher reached the boundary of the box 256^2 in 79% of the trials (to be contrasted with a fraction $\simeq 10^{-4}$ for methods (i), (iii) and (iv)). The mechanism leading to this behavior can be intuitively understood as follows. Let us consider a situation where the searcher is far from the source and has not received particles for some time. When particles are not received, the probabilities $P_t(\boldsymbol{r})$ are depleted in the neighborhood of the searcher, while they increase in the far region. If one compares neighboring points in the directions perpendicular and along the searcher line of motion, it is clear that the former are closer to the points where no particles were detected. The searcher will then tend to keep moving in the same direction. If the direction happens to be away from the source, the probability of detection will be further reducing and the whole process continues, leading to the loss of the searcher. Furthermore, Fig. S1 indicates that, even for those runs where the searcher is not lost, there is no particular advantage in terms of the speed of search and infotaxis is significantly faster.

One might consider alleviating the instability problem of the greedy strategy (ii) by considering multi-step algorithms (where trajectories of more than one step are evaluated) or stochastic strategies as those described previously. Still, the intrinsic instability of the strategy holds and might be traced back to the absence of any sense of the uncertainty and of the risks in the actions taken. The drawback is absent in infotaxis where the likelihood term in (11) is combined with terms sensing the far field via the hit rate. This prevents the searcher from being led astray when only scant information on the location of the source is available.

5 Learning about the medium and the source

The properties of propagation of the odor patches through the medium and their frequency of emission have hitherto been considered as known. If estimates are available, one can employ them in the model of the medium used during the searches. If no *a priori* information is available, the properties of the medium must be learnt before employing the robot on a regular basis. For example, for a sniffer to be employed for drug detection or to demine a field, one would set up a training set of runs and learn the parameters that will then be employed in extended series of searches. One approach towards learning the parameters of the medium and the source is described below. It shows that no particular stiffness in the estimation of the parameters is expected.

Robots designed to detect odors – sniffers – are typically endowed with sensors measuring wind velocity that provide them with relatively reliable estimates of *time averaged* quantities such as the mean wind and the root-mean-square V_{rms} of velocity fluctuations [10]. The sporadic nature of the signal detected by sniffers comes from the fact that odors are typically transported by strongly turbulent fields that efficiently mix and make the patches spread out and decay in intensity with time. Parameters of turbulent flows are notoriously hard to estimate but our major advantage here is that infotactic searches are *not* maximum likelihood strategies and are more robust to errors and incomplete information. Sophisticated modeling of the turbulent medium is therefore not crucial (and it is hard to imagine that birds and moths rely on fine-tuning of parameters in their searches).

The most effective approach to parameter estimation for the case of sniffers appears to be the following. In the absence of a precise value of the parameters of the medium, we can start searches with rough estimates which ensure that the rate function $R(\boldsymbol{r}|\boldsymbol{r}_0)$ is shallower than in reality. The purpose is that, in the absence of precise information, broad estimators of the rate function will avoid wrong estimates that could drive the searcher astray, e.g. to the boundaries of the box where the search is taking place. The searcher will thus typically arrive to the source, albeit in times much longer than with the correct estimation of $R(\boldsymbol{r}|\boldsymbol{r}_0)$. Once arrived to the source, and thus knowing the source position $\boldsymbol{r}_0$, the searcher can use the trace $\mathcal{T}$ of detections along its trajectory to estimate the parameters of the medium and the source.

For example, let us consider the problem of learning parameters for the model described in Section 2. Suppose that the mean velocity $\boldsymbol{V}$ is accessible via the sensors and, just for the sake of the argument, that the turbulent diffusivity D, the rms velocity fluctuations V_{rms} and the particle lifetime τ are related as $D = V_{rms}^2\tau$, a classical dimensional estimate of turbulent diffusivity [11] (see [12, 13] for more specific information in jet flows). We shall take the rate of emission of the source R and the effective diffusivity D as independent variables and try to learn the parameters of a medium having unit values for D, R and V and a turbulence level $V_{rms}/V = 20\%$, a typical value in turbulent jets. Choosing a shallow rate function $R(\boldsymbol{r}|\boldsymbol{r}_0)$ as initial guess corresponds to safe over-estimates of D and $1/R$, e.g. by two orders of magnitude $D_{est} = 1/R_{est} = 100$. Starting the searcher with these very rough estimates and from initial distances 50, i.e. twice the advection length $V\tau = 25$, slows down the search by about a factor 7. Still, more than half of the searches arrive to the source without ever touching the boundaries of the box (512×512). The resulting of the log-likelihoods for the trace $\mathcal{T}$ of detections along the corresponding trajectories have their optimum located at the real values $D = R = 1$ and in all cases that we have analyzed have been found to be convex. The parameters could then be correctly estimated by using any standard minimization algorithms, such as simplex method, simulated annealing or conjugate gradient.

The same learning procedure was employed for the experimental turbulent data on mixing flow [12] utilized in Fig. 3 in the paper. All free parameters in the expression (7) were estimated as described above and employed for the search. The estimated values of the mean velocity $\boldsymbol{V}$ and the injection rate R are very close to the real experimental values (5% being the maximum error).

6 Models of the turbulent medium

As stressed throughout this work, our major interest is in dilute conditions and devising a search strategy capable of dealing with wide voids where no cues are detected. That was the rationale for using a model with independent detections without entering into detailed modeling of the detection process. Fig. 3 in the paper indicates that this is indeed a sensible thing to do but also that there are obvious quantitative improvements one could make. The aim of this Section is to present a brief discussion on this issue. The important point to be stressed is that infotaxis is quite a general strategy and might be employed with any model of the medium where particles are propagated. What we have demonstrated on the specific examples in the previous Sections and in the article itself is that the strategy permits to deal with an uncertain environment. We expect that this property holds in general and this constitutes the crucial advantage of the method.

The first aspect where modeling of the medium might be improved is in accounting for the correlations among detection events. Detailed models of plumes have already been discussed in the literature [14, 15, 16, 17, 18, 19] and we shall therefore content ourselves with a brief discussion of a simple way to account for time-correlations among the detection events. The model is based on the following likelihood of experiencing a trace $\mathcal{T}_t$ of correlated odor encounters:

$$\mathcal{L}_{\boldsymbol{r}_0}(\mathcal{T}_t) = e^{-\sum_i \left[\int_{V_i} R(\boldsymbol{r}(t')|\boldsymbol{r}_0)\,dt' + \int_{D_i} Q(\boldsymbol{r}(t')|\boldsymbol{r}_0)\,dt'\right]} \prod_{i=1}^{H} R(\boldsymbol{r}(t_i)|\boldsymbol{r}_0) \prod_{j=1}^{H'} Q(\boldsymbol{r}(t_j)|\boldsymbol{r}_0)\,. \quad (12)$$

Here, the V_i's (D_i's) are the time intervals of absence (presence) of detections, H and the t_i's are the number and times of transition from no-detection to detection intervals ($V_i \to D_i$) and, finally, H' and t'_i are the number and times of the opposite transitions, from detection to no-detection. The posterior probability distribution $P_t\,(\boldsymbol{r}_0)$ is constructed as done previously, by using the Bayes formula: $P_t\,(\boldsymbol{r}_0) = \mathcal{L}_{\boldsymbol{r}_0}(\mathcal{T}_t)/\int \mathcal{L}_{\boldsymbol{x}}(\mathcal{T}_t)\,d\boldsymbol{x}$. The expression (12) amounts to assuming that patches of odors have finite extensions and thus the searcher will spend finite amounts of time within them. The extension of patches as a function of the distance to the source is controlled by the function $Q\,(\boldsymbol{r}|\boldsymbol{r}_0)$.

In the simplest possible setting where the function $Q(\boldsymbol{r}|\boldsymbol{r}_0)$ is taken to be constant in space, the expression for the posterior probability distribution simplifies as:

$$P_t\,(\boldsymbol{r}_0) = \frac{\exp\left[-\sum_i \int_{V_i} R\,(\boldsymbol{r}(t')|\boldsymbol{r}_0)\;dt'\right] \prod_{i=1}^{H} R(\boldsymbol{r}(t_i)|\boldsymbol{r}_0)}{\int \exp\left[-\sum_i \int_{V_i} R\,(\boldsymbol{r}(t')|\boldsymbol{x})\;dt'\right] \prod_{i=1}^{H} R(\boldsymbol{r}(t_i)|\boldsymbol{x})\,d\boldsymbol{x}}\,. \quad (13)$$

We thus recover the same structure as for independent hits but now consecutive detections are not overcounted. Note that there are no additional parameters to be estimated. Variations along the same lines might be considered by introducing additional parameters, for example treating the time distance to a previous hit in a patch as a free parameter and estimating it from the data.

The second point where the model could be made more realistic involves including, whenever available, the information on large-scale patterns of the flow. This for example is the case for the model of meandering plumes commonly employed to model the propagating turbulent flow (see, e.g., [20]) or the near-shore experimental oceanic flow recently employed for demonstrations of chemical plume tracing [16, 17, 19]. Both the mean and the noisy diffusive components of the velocity field are now time and space-dependent and the calculation of the probability of detection is thus more involved but still manageable, as shown in [19]. Probability maps were constructed in [19] but not coupled to the motion of the robot. The infotaxis strategy specifically implements this coupling and we have shown that this is realized in a robust and effective way. As we have previously mentioned, there is no particular limitation as to the model of the turbulent medium and the infotaxis strategy might be employed quite generally. It will then be interesting future work to implement this strategy on a robotic vehicle and to test its performances in atmospheric and oceanic flows.

References

[1] Berg, H.C., & Purcell, E.M. Physics of Chemoreception, *Biophys. Journal*, **20**, 193-219, (1977).

[2] Smoluchowski, M.V.Versuch einer mathematischen theorie des koagulationslinetic kolloider Lösungen. *Z. Phys. Chem.* **92**, 129-168 (1917).

[3] Shraiman, B.I. & Siggia, E.D. Scalar turbulence. *Nature*, **405**, 639-646 (2000).

[4] Falkovich, G., Gawędzki, K. & Vergassola, M. Particles and fields in fluid turbulence. *Rev. of Modern Physics*, **73** 913-975 (2001).

[5] Gradshteyn, I.S. & Ryzhik, I.M. Tables of Integrals, Series and Products. Acad. Press

[6] Shannon, C.E. A Mathematical Theory of Communication. *Bell Sys. Tech. Journ.* **27**, 379-423, 623-656, 1948.

[7] Cover, T.M. & Thomas, J.A. Elements of Information Theory. Wiley Series in Telecommunication, (1991).

[8] Sutton, R.S. & Barto, A.G.. Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning). MIT Press, Cambridge, MA (1998).

[9] Bertsekas, D.P. & Tsitsiklis, J.N. Neuro-Dynamic Programming. Athena Scientific, Belmont, MA, (1996).

[10] Russel, R.A. Odor Detection by Mobile Robots. World Scientific, Singapore, (1999).

[11] Tennekes, H. & Lumley, J.L. A First Course in Turbulence. The MIT Press, (1972).

[12] Villermaux, E.& Duplat, J. Mixing as an aggregation process. *Phys. Rev. Lett.* **91**, Art. No. 18 4501 (2003).

[13] Duplat, J., Innocenti, C. & Villermaux, E. Mixing by random stirring. *J. Fluid Mechanics*, to appear, (2006).

[14] Belanger, J.H. & Arbas, E. Behavioral strategies underlying pheromone-modulated flights in moths: Lessons from simulation studies. *J. of Comparative Physiology A* **183**, 345-360 (1998).

[15] Li, W., Farrell, J.A. & Cardé, R.T. Tracking of fluid-advected odor plumes: strategies inspired by insect orientation to pheromone. *Adaptive Behavior*, **9**, 143-170 (2001).

[16] Farrell, J.A., Pang, S., Li, W. Plume Mapping via Hidden Markov Methods. *IEEE Trans. SMC-B*, **33**, 850-863 (2003).

[17] Farrell, J.A., Pang, S., Li, W. Chemical plume tracing via an autonomous underwater vehicle *IEEE J. of Oceanic Eng. 30* 428-442 (2005).

[18] Ishida, H., Nakayama, G., Nakamoto, T., Moriizumi, T. Controlling a gas/odor plume-tracking robot based on transient responses of gas sensors. *IEEE Sensors J.* **5**, 537-545 (2005).

[19] Pang, S., Farrell, J.A. Chemical plume source localization. *IEEE Trans. SMC-B*, to appear in the December 2006 isssue.

[20] Murlis, J., Elkinton, J.S. & Cardé, R.T. Odor plumes and how insects use them. *Ann. Rev. of Entomology* **37**, 505-532 (1992).