Physica D 239 (2010) 477–484

Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

Optimizing chemotaxis by measuring unbound–bound transitions

Duncan Mortimer [a], Peter Dayan [d], Kevin Burrage [b,e], Geoffrey J. Goodhill [a,c,*]

[a] Queensland Brain Institute, The University of Queensland, St. Lucia, QLD 4072, Australia
[b] Institute for Molecular Bioscience, The University of Queensland, St. Lucia, QLD 4072, Australia
[c] School of Mathematics and Physics, The University of Queensland, St. Lucia, QLD 4072, Australia
[d] Gatsby Computational Neuroscience Unit, UCL, London, UK
[e] COMLAB, University of Oxford, Oxford, UK

ARTICLE INFO

Article history:
Available online 23 September 2009

Keywords:
Axon guidance
Bayesian model
Growth cone

ABSTRACT

The development of the nervous system requires nerve fibres to be guided accurately over long distances in order to make correct connections between neurons. Molecular gradients help to direct these growing fibres, by a process known as chemotaxis. However, this requires the accurate measurement of concentration differences by chemoreceptors. Here, we ask how the signals from a set of chemoreceptors interacting with a concentration gradient can best be used to determine the direction of this gradient. Prior models of chemotaxis have typically assumed that the chemoreceptors produce signals reflecting just the time-averaged binding state of those receptors. In this article, we show that in fact the optimal chemotaxis performance can be achieved when, in addition, each receptor also signals the number of unbound-to-bound transitions it experiences within the observation period. Furthermore, we show that this leads to an effective halving of the observation period required for a given level of performance. We also demonstrate that the degradation in performance observed to occur at high concentrations experimentally is likely to result not from noise intrinsic to receptor binding, but rather from noise in subsequent downstream signalling.

1. Introduction

The human nervous system is an incredibly complex structure: roughly 10^{11} neuronal cells, each 'wired' to on the order of 10^3–10^4 others. Furthermore, this structure constructs itself during development. Understanding how this is achieved is crucial for improving the treatment and diagnosis of neurological disorders, for developing our knowledge of biological computation in general, and perhaps for improving engineering techniques [1,2]. Wiring the nervous system requires the precise guidance of nerve fibres (axons) to make connections with their appropriate partner cells [3–6]. Steering a growing axon is the principal responsibility of the growth cone — a complex structure forming the tip of the axon, which has both sensory and motor functions [7]. Growth cones respond to chemical, electrical and mechanical cues in their immediate environment and transduce these into directed axonal growth [5].

Extracellular chemical gradients form an important class of such guidance cues [5]. To respond to a chemical gradient, a growth cone must be able to estimate how the external concentration varies across its spatial extent [8]. However, as with any sensory process, making such an estimate is subject to noise that potentially obscures any systematic concentration variation due to an external gradient [9–16]. More specifically, two critical issues that limit the accuracy with which the gradient can be estimated are as follows. The first is spatial uncertainty: a growth cone probes its environment with specialized chemoreceptors, which transduce external cues into intracellular signals — however, the position of a receptor on the surface of the growth cone must be estimated from the signals it produces, reducing its reliability. The second critical issue is temporal uncertainty associated with stochastic binding and unbinding of receptors. In other work, we have considered the effects of spatial uncertainty [17]; here, we focus on the temporal issues.

The influence of noise on chemotaxis has been studied in a number of situations other than the guidance of axons. In particular, the slime-mold *Dictyostelium discoideum* and leukocytes have been the subject of theoretical attention along these lines for a number of years [9,14,15,18–20]. A common assumption in these studies is that it is the time-averaged occupancy of receptors that is the relevant signal used by the cell to estimate the concentration. However, whether this is in fact the optimum information for receptors to 'pass on' for effective gradient detection has not been established.

* Corresponding author at: Queensland Brain Institute, The University of Queensland, St. Lucia, QLD 4072, Australia.
E-mail address: g.goodhill@uq.edu.au (G.J. Goodhill).

0167-2789/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.physd.2009.09.009

In this article, we explicitly model ligand–receptor binding over time, and determine the quantities of most value to a growth cone to extract from its receptors. We find that, in contrast to the typical assumption of receptor occupancy made in the literature, gradient direction can be better measured by estimating the length of time for which a receptor remains unbound, on average, before becoming bound: a calculation requiring more information from the receptors than simply their time-averaged occupancy. In particular, we show that when the chemoreceptors report how often they transition from unbound to bound, as well as their time-averaged binding state, the observation time is effectively doubled. These results give new insight into how fundamental physical limits constrain the ability of growth cones, and other small sensors, to measure external chemical gradients.

2. A general one-dimensional model for gradient sensing

We imagine that the growth cone senses its environment through a collection of spatially distributed "measurement devices". For example, these might represent chemoreceptors or clusters of chemoreceptors. Each of these devices interacts with its local environment – described entirely by the local average concentration of guidance cue, C – and produces an observation O with probability $P(O|C)$. Suppose the growth cone has a collection of such devices, located at positions $\vec{r} = (x_1, x_2, \ldots, x_n)$, and is immersed in a concentration field with spatial variation $C(x) = C(0) \times (1 + \mu x)$; then the joint probability of making the set of observations $\vec{O} = (O_1, O_2, \ldots, O_n)$ is

$$P(\vec{O}|C(0), \mu) = \prod_i P(O_i|C(0) \times (1 + \mu x_i)). \tag{1}$$

Here, μ is the relative change in concentration across the spatial extent of the growth cone.

Assuming shallow gradients (i.e. small μ), and expressing the concentration in dimensionless units, replacing $C(0)$ with $\gamma = C(0)/K_d$ (when we explicitly consider chemoreceptors, we measure the concentration in multiples of the receptor dissociation constant K_d), we can approximate this with

$$P(\vec{O}|\gamma, \mu) = P(\vec{O}|\gamma) \exp\left[\sum_i \log P(O_i|\gamma \times (1 + \mu x_i))\right] \tag{2}$$

$$\approx P(\vec{O}|\gamma)\left[1 + \mu \sum_i x_i \gamma \frac{\mathrm{d}}{\mathrm{d}\gamma} \log P(O_i|\gamma)\right]. \tag{3}$$

Applying Bayes rule and marginalizing over γ, we have

$$P(\mu|\vec{O}) \propto P(\mu) + \mu P(\mu) \sum_i x_i \gamma \frac{\mathrm{d}}{\mathrm{d}\gamma} \log P(O_i|\gamma), \tag{4}$$

where, to simplify the discussion, we have assumed that the growth cone's posterior uncertainty concerning γ is negligible.

By obtaining an expression for $P(O_i|\gamma)$, we can therefore find $P(\mu|\vec{O})$, the posterior probability that the gradient strength is μ given the observations $\vec{O}$. In the case where $P(\mu)$ is symmetric about 0, the growth cone can decide between turning left or right by comparing

$$\chi(\vec{O}) = \sum_i x_i \gamma \frac{\mathrm{d}}{\mathrm{d}\gamma} \log P(O_i|\gamma) \tag{5}$$

to zero: if χ is greater than zero then the gradient most likely points right, while if χ is less than zero, it most likely points left. We define χ_i as the coefficient of x_i in this sum; i.e.,

$$\chi_i = \gamma \frac{\mathrm{d}}{\mathrm{d}\gamma} \log P(O_i|\gamma). \tag{6}$$

3. Gradient sensing with complete information about receptor binding

We now derive an expression for the probability of observing a particular sequence of receptor binding states given a particular background concentration. We then substitute this into our formula for optimal gradient detection, Eq. (5), and give an intuitive explanation for the results.

We model receptor binding as a continuous time, two-state Markov process: one state represents the receptor in its bound configuration, while the other represents the receptor in its unbound configuration. The transition rates between the states are denoted r_- (for the transition from a bound state to an unbound state) and $r_+ = Ck_+$ (for a transition from an unbound to a bound state). Here, C is the local concentration of ligand at the receptor. Following the Michaelis–Menten model of receptor dynamics, the dissociation constant of the receptor is $K_d = r_-/k_+$. Writing $\gamma = C/K_d$ for the dimensionless concentration obtained by scaling C by K_d, we can also express the unbound-to-bound transition rate as $r_+ = \gamma r_-$.

As illustrated in Fig. 1, each observation O_i consists of a function $b_i(t)$, describing the binding state of the ith receptor for each time t during the interval for which the growth cone is making its observation, $[0, T]$. Since the transitions are assumed to be instantaneous, specifying $b(t)$ is equivalent to specifying a sequence of times $\vec{t}$ at which a receptor changes binding states, along with the initial state of the receptor $b \equiv b(t = 0)$; thus each observation can be described by the pair $(\vec{t}, b)$, where $\vec{t}$ is of some arbitrary length n. For example, $\vec{t}$ may be of length $n = 0$, describing the case in which no transitions occur in the interval $[0, T]$. For ease of presentation, we will simply refer to $\vec{t}$ and n in the following, it being understood that neither are fixed quantities.

To obtain an expression for $P(\vec{t}, b|\gamma, T)$ – the probability of observing binding state transitions at times $\vec{t}$ with initial binding state b, over a time T at background concentration γ – we find $P(\vec{t}|\gamma, T, b)$, the probability of observing a series of transition times $\vec{t}$ within a time T assuming that the receptor began in a state b, and then multiply by the probability of observing the receptor in state b at time $t = 0$:

$$P(\vec{t}, b|\gamma, T) = P(\vec{t}|\gamma, T, b)P(b|\gamma). \tag{7}$$

The likelihood of observing a particular trajectory of binding states $b(t) \equiv \{b, (t_1, \ldots, t_n)\}$ over a period of time T, given background concentration γ, can be obtained approximately by dividing $[0, T]$ into small intervals of length δt, which are small enough that the probability of more than one transition occurring within one interval is negligible:

$$\begin{aligned}\delta t^n P(t_1, t_2, \ldots, t_n|\gamma, b, T) &\approx \exp(-r_0 t_1) r_0 \delta t \times \exp(-r_1(t_2 - t_1)) r_1 \delta t \\ &\quad \times \exp(-r_{n-1}(t_n - t_{n-1})) r_{n-1} \delta t \exp(-r_n(T - t_n)),\end{aligned} \tag{8}$$

where r_m represents the appropriate transition rate for the mth transition; i.e.,

$$r_m = \begin{cases} \gamma r_- & \text{if } m + b \text{ is even} \\ r_- & \text{if } m + b \text{ is odd.} \end{cases} \tag{9}$$

Dividing both sides by δt^n and taking the limit as δt goes to zero, we obtain the exact result:

$$\begin{aligned}P(\vec{t}|\gamma, b, T) &= r_0 r_1 \ldots r_{n-1} \\ &\quad \times \exp(-r_0 t_1 - r_1(t_2 - t_1) - \cdots - r_n(T - t_n)) \\ &= r_-^{n_{bu}} (\gamma r_-)^{n_{ub}} \exp(-r_- T_b - \gamma r_- T_u),\end{aligned} \tag{10}$$

where n_{bu} and n_{ub} are the number of bound-to-unbound and unbound-to-bound transitions, respectively, and T_b and T_u are the

We can now use this to determine χ in Eq. (5), given $O_i = \{n^{(i)}, T_u^{(i)}, b^{(i)}\}$:

$$\chi_i = \gamma \frac{\mathrm{d}}{\mathrm{d}\gamma} \log P(T_u^{(i)}, n^{(i)}, b^{(i)}|\gamma, T) = \lfloor n^{(i)}/2 \rfloor + b^{(i)} - \frac{\gamma}{1+\gamma} - \gamma r_- T_u^{(i)} + \begin{cases} 1 - b^{(i)} & \text{if } n^{(i)} \text{ odd} \\ 0 & \text{if } n^{(i)} \text{ even.} \end{cases} \tag{21}$$

Since n_{ub} is just $\lfloor n/2 \rfloor$ if n is even, or $\lfloor n/2 \rfloor + 1 - b$ if n is odd, this is just

$$\chi_i = b^{(i)} - \frac{\gamma}{1+\gamma} + n_{ub}^{(i)} - \gamma r_- T_u^{(i)}. \tag{22}$$

This expression can be decomposed into two terms. The first characterizes the information provided by the initial state of the receptor:

$$\chi_{init}^{(i)} = b^{(i)} - \frac{\gamma}{1+\gamma}. \tag{23}$$

The second term corresponds to information obtained by following the receptor's state over time:

$$\chi_t^{(i)} = n_{ub}^{(i)} - \gamma r_- T_u^{(i)}. \tag{24}$$

This term is essentially equivalent to comparing an estimate of the local concentration at the ith receptor to an estimate of the concentration at the center of the growth cone. More specifically, given n_{ub} complete bound–unbound–bound sequences of total time T_u, we could estimate the local concentration γ_i by noting that the expected time t_u for the ith unbound receptor to become bound is $1/(r_-\gamma_i)$, and hence $\hat{\gamma}_i = 1/(r_- t_u)$. Given $T_u^{(i)}$ and $n_{ub}^{(i)}$, an obvious estimate of t_u is $T_u^{(i)}/n_{ub}^{(i)}$, and this gives us $\hat{\gamma}_i = n_{ub}^{(i)}/(r_- T_u^{(i)})$. We assumed that the growth cone has accurate knowledge of the average background concentration γ, so Eq. (24) is just

$$\chi_t^{(i)} = r_- T_u^{(i)} \times \left(\hat{\gamma}_i - \gamma\right). \tag{25}$$

If there was no gradient ($\mu = 0$), we would expect $\chi_t^{(i)}$ to be zero on average, as any systematic deviation in γ_i from γ is a result of a non-zero μ.

Finally, substituting Eq. (22) into Eq. (5), we find that the growth cone's optimal strategy is to compare

$$\chi = \sum_i x_i \times \left(b^{(i)} - \frac{\gamma}{1+\gamma} + n_{ub}^{(i)} - \gamma r_- T_u^{(i)}\right) \tag{26}$$

with zero. This is the main result of this section.

4. Gradient sensing with time-averaged occupancy

We now focus on the case in which the growth cone knows only the *average* occupancy f of each receptor, where $f = T_b/T$. Hence, we need an expression for $P(T_b|\gamma, T)$ independent of the initial binding state and the number of state-changes. We can obtain this by marginalizing over b and n in (20):

$$\begin{aligned} P(T_b|\gamma, T) &= \sum_b \sum_n P(T_b, n, b|\gamma, T) \\ &= \sum_n \frac{(\gamma r_-^2 T_u T_b)^{\lfloor n/2 \rfloor}}{\lfloor n/2 \rfloor!^2} \frac{\exp(-\gamma r_- T_u - r_- T_b)}{1+\gamma} \\ &\quad \times \begin{cases} \gamma\delta(T_u) + \delta(T_b) & n = 0 \\ 2r_-\gamma & n \text{ odd} \\ \lfloor n/2 \rfloor \left(\frac{\gamma}{T_u} + \frac{1}{T_b}\right) & n \text{ even and } n \neq 0 \end{cases} \\ &= \frac{\exp(-\gamma r_- T_u - r_- T_b)}{1+\gamma} \\ &\quad \times \left(\gamma\delta(T_u) + \delta(T_b) + 2r_-\gamma \sum_{l=0}^{\infty} \frac{(\gamma r_-^2 T_u T_b)^l}{l!^2} \right. \\ &\quad \left. + \left(\frac{\gamma}{T_u} + \frac{1}{T_b}\right) \sum_{l=0}^{\infty} \frac{l(\gamma r_-^2 T_u T_b)^l}{l!^2}\right). \end{aligned} \tag{27}$$

This can be expressed more succinctly by making use of the series expansion for integer-order modified Bessel functions [21],

$$I_\nu(x) = \left(\frac{1}{2}x\right)^\nu \sum_{k=0}^{\infty} \frac{(x^2/4)^k}{k!\Gamma(\nu + k + 1)}, \tag{28}$$

leading to

$$\begin{aligned} P(T_b|\gamma, T) &= \frac{\exp(-\gamma r_- T_u - r_- T_b)}{1+\gamma} \\ &\quad \times \left(\gamma\delta(T_u) + \delta(T_b) + 2r_-\gamma I_0\left(2\sqrt{\gamma r_-^2 T_u T_b}\right)\right. \\ &\quad \left. + \sqrt{\gamma r_-^2 T_u T_b}\left(\frac{\gamma}{T_u} + \frac{1}{T_b}\right) I_1\left(2\sqrt{\gamma r_-^2 T_u T_b}\right)\right). \end{aligned} \tag{29}$$

We can now determine the probability of observing an average occupancy f by recasting Eq. (29) in dimensionless units (with $T_b \to T_b/T = f$ and $T_u \to (T - T_b)/T = 1 - f$, and $\tau = r_- T$):

$$\begin{aligned} P(f|\gamma, \tau) &= \frac{\tau \exp\left(-\tau(\gamma(1-f) + f)\right)}{1+\gamma} \\ &\quad \times \left(\frac{1}{\tau}\delta(f) + \frac{\gamma}{\tau}\delta(1-f) + 2\gamma I_0\left(2\tau\sqrt{\gamma f(1-f)}\right)\right. \\ &\quad \left. + \sqrt{\gamma f(1-f)}\left(\frac{1}{f} + \frac{\gamma}{1-f}\right) I_1\left(2\tau\sqrt{\gamma f(1-f)}\right)\right). \end{aligned} \tag{30}$$

In making the change of variables $T_b \to f$, we have multiplied by T in Eq. (29), and divided by T in $\delta(T_u)$ and $\delta(T_b)$, so that the transformed formula is also a probability distribution.

Given Eq. (30), we find $\gamma \frac{\mathrm{d}}{\mathrm{d}\gamma} \log(P(f|\gamma, \tau))$, and approximate by substituting the first term in the asymptotic expansion of the Bessel function $I_\nu(x)$, which is

$$I_\nu(x) \approx \frac{\mathrm{e}^x}{\sqrt{2\pi x}}. \tag{31}$$

After some rearrangement, this gives

$$\gamma \frac{\mathrm{d}}{\mathrm{d}\gamma} \log P \approx \frac{1}{1+\gamma} + \frac{1}{2}\frac{\sqrt{\gamma f} - \sqrt{1-f}}{\sqrt{\gamma f} + \sqrt{1-f}} + (\sqrt{\gamma f(1-f)} - \gamma(1-f))\tau \tag{32}$$

which holds for large $\tau\sqrt{\gamma f(1-f)}$.

To explore further, we note that, for shallow gradients, f_i should be close to $\frac{\gamma}{1+\gamma}$ for large τ, and hence replace f_i with $\frac{\gamma}{1+\gamma} + \delta_i$. Under the additional assumptions that $\delta_i \ll \frac{\gamma}{1+\gamma}$ and $\delta_i \ll \frac{1}{1+\gamma}$, then for the ith receptor, Eq. (32) can be approximated by

$$\begin{aligned} \chi_i &= \frac{1}{1+\gamma} + \frac{1}{2}\frac{\gamma\sqrt{1 + \frac{1+\gamma}{\gamma}\delta_i} - \sqrt{1 - (1+\gamma)\delta_i}}{\gamma\sqrt{1 + \frac{1+\gamma}{\gamma}\delta_i} + \sqrt{1 - (1+\gamma)\delta_i}} \\ &\quad + \left(\sqrt{(1 + \frac{1+\gamma}{\gamma}\delta_i)(1 - (1+\gamma)\delta_i)}\right. \\ &\quad \left. - (1 - (1+\gamma)\delta_i)\right)\frac{\gamma}{1+\gamma}\tau \end{aligned}$$

$$\approx \frac{1}{2} + \frac{1}{2}\delta_i + \frac{1}{2}(1+\gamma)\delta_i\tau$$
$$= \frac{1}{2} + \frac{1}{2}(1+(1+\gamma)\tau)\left(f_i - \frac{\gamma}{1+\gamma}\right). \quad (33)$$

This approximate expression is the main result of this section, and it shows how the growth cone's optimal steering decision is related to the ith receptor's time-averaged binding state f_i, the total (dimensionless) observation time $\tau = r_- T$, and the mean concentration γ. We can compare Eq. (33) to the related expression for the full-knowledge case (Eq. (22)) by rewriting it in the following form:

$$\chi_i \approx \frac{1}{2}\left(1 + f_i - \frac{\gamma}{1+\gamma} + r_- T_b^{(i)} + r_- \gamma T_u^{(i)}\right) - r_- \gamma T_u^{(i)}. \quad (34)$$

The final term in this expression is identical to the final term in Eq. (22). We hypothesize that the first term estimates $n_{ub}^{(i)}$; however, more analysis is required to establish this definitively.

5. Comparing the performance of the full-knowledge and time-averaged occupancy cases

It may be that, when the growth cone is restricted to knowing only the time-averaged occupancy of its receptors, its performance is significantly impaired when compared with the situation in which it has full knowledge of its receptors' states; or it could be that its performance is barely affected. Thus, we now compare the performance of the two decision mechanisms developed in the previous sections, described by Eqs. (26) and (33).

We define the sensitivity of a particular strategy χ under particular gradient conditions (i.e., a single choice of (γ, μ)) as the probability $P_{correct}$ that the growth cone will choose the correct gradient direction using the χ strategy under those gradient conditions. In order to estimate this probability, we apply Monte Carlo simulation, using Gillespie's Stochastic Simulation Algorithm [22] in order to generate sequences of receptor binding states over time. We then estimate $P_{correct}$ based on the fraction of such trajectories for which a correct decision would be made.

For each concentration, we generated $N_{trajectories} = 10\,000$ receptor binding trajectories each of length $\tau = 80$. For each trajectory we determined the optimal decision the growth cone could make at various time-points given its assumed level of knowledge about the binding states of its receptors. The error bars shown reflect the standard deviation in the mean for a binomial distribution with success probability given by the estimated value of $P_{correct}$: $\sigma = \sqrt{P_{correct}(1-P_{correct})/N_{trajectories}}$. In each simulation, $\mu = 0.01$ (i.e. a 1% gradient), and the growth cone was assigned 200 receptors. Note that while real growth cones probably have many more receptors than this, our aim in these simulations was to explore the performance of the optimal strategies, not to compare them to real growth cones. Thus, to reduce the simulation time, we considered a reduced number of receptors. In any case, because the model assumes no interactions between receptors, changing the number of receptors can have only trivial effects on performance (i.e., performance scales with the square root of the number of receptors).

Fig. 2A illustrates the gradient sensing performance in the long time limit ($\tau \gg 1$) for the full-knowledge and average occupancy situations. Each curve is for a different value of τ. One interesting feature of Fig. 2A is that the performance plots of the full-knowledge and average occupancy situations both have very similar characteristics — in particular it appears that the average occupancy performance is essentially identical to the full-knowledge performance, assuming that half the time was allowed for observation.

Some intuition for this effect comes from assessing the Fisher information I_μ in the knowledge about the gradient steepness μ in the two cases. Expanding the exponential in Eq. (2) to second order, then substituting the definition of χ_i from Eq. (6), we obtain an approximate expression for the Fisher information in μ for the general case we considered in Section 2:

$$I_\mu = -\left\langle \frac{\partial^2}{\partial \mu^2} \log P(\vec{O}|\gamma,\mu)\right\rangle_{\vec{O}}$$
$$\approx -\sum_i \left\langle x_i^2\gamma^2 \frac{\mathrm{d}^2}{\mathrm{d}\gamma^2}\log P(O_i|\gamma)\right\rangle_{\vec{O}}$$
$$= -\sum_i x_i^2\gamma^2\left\langle \frac{\mathrm{d}}{\mathrm{d}\gamma}\frac{\chi_i}{\gamma}\right\rangle_{\vec{O}}. \quad (35)$$

It is now straightforward to approximate the Fisher information concerning μ in the full-knowledge and average occupancy situations. For the full-knowledge case we obtain

$$I_\mu \approx \sum_i \left(\langle n_{ub}^{(i)}\rangle_{\vec{O}} + \langle b_i\rangle_{\vec{O}} - \rho^2\right)$$
$$\approx N(\rho\tau + \rho(1-\rho)), \quad (36)$$

where we have used $\langle b_i\rangle_{\vec{O}} \approx \rho$ and $\langle n_{ub}^{(i)}\rangle \approx \rho\tau$ (justified below). Similarly, for the average occupancy case, combining Eqs. (33) and (35), we obtain

$$I_\mu \approx \frac{1}{2}\sum_i \left((1+\tau)\langle f_i\rangle_{\vec{O}} + (1-\rho)(1+\rho)\right)$$
$$\approx \frac{N}{2}(\rho\tau + \rho(1-\rho)) + \frac{N}{2}, \quad (37)$$

noting that $\langle f_i\rangle_{\vec{O}} \approx \rho$. When τ is large, we see that roughly twice the information is available in the full-knowledge case compared to the average occupancy case, as observed in Fig. 2A.

A second feature of Fig. 2A is that, for each value of τ, the gradient sensing performance increases monotonically with concentration. This contrasts with experimental observations, in which the performance is typically biphasic, peaked near the dissociation coefficient of the relevant receptors, as shown in Fig. 2B. In the case of our simulations, this unrealistic behaviour occurs because, under the current assumptions of the model, the growth cone can measure the interval of time for which a receptor is unbound with perfect accuracy. As a result, as the concentration increases and the length of time the receptor spends unbound decreases, the growth cone essentially reaches a limit to sensing accuracy related to the number of unbound-to-bound transitions that can occur within the observation period. In other words, each unbound-to-bound transition constitutes a "measurement" of the local concentration, and each individual measurement has the same degree of uncertainty associated with it, regardless of the background concentration, because we have assumed that the only uncertainty stems from the inherent stochasticity of receptor binding, not in downstream processing. As a result, the gradient sensing accuracy is limited by the number of such measurements that can occur in a given period of time, which is in turn limited by the time it takes to "reset" the receptor to its unbound state.

This means that the gradient detection performance of the growth cone should be determined solely by the number of such measurements that can be made. We can estimate this by noting that, on average, the time it takes a receptor to make a single measurement, and return to the unbound state, is given by the sum of the average time it takes the receptor to become bound, $1/\gamma r_-$, and the average time for which it remains bound, $1/r_-$. In an observation time T, the number of measurements a single receptor can make is roughly

$$n_{meas} \approx \frac{T}{\frac{1}{\gamma r_-} + \frac{1}{r_-}}$$

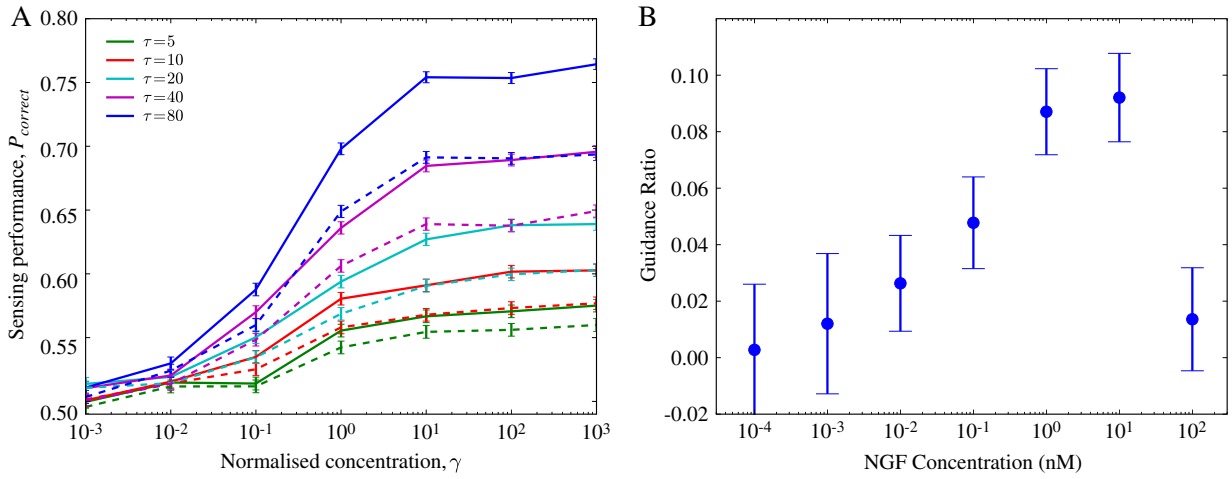

Fig. 2. Theoretical and experimental sensitivity curves. A: The solid lines show the theoretically optimal performance when it has full knowledge of its receptors' dynamics, while the dashed lines show the performance when only the time-averaged occupancy of each receptor is known. Note that the time-averaged occupancy performance closely matches the full-knowledge performance when the observation time is halved. B: Experimentally measured gradient sensing performance (data reproduced from [23]). Rat pup dorsal root ganglia were exposed to a gradient of nerve growth factor (NGF), allowed to grow for 48 hours, and then imaged. The degree of asymmetry in the resulting neurite growth was quantified and plotted against concentration (see [23] for details). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

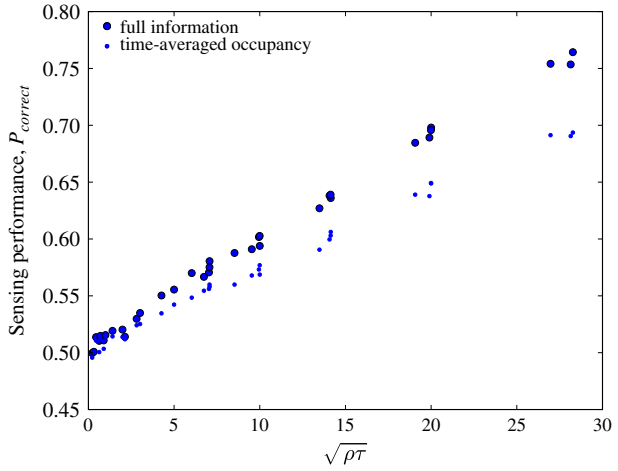

Fig. 3. The performance of both the full-knowledge and average occupancy optimal strategies is directly proportional to $\sqrt{\rho\tau}$ at long times. The data from Fig. 2A is replotted, showing direct proportionality between $P_{correct}$ and $\sqrt{\rho\tau}$.

$$= \frac{r_- T}{(1+\gamma)/\gamma} = \rho\tau, \tag{38}$$

and therefore we would expect the growth cone's gradient sensing performance to be proportional to $\sqrt{n_{meas}} = \sqrt{\rho\tau}$. Indeed, this is exactly what we observe in our simulation results for both the full-knowledge and average occupancy cases, as illustrated in Fig. 3.

6. The effect of downstream noise on gradient sensing

If we included a realistic mechanism for measuring time intervals – such as the production of second messenger molecules at a rapid, but finite rate – smaller time intervals would be measured with diminishing accuracy until, eventually, the limiting factor in determining the sensing accuracy at high concentrations would be the accuracy to which small time intervals can be measured, rather than the number of measurements that can be made. To illustrate this, we estimate the degradation in performance at high concentrations, supposing that unbound receptors produce downstream signals as a Poisson process with rate q.

Denoting the quantity of signalling molecules produced by a given receptor by m, we have

$$P(m|T_U) = \frac{(qT_U)^m \exp(-qT_U)}{m!}, \tag{39}$$

where T_U is a random variable defined by

$$T_U \approx \sum_{l=1}^{n_{meas}} t_l. \tag{40}$$

The t_l are independent, exponentially distributed random variables with mean $(\gamma r_-)^{-1}$, and n_{meas} is the number of "measurements" (i.e. unbound-to-bound transitions) that occur. Since we are particularly interested in the case where γ is large, we will assume that the number of measurements is determined by the time it takes a receptor–ligand complex to dissociate. Hence, n_{meas} is closely approximated by a Poisson random variable with expected value $r_- T = \tau$. We will focus on the limits where both τ and $q\langle T_U\rangle$ are large. Under these conditions, the Central Limit Theorem allows us to assume that $P(m|\gamma, T) = \int_0^T P(m|T_U)P(T_U|\gamma, T)\mathrm{d}T_U$ is roughly normally distributed. Thus, to approximate $P(m|\gamma, T)$, we need to find the mean $\langle m\rangle_{\gamma,T}$ and variance $\langle m^2\rangle_{\gamma,T} - \langle m\rangle^2_{\gamma,T}$ of m.

In order to calculate the mean and variance of m, we first calculate the moment generating function for m,

$$\begin{aligned} M_m(z) &= \sum_{m=0}^{\infty} P(m|\gamma, T)\mathrm{e}^{zm} \\ &= \int_0^T \mathrm{d}T_U P(T_U|\gamma, T) \sum_{m=0}^{\infty} P(m|T_U)\mathrm{e}^{zm} \\ &= \int_0^T \mathrm{d}T_U P(T_U|\gamma, T) \sum_{m=0}^{\infty} \frac{1}{m!}\left(qT_U\mathrm{e}^z\right)^m \exp(qT_U) \\ &= \int_0^T \mathrm{d}T_U P(T_U|\gamma, T) \exp\left(qT_U\left(\mathrm{e}^z - 1\right)\right) \\ &= M_{T_U}\left(q\left(\mathrm{e}^z - 1\right)\right), \end{aligned} \tag{41}$$

where M_{T_U} is the moment generating function for T_U. To calculate M_{T_U}, we use the fact that the probability distribution for a random variable which is a sum of independent, identically distributed

(i.i.d.) random variables (i.e., $T_U = \sum_{i=0}^{n_{meas}} t_i$, where each of the t_i are i.i.d.) is just the convolution of the distributions of the summed variables (i.e., $P(T_U|\gamma, T, n_{meas}) = [*^{n_{meas}} P(t_i|\gamma)]\,(T_U)$, where $*^n$ is defined recursively by $[*^{n+1} f(t)] = \int_0^t \mathrm{d}s f(s)[*^n f(t-s)]$), and then

$$\begin{aligned}
M_{T_U}(s) &= \int_0^T \mathrm{d}T_U P(T_U|\gamma, T)\mathrm{e}^{sT_U} \\
&= \sum_{n_{meas}} P(n_{meas}|\gamma, T) \int_0^T \mathrm{d}T_U P(T_U|\gamma, T, n)\mathrm{e}^{sT_U} \\
&= \sum_{n_{meas}} P(n_{meas}|\gamma, T) \int_0^T \mathrm{d}T_U \left[*^{n_{meas}} P(t_i|\gamma)\right](T_U)\mathrm{e}^{sT_U} \\
&= \sum_{n_{meas}} P(n_{meas}|\gamma, T) \left[\int_0^T \mathrm{d}t P(t|\gamma)\mathrm{e}^{st}\right]^{n_{meas}} \\
&= \sum_{n_{meas}} P(n_{meas}|\gamma, T) M_{t_i}(s)^{n_{meas}} \\
&= \sum_{n_{meas}} \frac{1}{n_{meas}!} \left(\tau M_{t_i}(s)\right)^{n_{meas}} \exp(-\tau) \\
&= \exp\left[\tau\left(M_{t_i}(s) - 1\right)\right], \qquad (42)
\end{aligned}$$

where we have used the fact that the Laplace transform of a convolution of functions is just the product of the Laplace transform for each function individually. The moment generating function for t_i can be found by an elementary integration:

$$\begin{aligned}
M_{t_i}(s) &= \int_0^\infty \mathrm{d}t P(t|\gamma)\mathrm{e}^{st} \\
&= \int_0^\infty \mathrm{d}t(\gamma r_-)\exp\left[(s - \gamma r_-)\,t\right] \\
&= \frac{\gamma r_-}{\gamma r_- - s}, \qquad (43)
\end{aligned}$$

and so, combining all of the above, we obtain

$$\begin{aligned}
M_m(z) &= \exp\left[\tau\left(\frac{\gamma r_-}{\gamma r_- - q\left(\mathrm{e}^z - 1\right)} - 1\right)\right] \\
&= \exp\left[\tau\frac{q\left(\mathrm{e}^z - 1\right)}{\gamma r_- - q\left(\mathrm{e}^z - 1\right)}\right]. \qquad (44)
\end{aligned}$$

We can easily determine the mean and variance of m from the moment generating function by noting that, given the moment generating function of a random variable X, the kth moment is given by

$$\begin{aligned}
\langle X^k\rangle &= \left\langle \frac{\mathrm{d}^k}{\mathrm{d}z^k}\mathrm{e}^{zX}\right\rangle\Big|_{s=0} \\
&= \frac{\mathrm{d}^k}{\mathrm{d}s^k} M_X(z=0) \qquad (45)
\end{aligned}$$

and hence

$$\begin{aligned}
\langle m\rangle &= \frac{\mathrm{d}}{\mathrm{d}z}\exp\left[\tau\frac{q\left(\mathrm{e}^z - 1\right)}{\gamma r_- - rq\left(\mathrm{e}^z - 1\right)}\right]\Big|_{z=0} \\
&= \tau\frac{\gamma r_- q\mathrm{e}^z}{\left(q\left(\mathrm{e}^z - 1\right) - \gamma r_-\right)^2}\exp\left[\tau\frac{q\left(\mathrm{e}^z - 1\right)}{\gamma r_- - q\left(\mathrm{e}^z - 1\right)}\right]\Big|_{z=0} \\
&= \tau\frac{\gamma r_- q}{\gamma^2 r_-^2} = \frac{q}{r_-\gamma}\tau \qquad (46)
\end{aligned}$$

and (without giving details of the differentiation)

$$\begin{aligned}
\langle m^2\rangle - \langle m\rangle^2 &= \frac{1}{\gamma}qT + \frac{q^2T^2}{\gamma^2} + \frac{2q^2T^2}{r_-\gamma^2} - \left(\frac{qT}{\gamma}\right)^2 \\
&= \frac{1}{\gamma}qT\left(1 + \frac{2q}{r_-\gamma}\right) \\
&= \frac{q}{r_-\gamma}\tau\left(1 + \frac{2q}{r_-\gamma}\right). \qquad (47)
\end{aligned}$$

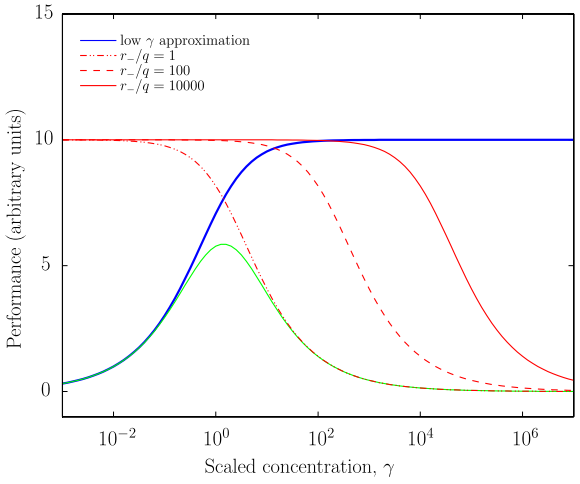

Fig. 4. The gradient sensing performance degrades at high concentration when downstream constraints on how accurately the period of time for which a receptor is unbound can be measured are included. The concentration at which performance is impacted depends on the relative timescales of receptor–ligand interaction, and downstream signalling. While our methods do not allow us to estimate the entire sensitivity curve, we might expect performance to behave somewhat like the product of the low-concentration (in blue) and high-concentration (in red) estimates. This is shown in green for the case in which downstream signalling occurs on the same timescale as receptor–ligand interaction. Note that now, the performance decreases at high concentrations, consistent with experimental data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Assuming that the receptors are distributed symmetrically around the center of the growth cone, then, for small μ, the gradient sensing performance should be proportional to the signal-to-noise ratio:

$$\begin{aligned}
|d'| &= \frac{\|\Delta\langle m\rangle\|}{\sqrt{2\left(\langle m^2\rangle - \langle m\rangle^2\right)}} \\
&= \frac{\mu\gamma\,\|d\langle m\rangle/d\gamma\|}{\sqrt{2\left(\langle m^2\rangle - \langle m\rangle^2\right)}} \\
&= \mu\frac{\frac{q}{r_-\gamma}\tau}{\sqrt{2\frac{q}{r_-\gamma}\tau\left(1 + \frac{2q}{r_-\gamma}\right)}} \\
&= \mu\sqrt{\frac{\tau}{\frac{r_-}{q}\gamma + 2}}. \qquad (48)
\end{aligned}$$

Thus, the concentration at which gradient sensing performance begins to decay is determined by the ratio of timescales of receptor signalling to receptor binding, $\epsilon = r_-/q$. This is illustrated in Fig. 4, in which the performance at high concentrations is compared for a range of different ϵ, with $\tau = 100$. As the rate of downstream signalling increases, the performance is maintained at higher concentrations; this is to be expected, as the accuracy with which a short interval can be estimated by counting the number of Poisson-distributed events that occur within it is proportional to the expected number of signalling events within that interval.

7. Discussion

We have introduced a general framework for modelling gradient detection in one dimension, and have applied it specifically

to the problem of extracting optimal information from individual chemoreceptors, given their binding states over time. We found that, for optimal detection of an unchanging external gradient, a receptor needs to provide two key quantities: the total time spent unbound within the observation period, and the number of transitions from unbound to bound occurring within that period. This contrasts with what is generally assumed in the chemotaxis modelling literature: namely, that a cell or growth cone knows only the time-averaged binding state of its receptors. We demonstrated that signalling in this manner effectively decreases the sampling time by a factor of a half — in other words, when only the time-averaged binding state is known, half the information available to the cell is discarded.

By simulating the gradient sensing performance achievable by these two signalling strategies, we found that the performance was monotonically increasing with concentration. This contrasts with experimental results, where the gradient sensing performance is biphasic with concentration. We proposed one possible explanation for why the performance might decay at high concentrations: as the concentration increases, it is increasingly difficult for the receptor to transmit accurate information about its binding trajectory, leading to a degradation in performance at high concentrations. The concentration at which this decay occurs is determined by the maximum rate at which downstream signalling molecules can be produced.

An interesting corollary of this suggestion is that it ought to be possible to independently manipulate the concentrations at which the gradient sensitivity initially rises from zero, and falls back to zero, by varying the ligand–receptor interaction kinetics and rate of downstream signalling, respectively. We might expect these to be at least partially decoupled in physical receptors, as presumably the rate of downstream signalling is primarily determined by the intracellular domain of the receptor, while the kinetics of ligand interaction may be primarily determined by the extracellular domain.

Finally, we note that while we have specifically focussed on temporal aspects of gradient sensing here, the interface between temporal and spatial uncertainty might provide a wealth of interesting problems. Ligand molecules diffuse in the surrounding medium and can interact with multiple receptors over a period of time (introducing correlations between adjacent receptors). The growth cone moves through its environment, and the environment is itself dynamic: the very quantities the growth cone is trying to estimate can be changing with time. Somehow, the growth cone must make reliable and useful decisions and estimates against this dynamic backdrop.

Acknowledgements

Funding comes from an Australian Postgraduate Award (DM), an Australian Research Council Federation Fellowship (KB), the Gatsby Charitable Foundation (PD), the Australian Research Council (Discovery Grant DP0666126) and the Australian National Health and Medical Research Council (Project Grant 456003).

References

[1] Hugh D. Simpson, Duncan Mortimer, Geoffrey J. Goodhill, Theoretical models of neural circuit development, Curr. Top. Dev. Biol. 87 (2009) 1–51.
[2] Johannes K. Krottje, Arjen van Ooyen, A mathematical framework for modeling axon guidance, Bull. Math. Biol. 69 (2007) 3–31.
[3] Marc Tessier-Lavigne, Corey S. Goodman, The molecular biology of axon guidance, Science 274 (1996) 1123–1133.
[4] H. Song, Mu-Ming Poo, The cell biology of neuronal navigation, Nat. Cell. Biol. 3 (2001) E81–E88.
[5] Barry J. Dickson, Molecular mechanisms of axon guidance, Science 298 (5600) (2002) 1959–1964.
[6] Celine Plachez, Linda J. Richards, Mechanisms of axon guidance in the developing nervous system, Curr. Top. Dev. Biol. 69 (2005) 267–346.
[7] Phillip R. Gordon-Weeks, Neuronal Growth Cones, Cambridge University Press, 2000.
[8] Duncan Mortimer, Thomas Fothergill, Zac Pujic, Linda J. Richards, Geoffrey J. Goodhill, Growth cone chemotaxis, Trends Neurosci. 31 (2) (2008) 90–98.
[9] H.C. Berg, E.M. Purcell, Physics of chemoreception, Biophys. J. 20 (2) (1977) 193–219.
[10] G.J. Goodhill, J.S. Urbach, Theoretical analysis of gradient detection by growth cones, J. Neurobiol. 41 (2) (1999) 230–241.
[11] G.J. Goodhill, M. Gu, J.S. Urbach, Predicting axonal response to molecular gradients with a computational model of filopodial dynamics, Neural Comput. 16 (2004) 2221–2243.
[12] Jun Xu, William J. Rosoff, Jeffrey S. Urbach, Geoffrey J. Goodhill, Adaptation is not required to explain the long-term response of axons to molecular gradients, Development 132 (20) (2005) 4545–4552.
[13] William Bialek, Sima Setayeshgar, Physical limits to biochemical signaling, Proc. Natl. Acad. Sci. USA 102 (29) (2005) 10040–10045.
[14] Masahiro Ueda, Tatsuo Shibata, Stochastic signal processing and transduction in chemotactic response of eukaryotic cells, Biophys. J. 93 (1) (2007) 11–20.
[15] Peter J.M. Van Haastert, Marten Postma, Biased random walk by stochastic fluctuations of chemoattractant-receptor interactions at the lower limit of detection, Biophys. J. 93 (5) (2007) 1787–1796.
[16] J.M. Kimmel, R.M. Salter, P.J. Thomas, An information theoretic framework for eukaryotic gradient sensing, in: Advances in Neural Information Processing Systems, 19, MIT Press, 2007, pp. 705–712.
[17] D. Mortimer, J. Feldner, T. Vaughan, I. Vetter, Z. Pujic, W.J. Rosoff, K. Burrage, P. Dayan, L.J. Richards, G.J. Goodhill, A Bayesian Model predicts the response of axons to molecular gradients, Proc. Natl. Acad. Sci. USA 106 (2009) 10296–10301.
[18] W.J. Rappel, H. Levine, Receptor noise and directional sensing in eukaryotic chemotaxis, PRL 100 (2008) 228101.
[19] R.T. Tranquillo, D.A. Lauffenburger, Stochastic model of leukocyte chemosensory movement, J. Math. Biol. 25 (3) (1987) 229–262.
[20] Burton W. Andrews, Pablo A. Iglesias, An information-theoretic characterization of the optimal gradient sensing response of cells, PLoS Comp. Biol. 3 (8) (2007) e153.
[21] M. Abramowitz, I.A. Stegun (Eds.), Handbook of Mathematical Functions, Dover publications, Inc., New York, 1970.
[22] D.T. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. 81 (25) (1977) 2340–2361.
[23] William J. Rosoff, Jeffrey S. Urbach, Mark A. Esrick, Ryan G. McAllister, Linda J. Richards, Geoffrey J. Goodhill, A new chemotaxis assay shows the extreme sensitivity of axons to molecular gradients, Nat. Neurosci. 7 (6) (2004) 678–682.