Optimality and Saturation in Axonal Chemotaxis

Jiajia Yuan
jiajia.yuan@uqconnect.edu.au
Stanley Chan
schan@physics.uq.edu.au
Duncan Mortimer
dmorti@gmail.com
Huyen Nguyen
huyen.vyvy@gmail.com
Queensland Brain Institute, University of Queensland, St. Lucia,
QLD 4072, Australia

Geoffrey J. Goodhill
g.goodhill@uq.edu.au
Queensland Brain Institute, University of Queensland, St. Lucia, and
School of Mathematics and Physics, QLD 4072, Australia

Chemotaxis (detecting and following chemical gradients) plays a crucial role in the function of many biological systems. In particular, gradient following by neuronal growth cones is important for the correct wiring of the nervous system. There is increasing interest in the constraints that determine how small chemotacting devices respond to gradients, but little quantitative information is available in this regard for neuronal growth cones. Mortimer et al. (2009) and Mortimer, Dayan, Burrage, and Goodhill (2011) proposed a Bayesian ideal observer model that predicts chemotactic performance for shallow gradients. Here we investigated two important aspects of this model. First, we found by numerical simulation that although the analytical predictions of the model assume shallow gradients, these predictions remain remarkably robust to large deviations in gradient steepness. Second, we found experimentally that the chemotactic response increased linearly with gradient steepness for very shallow gradients as predicted by the model; however, the response saturated for steeper gradients. This saturation could be reproduced in simulations of a growth rate modulation response mechanism. Together these results illuminate the domain of validity of the Bayesian model and give further insight into the biological mechanisms of axonal chemotaxis.

1 Introduction

A large part of the computational power of biological nervous systems lies in their intricate connectivity. During the development of complex nervous systems, billions of nerve fibers (axons) must grow over potentially long distances to find their appropriate targets. One of the crucial mechanisms axons use to achieve this feat is guidance by chemical gradients. This is an example of chemotaxis, and it is believed that axons may share many features in this regard with other chemotacting systems at a similar scale such as leukocytes and *Dictyostelium* (von Philipsborn & Bastmeyer, 2007; Mortimer, Fothergill, Pujic, Richards, & Goodhill, 2008).

Gradients for axon guidance in vivo can be set up in a variety of ways, including diffusion from the target tissue and differential expression on the surfaces of cells over which the axons grow. Due to the very small concentrations of guidance factors involved, it has seldom been possible to obtain quantitative information regarding the shape of these gradients in vivo (Kennedy, Wang, Marshall, & Tessier-Lavigne, 2006). However, it is likely that both gradient steepness and concentration will vary considerably between different guidance paths and also over the length of an individual path and that these variations will influence axon behavior (Isbister, Mackenzie, To, & O'Connor, 2003; Legg & O'Connor, 2003; Mortimer, Pujic et al., 2010; Thompson, Pujic, Richards, & Goodhill, 2011). It is thus important to gain a more quantitative understanding of how gradient parameters affect axonal growth.

At the most fundamental level, chemotaxis is limited by unavoidable noise due to the limited number of individual ligand and receptor molecules involved (Berg & Purcell, 1977; Bialek & Setayeshgar, 2005; Andrews & Iglesias, 2007; Ueda & Shibata, 2007; Endres & Wingreen, 2008; Hu, Rappel, & Levine, 2010). One source of such noise is that receptor binding is a fundamentally stochastic process, so that determining gradient direction is a form of reasoning under sensory uncertainty. Mortimer et al. (2009) and Mortimer, Dayan, Burrage, and Goodhill (2011) analyzed this problem using a Bayesian ideal observer approach. In particular, they calculated the form of the posterior probability distribution for gradient direction, given the binding state of a set of receptors spatially distributed over a small chemotacting device and prior assumptions regarding the concentrations and steepnesses likely to be encountered (see Figure 1A). Assuming shallow gradients allowed a simple analytic form for the chemotactic performance of such devices as a function of concentration and steepness to be derived. In general this prediction matches well with quantitative measurements of the behavior of axons (Mortimer et al., 2009; see Figures 1B, and 1C). However, this work left two crucial questions regarding the model and real axonal performance unanswered.

First, to derive a closed-form solution for the chemotactic performance from the model, shallow gradients were assumed in terms of both the

A

Decision strategy

↓

OR

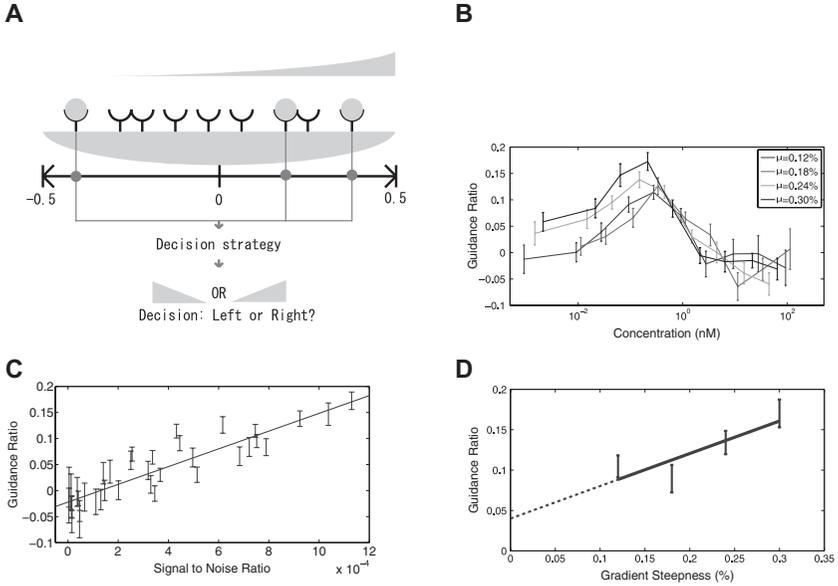Decision: Left or Right?

B

C

D

Figure 1: Summary of the Bayesian model of Mortimer et al. (2009, 2011) and previous experimental work. (A) Receptors are assumed to be randomly distributed across a one-dimensional chemotactic sensing device. These receptors are exposed to an external gradient, and the positions of bound receptors provide signals for the sensing device to decide the gradient direction. (B) Experimentally measured chemotactic response (guidance ratio, GR) of dorsal root ganglion explants for 38 different combinations of concentration and gradient steepness. (C) Correlation of response predicted by the Bayesian model and measured response across all gradients conditions (B, C replotted from Mortimer et al., 2009). (D) Closer examination of the response interpolated at 0.2 nM for each gradient steepness suggests that response may not decline linearly to zero as predicted by the Bayesian model: extrapolating a linear fit to zero gradient steepness gives a nonzero GR.

gradients presented and the assumptions encoded in the prior distributions. It is thus unclear to what extent the analytic predictions are robust to relaxing these shallow gradient assumptions. This is important because gradients for axon guidance in vivo are expected to range from shallow to steep. We therefore used numerical calculation based on the exact expression for the posterior distribution to determine chemotactic performance in steep gradients and in a range of gradient steepnesses when the prior expectation was steep gradients. Comparing these exact predictions with the analytic expression derived from the shallow-gradient approximations, we found that the analytical approximation is robust to such variations, and

matches very closely with the exact solution across a wide range of gradient parameters.

Second, the model predicts that chemotactic performance should scale linearly with gradient steepness. Mortimer et al. (2009) tested four different gradient steepnesses experimentally, and although the measured response increased with gradient steepness, it appeared that the response might be saturating at higher gradient steepnesses. In addition, it was unclear whether, as predicted by the model, there was a smooth decline to zero response as gradient steepness decreased to zero (see Figure 1D). However due to the way in which these experiments were performed (see section 2), it was hard to reach definitive conclusions on these points. We therefore conducted more carefully controlled experiments to measure the precise form of the relationship between gradient steepness and chemotactic response. We found that under these conditions, the response did indeed decline linearly to zero with gradient steepness, arguing against the presence of an absolute threshold in gradient steepness below which gradient detection is impossible. However, for higher gradient steepnesses, the response clearly saturated. A simple computational investigation revealed that a possible explanation for this saturation lies in the mechanism by which axons respond to shallow gradients.

2 Methods

2.1 Bayesian Model of Chemotaxis. We first briefly recapitulate the model presented in Mortimer et al. (2009, 2011) for shallow gradients and then present the generalization to steep gradients.

2.1.1 Shallow Gradient Model. We first briefly recapitulate the model presented in Mortimer et al. (2009, 2011; for further details see those papers). We consider the simplified growth cone shown in Figure 1A, which consists of a one-dimensional array of a uniformly distributed set of receptors. An external ligand gradient is defined by

$$\gamma(r) = \gamma(0) \times (1 + \mu r), \tag{2.1}$$

where γ is concentration C divided by the dissociation constant K_D, r is the position x of the receptor divided by the growth cone's diameter l, μ is the dimensionless gradient steepness $\mu = \frac{1}{C(0)}\frac{dC}{dx}l$, and $C(0)$ is the concentration at the center of the growth cone. Assuming that the probability of a receptor being bound at this instant is $\gamma/(1+\gamma)$, we have

$$P(\mathcal{B}|\gamma, \mu, \mathcal{R}) = \chi(\mathcal{B} \subset \mathcal{R}) \prod_{r \in \mathcal{B}} \frac{\gamma(1 + \mu r)}{1 + \gamma(1 + \mu r)} \prod_{r \in \mathcal{R}/\mathcal{B}} \frac{1}{1 + \gamma(1 + \mu r)},$$

$$\tag{2.2}$$

where $P(\mathcal{B}|\gamma, \mu, \mathcal{R})$ is the probability of observing the multiset $\mathcal{B}$ giving the positions of the bound receptors, given γ, μ, and the multiset $\mathcal{R}$ giving the positions of all receptors. $\chi(\mathcal{B} \subset \mathcal{R})$ is one when $\mathcal{B}$ is a subset of $\mathcal{R}$ and zero otherwise. This can be rewritten as

$$P(\mathcal{B}|\gamma, \mu, \mathcal{R}) = \chi(\mathcal{B} \subset \mathcal{R})\frac{\gamma^n}{(1+\gamma)^N}\prod_{r\in\mathcal{B}}\frac{1+\mu r}{1+\mu\rho r}\prod_{r\in\mathcal{R}/\mathcal{B}}\frac{1}{1+\mu\rho r}, \quad (2.3)$$

where $\rho = \gamma/(1+\gamma)$, $n = |\mathcal{B}|$ is the number of bound receptors and $N = |\mathcal{R}|$ is the total number of receptors. If we now assume independence of γ, μ, and $\mathcal{R}$ and that receptors are distributed independently and only bound receptors contribute to the signaling, we can apply Bayes' theorem and marginalize over γ and $\mathcal{R}$ to obtain

$$P(\mu|\mathcal{B}) \propto P(\mu)\int_0^\infty \int d\gamma P(\gamma)\frac{\gamma^n}{(1+\gamma)^N}P(\mathcal{R})\chi(\mathcal{B} \subset \mathcal{R})$$
$$\times \prod_{r\in\mathcal{B}}\frac{1+\mu r}{1+\mu\rho r}\prod_{r\in\mathcal{R}/\mathcal{B}}\frac{1}{1+\mu\rho r}d^N\mathcal{R}. \quad (2.4)$$

Reexpressing the products in this expression as exponentials and assuming that the gradient is sufficiently shallow that we can approximate the resulting exponent to first order in μ, we obtain

$$P(\mu|\mathcal{B}) \propto P(\mu)\int d\gamma P(\gamma)\frac{\gamma^n}{(1+\gamma)^N}\exp\left[(1-\rho)\mu R_b\right], \quad (2.5)$$

where $R_b = \sum_{r\in\mathcal{B}} r$. Assuming that the growth cone has a large number of receptors, the integral over γ can be approximated by

$$P(\mu|\mathcal{B}) \propto P(\mu)\exp\left[(1-\hat{\rho})\mu R_b\right], \quad (2.6)$$

where $\hat{\rho}$ is the a posteriori mean estimate for ρ assuming that no gradient is present.

The best estimate for the gradient direction at this instant can now be determined by comparing the probability that the gradient points to the right against the probability that it points to the left:

$$\Delta P = P(\mu > 0|\mathcal{B}) - P(\mu < 0|\mathcal{B}). \quad (2.7)$$

If ΔP is less than zero, that indicates that it is more probable that μ is less than zero, while if ΔP is greater than zero, it is more probable that μ is greater than zero. If we assume that the growth cone is initially agnostic about the direction of the gradient so that $P(\mu)$ is symmetric around $\mu = 0$,

then the sign of ΔP is entirely determined by the sign of $(1 - \hat{\rho})R_b$. The growth cone's optimal estimate for the sign of μ is thus

$$\text{sign}[(1 - \hat{\rho})R_b] = \text{sign}R_b. \tag{2.8}$$

This corresponds to simply summing the positions of bound receptors relative to the center of the growth cone.

In order to estimate the performance of this strategy, we consider $P_{correct}(\gamma, \mu)$, the probability that the strategy makes the correct decision about the gradient direction for a given γ and μ. $P_{correct}(\gamma, \mu)$ can be approximated analytically (Mortimer et al., 2009).

$$P_{correct}(\gamma, \mu) \approx \frac{1}{2}\text{erfc}\left(-|\mu|\sqrt{\frac{N}{24}\frac{\gamma}{(1+\gamma)^3}}\right)$$

$$\approx \frac{1}{2} + |\mu|\sqrt{\frac{N}{24\pi}}\sqrt{\rho}\,(1-\rho), \tag{2.9}$$

where erfc is the complementary error function, from which we identify the signal-to-noise ratio (SNR):

$$SNR \propto |\mu|\sqrt{\frac{\gamma}{(1+\gamma)^3}}. \tag{2.10}$$

Mortimer et al. (2009) showed that the SNR correlated well with the chemotactic sensitivity of axons under a wide range of γ and small μ. Mortimer et al. (2011) then extended the model to 2D and showed that the basic performance characteristics of the model were unchanged between 1D and 2D. Time is not explicitly considered in this model; however, one could imagine an additional simple temporal averaging process to gain better statistics regarding the patterns of receptor binding. Mortimer, Dayan et al. (2010) explicitly considered the temporal case, where information about the dynamics of unbound-bound transitions is used directly.

2.1.2 Steep Gradient Model. For shallow gradients, it is reasonable to approximate the gradient as linear across the receptor array. However, this assumption fails for steep gradients as it can potentially lead to negative concentrations at the low-concentration end of the array. We therefore avoid this problem for steep gradients by assuming an exponential gradient:

$$\gamma(r) = \gamma \exp(\mu r). \tag{2.11}$$

As in the shallow gradient case, the optimal estimate of the gradient direction is determined by

$$\Delta P = P(\mu > 0|\mathcal{B}) - P(\mu < 0|\mathcal{B}). \tag{2.12}$$

However, in this case, we cannot obtain a simple analytic expression for $\text{sign}(\Delta P)$. Rather, we must evaluate $P(\mu > 0|\mathcal{B})$ and $P(\mu < 0|\mathcal{B})$ numerically. We have

$$
\begin{aligned}
P(\mu > 0|\mathcal{B}) = V \int_0^\infty d\mu P(\mu) \int_0^\infty d\gamma P(\gamma) \int d^N \mathcal{R} P(\mathcal{R}) \chi (\mathcal{B} \subset \mathcal{R}) \\
\times \prod_{r \in \mathcal{B}} \frac{\gamma \exp(\mu r)}{1 + \gamma \exp(\mu r)} \prod_{r \in \mathcal{R}/\mathcal{B}} \frac{1}{1 + \gamma \exp(\mu r)}, \\
= V \int_0^\infty d\mu P(\mu) \int_0^\infty d\gamma P(\gamma) \prod_{r \in \mathcal{B}} \frac{\gamma \exp(\mu r)}{1 + \gamma \exp(\mu r)} \int d^{N-n} \mathcal{U} P(\mathcal{U}) \\
\times \prod_{r \in \mathcal{U}} \frac{1}{1 + \gamma \exp(\mu r)}, \quad\quad\quad (2.13)
\end{aligned}
$$

where we have used the fact that the positions of the bound receptors are known to partially evaluate the integral over $\mathcal{R}$ and V is an unknown normalizing constant. A similar expression is obtained for the integral over negative values of μ.

In order to estimate the performance of the decision strategy for a given γ and μ, we generated 1000 different receptor binding configurations and for each configuration evaluated the $\mu > 0$ and $\mu < 0$ integrals numerically. Specifically, for each binding distribution, we generated five runs of 10,000 samples using the Metropolis-Hastings algorithm, discarding the first 1500 samples in order to remove bias due to the seed for the run ($\mu = 0$, $\log \gamma = 0$ in each case). For each binding configuration and for each Metropolis-Hastings run, we estimated ΔP by subtracting the fraction of samples for which $\mu < 0$ from the fraction of samples for which $\mu > 0$. Our performance estimate for a given γ and μ was then taken to be $2q - 1$, where q was the proportion of the 1000 binding configurations for which the sign of ΔP was the same as that of the gradient (since $q = 0.5$ corresponds to a chance level of performance).

We assumed flat priors for $P(\mathcal{R})$ and $P(\gamma)$—that receptors are distributed uniform randomly across the growth cone and that the growth cone has no prior expectation as to the concentrations it will encounter. For $P(\mu)$, we considered two cases. First, we assumed a gaussian probability centered on steepness 0, where the variance represents the range of steepness the growth cone is expecting. We investigated variances from 0.1 (corresponding to $\approx 3\%$ change in concentration across the array) to 10 (i.e., $\approx 300\%$ change in concentration across the array). Second, we considered a sum of two equal gaussians centered at steepness ± 0.5 (i.e., a 50% change in concentration across the array) with variance 0.01. This corresponds to an expectation on the part of the growth cone that the steepness will be close to 50%. However, we show in section 3 that the chemotactic sensitivity predicted

by the Bayesian model is quite insensitive to these assumptions except in extreme cases.

2.2 Experiments

2.2.1 Overview. Gradients of nerve growth factor (NGF) were produced in three-dimensional collagen gels in a manner similar to that previously described (Rosoff, McAllister, Esrick, Goodhill, & Urbach, 2005; Mortimer et al., 2009; Vetter, Pujic, & Goodhill, 2010). Briefly, a Gesim nanoplotter was used to print precisely controlled volumes of NGF stock solution in a grid pattern on the surface of the gel. Rosoff et al. (2005) showed that these spots diffuse into the collagen, creating a gradient that remains relatively stable for many hours. Explants of early postnatal rat dorsal root ganglia (DRGs) were embedded in the collagen prior to gradient printing and then allowed to grow for 40 hours in the NGF gradient. At the end of this time, response to the gradient was measured by comparing the number of pixels representing neurite growth up the gradient compared to neurite growth down the gradient as described in Mortimer et al. (2009). This yielded a guidance ratio (GR) that is zero for radially symmetric neurite outgrowth, with increasingly positive values representing an increasingly strong response to the gradient.

2.2.2 Modifications from Previous Work. Following the study presented in Mortimer et al. (2009), we identified several ways to potentially reduce the variability in these types of experiments. First, our subsequent work revealed that DRGs from different spinal levels show different degrees of response to the same NGF gradients (Vetter et al., 2010). For this work, we therefore used DRGs only from the lumbar region rather than from lumbar and thoracic regions mixed together, as in Mortimer et al. (2009). Second, expression of the NGF receptor TrkA is known to change with age (Phillips & Armanini, 1996; Molliver & Snider, 1997), and we therefore used a narrower range of ages of rat pups—postnatal day 0 to 1 (P0-P1), as compared to P0-P3 in Mortimer et al. (2009). Third, we printed gradients consisting of 16 lines of different concentrations of NGF rather than 12 lines plus 4 plateau lines as in Mortimer et al. (2009), which leads to greater stability of the gradient. Fourth, Mortimer et al. (2009) showed by finite element modeling that the parameters of the gradient in the gel do not exactly reflect the parameters of the gradient produced in the collagen. While Mortimer et al. (2009) applied correction factors to take this into account, this had the consequence that the response to different gradient steepnesses was never compared for exactly the same concentration. In this work, we applied analogous correction factors directly to the gradient printing, ensuring that the same concentration was produced at the position of the explants for each gradient steepness. Finally, the experiments in Mortimer et al. (2009) were performed by several different people, introducing some inevitable

variability in tissue preparation. Here only one person (J.Y.) performed the experiments.

2.2.3 Detailed Methods. All experiments were approved by the University of Queensland Animal Ethics Committee. DRGs were dissected and trimmed from the lumbar regions of P0–P1 Wistar rat pups, and stored in Hibernate E (Brainbits) at 4°C overnight. The next day, the DRGs were incubated in 5 ml Hanks balanced salt solution with 0.25% trypsin at 37°C for 6 minutes to remove the outer capsule. Then 20 ml of Leibvitz with 1% glucose was added to halt the enzyme activity. After removing liquid, the process was repeated twice. Six explants in a row were then embedded in a thin layer of collagen gel in the middle of a 35 mm tissue culture dish. The collagen gel contained 0.2 mg/ml collagen, 1× optimem, 1 mg/ml sodium bicarbonate, and 1× antibiotic-antimycotic (Invitrogen).

A Gesim nanoplotter 2.0 was then used to "print" the desired gradient of NGF (Biosensis). Sixteen NGF stock solutions with exponentially increasing NGF concentration were deposited on the surface of the collagen gels in the form of 16 parallel lines, with the explants located between the eighth and nineth lines. Three dishes were printed in parallel. The concentration and gradient steepness were chosen such that after the application of correction factors analogous to those in Mortimer et al. (2009), gradients were produced with a final concentration 0.2 nM at the explants, with steepnesses of 0.0%, 0.03%, 0.06%, 0.09%, 0.12%, 0.15%, 0.18%, 0.21%, and 0.24% fractional change across 10 microns, the approximate width of a DRG growth cone. After printing, dishes were incubated at 37°C with 5%CO_2 for 40 hours.

Explants were fixed with 10% formaldehyde 0.1% Triton X-100 in phosphate buffered solution (PBS) overnight. The dishes were washed 6 times with PBS at 1 hour intervals, and then the neuronal tubulin antibody TUJ1 (1μg/ml) was added and the dishes left overnight. Following five washes with PBS for 1 hour each, the dishes were left overnight in the secondary antibody Alexa Fluor 488-conjugated goat antimouse IgG (1:1000). After another five washes the next day, explants were photographed with an AxioCam HRm camera on a Zeiss Image Z1 fluorescence microscope. After manually masking the explant in the resulting image file, we used custom Matlab code to trace the neurites in each image. We identified the centroid of the explant body and then counted the number of neurite pixels on the up- and down-gradient sides of this center point. Outgrowth asymmetry was quantified using the guidance ratio $GR = (H - L)/(H + L)$, where H and L are the number of neurite pixels on the high and low NGF concentration sides of the explant, respectively. As Mortimer, Pujic, et al. (2010) explained, there was no obvious turning of neurites; therefore, the readout of chemotactic response was the asymmetry of growth up versus down the gradient.

The experiments were conducted in two groups: first steepnesses 0.03% to 0.12% and then steepnesses 0.15% to 0.24%, each with its own 0% control. For the first group, we found a small negative offset to the GR in the control condition of -0.025 (compared to typical values of ≈ 0.1 for robust guidance responses). However, for the second group of experiments, we turned the dishes by 90 degrees on the nanoplotter and printed the NGF lines in a similarly rotated direction. In this case, the GR in the control condition fell to -0.007. To match the results from the two groups, we therefore added the negative offset value of the relevant control group to all the GRs obtained, so that the GR in the control conditions was exactly zero. The experiments of Mortimer et al. (2009) also showed a slight negative offset in the control condition of ≈ -0.01, which was not corrected for in that work. We do not currently understand the reason for this small effect on the GR of the orientation of the dishes on the nanoplotter. One possibility is that there could perhaps have been a small electrical potential gradient present in the nanoplotter for the current experiments that also affected neurite growth by galvanotaxis (McCaig, 1986).

3 Results

3.1 Chemotactic Sensitivity Predictions for Steep Gradients. Mortimer et al. (2009) showed previously that the analytical approximation for chemotactic performance derived assuming shallow gradients fit well with experimental measurements of the chemotactic sensitivity of axons using shallow gradients. We now tested how well this analytical approximation matches the behavior of the full model under less restrictive conditions regarding gradient steepness.

First, we compared the analytical approximation with the numerically simulated results over a wide range of steepnesses and concentrations for three different variances σ^2 of the prior distribution over gradient steepnesses (see Figure 2). There was generally a very good fit between the analytical approximation and simulated results, even for gradient steepnesses and prior distributions well outside the expected range of validity of the analytical approximation. A slight difference did occur for high steepnesses at high concentrations, where the analytical approximation slightly overestimated the chemotactic performance. However, in general, the analytical prediction is surprisingly robust.

This comparison was of performance averaged over many different receptor distributions. We therefore next compared the analytical and simulated results for exactly the same receptor distributions (see Figures 3A to 3C). In particular, we defined the common decision fraction as the ratio of times in each condition that the shallow gradient and estimated Bayes-optimal predictions were the same. At least in the middle range of concentrations ($10^{-2} \leq \gamma \leq 10^1$), there was close to perfect agreement for the larger variances in the gradient-steepness prior ($\sigma^2 = 0.1, 1.0, 10.0$).
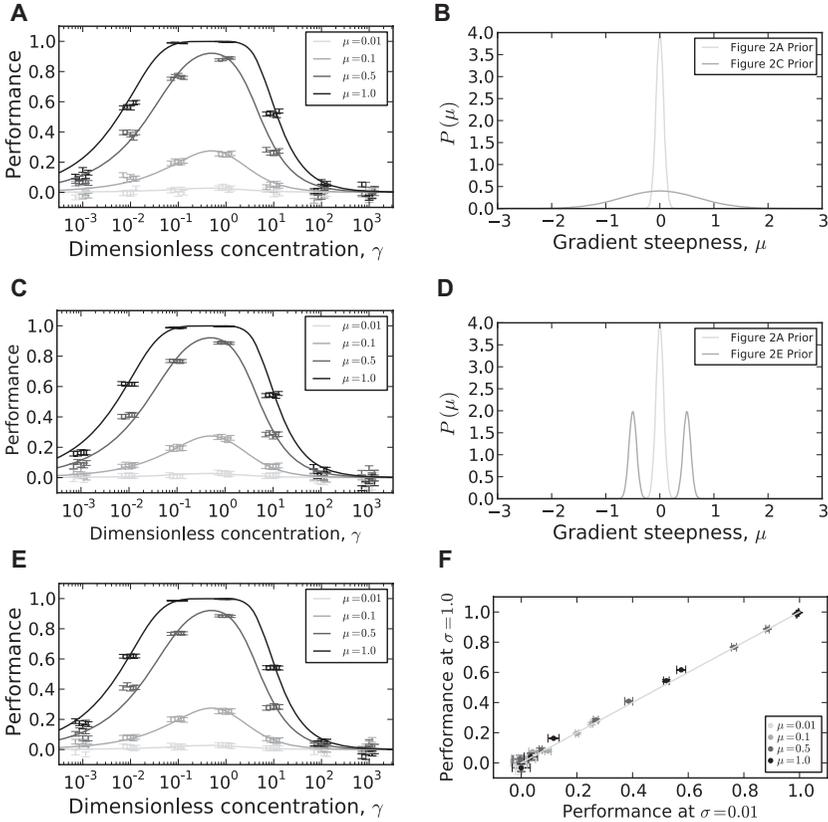
Figure 2: Comparison of the Monte Carlo estimation for the performance of the optimal strategy to the analytic approximation for the performance of the shallow gradient strategy. (A, C, E) Chemotactic sensitivity as a function of concentration for $P(\mu)$. (A) $P(\mu)$ normally distributed with variance $\sigma^2 = 0.01$. (C) $P(\mu)$ normally distributed with variance $\sigma^2 = 1.0$. (E) $P(\mu)$ an equal mixture of normal distributions at $\mu = \pm0.5$, each with variance $\sigma^2 = 0.01$. In each case, curves are shown for four different values of μ, the actual gradient steepness presented. The curves were calculated analytically using the shallow gradient approximation for ΔP. For each gradient steepness and concentration, 1000 binding distributions were generated, and for each such distribution, ΔP was calculated using five independent Markov chain Monte Carlo runs (see section 2). Each data point shows the result of one such run, with error bars shown according to the expected standard error in the mean assuming calculated ΔP was correct. Data points in panels A, C, and E are horizontally offset for visibility; in each case, the center point is at the true concentration value. Panel F illustrates that there is no apparent systematic dependence of ΔP on the prior distribution for μ.
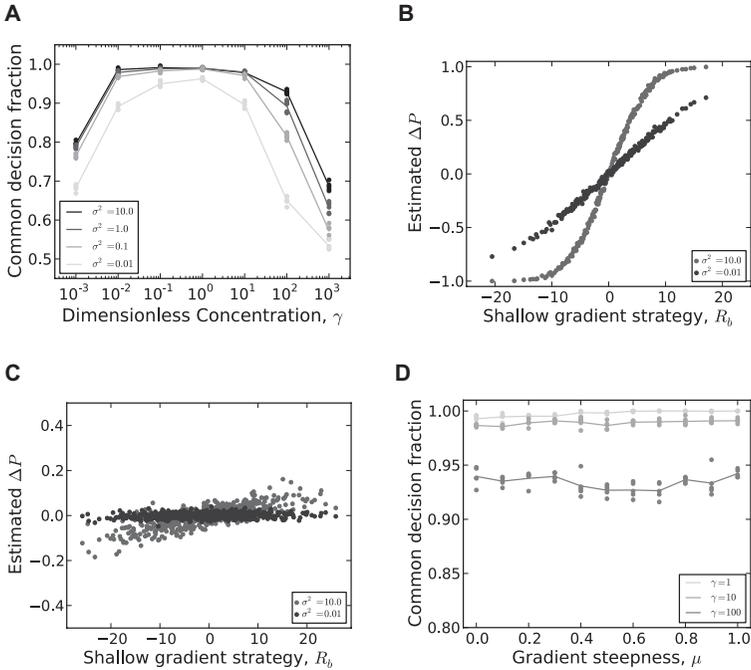
Figure 3: Difference in decision making between the analytic shallow-gradient strategy and the estimated Bayes-optimal strategy. (A) Common decision fraction between the shallow gradient and estimated Bayes optimal strategies as a function of concentration γ. The different curves show different values for the variance σ^2 of the prior gradient steepness distribution. The actual gradient steepness presented in each case was $\mu = 0.01$. There is a drop in similarity of decision making at high and low concentrations and noticeably less agreement for the gradient steepness prior with the smallest variance ($\sigma^2 = 0.01$). Panels B and C compare R_b (i.e., the sum of the positions of the bound receptors, which determines the shallow gradient strategy) to the corresponding Markov chain Monte Carlo estimate of ΔP for 1000 different sets of receptor positions for (B) $\gamma = 1$ and (C) $\gamma = 1000$, with $\sigma^2 = 0.01$, $\sigma^2 = 10.0$. (D) Common decision fraction as a function of μ for $\gamma = 1, 10, 100$, with the mixed prior in Figure 2D. The common decision ratio is close to one for $\gamma = 1$ but drops for higher γ. However, for each γ, the common decision ratio is largely independent of the gradient steepness μ, even when μ differs substantially from $\mu = 0.5$, the steepness expected according to the prior distribution.

Counterintuitively, agreement between the strategies was lower for the smallest variance in the gradient-steepness prior ($\sigma^2 = 0.01$; Figure 3A). One would expect that when the prior gradient steepness was concentrated closer to the origin, the full Bayesian strategy and approximate

shallow-gradient strategy would agree more closely. Figures 3B and 3C plot the Monte Carlo estimate of ΔP against R_b and illustrate why this reduction in agreement occurs. When the signs of these quantities are the same, the shallow-gradient strategy and estimated optimal strategy agree. It is evident in these plots that for smaller values of σ^2, the slope of the ΔP versus R_b line also decreases near $R_b = 0$. For extremely shallow gradients, the binding probability on the growth cone is essentially uniform, and measuring the difference in binding probability by Monte Carlo simulations can become a very noisy process. This means that for lower values of σ^2, there are comparatively more cases in which R_b and the estimate of ΔP differ in sign. For extreme values of γ (see Figure 3C), the slope of the ΔP versus R_b line is even further reduced, magnifying this effect.

Next, we investigated the case when the expected magnitude of $P(\mu)$ is no longer centered at zero, that is, the growth cone is expecting gradients of a specific steepness. In particular we now took $P(\mu)$ to be the sum of two gaussians centered at ±0.5 with variance 0.01 (see Figure 3D). For $\gamma = 1$, the common decision fraction was close to unity over a range of μ, though again this fraction dropped as γ increased beyond 1.

The surprising agreement in performance between the first-order approximation and the full calculation can be explained by noticing that even without the shallow gradient approximation, the term ΔP almost always has the same sign as R_b, except for extreme values of γ or μ. For any value of the gradient steepness, we can expand the expectation value of R_b, $E[R_b]$, as follows:

$$E[R_b] = E\left[\sum_{i=1}^{n} r_i b_i\right] = NE\left[rE[b]\right] = NE\left[r\frac{\gamma \exp(\mu r)}{1 + \gamma \exp(\mu r)}\right] \quad (3.1)$$

$$= NE\left[\frac{r\gamma}{\gamma + 1} + \frac{\mu \gamma r^2}{(\gamma + 1)^2} + \frac{\mu^2 \gamma (1 - \gamma)r^3}{2(\gamma + 1)^3}\right.$$

$$\left. + \frac{\mu^3 \gamma (\gamma^2 - 4\gamma + 1)r^4}{6(\gamma + 1)^4} + \cdots\right]. \quad (3.2)$$

Since we have assumed $P(r)$ to be a uniform distribution between $[-1/2; 1/2]$, in the expression above, all the terms containing an odd function of r will disappear after the integration. Meanwhile, the higher even-order moments of r become increasingly small ($E[r^2] = 1/12$, $E[r^4] = 1/80$, $E[r^6] = 1/448$), and in practice, μ is unlikely to be larger than 1 (corresponding to a doubling in concentration over 10 microns). Therefore, we can ignore any terms higher than the second order of r, yielding

$$E[R_b] \approx NE\left[\frac{r\gamma}{\gamma + 1}\right]. \quad (3.3)$$

Similarly, the variance can be approximated as

$$\text{Var}[R_b] \approx \frac{\gamma}{\gamma + 1} NE[r^2].$$

(3.4)

We can then approximate the performance as

$$\Delta P \approx sign(\mu)\text{erf}\left(\frac{E[R_b]}{\sqrt{2Var[R_b]}}\right)$$

(3.5)

$$\approx |\mu|\sqrt{\frac{N}{6\pi}}\sqrt{\frac{\gamma}{(\gamma + 1)^3}}$$

(3.6)

and thus

$$P_{correct} \approx \frac{1}{2}(1 + \Delta P) = \frac{1}{2} + \mu\sqrt{\frac{N}{24\pi}}\sqrt{\frac{\gamma}{(\gamma + 1)^3}}.$$

(3.7)

This analysis in the general case leads to the same formula for the performance as the first order approximation in equation 2.9, thus explaining the robustness of the analytically derived SNR expression. It is also evident that when μ is extremely shallow, or γ too high or too low, the SNR becomes very small. The Monte Carlo computation of the probability difference then shows large fluctuations, leading to the poor agreement between R_b and ΔP as demonstrated above.

3.2 Measured Variation in Response with Gradient Steepness. The theoretically derived chemotactic sensitivity (see equation 2.9) predicts that the chemotactic response should decline linearly to zero with gradient steepness. However, this prediction has not yet been explicitly tested experimentally, and previous data hinted that there could instead be a threshold gradient steepness below which the response drops to zero (see Figure 1D). To investigate the variation in guidance response with gradient steepness, rat DRGs were grown in precisely controlled gradients of NGF with a concentration of 0.2 nM at the explants and steepnesses varying from 0 to 0.24% change in concentration per 10 microns, using more controlled conditions than in previous work (see section 3). Typical explants are shown in Figures 4A and 4B. Total neurite outgrowth from the explants was roughly constant across different gradient steepnesses (see Figure 4C). The overall variation in chemotactic response (measured by the guidance ratio, GR; see section 3) with steepness is shown in Figure 4D. Also shown for comparison is the roughly corresponding data from Mortimer et al. (2009). The response first increases roughly linearly with gradient steepness, agreeing with the prediction of the Bayesian model. In particular there was no indication of a
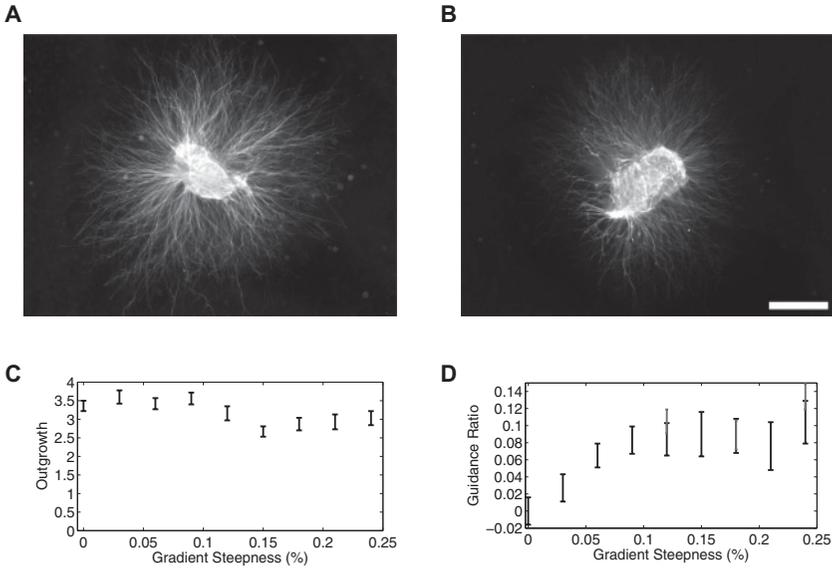
Figure 4: Experimental analysis of decline in chemotactic response with gradient steepness. (A, B) Example DRG explants after two days of growth in precisely controlled NGF gradients (concentration increasing upward). (A) GR = 0.0. (B) GR = 0.1 (scale bar = 500 μm). (C) There is little variation in total neurite outgrowth with gradient steepness. The experiments were performed in two groups at different times (0.03%–0.12% and 0.15%–0.24%, defined as fractional change over 10 microns), which accounts for the small difference between these two groups. (D) The GR increases linearly with gradient steepness and then saturates. $n = 40$–83 explants per condition. p-values compared to control are 0.1 (0.03%) and 0.001 (0.06%), and the p value for 0.03% compared to 0.06% is 0.04 (one-tailed t-tests). p values for all other conditions compared to control are less than 0.004, except for 0.21%, which is 0.024. Also plotted are the roughly corresponding data points from the study of Mortimer et al. (2009). The correspondence is only approximate since none of the data points in Mortimer et al. (2009) were at exactly 0.2 nM.

threshold. Beyond a steepness of $\approx$0.1%, however, the response saturated and did not increase with steepness.

Why does this saturation occur? One potential explanation is that in a gradient steepness of $\approx$0.1% over 10 microns, all axons are making guidance decisions with close to 100% fidelity, so that increases in gradient steepness confer no further improvement. This initially appears implausible due to the tiny change in concentration, and thus high noise in measurement, that exists across the width of a growth cone in a 0.1% gradient. However, recent evidence (Mortimer, Pujic, et al., 2010; Thompson

et al., 2011) suggests that the way in which axons convert a graded pattern of receptor binding into directed movement is fundamentally different in shallow gradients as compared to steep gradients (e.g., ≈10% across 10 microns, as exists in the standard growth cone turning assay (Zheng, Felder, Conner, & Poo, 1994; Song, Ming, & Poo, 1997). In particular, axons turn in steep gradients, but in shallow gradients they modulate their growth rate so that they move faster when pointed up the gradient than down the gradient (Mortimer, Pujic, et al., 2010). Turning requires a comparison of concentrations across the width of the growth cone, while growth rate modulation requires a comparison between the tip of the axon and some point farther down the axon shaft, potentially much farther than the width of a growth cone, leading to a more reliable decision. Furthermore the two-day timescale of our experiments allows multiple measurements to be averaged, again increasing the reliability.

To investigate further the plausibility of this as an explanation for saturation, we simulated axonal growth, from explants using a simple growth rate modulation model similar to that of Mortimer, Pujic, et al. (2010), where neurites were able to compare concentrations along their entire length. In particular, we assumed that the amount grown at each time step by each neurite is slightly larger (smaller) when it has decided (based on the pattern of bound receptors along its length) that it is growing up (down) the gradient. The percentage change in concentration along the length of a neurite is then of order 10%, which, combined with the large number of receptors, potentially available, leads to highly reliable measurements. However, we assume that changes in growth rate are nonetheless constrained by the limited resources for growth available to the neurite. As in Mortimer, Pujic, et al. (2010) neurites also made a small turn at each time step, but the direction of this turn was random. Figure 5 shows plots of the guidance ratio versus gradient steepness and guidance ratio versus concentration for these simulations, which match those seen experimentally. Although this does not of course definitively establish that this is the mechanism underlying saturation in our experiments, it provides at least one possible explanation.

4 Discussion

By extending the previous theoretical and experimental work of Mortimer et al. (2009, 2011), and Mortimer, Pujic, et al. (2010), we have shown that the optimal strategy for chemotaxis approximated for shallow gradients remains valid for very steep gradients, that the predicted linear increase in chemotactic response with concentration matches with the behavior of real neurites, that the chemotactic response saturates rapidly with gradient steepness in our shallow-gradient assay, and that this saturation can be explained within the context of a growth rate modulation model of the chemotactic response.
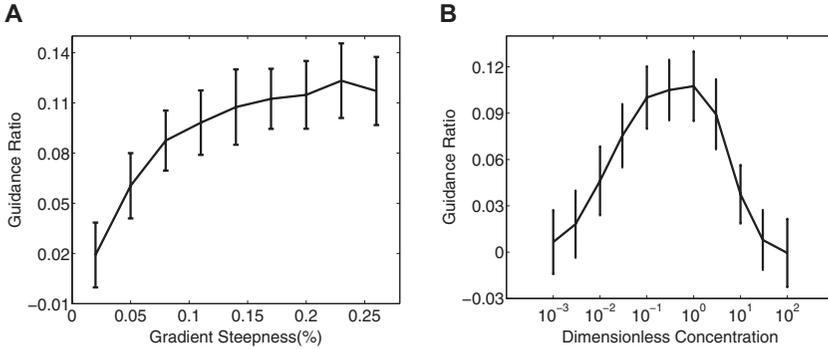
A

B

Figure 5: Modeling saturation. (A) Guidance ratio of 100 simulated explants per condition as a function of gradient steepness ($\gamma = 1$), showing saturation similar to that we observed experimentally. Neurites responded to the gradient by growth rate modulation. (B) Guidance ratio as a function of concentration for the same model (steepness = 0.14%), reproducing the general shape measured in Mortimer et al. (2009).

An interesting consequence of the robustness of our Bayesian framework to (symmetric) variations in the priors is that this implies that similar results would be obtained from a maximum likelihood (ML) approach. In recent work, Hu, Chen, Rappel, and Levine (2011) combined ML with an information-theoretic framework to derive estimates for the accuracy of gradient sensing. This yielded a formula for the SNR very similar to our equation 2.10:

$$SNR \approx \mu \sqrt{\frac{N\gamma}{8(\gamma + 1)^2}},$$

though this result was not compared to experimental data. Working within the full Bayesian framework maintains the generality required to investigate situations in which the growth cone has a biased (asymmetric) expectation for the gradient, as Hu et al. (2011) have also explored.

It is remarkable that the approximation to the performance (see equation 2.9) derived for shallow gradients is almost indistinguishable from the performance of the Monte Carlo estimation of the exact Bayesian strategy, even at very high gradient steepnesses and for prior distributions of the gradient steepness incorporating steep gradients (see Figures 2C and 2E). In the model we used, background concentration was defined to be the concentration at the center of the growth cone and was assumed to vary exponentially across the width of the growth cone. However, one could imagine alternative formulations consistent with a linear gradient at low

steepnesses. For example, we could instead consider a more or less sharp boundary between two different concentrations, with a linear transition region (with the boundary not necessarily occurring at $x = 0$). Such a situation might be more appropriate for modeling a growth cone's response to short-range or contact-mediated guidance cues. In this case, a steep gradient would correspond to a very sharp boundary, while a shallow gradient would be associated with a gradual transition from low to high concentration (perhaps over a distance larger than that of the growth cone width). We would still expect the shallow gradient strategy to correspond to that derived in Mortimer et al (2009, 2011); however, in steep gradients, the growth cone might perform better by searching for a sharp transition in the density of receptor binding.

At low concentrations and high gradient steepnesses, we did observe slightly lower performance than that predicted by equation 2.9. This observation can be explained by noting that equation 2.9 is valid only when a sufficiently large number of receptors are bound for the central limit theorem to hold. This assumption begins to break down at low concentrations. In particular, equation 2.9 then tends to understimate the level of noise in the growth cone's estimate of the gradient direction.

While our experimental confirmation of a linear decline in the chemotactic response as the gradient steepness decreased to zero provides further support for the Bayesian model, it does not strongly privilege the Bayesian explanation over other models, since presumably some alternative models would also predict (at least approximately) a linear relationship for very low steepness. The more interesting finding from a biological perspective is the apparent lack of an absolute threshold below which gradient detection is impossible, which could have been an alternative outcome of the experiment. This suggests that at least at a concentration of roughly K_D, there is no step in the signaling cascade that has a fixed absolute threshold. However, our data could still be consistent with absolute thresholds that vary between different neurites: since our measure of chemotactic response averages over many neurites, the smooth decline to zero could represent the combined effect of many all-or-nothing switches if these thresholds vary across the neurite population. Although the response we measured for a 0.03% gradient is close to exactly half that of a 0.06% gradient, the p-value for the 0.03% gradient response is only 0.1 compared to control. Thus, it is still possible that a sharp threshold might exist somewhere below 0.06% steepness. This issue is discussed in more detail in Fuller, Chen, Adler, Groisman, Levine, and Rappel (2010), where the authors compute the mutual information between the gradient and the spatial distribution of bound receptors and the mutual information between the input gradient and the motility direction of a chemotacting cell. They suggest that for shallow gradients, the information loss due to downstream signaling is insignificant. One way a threshold could potentially arise in our model is due to the finite precision with which a growth cone could know the positions of bound receptors. Unfortunately,

given the high degree of natural variability displayed in the responses, the n-values required to investigate this would be extremely large. An additional possibility is that a threshold could appear if the growth cone were compelled to make a decision regarding gradient direction on a significantly shorter timescale than that of the experimental assay used here.

One possibility for the saturation we observed is that only a subset of neurites is competent to respond to the gradient (i.e., express the NGF receptor TrkA). TrkA receptor levels are known to decline with age in embryonic and early postnatal rats (Phillips & Armanini, 1996; Molliver & Snider, 1997), and at the age of the rats we used for these experiments, only about 40% of neurites express TrkA (though it is possible that in our cultures, non-TrkA positive neurons will already have died after 2 days in vitro due to a lack of trophic support). Thus, if 60% of neurites are unaffected by the gradient, one would expect that the GR could not exceed a certain level, even if the remaining 40% of neurites all grew straight up the gradient. However, simulation results based on the neurite growth models of Xu, Rosoff, Urbach, and Goodhill (2005) and Mortimer, Pujic, et al. (2010) suggest that the upper bound on GR this implies is much larger than that observed experimentally (data not shown). Instead, here we showed that saturation could naturally arise from the assumption that gradient detection on a scale of an entire neurite, as might occur in a growth rate modulation mechanisms (Mortimer, Pujic, et al., 2010), leads to high gradient detection fidelity even in shallow gradients. One potential experimental test would be to grow neurites in compartmentalized cultures. Here the cell body is maintained at a different ligand concentration from the growth cone, producing a step-like gradient, which could potentially modulate neurite growth rate depending on the magnitude of the concentration difference between the two compartments.

However, to more thoroughly address these questions requires new in vitro technologies for examining the long-term growth of neurites, both dissociated and in the form of explants, in gradients that can be varied in steepness over a broad range and precisely controlled for long periods of time. While technologies satisfying all these constraints do not currently exist, a promising possible direction is microfluidics (Whitesides, 2006). Although a number of microfluidic approaches have been developed for cells in general (reviewed in Kim, Kim, & Jeon, 2010), the delicacy of neurons, and especially their growth cones, provides a particular challenge (Wang et al., 2008; Bhattacharjee, Li, Keenan, & Folch, 2010). However, these types of new technologies, combined with theories such as we have presented here, have the potential to greatly increase our understanding of axon guidance by gradients in vitro and ultimately also in vivo.

Acknowledgments

References

Andrews, B. W., & Iglesias, P. A. (2007). An information-theoretic characterization of the optimal gradient sensing response of cells. *PLoS Comp. Biol.*, *3*, e153.

Berg, H. C., & Purcell, E. M. (1977). Physics of chemoreception. *Biophysical Journal*, *20*, 193–219.

Bhattacharjee, N., Li, N., Keenan, T. M., & Folch, A. (2010). A neuron-benign microfluidic gradient generator for studying the response of mammalian neurons towards axon guidance factors. *Integr. Biol.*, *2*, 669–679.

Bialek, W., & Setayeshgar, S. (2005). Physical limits to biochemical signaling. *Proc. Natl. Acad. Sci. USA*, *102*, 10040–10045.

Endres, R. G., & Wingreen, N. S. (2008). Accuracy of direct gradient sensing by single cells. *Proc. Natl. Acad. Sci, USA*, *105*, 15749–15754.

Fuller, D., Chen, W., Adler, M., Groisman, A., Levine, H., & Rappel, W.-J. (2010). External and internal constraints on eukaryotic chemotaxis. *Proc. Natl. Acad. Sci, USA*, *107*, 9656–9659.

Hu, B., Chen, W., Rappel, W.-J., & Levine, H. (2011). How geometry and internal bias affect the accuracy of eukaryotic gradient sensing. *Phys. Rev. E.*, *83*, 021917.

Hu, B., Rappel, W-J., & Levine, H. (2010). Physical limits on cellular sensing of spatial gradients. *Phys. Rev. Lett.*, *105*, 048104.

Isbister, C. M., Mackenzie, P. J., To, K. C., & O'Connor, T. P. (2003). Gradient steepness influences the pathfinding decisions of neuronal growth cones in vivo. *J. Neurosci.*, *23*, 193–202.

Kennedy, T. E., Wang, H., Marshall, W., & Tessier-Lavigne, M. (2006). Axon guidance by diffusible chemoattractants: A gradient of netrin protein in the developing spinal cord. *J. Neurosci.*, *26*, 8866–8874.

Kim, S., Kim, H. J., & Jeon, N. L. (2010). Biological applications of microfluidic gradient devices. *Integr. Biol.*, *2*, 584–603.

Legg, A. T., & O'Connor, T. P. (2003). Gradients and growth cone guidance of grasshopper neurons. *Journal of Histochemistry and Cytochemistry*, *51*, 445–454.

McCaig, C. D. (1986). Electric fields, contact guidance and the direction of nerve growth. *J. Embryol. Exp. Morph.*, *94*, 245–255.

Molliver, D. C., & Snider, W. D. (1997). Nerve growth factor receptor TrkA is downregulated during postnatal development by a subset of dorsal root ganglion neurons. *J. Comp. Neurol.*, *381*, 428–438.

Mortimer, D., Dayan, P., Burrage, K., & Goodhill, G. J. (2010). Optimizing chemotaxis by measuring unbound-bound transitions. *Physica D*, *239*, 477–484.

Mortimer, D., Dayan, P., Burrage, K., & Goodhill, G. J. (2011). Bayes-optimal chemotaxis. *Neural Comput.*, *23*, 336–373.

Mortimer, D., Feldner, J., Vaughan, T., Vetter, I., Pujic, Z., Rosoff, W. J., et al. (2009). A Bayesian model predicts the response of axons to molecular gradients. *Proc. Natl. Acad. Sci., USA*, *106*, 10296–10301.

Mortimer, D., Fothergill, T., Pujic, Z., Richards, L. J., & Goodhill, G. J. (2008). Growth cone chemotaxis. *Trends Neurosci.*, *31*, 90–98.

Mortimer, D., Pujic, Z., Vaughan, T., Thompson, A., Feldner, J., Vetter, I., & Goodhill, G. J. (2010). Axon guidance by growth-rate modulation *Proc. Natl. Acad. Sci, USA*, *107*, 5202–5207.

Phillips, H. S., & Armanini, M. P. (1996). Expression of the Trk family of neurotrophin receptors in developing and adult dorsal root ganglion neurons. *Phil. Trans. R. Soc. Lond. B*, *351*, 413–416.

Rosoff, W. J., McAllister, R. G., Esrick, M. A., Goodhill, G. J., & Urbach, J. S. (2005). Generating controlled molecular gradients in 3D gels. *Biotechnology and Bioengineering*, *91*, 754–759.

Song, H. J., Ming, G. L., & Poo, M-m. (1997). cAMP-induced switching in turning direction of nerve growth cones. *Nature*, *388*, 275–279.

Thompson, A. W., Pujic, Z., Richards, L. J., & Goodhill, G. J. (2011). Cyclic nucleotide-dependent switching of mammalian axon guidance depends on gradient steepness. *Mol. Cell Neurosci.*, *47*, 45–52.

Ueda, M., & Shibata, T. (2007). Stochastic signal processing and transduction in chemotactic response of eukaryotic cells. *Biophys. J.*, *93*, 11–20.

Vetter, I., Pujic, Z., & Goodhill, G. J. (2010). Response of dorsal root ganglion axons to NGF gradients depends on spinal level. *J. Neurotrauma*, *27*, 1379–1386.

von Philipsborn, A., & Bastmeyer, M. (2007). Mechanisms of gradient detection: A comparison of axon pathfinding with eukaryotic cell migration. *Int. Rev. Cytology*, *263*, 1–62.

Wang, C. J., Li, X., Lin, B., Shim, S., Ling, G-L., & Levchenko, A. (2008). A microfluidics-based turning assay reveals complex growth cone responses to integrated gradients of substrate-bound ECM molecules and diffusible guidance cues. *Lab on a Chip*, *8*, 227–237.

Whitesides, G. M. (2006). The origins and future of microfluidics. *Nature*, *442*, 368–373.

Xu, J., Rosoff, W. J., Urbach, J. S., & Goodhill, G. J. (2005). Adaptation is not required to explain the long-term response of axons to molecular gradients. *Development*, *132*, 4545–4552.

Zheng, J. Q., Felder, M., Conner, J. A., & Poo, M-m. (1994). Turning of growth cones induced by neurotransmitters. *Nature*, *368*, 140–144.