Chemotaxis in uncertain environments: Hedging bets with multiple receptor types

Austin Hopkins[1] and Brian A. Camley[2]

[1]*Department of Physics & Astronomy, Johns Hopkins University, Baltimore, Maryland 21218, USA*
[2]*Department of Physics & Astronomy, Department of Biophysics, Johns Hopkins University, Baltimore, Maryland 21218, USA*

Eukaryotic cells are able to sense chemical gradients in a wide range of environments. We show that, if a cell is exposed to a highly variable environment, it may gain chemotactic accuracy by expressing multiple receptor types with varying affinities for the same signal, as found commonly in chemotaxing cells like Dictyostelium. As environment uncertainty is increased, there is a transition between cells preferring a single receptor type and a mixture of types—hedging their bets against the possibility of an unfavorable environment. We predict the optimal receptor affinities given a particular environment. In chemotaxing, cells may also integrate their measurement over time. Surprisingly, time integration with multiple receptor types is qualitatively different from gradient sensing by a single type—cells may extract orders of magnitude more chemotactic information than expected by naive time integration. Our results show when cells should express multiple receptor types to chemotax, and how cells can efficiently interpret the data from these receptors.

I. INTRODUCTION

As a white blood cell finds a wound, or an amoeba finds nutrition, they chemotax, sensing and following chemical gradients. Eukaryotic cells sense gradients of chemical ligands by measuring ligand binding to receptors on the cell's surface. In eukaryotic chemotaxis in shallow gradients, accuracy is limited by unavoidable stochasticity arising from randomness in ligand-receptor binding and diffusion [1–4]. Eukaryotic chemotaxis has been extensively modeled [1,5–8], including extensions to collective chemotaxis [9–12] and stochastic simulation [13,14]. Modeling and experiment show eukaryotic chemotaxis is most accurate at ligand concentrations near the receptor dissociation constant K_D [2,3,5]. Cells crawling through tissue and searching for targets at variable concentrations are exposed to a huge variation in environmental signals. Cells often express multiple receptors for the same signal, with K_D values ranging over orders of magnitude [15,16]. For instance, during Dictyostelium's life cycle, Dicty expresses multiple different combinations of cAMP receptors CAR1-CAR4 [17], with ranges of K_D from 25 nM to >5000 nM. Larger-K_D receptors are expressed later in development, when the cAMP background level rises; the change in receptor expression has been suggested to allow Dicty to deal with the new environments [18,19]. In addition, Segota *et al.* recently showed that to explain the high accuracy of Dictyostelium chemotaxis to folic acid over a broad range of folic acid concentrations, multiple receptor types (with K_D values ranging from 2 to 450 nM [15]) and multiple measurements over time must be accounted for [4].

We argue that if a cell is sufficiently uncertain about its chemical environment, it should express multiple receptor types. We provide results for the optimal receptor K_D depending on environmental uncertainty. In addition, we show that integrating information from multiple measurements of the receptor binding state is more complicated in the many-receptor-type case, and show that a standard approach significantly underestimates gradient sensing accuracy.

II. MODEL AND RESULTS

We generalize the model of [5,20], considering a cell with receptors of R types with dissociation constants K_D^i, $i = 1 \cdots R$, with receptors spread evenly over a circular cell. We find the fundamental limit set by the Cramér-Rao bound with which this cell can measure a shallow gradient **g** using a snapshot of current receptor occupation. If the concentration near the cell is locally $c(\mathbf{r}) = c_0 e^{\mathbf{r} \cdot \mathbf{g}/L}$, i.e., g is the percentage change across the cell diameter L, this uncertainty is (Appendix B)

$$\sigma_{\mathbf{g}}^2 = \sigma_{g_x}^2 + \sigma_{g_y}^2 = \frac{16}{N \sum_{i=1}^{R} f_i \frac{c_0 K_D^i}{(c_0 + K_D^i)^2}} \quad \text{(snapshot)}, \quad (1)$$

where N is the total receptor number, and f_i is the fraction of receptors that are type i.

When can a cell improve its accuracy by expressing multiple receptor types? If we try to maximize signal-to-noise ratio SNR $= g^2/\sigma_g^2$, we see from Eq. (1) that SNR is a linear function of the f_i, and will be maximized by choosing $f = 1$ for the type with the largest value of $\frac{c_0 K_D^i}{(c_0 + K_D^i)^2}$, and $f = 0$ for the others. For the simplest case of two types A and B with dissociation constants K_A, K_B ($K_B > K_A$) this means that the best accuracy occurs with all A receptors when $c_0 < \sqrt{K_A K_B}$, and with all B receptors when $c_0 > \sqrt{K_A K_B}$ [Fig. 1(a)].
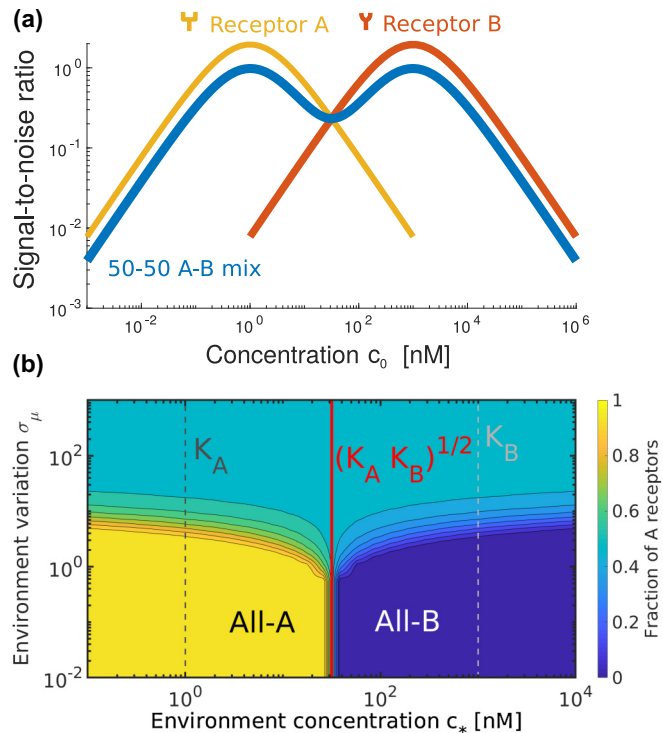
(a)

(b)

FIG. 1. Cells benefit from mixed receptor expression when their environment is uncertain. (a) The signal-to-noise ratio g^2/σ_g^2 from Eq. (1) for 100% A receptors (yellow), 100% B receptors (red), or both in a 50-50 combination (blue). (b) Fraction of A receptors that maximizes $\overline{\text{CI}}$ for snapshot measurements [Eq. (1)] as a function of $p(c_0)$ where $p(\ln c_0) \sim e^{-(\ln c_0 - \ln c_*)^2/2\sigma_\mu^2}$, characterized by the environment concentration c_* and the standard deviation of the log of the concentration σ_μ. The transition between all-A and all-B occurs at $\sqrt{K_A K_B}$ at low uncertainty. Dashed lines indicate receptor K_D. $K_A = 1$ nM, $K_B = 1000$ nM, $N = 5 \times 10^4$, and $g = 0.05$ in both plots.

If the background concentration is completely known, it is never beneficial for a cell to express multiple receptor types simultaneously. However, if a cell is uncertain about the concentration it is likely to encounter, it may hedge its bets by expressing multiple receptor types, allowing it to chemotax effectively in more environments. Does a cell in concentration c_0 with probability $p(c_0)$ benefit from multiple receptor types? What metric is appropriate? We could compute average signal-to-noise ratio $\overline{\text{SNR}} = \int dc_0 p(c_0) g^2/\sigma_g^2$, but increasing SNR at one concentration is little consolation to the cell that finds itself completely lost at another concentration—the utility of SNR saturates. We therefore optimize the mean chemotactic index $\overline{\text{CI}}$, where $\text{CI} = f(\text{SNR})$, with $f(x)$ saturating at 1 as $x \to \infty$. Here we use $f(x) = \sqrt{2x/\pi}/L_{1/2}(-x/2)$, where $L_{1/2}(x)$ is a generalized Laguerre polynomial, as in [21]; alternate definitions of CI lead to similar results.

In Fig. 1 we consider two receptor types with dissociation constants K_A and K_B, and numerically determine the fraction of A receptors that maximizes $\overline{\text{CI}}$. We choose $p(c_0)$ to be log-normal, $p(\ln c_0) \sim e^{-(\ln c_0 - \ln c_*)^2/2\sigma_\mu^2}$—a generic option for large variability. When environmental uncertainty σ_μ is small, we see the behavior predicted above—at small c_*, the cell should express all-A receptors, while for $c_* > \sqrt{K_A K_B}$ the cell

switches to all-B. However, at larger σ_μ, cells optimize $\overline{\text{CI}}$ by expressing equal amounts of A and B receptors [Fig. 1(b)].

Why is a 50-50 mix optimal even when $c_* \approx K_A$? This may seem like a natural response to uncertainty, but it is not obvious why, if the typical concentration c_* is close to K_A, the cell would not prefer A-type receptors. We argue that in the limit of large σ_μ, where $p(c_0)$ becomes very broad, but $\text{CI}(c_0)$ is locally peaked, the optimal fraction should not depend on c_* or σ_μ. The chemotactic index $\text{CI}(c_0)$ for snapshot sensing is generally peaked when c_0 is around K_A or K_B, because the SNR decays when $c_0 \ll K_A$ or $c_0 \gg K_B$ [Fig. 1(a)]. As σ_μ increases and $p(c_0)$ becomes more weakly dependent on c_0, we can approximate the integral defining $\overline{\text{CI}}$ as

$$\overline{\text{CI}} = \int dc_0 p(c_0) \text{CI}(c_0)$$

$$\approx p(\sqrt{K_A K_B}) \int dc_0 \text{CI}(c_0). \qquad (2)$$

(We have chosen $\sqrt{K_A K_B}$ as a typical value in the range $K_A \cdots K_B$.) In Eq. (2), the parameters c_* and σ_μ of the environment distribution $p(c_0)$ only appear in $p(\sqrt{K_A K_B})$, and the receptor fractions only appear in the term $\int dc_0 \text{CI}(c_0)$. In this limit, $p(\sqrt{K_A K_B})$ becomes an irrelevant prefactor—the same fraction will optimize $\overline{\text{CI}}$ independent of c_* and σ_μ, and so we see a 50-50 mix for a broad range of parameters. We will see later that the 50-50 mix is no longer optimal when cells time average and $\text{CI}(c_0)$ is no longer locally peaked.

The 50-50 mix between A and B receptors in Fig. 1 emerges when $p(c_0)$ is so broad it is slowly varying on the scale of $\text{CI}(c_0)$. We caution that at these large levels of environmental variation, the difference between the optimal receptor configuration and simply choosing all-A or all-B receptors is small (Appendix F)—at sufficiently high uncertainties, no configuration is particularly successful.

Figure 1 shows when a cell should choose to express a combination of receptor types. Can we also find which receptors a cell would evolve to maximize gradient-sensing ability in a given $p(c_0)$? We optimize $\overline{\text{CI}}$ by varying K_D^i and receptor fraction f_i for different numbers of receptor types R and different widths σ_μ, holding total receptor number N constant. We choose the configuration that maximizes $\overline{\text{CI}}$ with an important caveat. By adding more types with arbitrary K_D, we can always at least match the performance of a single type. If many configurations generate roughly the same near-optimal $\overline{\text{CI}}$ (all within $\Delta_{\text{CI}} = 0.01$), we choose from these the configuration with the fewest receptor types R. To reduce the number of variables we vary, we use the symmetry of $p(c_0)$, assuming $\ln K_D^i$ and f_i are mirror symmetric around $\ln c^*$. The resulting optimal K_D^i are shown in Fig. 2. We see that as the environment uncertainty σ_μ increases, there is a transition between preferring a single receptor type and multiple receptor types, with the K_D values for the multiple types being spread over the likely range of concentrations observed.

Figures 1 and 2 are based on Eq. (1), which gives the fundamental uncertainty for a cell sensing a gradient *from a single snapshot of its receptors*. If cells integrate measurements over time [3,5,6,20,22,23], they can improve gradient sensing. Our results in Appendix B give the estimator $\hat{\mathbf{g}}$ for the gradient vector $\mathbf{g}$ given the snapshot data. Defining
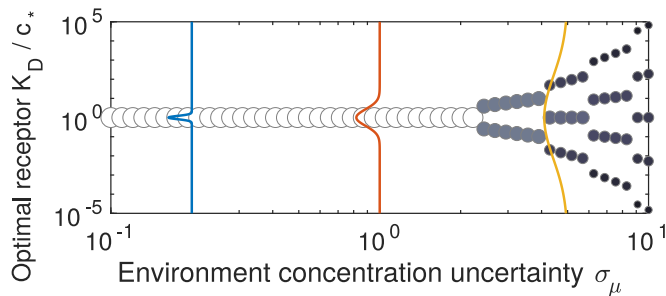
FIG. 2. Optimal receptor configuration. Best K_D^i values for receptors as a function of environmental uncertainty σ_μ. Marker areas are scaled to the fraction of that receptor type. Solid lines illustrate $p(\ln c_0)$. $c_* = 1$ nM, $N = 5 \times 10^4$ and $g = 0.05$, $\Delta_{\text{CI}} = 0.01$, and a maximum of $R = 7$ types are considered.

a time-integrated estimator $\hat{\mathbf{g}}_T = \frac{1}{T} \int_0^T \hat{\mathbf{g}}(t) dt$, as in earlier work [5,10,20], we find its variance $\sigma_{g,T}^2 = \langle |\mathbf{g}_T - \mathbf{g}|^2 \rangle$. To do this, we must consider the kinetics of binding and unbinding at each receptor, which happens with a receptor correlation time of $\tau_i = 1/(k_-^i + c_0 k_+^i)$. In the limit of large averaging times $T \gg \tau_i$, we find (Appendix C)

$$\sigma_{g,T}^2 = \frac{32 \sum_i^R f_i \beta_i \tau_i}{NT \left(\sum_i^R f_i \beta_i \right)^2} \quad \text{(naive average)}, \quad (3)$$

where $\beta_i = c_0 K_D^i / (c_0 + K_D^i)^2$ reflects the accuracy of measuring by type i—maximized at $c_0 = K_D^i$. For a single type, error is reduced by the effective number of measurements $T/2\tau$, $\sigma_{g,T}^2 = \sigma_g^2 \times 2\tau/T$ [5,20]. Equation (3) can be cast in a similar form as

$$\sigma_{g,T}^2 = \sigma_g^2 \sum_i^R \alpha_i \frac{2\tau_i}{T},$$

where $\alpha_i = \beta_i f_i / \sum_i^R \beta_i f_i$ is the weight given to receptor type i. This equation shows that the variance—in a naive time average—is reduced by a weighted sum that depends on the receptor correlation times. In the single receptor type case, $\sigma_{g,T}^2$ becomes arbitrarily small as $\tau \to 0$. However, because $\sigma_{g,T}^2$ is proportional to a weighted *sum* of τ_i/T, when receptor correlation times τ_i decrease, error is limited by the slowest correlation time. If one receptor correlation time is significantly faster than another, Eq. (3) predicts reduced error merely by removing the slow receptors (Fig. 3). The naive time average, therefore, does not efficiently use the information available—if it did, the cell would not be able to gain accuracy by throwing away measurements. The core reason for the failure of naive time averaging is that the snapshot estimator weights receptors equally, which is appropriate to the amount of information they provide *at that moment*. Naively averaging this estimator weighs information from fast receptors (which gain more information as T increases) and slow receptors (which gain less information) similarly.

The failure of naive time averaging is reminiscent of a well-known result for concentration sensing: it is not optimal for a single receptor to estimate c from a simple time average of its occupation $T^{-1} \int_0^T n(t) dt$. If the whole history of binding and unbinding events is used in a maximum
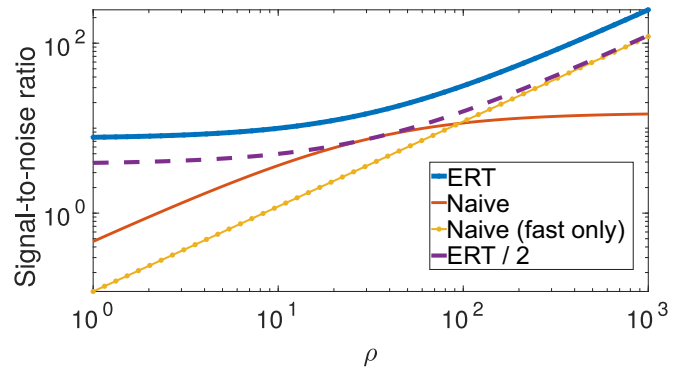

FIG. 3. Estimation using the entire receptor trajectory is much more accurate than naive averaging. SNR g^2/σ_g^2 as a function of $\rho = k_-^B / k_-^A$; $\rho > 1$ states B receptors have faster off rates. SNR is much larger using the entire receptor trajectory (ERT) method; even in its best case, the naive average only reaches half of the ERT SNR. We use two receptor types, $K_A = 1$ nM, $K_B = 1000$ nM and $N = 5 \times 10^4$, $g = 0.05$, with fixed $c_0 = \sqrt{K_A K_B}$, and $f_A = 0.5$, with $k_-^A T = 2$.

likelihood estimate, the error σ_c^2 is reduced by two [24]. We compute the accuracy limit for gradient sensing $\sigma_{g,T}^2$ using the entire receptor trajectory (Appendix D), finding (again in the limit $T \gg \tau_i$)

$$\sigma_{g,T;\text{ERT}}^2 = \frac{16}{N \sum_i^R f_i \beta_i \frac{T}{\tau_i}} \quad \text{(entire receptor trajectory)} \quad (4)$$

or more intuitively

$$\sigma_{g,T;\text{ERT}}^2 = \sigma_g^2 \frac{1}{\sum_i^R \alpha_i \frac{T}{\tau_i}}.$$

For a single receptor type, Eq. (4) is a factor of 2 smaller than the naive time average Eq. (3), precisely as in concentration sensing. However, for multiple types, ERT error can be orders of magnitude better, as the time correlation factors τ_i/T add "in parallel"—error is no longer limited by the slowest type.

We illustrate the differences between these two errors in Fig. 3, computing SNR g^2/σ_g^2 for two receptor types. Which type provides more information depends on the relative off rates of the two types:

$$\text{SNR}_{\text{ERT}} = \frac{Ng^2 k_-^A T}{16} \left[f_A \frac{c_0}{c_0 + K_A} + (1 - f_A)\rho \frac{c_0}{c_0 + K_B} \right], \quad (5)$$

where $\rho = k_-^B / k_-^A$.

Figure 3 shows that as ρ is varied, the naive time average SNR is always *at least* a factor of 2 lower than ERT. When the B receptor off rate is large ($\rho \gg 1$), the naive average is worse than if only the $N/2$ B receptors were used [the "Naive (fast only)" yellow dotted line]. In this limit, most information is from the B receptors, and using only B receptors reaches half the ERT SNR (Fig. 3).

How does time averaging affect bet hedging? For a single environmental concentration, all-A is optimal when Eq. (5) increases with increasing f_A, i.e., $c_0/(c_0 + K_A) > \rho c_0/(c_0 + K_B)$, or $c_0 < c_{\text{bal}} \equiv \frac{K_B - K_A \rho}{\rho - 1}$. For $\rho \leqslant 1$, it is always best to use the lower-K_D receptor, but trade-offs are more complex when B receptors are faster. The balancing point c_{bal} varies from $c_{\text{bal}} \to \infty$ at $\rho = 1$ to $c_{\text{bal}} = 0$ at $\rho = K_B/K_A$. $\rho = K_B/K_A$
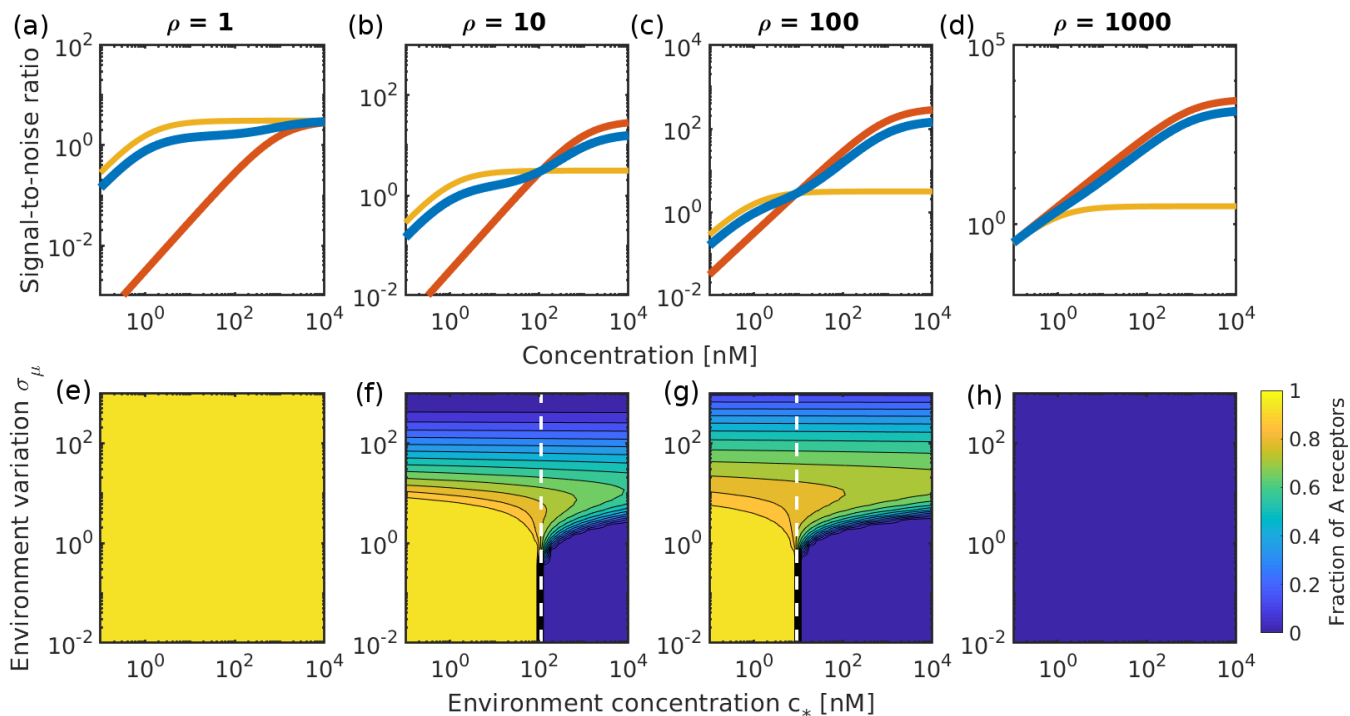
FIG. 4. Trade-off between two receptor types depends on their dynamic properties. Top row: (a)–(d) SNR calculated for differing values of ρ at different fixed concentrations c_0 using Eq. (5). As in Fig. 1, yellow indicates all-A, red all-B, and blue a 50-50 mix. Bottom row: (e)–(h) Optimal receptor fraction with time averaging, with ρ corresponding to the figure in the top row. Dashed white lines show the analytical c_{bal}. Parameters are as in Fig. 1 except $N = 10^4$ (chosen as cells have larger SNR with fewer receptors when time averaging) and $k_-^A T = 2$.

corresponds to the condition where the on rates of the two types are equal, $k_+^A = k_+^B$, as $k_-^B/k_-^A = K_B/K_A = \frac{k_-^B}{k_-^A}\frac{k_+^A}{k_+^B}$, this would be the case if the on rates were diffusion limited.

Dependence on receptor off rates is preserved when we study optimal receptor configurations in an uncertain environment $p(\ln c_0) \sim e^{-(\ln c_0 - \ln c_*)^2/2\sigma_\mu^2}$. In Fig. 4 we show how the optimal share of A receptors depends on $\rho = k_-^B/k_-^A$. Even with significant uncertainty, when $\rho = 1$, all-A is optimal. However, at higher values of ρ, a transition between all-A and all-B like Fig. 1 occurs when $c_* = c_{\text{bal}}$ (white dashed line). By contrast with the snapshot results, in the time average at large uncertainties $\sigma_\mu > 1$, the 50-50 mixture is not optimal. Instead, at high uncertainties, for $\rho = 10, 100$, optimal receptor fractions are similar above and below c_{bal}, and the fraction of A receptors exceeds 0.5, with f_A decreasing as $\sigma_\mu \gtrsim 10$. Why? When cells time average, CI is nonzero over a broad range of c_0 (see Fig. 4, top row), and our argument from the snapshot case fails.

Can we produce a plot such as Fig. 2 showing the optimal receptor configuration if time averaging is performed? No. When time averaging, error is minimized by making the correlation times as small as possible—taking $k_- \to \infty$. We cannot find a consistent set of optimal receptors unless k_- is restricted by some biochemical constraint. This is because, as recognized for concentration sensing [24], only binding rates are sensitive to c—bound times should be minimized.

III. DISCUSSION

Our results show that cells can hedge their bets against an uncertain environment by expressing multiple receptor types,

but that this behavior is only reasonable if the uncertainty in c_0 spans the range of observed K_D (Fig. 2), i.e., if cells typically explore environments where concentration varies over orders of magnitude. The idea that signal processing should be adapted to the likely range of concentrations is similar to classical results showing information transmission is maximized by tuning input-output relationships to the input probability distribution [25–28]. When cells chemotax to hunt bacteria, as Dictyostelium uses folic acid chemotaxis, it is intuitively plausible that observed c_0 span orders of magnitude, as bacterial hunting must function over both sparse and concentrated solutions, and over many distances to bacteria. However, at any fixed concentration, expressing multiple receptor types is always suboptimal to choosing the receptor that best fits your current concentration c_0—so hedging is plausible in circumstances when the environment is uncertain *over the timescale on which the receptor affinity is fixed*. What other timescales could appear in the problem? Dictyostelium receptors are internalized in response to large increases in cAMP concentration [29,30], but this process takes several minutes—much longer than it would typically take Dicty to chemotax to a new mean concentration level. We also show that even if receptor numbers change in different environments, we see very similar results (Appendix E).

Our work so far has not distinguished between truly different receptors and receptors that are phosphorylated or otherwise modified to change their K_D [31,32], which play a role in adaptation to different signal levels in bacteria [33]. If receptor modification is fast compared to the environment's change in concentration c_0, i.e., can occur before the cell samples a new concentration from $P(c_0)$, hedging will be less

effective. Our work thus suggests interesting future directions for extension, based on recent studies of concentration sensing in time-varying environments [34]. Future work could also consider multiple ligand types, which could also limit accuracy by competing for receptors [35,36]. For cells to reach the lower bound of Eq. (4), they must compute estimates with some reaction network, possibly extending recent work showing how to compute the ERT estimate in concentration sensing [36,37]. One important difference here is that finding the ERT $\hat{g}$ requires spatially resolved measurements of bound and unbound times separately for each receptor type. Computation of the ERT estimate for concentration requires additional free energy expenditure [37]—it would be interesting to determine if the extravagant benefits of the ERT approach for gradient sensing with multiple receptor types (Fig. 3) comes with a commensurate cost. Although these are significant complexities, the huge gap between the fundamental bound of Eq. (4) and the naive average of Eq. (3) shows even very rough approximations to the ERT provide significant gains over a naive average.

NOTATION FOR APPENDICES

This Appendix includes calculations where we have to distinguish the type i and the index n of each receptor. The intermediate calculations are more complex than the final results in the main paper, and to help keep the details straight, we will denote receptor types i with superscripts and receptor indices n with subscripts. A superscript to a power i like C_n^i never denotes exponentiation. To keep consistent with this, we have included formulas with β^i where in the main text we have written β_i.

APPENDIX A: NUMERICAL DETAILS

To evaluate the integral $\overline{\text{CI}} = \int dc_0 p(c_0)\text{CI}(c_0)$ numerically, we found some difficulties that arise because of how broad $p(c_0)$ is. In particular, because the SNR most naturally varies on the log scale [see Fig. 1(a)], it is easiest for us to evaluate this expectation in terms of $\ln c_0$,

$$\overline{\text{CI}} = \int d\ln c_0 p(\ln c_0)\text{CI}(e^{\ln c_0}). \tag{A1}$$

For our log-normal distribution, $p(\ln c_0)$ has the simple Gaussian form

$$p(\ln c_0) = \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp\left[-\frac{(\ln c_0 - \ln c_*)^2}{2\sigma_\mu^2}\right]. \tag{A2}$$

However, we note that even Eq. (A1) can be tricky to evaluate when CI is nonzero only for a range of $\ln c_0$ compared with the scale σ_μ. We evaluate this integral using Matlab's Gauss-Kronrod quadrature (quadgk), setting waypoints to ensure that all nonzero ranges of both functions should be included.
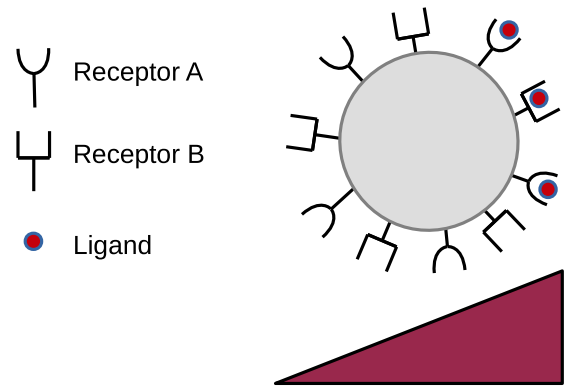

FIG. 5. Illustration of data used for gradient estimate using a snapshot of receptor state; only two receptor types are illustrated.

To find the receptor type fractions that optimize $\overline{\text{CI}}$ we use Matlab's fminbnd (golden section search) and fminsearch (Nelder-Mead), depending on the number of variables to be optimized over.

Code to reproduce the results in the paper can be found at [45].

APPENDIX B: SNAPSHOT SENSING

How precisely can a cell make a measurement of a chemical gradient using only its current information about which receptors on its surface are occupied? We extend results from [5,20] on this accuracy to include multiple receptor types. We assume that the cell is in a shallow exponential gradient with direction ϕ and steepness $g = \frac{L|\nabla C|}{C_0}$, where L is the diameter of the cell and C_0 is the concentration at the cell center. The gradient $\mathbf{g}$ can also be written as $\mathbf{g} = (g_x, g_y) = [g\cos(\phi), g\sin(\phi)]$. The concentration at a cell receptor with angular coordinate φ can then be written as

$$C(\varphi) = C_0 \exp\left[\frac{g}{2}\cos(\varphi - \phi)\right]. \tag{B1}$$

Let there be R receptor types, where there are N^i receptors of type i and $N = \sum_i^R N^i$ total receptors. We describe the receptors as being uniformly spread across the cell, with angular positions φ_n^i (Fig. 5). Then each receptor of type i can be represented as a Bernoulli trial, i.e., we define a variable x_n^i that is one if a receptor n of type i is occupied, and zero otherwise. The probability of $x_n^i = 1$ is the probability of that receptor being occupied,

$$P_n^i = C_n^i/(C_n^i + K_D^i), \tag{B2}$$

where C_n^i is the concentration at the nth receptor of type i and K_D^i is the dissociation constant of receptor type i. The dissociation constant K_D^i is the ratio of unbinding and binding rates of the receptors of type i, i.e., $K_D^i = \frac{k_-^i}{k_+^i}$. Assuming that all the receptors are independent of one another, we have the following likelihood function giving the probability of seeing receptor occupations x_n^i given gradient $\mathbf{g}$,

$$\mathcal{L}(\mathbf{g}|x_1^0,\ldots,x_{N^R}^R) = \prod_i^R \prod_n^{N^i} \left[(P_n^i)^{x_n^i}(1-P_n^i)^{1-x_n^i}\right].$$

The log-likelihood function is

$$
\begin{aligned}
\ln\mathcal{L} &= \sum_i^R \sum_n^{N^i} \left[x_n^i \ln\left(\frac{C_n^i}{C_n^i + K_D^i} \right) + (1 - x_n^i)\ln\left(\frac{K_D^i}{C_n^i + K_D^i} \right) \right] \\
&= \sum_i^R \sum_n^{N^i} \left[x_n^i \ln\left(\frac{C_n^i}{K_D^i} \right) + \ln\left(\frac{K_D^i}{C_n^i + K_D^i} \right) \right].
\end{aligned}
\tag{B3}
$$

In the second term in this equation, we then assume that the receptors are numerous enough that we can replace the sum over receptor position by a continuous integral $\sum_n^{N^i} \rightarrow \frac{N^i}{2\pi} \int_0^{2\pi} d\varphi$:

$$
\begin{aligned}
\ln\mathcal{L} = \sum_i^R \Bigg\{ &\sum_n^{N^i} \left[\frac{1}{2} x_n^i g \cos(\varphi_n^i - \phi) + \ln\left(\frac{C_0}{K_D} \right) x_n^i \right] \\
&+ \int_0^{2\pi} \frac{N^i}{2\pi} \ln\left(\frac{K_D^i}{C_0 \exp\left[\frac{g}{2}\cos(\varphi - \phi) \right] + K_D^i} \right) d\varphi \Bigg\}.
\end{aligned}
\tag{B4}
$$

We define $z_1^i = \sum_n^{N^i} x_n^i \cos\varphi_n^i$ and $z_2^i = \sum_n^{N^i} x_n^i \sin\varphi_n^i$, which measure the spatial asymmetry in the occupancy of receptors of type i. $Z_1 = \sum_i z_1^i$ and $Z_2 = \sum_i z_2^i$ measure the total spatial asymmetry in receptor occupancy of the cell. In a shallow exponential gradient, we can neglect terms $O(g^4)$ and higher for an estimation of the gradient $\vec{g} = (g_x, g_y)$. Then, the log-likelihood function becomes

$$
\begin{aligned}
\ln\mathcal{L} = \sum_i^R \Bigg[&\frac{g_x z_1^i + g_y z_2^i}{2} + \ln\frac{C_0}{K_D^i}\sum_n^{N^i} x_n^i \\
&- \frac{NC_0 f^i K_D^i (g_x^2 + g_y^2)}{16(C_0 + K_D^i)^2} + \ln\frac{C_0}{C_0 + K_D^i} \Bigg].
\end{aligned}
\tag{B5}
$$

(We note that past papers with similar derivations [12,20] have not always written the last term in this log-likelihood, which is an irrelevant constant). Taking the derivative with respect to $g_{x,y}$ gives

$$
\frac{\partial}{\partial g_{x,y}} \ln\mathcal{L} = \sum_i^R \left[\frac{z_{1,2}^i}{2} - \frac{NC_0 f^i K_D^i g_{x,y}}{8(C_0 + K_D^i)^2} \right].
\tag{B6}
$$

Because the log function is monotonic, we can set Eq. (B6) equal to zero to find $\hat{g}_x$ and $\hat{g}_y$, the parameters which for g_x and g_y which maximize the likelihood function. Carrying out this procedure we find

$$
\sum_i^R \left[\frac{z_{1,2}^i}{2} - \frac{NC_0 f^i K_D^i}{8(C_0 + K_D^i)^2}\hat{g}_{x,y} \right] = 0,
\tag{B7}
$$

$$
\sum_i^R \frac{z_{1,2}^i}{2} = \sum_i^R \left[\frac{NC_0 f^i K_D^i}{8(C_0 + K_D^i)^2}\hat{g}_{x,y} \right],
\tag{B8}
$$

$$
\frac{Z_{1,2}}{2} = \frac{NC_0}{8} \sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2}\hat{g}_{x,y} \right],
\tag{B9}
$$

which be solved for $\hat{g}_{x,y}$ to determine estimators $\hat{g}_{x,y}$ as

$$
\hat{g}_x = \frac{4Z_1}{NC_0 \sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2} \right]} \quad \text{and} \quad \hat{g}_y = \frac{4Z_2}{NC_0 \sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2} \right]}.
\tag{B10}
$$

To determine the asymptotic variance on these estimators, we will need to compute the second derivative of the log-likelihood function. Applying an additional derivative to Eq. (B6) gives

$$
\begin{aligned}
\frac{\partial^2}{\partial g_{x,y}^2} \ln\mathcal{L} &= -\frac{NC_0}{8} \sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2} \right], \\
\frac{\partial^2}{\partial g_x \partial g_y} \ln\mathcal{L} &= 0.
\end{aligned}
\tag{B11}
$$

From the log-likelihood function, we can also determine the Fisher information matrix, which controls the best possible measurement that the cell can make of the uncertain $\mathbf{g}$ [5,38]. In this case it is diagonal, and its inverse gives the variances of $\hat{g}_x$ and $\hat{g}_y$ in the limit of many samples. As a result, we have expressions for the asymptotic variances for $\hat{g}_x$ and $\hat{g}_y$,

$$
\begin{aligned}
\frac{1}{\sigma_{g_{x,y}}^2} &= \left\langle \left(\frac{\partial \ln\mathcal{L}}{\partial g_{x,y}} \right)^2 \right\rangle = -\left\langle \frac{\partial^2 \ln\mathcal{L}}{\partial g_{x,y}^2} \right\rangle \\
&= \frac{NC_0}{8} \sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2} \right],
\end{aligned}
$$

and so

$$
\sigma_{g_{x,y}}^2 = \frac{8}{NC_0 \sum_i^R \frac{f^i K_D^i}{(C_0 + K_D^i)^2}}.
\tag{B12}
$$

The important parameter is $\sigma_{\mathbf{g}}^2$, which is just the sum of the component variances

$$
\sigma_{\mathbf{g}}^2 = \sigma_{g_x}^2 + \sigma_{g_x}^2 = \frac{16}{NC_0 \sum_i^R \frac{f^i K_D^i}{(C_0 + K_D^i)^2}}.
\tag{B13}
$$

As the sample size becomes large, the distribution of $\hat{g}_{1,2}$ converges to a normal distribution with means $g_{x,y}$ and variance σ_{g_x,g_y}^2. This also implies that the mean values of Z_1 and Z_2 are

$$
\langle Z_{1,2} \rangle = \frac{1}{4}NC_0 \sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2} \right] g_{x,y}.
\tag{B14}
$$

APPENDIX C: NAIVE TIME AVERAGING

A cell may improve its estimation of the gradient by time averaging. In the previous Appendix we determined an estimator $\hat{g}$ that is the best estimate of a cell's gradient, given a snapshot of its receptor information. Naively, a cell could improve its accuracy by making a measurement over a time T and determining the average of these estimates

$$
\hat{\mathbf{g}}_T = \frac{1}{T} \int_0^T dt \ \hat{\mathbf{g}}(t).
\tag{C1}
$$

Then the variance of this new estimator will be reduced,

$$\sigma_{g,T}^2 = \langle |\hat{\mathbf{g}}_T|^2 \rangle - \langle \hat{\mathbf{g}}_T \rangle^2$$

$$= \frac{1}{T^2} \int_0^T dt \int_0^T ds [\langle \hat{\mathbf{g}}(s) \cdot \hat{\mathbf{g}}(t) \rangle - \langle \hat{\mathbf{g}} \rangle^2]. \quad \text{(C2)}$$

To understand how time averaging improves the cell's sensing accuracy, we need to compute $\langle \hat{\mathbf{g}}(s) \cdot \hat{\mathbf{g}}(t) \rangle$, the correlation function of $\hat{\mathbf{g}}$. This correlation function is related to the correlation functions in the estimates of each component

of the gradient as

$$\langle \hat{\mathbf{g}}(s) \cdot \hat{\mathbf{g}}(t) \rangle = \langle \hat{g}_x(s)\hat{g}_x(t) \rangle + \langle \hat{g}_y(s)\hat{g}_y(t) \rangle. \quad \text{(C3)}$$

And, by Eq. (B10), the correlation functions for $g_{x,y}$ can be related to the correlation functions for $Z_{1,2}$ as

$$\langle \hat{g}_{x,y}(s)\hat{g}_{x,y}(t) \rangle = \frac{16\langle Z_{1,2}(s)Z_{1,2}(t) \rangle}{N^2 C_0^2 \left(\sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2} \right] \right)^2}. \quad \text{(C4)}$$

The correlation functions for Z_1 can be written in terms of the single receptor correlation function:

$$\langle Z_1(s)Z_1(t) \rangle = \left\langle \left(\sum_i^R \left[\sum_n^{N^i} x_n^i(s)\cos(\varphi_n^i) \right] \right) \left(\sum_i^R \left[\sum_n^{N^i} x_n^i(t)\cos(\varphi_n^i) \right] \right) \right\rangle \quad \text{(C5)}$$

$$= \left\langle \sum_i^R \sum_j^R \sum_n^{N^i} \sum_m^{N^j} x_n^i(s)\cos(\varphi_n^i) x_m^j(t)\cos(\varphi_m^j) \right\rangle \quad \text{(C6)}$$

$$= \sum_i^R \sum_j^R \sum_n^{N^i} \sum_m^{N^j} \langle x_n^i(s)x_m^j(t) \rangle \cos(\varphi_n^i)\cos(\varphi_m^j). \quad \text{(C7)}$$

The kinetics of receptor binding and unbinding with multiple receptor types can be quite complicated [39,40], with ligands potentially diffusing from one receptor to another. However, if the binding and unbinding process is slow with respect to this diffusion, i.e., binding is reaction limited, as is believed to be the case in eukaryotic chemotaxis [20,39], it is appropriate to think of the ligand-receptor binding having two states—one bound and one with ligand in the bulk. Then, there are two relevant rates, that of ligand binding to a receptor of type i exposed to concentration C is $k_+^i C$ and the off rate is k_-^i—which results in an exponential single receptor correlation function

$$\langle x_n^i(s)x_n^i(t) \rangle = \sigma_{x_n^i}^2 e^{-|t-s|/\tau_n^i} + \langle x_n^i \rangle^2 \quad \text{(C8)}$$

for receptor n of type i. This limit is also appropriate if all ligand is internalized, as discussed by [24]. The parameter $\sigma_{x_n^i}^2$ characterizes the fluctuations in the occupancy of the receptor,

and is given by

$$\sigma_{x_n^i}^2 = \frac{C_n^i K_D^i}{(C_n^i + K_D^i)^2}, \quad \text{(C9)}$$

the variance of a Bernoulli trial. τ_n^i is the single receptor correlation time

$$\tau_n^i = 1 / (k_-^i + C_n^i k_+^i) \quad \text{(C10)}$$

in the reaction-limited case. (Generalization to other limits is possible but not straightforward [40–42]). Because different receptors are independent, the mean of their product is just the product of their means

$$\langle x_n^i(s)x_m^j(t) \rangle = \langle x_n^i(s) \rangle \langle x_m^j(t) \rangle \quad \text{if } i \neq j \text{ or } n \neq m. \quad \text{(C11)}$$

Using Eq. (C8) for terms where $i = j$ and $n = m$ and Eq. (C11) otherwise, we can expand the correlation function of Z_1 in Eq. (C7) as

$$\langle Z_1(s)Z_1(t) \rangle = \sum_i^R \sum_j^R \sum_n^{N^i} \sum_m^{N^j} \langle x_n^i(s)x_m^j(t) \rangle \cos(\varphi_n^i)\cos(\varphi_m^j) \quad \text{(C12)}$$

$$= \sum_i^R \sum_n^{N^i} \sigma_{x_n^i}^2 e^{-|t-s|/\tau_n^i} \cos^2(\varphi_n^i) + \sum_i^R \sum_j^R \sum_n^{N^i} \sum_m^{N^j} \langle x_n^i(s) \rangle \langle x_m^j(t) \rangle \cos(\varphi_n^i)\cos(\varphi_m^j). \quad \text{(C13)}$$

The second term in Eq. (C13) is $\langle Z_1 \rangle^2$, which can be solved as

$$\langle Z_1 \rangle^2 = \frac{1}{16} N^2 C_0^2 \left(\sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2} \right] \right)^2 g_x^2 \quad \text{(C14)}$$

by Eq. (B14). For the first term in Eq. (C13), taking the sum to an integral, gives

$$\sum_i^R \sum_n^{N^i} \sigma_{x_n^i}^2 e^{-|t-s|/\tau_n^i} \cos^2(\varphi_n^i) = \frac{NC_0}{2} \sum_i^R f^i \frac{K_D^i}{(C_0 + K_D^i)^2} e^{-|t-s|/\tau^i} + O(g^2), \quad \text{(C15)}$$

where $\tau^i = 1/(k_-^i + C_0 k_+^i)$, i.e., Eq. (C10) for a receptor in the ambient concentration C_0. Therefore, for shallow gradients, the correlation function for Z_1 is

$$\langle Z_1(s) Z_1(t) \rangle = \frac{NC_0}{2} \sum_i^R f^i \frac{K_D^i}{(C_0 + K_D^i)^2} e^{-|t-s|/\tau^i}$$
$$+ \frac{1}{16} N^2 C_0^2 \left(\sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2} \right] \right)^2 g_x^2. \quad (C16)$$

Using the relation in Eq. (C4), the correlation function of the estimator $\hat{g}_1$ can be found from Eq. (C16):

$$\langle \hat{g}_1(s) \hat{g}_1(t) \rangle = \frac{8 \sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2} e^{-|t-s|/\tau^i} \right]}{NC_0 \left(\sum_i^R \frac{f^i K_D^i}{(C_0 + K_D^i)^2} \right)^2} + g_x^2. \quad (C17)$$

Similar expressions for the correlation functions of Z_2 and $\hat{g}_2$ can be derived as

$$\langle Z_2(s) Z_2(t) \rangle = \frac{NC_0}{2} \sum_i^R f^i \frac{K_D^i}{(C_0 + K_D^i)^2} e^{-|t-s|/\tau^i}$$
$$+ \frac{1}{16} N^2 C_0^2 \left(\sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2} \right] \right)^2 g_y^2 \quad (C18)$$

and

$$\langle \hat{g}_2(s) \hat{g}_2(t) \rangle = \frac{8 \sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2} e^{-|t-s|/\tau^i} \right]}{NC_0 \left(\sum_i^R \frac{f^i K_D^i}{(C_0 + K_D^i)^2} \right)^2} + g_y^2. \quad (C19)$$

Thus, the correlation in $\hat{\mathbf{g}}$ for shallow gradients is

$$\langle \hat{\mathbf{g}}(s) \cdot \hat{\mathbf{g}}(t) \rangle = \frac{16 \sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2} e^{-|t-s|/\tau^i} \right]}{NC_0 \left(\sum_i^R \frac{f^i K_D^i}{(C_0 + K_D^i)^2} \right)^2} + g^2. \quad (C20)$$

This now gives us enough information to compute the time-averaged variance,

$$\sigma_{g,T}^2 = \frac{1}{T^2} \int_0^T dt \int_0^T ds [\langle \hat{\mathbf{g}}(s) \cdot \hat{\mathbf{g}}(t) \rangle - \langle \hat{\mathbf{g}} \rangle^2]. \quad (C21)$$

Then, using the result in Eq. (C20), we have

$$\sigma_{g,T}^2 = \frac{16 \sum_i^R \left[\frac{f^i K_D^i}{(C_0 + K_D^i)^2} \int_0^T dt \int_0^T ds e^{-|t-s|/\tau^i} \right]}{T^2 NC_0 \left(\sum_i^R \frac{f^i K_D^i}{(C_0 + K_D^i)^2} \right)^2} \quad (C22)$$

$$= 2\sigma_{\mathbf{g}}^2 \frac{\sum_i^R \frac{f^i K_D^i}{(C_0 + K_D^i)^2} \tau^i [T - \tau^i (1 - e^{-T/\tau^i})]}{T^2 \sum_i^R \frac{f^i K_D^i}{(C_0 + K_D^i)^2}}. \quad (C23)$$

In the limit where $T \gg \tau^i$ for all τ^i, Eq. (C23) becomes

$$\sigma_{g,T}^2 = 2\sigma_{\mathbf{g}}^2 \frac{\sum_i^R \frac{f^i K_D^i}{(C_0 + K_D^i)^2} \tau^i}{T \sum_i^R \frac{f^i K_D^i}{(C_0 + K_D^i)^2}} = \frac{32 \sum_i^R f^i \beta^i \tau^i}{NT \left(\sum_i^R f^i \beta^i \right)^2}, \quad (C24)$$
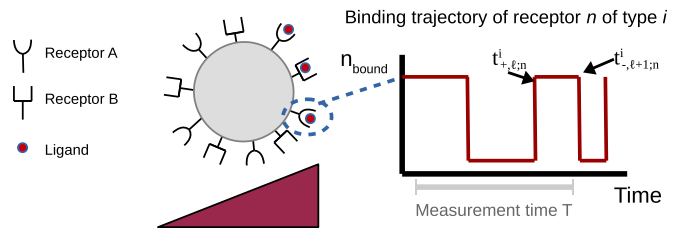

FIG. 6. Illustration of data used for maximum likelihood estimate from entire receptor trajectories (ERT).

where the parameter $\beta^i = C_0 K_D^i / (C_0 + K_D^i)^2$ reflects the accuracy of measuring only using receptor i, as in the main text.

APPENDIX D: MAXIMUM LIKELIHOOD USING ENTIRE RECEPTOR BINDING TRAJECTORY (ERT)

Instead of simply performing a naive average, a cell could also improve its sensing of the gradient by determining an estimate of the gradient from the history of its receptors over the measurement time—when they are bound and unbound (Fig. 6). We describe a time interval for receptor n of type i as time series $\{t_{+;n}^i, t_{-;n}^i\}$, where particles bind at times $t_{+;n,\ell}^i$ and unbind at times $t_{-;n,\ell}^i$, where ℓ indexes the binding and unbinding events. Following [24], we compute the probability for a time series of binding and unbinding events. Define a function

$$f_{-;n}^{*;i}\left(t_{-,\ell;n}^i\right) = p_{-,\ell;n}^i\left(t_{-,\ell;n}^i | t_{+,1;n}^i, t_{-,1;n}^i, \ldots, t_{-,\ell-1;n}^i, t_{+,\ell;n}^i\right),$$

which is the probability density for the event that receptor n of type i experiences an unbinding at time $t_{-,\ell;n}^i$ given the previous time series data $\{t_{+,1;n}^i, t_{-,1;n}^i, \ldots, t_{-,\ell-1;n}^i, t_{+,\ell;n}^i\}$. Here the time series has been written for a receptor that is initially unbound, and the indexing of the time series in the following equations will follow that notation. The procedure is the same for a receptor that starts in the bound state. Define the analogous function

$$f_{+;n}^{*;i}\left(t_{+,\ell;n}^i\right) = p_{+,\ell;n}^i\left(t_{+,\ell;n}^i | t_{+,1;n}^i, t_{-,1;n}^i, \ldots, t_{+,\ell-1;n}^i, t_{-,\ell-1;n}^i\right)$$

for binding events. Then the probability of observing a time series $\{t_{+;n}^i, t_{-;n}^i\}$ is given by

$$p\left(\{t_{+;n}^i, t_{-;n}^i\}\right) = \prod_\ell^{\eta_{b;n}^i} f_{-;n}^{*;i}\left(t_{-,\ell;n}^i\right) \prod_{\ell'}^{\eta_{u;n}^i} f_{+;n}^{*;i}\left(t_{+,\ell';n}^i\right), \quad (D1)$$

where $\eta_{b;n}^i$, $\eta_{u;n}^i$ are the numbers of binding events and unbinding events, respectively. If a cell measures for a time interval T that is long compared to the relevant timescales (i.e., $T \gg 1/k_-^i$, $T \gg 1/C_0 k_+^i$ for all receptor types), then $\eta_{b;n}^i \approx \eta_{u;n}^i$ because the number of binding events can differ by at most one from the number of unbinding events. In this limit, the information about the gradient is dominated by the observed time series, and not the initial snapshot state of the receptors.

We assume that we are in the reaction-limited case, where we can treat the rate of binding to a receptor of type i exposed to concentration C as $k_+^i C$ and the off rate as k_-^i—neglecting rebinding. (Neglecting rebinding, as discussed in more detail

by [24], is also the appropriate limit to find the fundamental bound to accuracy, as cells may prohibit rebinding by degrading or internalizing ligand.) The functions $f_{+;n}^{*;i}(t_{-,\ell;n}^i)$ and $f_{-;n}^{*;i}(t_{-,\ell;n}^i)$, with this simple Markovian kinetics, do not depend on the whole time series, they only depend on the time of previous unbinding/binding event:

$$f_{-;n}^{*;i}\left(t_{-,\ell;n}^i\right) = k_-^i e^{-k_-^i(t_{-,\ell;n}^i - t_{+,\ell-1;n}^i)}, \tag{D2}$$

$$f_{+;n}^{*;i}\left(t_{+,\ell;n}^i\right) = k_+^i C_n^i e^{-k_+^i C_n^i(t_{+,\ell;n}^i - t_{-,\ell;n}^i)}. \tag{D3}$$

These are probability density functions for exponential distributions with rates k_-^i and $k_+^i C_n^i$ for unbinding and binding, respectively. From these equations, we can determine the likelihood function for the gradient parameters g_x and g_y given the observed time series $\{t_{+;n}^i, t_{-;n}^i\}$ at receptors n of types i. Because each receptor is independent, the likelihood function is the product of the probability function in Eq. (D1):

$$\mathcal{L} = \prod_i^R \prod_n^{N^i} \left(\prod_\ell^{\eta_{b;n}^i} k_-^i e^{-k_-^i(t_{-,\ell;n}^i - t_{+,\ell-1;n}^i)} \prod_{\ell'}^{\eta_{u;n}^i} k_+^i C_n^i e^{-k_+^i C_n^i(t_{+,\ell;n}^i - t_{-,\ell';n}^i)} \right). \tag{D4}$$

Then, define the total time bound $\mathcal{T}_{b;n}^i$ and the total time unbound $\mathcal{T}_{u;n}^i$ for receptor n of type i (the indexing in these definitions assume the receptor starts unbound, but analogous

definitions can be written for a receptor that is bound at $t = 0$, and in the limit of long times treated here, this assumption does not matter):

$$\mathcal{T}_{b;n}^i = \sum_\ell^{\eta_{b;n}^i} t_{-,\ell;n}^i - t_{+,\ell;n}^i, \tag{D5}$$

$$\mathcal{T}_{u;n}^i = \sum_\ell^{\eta_{u;n}^i} t_{+,\ell;n}^i - t_{-,\ell-1;n}^i. \tag{D6}$$

With these definitions, the likelihood function in Eq. (D4) becomes

$$\mathcal{L} = \prod_i^R \prod_n^{N^i} (k_-^i)^{\eta_{b;n}^i} \left(k_+^i C_n^i \right)^{\eta_{u;n}^i} e^{-k_-^i \mathcal{T}_{b;n}^i} e^{-k_+^i C_n^i \mathcal{T}_{u;n}^i}.$$

The log-likelihood function is then

$$\ln\mathcal{L} = \sum_i^R \sum_n^{N^i} \left[\eta_{b;n}^i \ln(k_-^i) + \eta_{u;n}^i \ln\left(k_+^i C_n^i\right) \right.$$
$$\left. - k_-^i \mathcal{T}_{b;n}^i - k_+^i C_n^i \mathcal{T}_{u;n}^i \right]. \tag{D7}$$

Substituting the expression $C_n^i = C_0 \exp\{\frac{1}{2}[g_x\cos(\phi_n^i) + g_y \sin(\phi_n^i)]\}$ in Eq. (D7) gives

$$\ln\mathcal{L} = \sum_i^R \sum_n^{N^i} \left\{ \eta_{b;n}^i \ln(k_-^i) + \eta_{u;n}^i \ln\left(k_+^i C_0\right) + \eta_{u;n}^i \left[\frac{1}{2}\left[g_x\cos(\phi_n^i) + g_y\sin(\phi_n^i)\right] \right] \right.$$
$$\left. - k_-^i \mathcal{T}_{b;n}^i - k_+^i C_0 \exp\left[\frac{1}{2}\left[g_x\cos(\phi_n^i) + g_y\sin(\phi_n^i)\right] \right] \mathcal{T}_{u;n}^i \right\}. \tag{D8}$$

For shallow gradients, we can approximate the log-likelihood function by expanding to second order in the magnitude of the gradient. This results in

$$\ln\mathcal{L} \approx \sum_i^R \sum_n^{N^i} \left\{ \eta_{b;n}^i \ln(k_-^i) + \eta_{u;n}^i \ln\left(k_+^i C_0\right) + \eta_{u;n}^i \left[\frac{1}{2}\left[g_x\cos(\phi_n^i) + g_y\sin(\phi_n^i)\right] \right] - k_-^i \mathcal{T}_{b;n}^i \right.$$
$$\left. - k_+^i C_0 \left[1 + \frac{1}{2}\left[g_x\cos(\phi_n^i) + g_y\sin(\phi_n^i)\right] + \frac{1}{8}\left[g_x\cos(\phi_n^i) + g_y\sin(\phi_n^i)\right]^2 \right] \mathcal{T}_{u;n}^i \right\}. \tag{D9}$$

Differentiating Eq. (D9) with respect to g_x and g_y, we get

$$\frac{\partial}{\partial g_x}\ln\mathcal{L} = \sum_i^R \sum_n^{N^i} \left\{ \frac{\eta_{u;n}^i}{2}\cos(\phi_n^i) - k_+^i C_0 \left[\frac{1}{2}\cos(\phi_n^i) + \frac{1}{4}\left[g_x\cos(\phi_n^i) + g_y\sin(\phi_n^i)\right]\cos(\phi_n^i) \right] \mathcal{T}_{u;n}^i \right\}, \tag{D10}$$

$$\frac{\partial}{\partial g_y}\ln\mathcal{L} = \sum_i^R \sum_n^{N^i} \left\{ \frac{\eta_{u;n}^i}{2}\sin(\phi_n^i) - k_+^i C_0 \left[\frac{1}{2}\sin(\phi_n^i) + \frac{1}{4}\left[g_x\cos(\phi_n^i) + g_y\sin(\phi_n^i)\right]\sin(\phi_n^i) \right] \mathcal{T}_{u;n}^i \right\}. \tag{D11}$$

Because g_x and g_y here do not depend on i and n, Eq. (D11) can be equated to zero and solved to find the maximum likelihood estimator in terms of sums over functions of $\mathcal{T}_{u;n}^i$ and ϕ_n^i. However, we have not found the precise form very useful. From the derivatives of the log-likelihood function, we can compute the Fisher information matrix:

$$\mathcal{I}_{a,b} = -\left\langle \frac{\partial^2}{\partial a \partial b}\ln\mathcal{L} \right\rangle. \tag{D12}$$
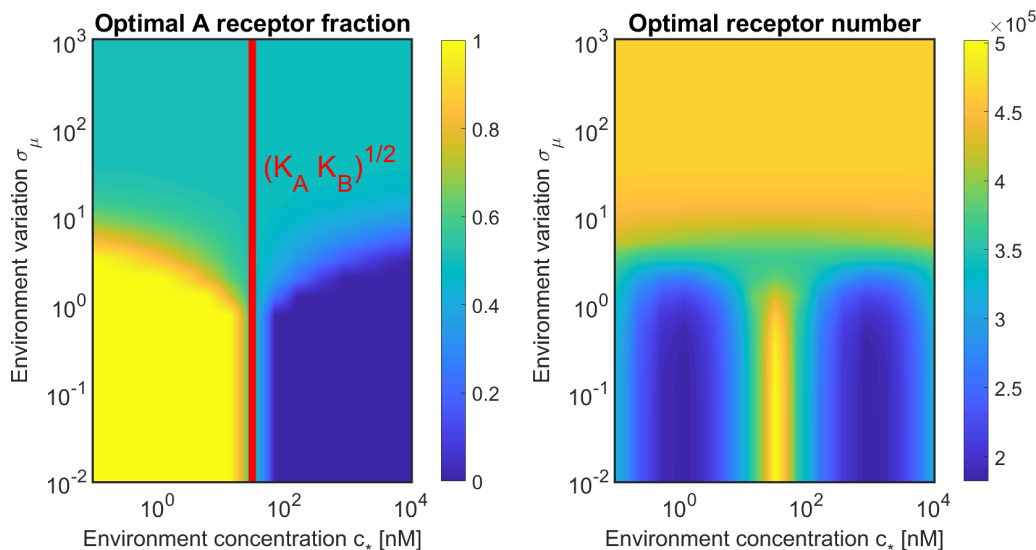
FIG. 7. Transition between all-*A* and all-*B* is preserved in a model variant where the number of receptors is allowed to vary. Parameters are the same as Fig. 1 in the main paper, except for the penalty for increasing the number of receptors, which is $N^{\text{penalty}} = 50N^{\text{basal}}$ (see text).

Differentiating Eqs. (D10) and (D11) with respect to combinations of g_x and g_y gives the following matrix:

$$\mathcal{I} = \begin{bmatrix} \sum_i^R \sum_n^{N^i} \frac{k_+^i C_0 \langle \mathcal{T}_{u;n}^i \rangle}{4} \cos^2(\phi_n^i) & \sum_i^R \sum_n^{N^i} \frac{k_+^i C_0 \langle \mathcal{T}_{u;n}^i \rangle}{4} \cos(\phi_n^i) \sin(\phi_n^i) \\ \sum_i^R \sum_n^{N^i} \frac{k_+^i C_0 \langle \mathcal{T}_{u;n}^i \rangle}{4} \cos(\phi_n^i) \sin(\phi_n^i) & \sum_i^R \sum_n^{N^i} \frac{k_+^i C_0 \langle \mathcal{T}_{u;n}^i \rangle}{4} \sin^2(\phi_n^i) \end{bmatrix}. \tag{D13}$$

The expectation value $\langle \mathcal{T}_{u;n}^i \rangle$ can be found in terms of the measurement time T and the probability P_n^i that a receptor is occupied [Eq. (B2)],

$$\langle \mathcal{T}_{u;n}^i \rangle = T(1 - P_n^i) = T\frac{K_d^i}{C_n^i + K_d^i}. \tag{D14}$$

By substituting Eq. (D14) into Eq. (D13) and taking the inner sums to an integral, we have the following expressions for each matrix element:

$$\sum_i^R \sum_n^{N^i} \frac{k_+^i C_0 \langle \mathcal{T}_{u;n}^i \rangle}{4} \cos^2(\phi_n^i) = \frac{1}{4} \sum_i^R \frac{N^i}{2\pi} \int_0^{2\pi} \frac{k_+^i C_0 T K_d^i}{\left\{ C_0 \exp\left[\frac{g}{2}\cos(\varphi - \phi)\right] + K_d^i \right\}} \cos^2(\phi) d\phi$$

$$= N \sum_i^R \frac{f^i k_-^i C_0 T}{8(C_0 + K_d^i)} + O(g^2), \tag{D15}$$

$$\sum_i^R \sum_n^{N^i} \frac{k_+^i C_0 \langle \mathcal{T}_{u;n}^i \rangle}{4} \sin^2(\phi_n^i) = \frac{1}{4} \sum_i^R \frac{N^i}{2\pi} \int_0^{2\pi} \frac{k_+^i C_0 T K_d^i}{\left\{ C_0 \exp\left[\frac{g}{2}\cos(\varphi - \phi)\right] + K_d^i \right\}} \sin^2(\phi) d\phi$$

$$= N \sum_i^R \frac{f^i k_-^i C_0 T}{8(C_0 + K_d^i)} + O(g^2), \tag{D16}$$

$$\sum_i^R \sum_n^{N^i} \frac{k_+^i C_0 \langle \mathcal{T}_{u;n}^i \rangle}{4} \cos(\phi_n^i) \sin(\phi_n^i) = \frac{1}{4} \sum_i^R \frac{N^i}{2\pi} \int_0^{2\pi} \frac{k_+^i C_0 T K_d^i}{\left\{ C_0 \exp\left[\frac{g}{2}\cos(\varphi - \phi)\right] + K_d^i \right\}} \cos(\phi_n^i) \sin(\phi_n^i) d\phi = 0. \tag{D17}$$

Therefore, the Fisher information matrix is diagonal, and in shallow gradients it is

$$\mathcal{I} = \begin{bmatrix} NC_0 T \sum_i^R \frac{f^i k_-^i}{8(C_0 + K_d^i)} & 0 \\ 0 & NC_0 T \sum_i^R \frac{f^i k_-^i}{8(C_0 + K_d^i)} \end{bmatrix}. \tag{D18}$$

We note that Eq. (D18) has only been calculated in the large-T limit; in the limit of $T \to 0$, we would expect the Fisher information to limit to the estimate from a single snapshot.

For cells with a single receptor type, Eq. (D18) implies that the asymptotic variances on g_x and g_y are 1/2 of their value determined from time averaging—the same factor as in
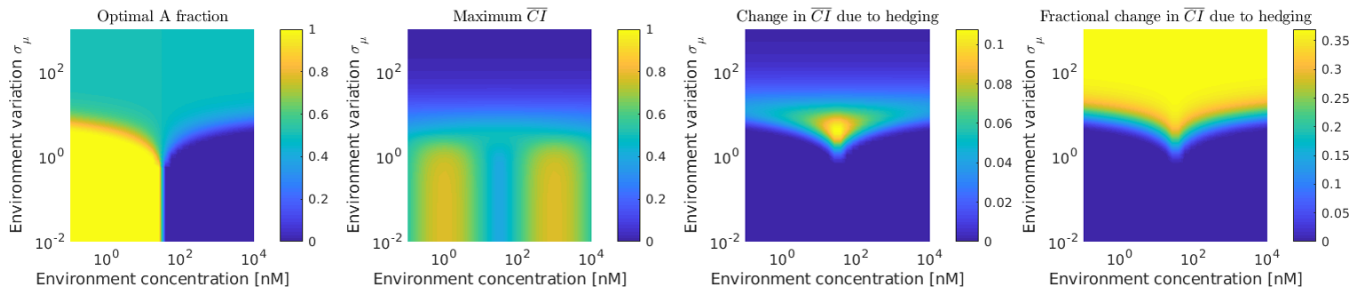
FIG. 8. Trade-offs in snapshot sensing. This figure complements Fig. 1 in the main text, showing how, with the same parameters, the maximum $\overline{CI}$ depends on the uncertainty (second panel). The third and fourth panel show the increase in the mean CI due to hedging, i.e., the change vs all-A or all-B (whichever of these is better). The largest absolute improvements in $\overline{CI}$ due to hedging are at intermediate uncertainties; in the limit of truly high uncertainties, no configuration creates a large $\overline{CI}$.
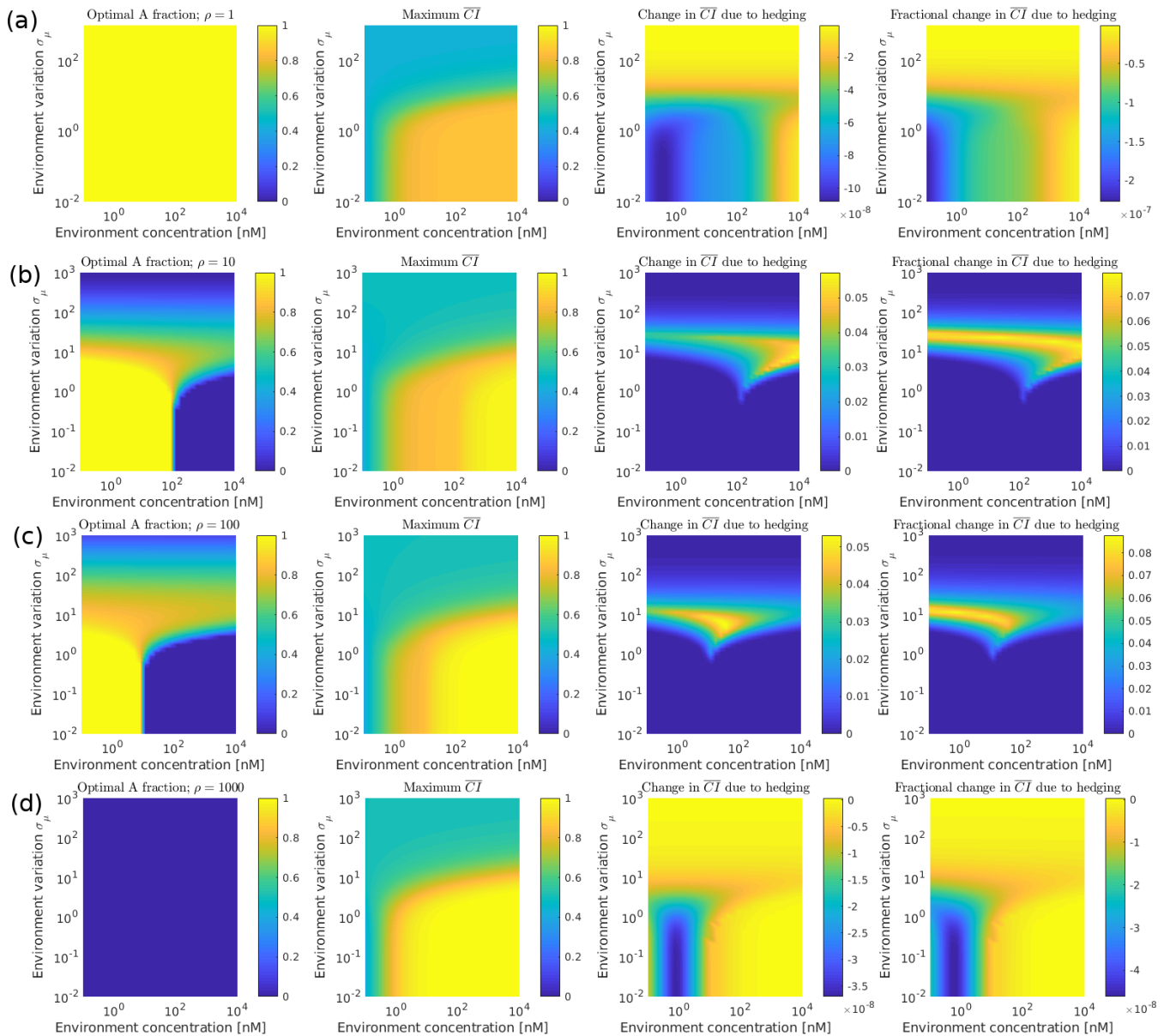

FIG. 9. Trade-offs in time-averaged sensing. This figure shows the maximum $\overline{CI}$ and increase in $\overline{CI}$ due to hedging for the time-average case. This corresponds to Fig. 4 in the main paper, with the left panels redrawing that data. We show these values for (a) $\rho = 1$, (b) $\rho = 10$, (c) $\rho = 100$, and (d) $\rho = 1000$. Note that for (a) and (d), the change in mean CI due to hedging is slightly negative—the optimal configuration is all-A or all-B, but our optimization does not recover $f_A = 0, 1$ with numerical precision.

concentration sensing [24]. However, in the multiple receptor type case, there is a more significant difference. The variance in $\hat{\mathbf{g}}$ determined from Eq. (D18) is the sum of the inverses of the diagonal elements

$$
\begin{aligned}
\sigma_{g,T;\text{ERT}}^2 &= \frac{16}{NC_0 T \sum_i^R \frac{f^i k_-^i}{(C_0+K_d^i)}} \\
&= \frac{16}{NC_0 \sum_i^R \frac{f^i K_d^i}{(C_0+K_d^i)^2} \frac{T}{\tau^i}}.
\end{aligned} \tag{D19}
$$

As discussed in the main text, this shows that a slow receptor correlation does not act as a limiting factor when the entire receptor trajectory is considered.

APPENDIX E: HEDGING ALLOWING THE NUMBER OF RECEPTORS TO CHANGE

Within the main text we have followed earlier work in keeping the number of receptors on the cell fixed [2–5,14,20,43]. However, it is possible that when cells explore more complex environments, they should express different numbers of receptors depending on the typical concentration c_* and the level of uncertainty σ_μ. We address this possibility in Fig. 7.

Within the framework we have applied in this paper, accuracy always increases with increasing N—there are more measurements of the gradient, leading to increased accuracy [see Eqs. (1), (3), and (4) in the main text]. If we allow the number of receptors to freely vary, and choose the number of receptors N and receptor fractions f, we would find that the receptor number would increase without bound. This is obviously unphysical. Cells are under many restraints in controlling how many receptors they have, both in terms of the energetic cost of synthesizing them, and in the opportunity cost in taking up space on the cell surface.

In modeling cells with varying receptor number, we chose to find the receptor configuration that maximized $\overline{\text{CI}} \times e^{-N^{\text{added}}/N^{\text{penalty}}}$, where N^{added} is the number of receptors

expressed beyond the typical value $N^{\text{basal}} = 5 \times 10^4$, and $N^{\text{penalty}} = 50 N^{\text{basal}}$. This choice ensured that cells could easily express more than the basal level of receptors, but that expressing multiple orders of magnitude more receptors would be implausible—consistent with the observed variation in receptor number on the membrane. We found that, though the optimal receptor numbers varied depending on the environment (Fig. 7), the optimal receptor fractions closely agreed with those found assuming a constant number of receptors (Fig. 1).

Other choices for the penalty (e.g., optimizing $\overline{\text{CI}} + \alpha N^{\text{added}}$) gave different optimal receptor numbers but preserved the optimal receptor fractions and the transition between all-*A*, all-*B*, and the 50-50 mix. This suggests that the receptor fractions and the transition are highly robust to allowing the number of receptors to change. This may reflect that the optimal fractions are only very weakly dependent on the total number of receptors.

Experimental measurements on Dictyostelium do see that receptors are internalized in response to saturating levels of chemoattractant; however, this happens on a long timescale ($\sim$5–10 min) and results in a change of about 50% of the receptors being internalized [29,30]. For Dictyostelium cells, which travel about a body length in a minute, we would expect that crawling cells would likely explore another concentration level c_0 before the receptor numbers adapt. Adaptation in eukaryotic chemotaxis is generally thought to occur on a post-receptor level [33,44].

APPENDIX F: EXTENDED DATA ON HEDGING

Within the main paper we have presented the optimal configuration of receptors as a function of the environment. However, at large environmental uncertainties, the benefit from hedging bets may not be as large. We show extended data corresponding to Figs. 1 and 4 in the main paper in Figs. 8 and 9.

[1] H. Levine and W.-J. Rappel, The physics of eukaryotic chemotaxis, Phys. Today **66**(2), 24 (2013).

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

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

[4] I. Segota, S. Mong, E. Neidich, A. Rachakonda, C. J. Lussenhop, and C. Franck, High fidelity information processing in folic acid chemotaxis of Dictyostelium amoebae, J. R. Soc. Interface **10**, 20130606 (2013).

[5] B. Hu, W. Chen, W.-J. Rappel, and H. Levine, Physical Limits on Cellular Sensing of Spatial Gradients, Phys. Rev. Lett. **105**, 048104 (2010).

[6] R. G. Endres and N. S. Wingreen, Accuracy of direct gradient sensing by single cells, Proc. Natl. Acad. Sci. USA **105**, 15749 (2008).

[7] C. Shi, C.-H. Huang, P. N. Devreotes, and P. A. Iglesias, Interaction of motility, directional sensing, and polarity modules recreates the behaviors of chemotaxing cells, PLoS Comput. Biol. **9**, e1003122 (2013).

[8] I. Hecht, M. L. Skoge, P. G. Charest, E. Ben-Jacob, R. A. Firtel, W. F. Loomis, H. Levine, and W.-J. Rappel, Activated membrane patches guide chemotactic cell motility, PLoS Comput. Biol. **7**, e1002044 (2011).

[9] B. A. Camley, Collective gradient sensing and chemotaxis: Modeling and recent developments, J. Phys.: Condens. Matter **30**, 223001 (2018).

[10] B. A. Camley and W.-J. Rappel, Cell-to-cell variation sets a tissue-rheology–dependent bound on collective gradient sensing, Proc. Natl. Acad. Sci. USA **114**, E10074 (2017).

[11] A. Mugler, A. Levchenko, and I. Nemenman, Limits to the precision of gradient sensing with spatial communication and temporal integration, Proc. Natl. Acad. Sci. USA **113**, E689 (2016).

[12] A. Hopkins and B. A. Camley, Leader cells in collective chemotaxis: Optimality and trade-offs, Phys. Rev. E **100**, 032417 (2019).

[13] R. Sharma and E. Roberts, Gradient sensing by a bistable regulatory motif enhances signal amplification but decreases accuracy in individual cells, Phys. Biol. **13**, 036003 (2016).

[14] V. Lakhani and T. C. Elston, Testing the limits of gradient sensing, PLoS Comput. Biol. **13**, e1005386 (2017).

[15] R. J. W. de Wit and P. J. M. van Haastert, Binding of folates to Dictyostelium discoideum cells. Demonstration of five classes of binding sites and their interconversion, Biochim. Biophys. Acta **814**, 199 (1985).

[16] R. L. Johnson, P. J. Van Haastert, A. R. Kimmel, C. L. Saxe, B. Jastorff, and P. N. Devreotes, The cyclic nucleotide specificity of three cAMP receptors in Dictyostelium, J. Biol. Chem. **267**, 4600 (1992).

[17] D. Hereld and P. N. Devreotes, The cAMP receptor family of *Dictyostelium*, Int. Rev. Cytol. **137**, 35 (1993).

[18] J. Y. Kim, J. A. Borleis, and P. N. Devreotes, Switching of chemoattractant receptors programs development and morphogenesis in Dictyostelium: Receptor subtypes activate common responses at different agonist concentrations, Dev. Biol. **197**, 117 (1998).

[19] D. Dormann, J.-Y. Kim, P. N. Devreotes, and C. J. Weijer, cAMP receptor affinity controls wave dynamics, geometry and morphogenesis in *Dictyostelium*, J. Cell Sci. **114**, 2513 (2001).

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

[21] B. A. Camley, J. Zimmermann, H. Levine, and W.-J. Rappel, Emergent Collective Chemotaxis without Single-Cell Gradient Sensing, Phys. Rev. Lett. **116**, 098101 (2016).

[22] H. C. Berg and E. M. Purcell, Physics of chemoreception, Biophys. J. **20**, 193 (1977).

[23] P. R. ten Wolde, N. B. Becker, T. E. Ouldridge, and A. Mugler, Fundamental limits to cellular sensing, J. Stat. Phys. **162**, 1395 (2016).

[24] R. G. Endres and N. S. Wingreen, Maximum Likelihood and the Single Receptor, Phys. Rev. Lett. **103**, 158101 (2009).

[25] W. Bialek, *Biophysics: Searching for Principles* (Princeton University Press, Princeton, NJ, 2012).

[26] H. B. Barlow, Possible principles underlying the transformation of sensory messages, in *Sensory Communication*, edited by W. A. Rosenblith (MIT Press, Cambridge, MA, 1961), Chap. 13, pp. 217–234.

[27] G. Tkačik, C. G. Callan, and W. Bialek, Information flow and optimization in transcriptional regulation, Proc. Natl. Acad. Sci. USA **105**, 12265 (2008).

[28] S. Laughlin, A simple coding procedure enhances a neuron's information capacity, Z. Naturforsch. C **36**, 910 (1981).

[29] A. Serge, S. de Keijzer, F. Van Hemert, M. R. Hickman, D. Hereld, H. P. Spaink, T. Schmidt, and B. E. Snaar-Jagalska,

Quantification of GPCR internalization by single-molecule microscopy in living cells, Integrative Biol. **3**, 675 (2011).

[30] M. Wang, P. J. M. Van Haastert, P. N. Devreotes, and P. Schaap, Localization of chemoattractant receptors on Dictyostelium discoideum cells during aggregation and down-regulation, Dev. Biol. **128**, 72 (1988).

[31] A. F. M. Tariqul Islam, H. Yue, M. Scavello, P. Haldeman, W.-J. Rappel, and P. G. Charest, The cAMP-induced G protein subunits dissociation monitored in live Dictyostelium cells by BRET reveals two activation rates, a positive effect of caffeine and potential role of microtubules, Cell. Signal. **48**, 25 (2018).

[32] Z. Xiao, Y. Yao, Y. Long, and P. Devreotes, Desensitization of G-protein-coupled receptors, J. Biol. Chem. **274**, 1440 (1999).

[33] Y. Tu and W.-J. Rappel, Adaptation in living systems, Annu. Rev. Condens. Matter Phys. **9**, 183 (2018).

[34] T. Mora and I. Nemenman, Physical Limit to Concentration Sensing in a Changing Environment, Phys. Rev. Lett. **123**, 198101 (2019).

[35] T. Mora, Physical Limit to Concentration Sensing Amid Spurious Ligands, Phys. Rev. Lett. **115**, 038102 (2015).

[36] V. Singh and I. Nemenman, Universal Properties of Concentration Sensing in Large Ligand-Receptor Networks, Phys. Rev. Lett. **124**, 028101 (2020).

[37] A. H. Lang, C. K. Fisher, T. Mora, and P. Mehta, Thermodynamics of Statistical Inference by Cells, Phys. Rev. Lett. **113**, 148103 (2014).

[38] S. M. Kay, *Fundamentals of Statistical Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1993).

[39] K. Wang, W.-J. Rappel, R. Kerr, and H. Levine, Quantifying noise levels of intercellular signals, Phys. Rev. E **75**, 061905 (2007).

[40] A. M. Berezhkovskii and A. Szabo, Effect of ligand diffusion on occupancy fluctuations of cell-surface receptors, J. Chem. Phys. **139**, 121910 (2013).

[41] K. Kaizu, W. de Ronde, J. Paijmans, K. Takahashi, F. Tostevin, and P. R. ten Wolde, The Berg-Purcell limit revisited, Biophys. J. **106**, 976 (2014).

[42] W. Bialek and S. Setayeshgar, Physical limits to biochemical signaling, Proc. Nat. Acad. Sci. USA **102**, 10040 (2005).

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

[44] K. Takeda, D. Shao, M. Adler, P. G. Charest, W. F. Loomis, H. Levine, A. Groisman, W.-J. Rappel, and R. A. Firtel, Incoherent feed forward control governs adaptation of activated Ras in a eukaryotic chemotaxis pathway, Sci. Signal. **5**, ra2 (2012).

[45] https://github.com/bcamley/hedging_reproduce.