How geometry and internal bias affect the accuracy of eukaryotic gradient sensing

Bo Hu, Wen Chen, Wouter-Jan Rappel, and Herbert Levine

How geometry and internal bias affect the accuracy of eukaryotic gradient sensing

Bo Hu, Wen Chen, Wouter-Jan Rappel and Herbert Levine[1]

[1] *Department of Physics and Center for Theoretical Biological Physics,*
University of California, San Diego, La Jolla, CA 92093-0374

Many motile eukaryotic cells determine their direction by measuring external chemical gradients through the binding of ligands to membrane bound receptors. This process is limited by fluctuations arising from the binding process and from the diffusion of the ligand molecules. Here, we apply estimation-theoretic methods to determine the physical limits of gradient sensing for cells that are non-circular and for cells that have an internal bias. Specifically, we derive theoretical expressions for the accuracy of gradient sensing in elliptical cells. This accuracy for highly elliptical cells can significantly deviate from the gradient sensing limits derived for circular cells. Furthermore, we find that a cell cannot improve its sensing of the gradient steepness and direction simultaneously by elongating its cell body. Finally, we derive a lower bound on the accuracy of gradient sensing for cells that possess an internal bias and compare our analytical results with recent experimental findings.

I. INTRODUCTION

The response in biological systems to external stimuli is often limited by the inherent stochasticity of these stimuli. For example, the accuracy of human (and other vertebrate) vision at low light intensities approaches a fundamental limit set by statistical fluctuations of the number of absorbed photons [1, 2]. Another example includes the embryo patterning along the anterior-posterior axis of the fly *Drosophila melanogaster* which is primarily determined by concentration profiles of certain morphogens [3]. The accuracy of the resulting pattern is limited by the noise levels in these profiles. A final example is provided by chemotaxis, the directed movement of cells up or down a chemical gradient. In both prokaryotic chemotaxis, where cells measure and compare concentration signals over time [4, 5], and in eukaryotic chemotaxis, where cells measure concentration differences in space [6, 7], the precision of gradient sensing is limited by the stochastic binding of diffusing chemical molecules (ligands) to specific chemoreceptors on the cell membrane.

In this paper, we will apply techniques from estimation theory, which attempts to estimate the values of parameters based on measured data that contain a random component [8], to eukaryotic gradient sensing. Even though our approach and results are applicable in a general way, we will focus mainly on the social amoeba *Dictyostelium discoideum*, a well-characterized model system [9, 10]. Specific G-protein coupled receptors on the *Dictyostelium* cell membrane bind and detect the chemoattractants in the surrounding medium. This results in an asymmetric distribution of ligand-occupied receptors, which further activates multiple second-messenger pathways inside the cell and drive the extension of pseudopods preferentially in the direction of the chemoattractant gradient. Due to fluctuations in the ligand binding to chemoreceptors, the receptor signal is inherently noisy, as demonstrated by recent single-molecule imaging experiment [11]. Surprisingly, *Dictyostelium* cells exhibit extremely high sensitivity to gradients, as they have been observed to be able to detect a $1 - 2\%$ difference in chemical concentration across the cell length [12, 13]. The difference in receptor occupancy between the front and back halves of a cell in these shallow gradients can be calculated to be 10-30. This raises a puzzle about eukaryotic chemotaxis: how can cells reliably acquire the gradient information from such a noisy receptor signal?

In 1977, Berg and Purcell analyzed bacterial chemotaxis and demonstrated that the limit of uncertainty of concentration sensing is set by the diffusion of ligand particles [14]. Their seminal work has been extended by many others [15–19]. The results of Berg and Purcell, however, do not completely carry over to eukaryotic cells, which employ a spatial sensing mechanism. In this mechanism, and in contrast to prokaryotic chemotaxis, cells use the spatial asymmetry (including the gradient steepness and direction) to direct their motion. A number of studies have been carried out to reveal the limits to spatial gradient sensing, but are either only applicable to idealized mechanisms that ignore the receptor kinetics [18], adopt heuristic signaling models [20, 21] or use a simplified geometry [22]. We recently addressed this problem for circular cells using a general statistical mechanical approach, where we view the surface receptors as a (possibly coupled) spin chain and treat the chemical gradient as a perturbation field [23]. By calculating the system's partition function, we were able to derive the gradient sensing limits for either independent receptors or receptors exhibiting cooperativity. Furthermore, using information theoretical concepts and comparing theoretical results with experiments of chemotaxing *Dictyostelium* cells, we have shown that for shallow gradients and low background concentrations, the accuracy of gradient sensing is upper-bounded by fluctuations at the receptor

level [13, 24].

In this paper, we will first revisit the spatial gradient sensing problem for circular cell shapes using purely estimation-theoretic methods (Section II). This will be the foundation for our extension to the elliptical cell shapes in the next section (Section III). Finally, we will examine how a possible intracellular bias can affect the perceived gradient and compare our results with recent experimental data (Section IV).

II. GRADIENT SENSING FOR A CIRCULAR CELL

We revisit the gradient sensing problem which we have examined in our previous paper using the method of *maximum likelihood estimate* (MLE) [8, 19, 23]. Consider a circular cell with diameter L placed in a chemoattractant gradient and assume that there are N receptors uniformly distributed on the cell perimeter. The angular coordinates of these receptors are denoted by φ_n for $n = 1, ..., N$, which satisfy the uniform distribution $P(\varphi_n) = 1/(2\pi)$. We further assume that the gradient field takes an exponential profile as has recently been realized in experiments [13, 25]. Then the local concentration at the n-*th* receptor can be expressed as $C_n = C_0 \exp\left[\frac{p}{2}\cos(\varphi_n - \phi)\right]$, where C_0 is the ambient (mean) concentration, $p \equiv \frac{L}{C_0}|\vec{\nabla}C|$ defines the gradient steepness that quantifies the percentage concentration change across the cell length L, and ϕ denotes the gradient direction. Each receptor switches independently between two states, either empty (0) or occupied (1), with transition rates determined by the local concentration and the relevant chemical kinetics. Therefore, these receptors in a single snapshot constitute a series of independent Bernoulli random variables, represented by

$$x_n = \begin{cases} 1, & \text{with probability} \quad P_n, \\ 0, & \text{with probability} \quad 1 - P_n, \end{cases} \tag{1}$$

for $n = 1, ..., N$. For simple ligand-receptor kinetics, the occupancy probability of the n_{th} receptor is $P_n = C_n/(C_n + K_d)$, where $K_d = k_-/k_+$ is the dissociation constant. The probability mass distribution for Eq. (1) can be expressed as:

$$f_n(x_n|\boldsymbol{\Theta}) = P_n^{x_n}(1 - P_n)^{1-x_n}, \quad \text{for} \quad x_n \in \{0, 1\}, \tag{2}$$

where $\boldsymbol{\Theta} \equiv \{p, \phi\}$ represents the parameters to estimate. Therefore, the likelihood function for a sample of N independent receptors is given by

$$\mathcal{L}(\boldsymbol{\Theta}|x_1, ..., x_N) = f(x_1, ..., x_N|\boldsymbol{\Theta}) = \prod_{n=1}^{N} f_n(x_n|\boldsymbol{\Theta}), \tag{3}$$

and the log-likelihood function is

$$\begin{aligned} \ln\mathcal{L} &= \sum_n \left[x_n \ln\frac{C_n}{C_n + K_d} + (1 - x_n)\ln\frac{K_d}{C_n + K_d}\right] \\ &= \sum_n x_n \ln\frac{C_n}{K_d} + \sum_n \ln\frac{K_d}{C_n + K_d} \\ &= \frac{1}{2}\sum_n x_n p\cos(\varphi_n - \phi) + \ln\frac{C_0}{K_d}\sum_n x_n \\ &\quad + \int_0^{2\pi} \frac{N}{2\pi} \ln\left[\frac{K_d}{C_0\exp\left[\frac{p}{2}\cos(\varphi - \phi)\right] + K_d}\right] d\varphi \\ &= \frac{p\cos\phi}{2}\sum_n x_n \cos\varphi_n + \frac{p\sin\phi}{2}\sum_n x_n \sin\varphi_n \\ &\quad + \ln\frac{C_0}{K_d}\sum_n x_n - \frac{NC_0 K_d p^2}{16(C_0 + K_d)^2} + \mathcal{O}(p^4). \end{aligned} \tag{4}$$

Here, we introduce the transformation $\boldsymbol{\Theta}_\alpha = (\alpha_1, \alpha_2)^T \equiv (p\cos\phi, p\sin\phi)^T$ and define $(z_1, z_2) \equiv (\sum_n x_n \cos\varphi_n, \sum_n x_n \sin\varphi_n)$ which measures the spatial asymmetry in the receptor occupancy. Then, we have $p^2 = \alpha_1^2 + \alpha_2^2$, and for shallow gradients the log-likelihood function becomes

$$\ln\mathcal{L} \approx \frac{\alpha_1 z_1 + \alpha_2 z_2}{2} + \ln\frac{C_0}{K_d}\sum_n x_n - \frac{NC_0 K_d(\alpha_1^2 + \alpha_2^2)}{16(C_0 + K_d)^2}. \tag{5}$$

The method of *maximum likelihood* estimates the unknown parameters by finding a value of $\boldsymbol{\Theta}_\alpha$ that maximizes $\mathcal{L}(\boldsymbol{\Theta}_\alpha|x_1,...,x_N)$, i.e.,

$$\widehat{\boldsymbol{\Theta}}_{\alpha,\mathbf{mle}} = \arg\max_{\boldsymbol{\Theta}_\alpha} \mathcal{L}(\boldsymbol{\Theta}_\alpha|x_1,...,x_N), \tag{6}$$

where the subscript "mle" denotes the maximum likelihood estimator. Since the logarithm is a continuous strictly increasing function over the range of the likelihood, the values which maximize the likelihood will also maximize its logarithm. Thus, the MLE can be found from $\partial_{\alpha_1}\ln\mathcal{L} = 0$ and $\partial_{\alpha_2}\ln\mathcal{L} = 0$, with the following solution

$$\widehat{\boldsymbol{\Theta}}_{\alpha,\mathbf{mle}} = \begin{pmatrix} \widehat{\alpha}_1 \\ \widehat{\alpha}_2 \end{pmatrix} = \frac{1}{\mu}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \quad \text{where} \quad \mu \equiv \frac{NC_0K_d}{4(C_0+K_d)^2}. \tag{7}$$

This solution is indeed the maximum of the likelihood function since it is the only turning point in $\boldsymbol{\Theta}_\alpha$ and the second derivative is strictly less than zero. By the properties of MLE, both $\widehat{\alpha}_1$ and $\widehat{\alpha}_2$ are asymptotically unbiased and normal as the sample size N goes to infinity, i.e., $\widehat{\alpha}_1 \xrightarrow{d} \mathcal{N}(\alpha_1,\sigma_1^2)$ and $\widehat{\alpha}_2 \xrightarrow{d} \mathcal{N}(\alpha_2,\sigma_2^2)$, where "$\xrightarrow{d}$" denotes convergence in distribution. The asymptotic variances σ_1^2 and σ_2^2 can be derived from the inverse of the Fisher information matrix [8, 19]. This matrix has to be diagonal here as α_1 and α_2 are independent of each other. Thus, we have

$$\frac{1}{\sigma_{1,2}^2} = \left\langle \left(\frac{\partial\ln\mathcal{L}}{\partial\alpha_{1,2}}\right)^2 \right\rangle = -\left\langle \frac{\partial^2\ln\mathcal{L}}{\partial\alpha_{1,2}^2} \right\rangle = \frac{\mu}{2}, \tag{8}$$

where $<>$ represents the expectation and the proof of the second equality can be found in [8]. This equality holds because $\ln\mathcal{L}$ is twice differentiable with respect to α_1 and α_2. From the relation $\widehat{\alpha}_{1,2} = z_{1,2}/\mu$, we can see that $<z_{1,2}> = \mu\alpha_{1,2}$ and that the asymptotic variances of z_1 and z_2 are equal:

$$\sigma^2 = \mu^2\sigma_{1,2}^2 = 2\mu^2/\mu = 2\mu. \tag{9}$$

In addition, one can check that the covariance is $\text{Cov}[z_1,z_2] = 0$, as reflected in the affine form of the log-likelihood: $\ln\mathcal{L} = \alpha_1 z_1/2 + \alpha_2 z_2/2 + ...$ Thus, for small p, the joint probability density of z_1 and z_2 is [23, 25]:

$$P(z_1,z_2|\boldsymbol{\Theta}_\alpha) = \frac{1}{2\pi\sigma^2}\exp\left[-\frac{(z_1-\mu\alpha_1)^2 + (z_2-\mu\alpha_2)^2}{2\sigma^2}\right]. \tag{10}$$

As the one-to-one transformation of $\widehat{\boldsymbol{\Theta}}_{\alpha,\mathbf{mle}}$, the MLE of $\boldsymbol{\Theta} = (p,\phi)^T$ is given by

$$\widehat{\boldsymbol{\Theta}}_{\mathbf{mle}} = \begin{pmatrix} \widehat{p} \\ \widehat{\phi} \end{pmatrix} = \begin{pmatrix} \mu^{-1}\sqrt{z_1^2+z_2^2} \\ \arctan(z_2/z_1) \end{pmatrix}. \tag{11}$$

If we introduce the complex random variable $\mathbf{Z} \equiv z_1 + iz_2$, then $\mathbf{Z}$ follows the complex Gaussian distribution, and in the polar coordinates $\mathbf{Z} = \mu\widehat{p}\exp(i\widehat{\phi})$. Eq. (10) suggests that we can define a *signal-to-noise* (SNR) ratio

$$\kappa \equiv \frac{\mu^2 p^2}{\sigma^2} = \frac{\mu p^2}{2} = \frac{Np^2 C_0 K_d}{8(C_0+K_d)^2}. \tag{12}$$

Again, by the properties of MLE, both $\widehat{p}$ and $\widehat{\phi}$ tend to be unbiased and normal in the large N limit, i.e., $\widehat{p} \xrightarrow{d} \mathcal{N}(p,\sigma_p^2)$ and $\widehat{\phi} \xrightarrow{d} \mathcal{N}(\phi,\sigma_\phi^2)$. It has been found numerically that the asymptotic normality becomes excellent when the SNR is larger than 9 [25, 26]. Similarly, the asymptotic variances σ_p^2 and σ_ϕ^2 can be computed from the inverse of the Fisher information matrix with respect to $\boldsymbol{\Theta} = (p,\phi)$. The matrix is still diagonal due to the orthogonality of p and ϕ. Thus, we have

$$\frac{1}{\sigma_p^2} = \left\langle \left(\frac{\partial\ln\mathcal{L}}{\partial p}\right)^2 \right\rangle = -\left\langle \frac{\partial^2\ln\mathcal{L}}{\partial p^2} \right\rangle = \frac{\mu}{2}, \tag{13}$$

$$\frac{1}{\sigma_\phi^2} = \left\langle \left(\frac{\partial\ln\mathcal{L}}{\partial\phi}\right)^2 \right\rangle = -\left\langle \frac{\partial^2\ln\mathcal{L}}{\partial\phi^2} \right\rangle = \frac{\mu p^2}{2}, \tag{14}$$

or equivalently

$$\sigma_p^2 = \frac{8(C_0+K_d)^2}{NC_0K_d} \quad \text{and} \quad \sigma_\phi^2 = \frac{8(C_0+K_d)^2}{Np^2 C_0 K_d} = \frac{1}{\kappa}. \tag{15}$$

The above results are identical to the ones we have recently derived in [23]. According to the Cramér-Rao inequality, the variances σ_p^2 and σ_ϕ^2 represent the minimal uncertainties of gradient measurements from an instantaneous sampling of the receptor states [8].

If the cell in question integrates receptor signals over some time interval $\mathcal{T}$, then averaging over multiple measurements should obviously reduce the errors of gradient sensing. However, the error-reduction via temporal averaging is limited by the expected time it takes to perform an independent measurement. As shown in [17, 27], the time to make a single measurement is roughly twice the receptor correlation time τ. Hence, the number of independent measurements a cell can make within $\mathcal{T}$ is about $\mathcal{T}/(2\tau)$. The correlation time can be decomposed as $\tau = \tau_{\text{rec}} + \tau_{\text{diff}}$, where $\tau_{\text{rec}} = 1/(k_- + C_0 k_+)$ is the time-scale of receptor-ligand reaction and $\tau_{\text{diff}} \simeq N/(2\pi DLK_d)$ describes the diffusion noise correlation time [15, 17, 20, 21, 23]. Let $\eta \equiv \tau_{\text{diff}}/\tau_{\text{rec}}$, then the measurement is said to be reaction-limited if $\eta \ll 1$ and diffusion-limited if $\eta \gg 1$ [28]. Combining the above arguments, we find that integrating signals over $\mathcal{T}$ yields better gradient estimate,

$$\sigma_{p,\mathcal{T}}^2 \simeq \frac{2\tau}{\mathcal{T}}\sigma_p^2 = \frac{4\tau_{rec}(1+\eta)}{\mu\mathcal{T}} = \frac{16(1+\eta)(C_0+K_d)}{N\mathcal{T}k_-C_0}. \tag{16}$$

We can derive similar results for the direction inference, since $\sigma_{\phi,\mathcal{T}}^2 = \sigma_{p,\mathcal{T}}^2/p^2$. For typical eukaryotic cells, it has been estimated [17, 28] that $\eta \ll 1$, which implies that $\sigma_{\phi,\mathcal{T}}^2 \simeq 16(C_0+K_d)/(Np^2\mathcal{T}k_-C_0)$.

An implicit assumption in the above derivation as well as in our previous paper [23] is that a cell can tell the location of each individual receptor. This is reflected in $(z_1, z_2) \equiv (\sum_n x_n \cos\varphi_n, \sum_n x_n \sin\varphi_n)$ which keeps track of every individual receptor and its spatial position. Since the density of receptors on the cell surface can be high, one may argue that the cell cannot distinguish between receptors that are very close to each other. We can, however, easily relax this assumption by taking advantage of the *Central Limit Theorem* (CLT). To see this, we divide the cell surface into M small sensory sectors such that receptors in the same sector are responding to an almost identical chemical concentration. Therefore, receptors in the same sector can be regarded as independent and identically distributed (Bernoulli) random variables, and the state of each sector is represented by its receptor occupancy number which follows the binomial distribution. One can think of M as a quantity that reflects to what extent the cell can spatially distinguish its receptors. Of course, we want to ensure that the number of receptors in each sector $N_s = N/M$ is large enough for the application of CLT. Here, we will choose $M = 1000$ and $N = 40,000$, leading to $N_s = 40$. Now the local concentration at the m_{th} sector with angular position $\vartheta_m = 2\pi m/M$ is given by $C_m = C_0 \exp\left[\frac{p}{2}\cos(\vartheta_m - \phi)\right]$ for $m = 1, ..., M$. Using the CLT, the number of occupied receptors in the m_{th} sector in a single snapshot, y_m, is approximately $y_m = N_s C_m/(C_m + K_d) + \eta_m$ $(m = 1, ..., M)$, where the Gaussian random component η_m satisfies $\langle \eta_m \rangle = 0$ and $\langle \eta_m \eta_n \rangle = \delta_{mn} N_s C_m K_d/(C_m + K_d)^2 \approx \delta_{mn} N_s C_0 K_d/(C_0 + K_d)^2$ [20, 21, 28]. For small gradients, we can expand y_m around p:

$$y_m \approx \frac{N_s C_0}{K_d + C_0} + \frac{N_s K_d C_0}{2(C_0 + K_d)^2}p\cos(\vartheta_m - \phi) + \eta_m. \tag{17}$$

We can see that the sector states, denoted by $\mathbf{Y} = \{y_1, y_2, ..., y_M\}^T$, constitute a vector of independent Gaussian random variables with space-dependent means but approximately identical variance $\sigma_s^2 \equiv N_s K_d C_0/(C_0 + K_d)^2$. Hereafter, the superscript symbol ^{T} means transpose. Eq. (17) suggest that one can view $\mathbf{Y} = \{y_1, y_2, ..., y_M\}^T$ as observations of a sinusoidal signal corrupted by some white Gaussian noise. The likelihood function of $\mathbf{Y}$ conditional on $\mathbf{\Theta}$ reads,

$$P(\mathbf{Y}|\mathbf{\Theta}) \approx \frac{1}{(2\pi\sigma_s^2)^{M/2}}\exp\left[-\frac{J(\mathbf{Y}|\mathbf{\Theta})}{2\sigma_s^2}\right], \tag{18}$$

where

$$J(\mathbf{Y}|\mathbf{\Theta}) = \sum_{m=1}^{M}\left(y_m - \frac{N_s C_0}{C_0 + K_d} - \frac{\sigma_s^2}{2}p\cos(\vartheta_m - \phi)\right)^2. \tag{19}$$

To maximize $P(\mathbf{Y}|\mathbf{\Theta})$ is equivalent to minimize $J(\mathbf{Y}|\mathbf{\Theta})$ which can be converted to a quadratic function by the one-to-one transformation $\mathbf{\Theta}_\alpha = (\alpha_1, \alpha_2)^T = (p\cos\phi, p\sin\phi)^T$:

$$J(\mathbf{Y}|\mathbf{\Theta}_\alpha) = (\mathbf{Y} - \mathbf{Y_0} - \mathbf{H}\mathbf{\Theta}_\alpha)^T(\mathbf{Y} - \mathbf{Y_0} - \mathbf{H}\mathbf{\Theta}_\alpha), \tag{20}$$

with

$$\mathbf{Y_0} = \frac{N_s C_0}{C_0 + K_d}\begin{pmatrix} 1 \\ 1 \\ ... \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{H} = \frac{\sigma_s^2}{2}\begin{pmatrix} \cos\vartheta_1 & \sin\vartheta_1 \\ \cos\vartheta_2 & \sin\vartheta_2 \\ ... & ... \\ \cos\vartheta_M & \sin\vartheta_M \end{pmatrix}$$
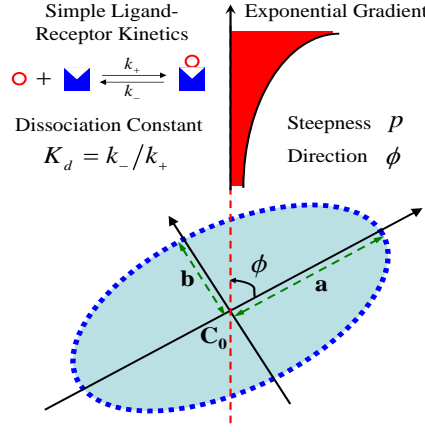
FIG. 1: (Color online) Schematic representation of our model: an elliptical cell, covered with receptors, is placed in an exponential gradient. The angle between the direction of the gradient and the major axis of the elliptical cell is denoted by ϕ. The forward and backward rates $k_{\pm}$ control the transition between the bound and unbound states for the receptors.

We can apply the formula of *ordinary least squares* (OLS) estimate to find the MLE solution that minimizes the quadratic objective function $J(\mathbf{Y}|\mathbf{\Theta}_\alpha)$ and hence maximizes the likelihood function $P(\mathbf{Y}|\mathbf{\Theta}_\alpha)$:

$$\widehat{\mathbf{\Theta}}_{\alpha,\mathbf{mle}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T(\mathbf{Y}-\mathbf{Y_0}). \tag{21}$$

For large M, we have $\frac{1}{M}\sum_m \cos\vartheta_m \approx \frac{1}{M}\sum_m \sin\vartheta_m \approx 0$, $\frac{1}{M}\sum_m \cos\vartheta_m \sin\vartheta_m \approx 0$, and $\frac{1}{M}\sum_m \cos^2\vartheta_m \approx \frac{1}{M}\sum_m \sin^2\vartheta_m \approx 1/2$. Thus,

$$\begin{aligned}\widehat{\mathbf{\Theta}}_{\alpha,\mathbf{mle}} &= \begin{pmatrix}\widehat{\alpha}_1 \\ \widehat{\alpha}_2\end{pmatrix} \approx \frac{4}{\sigma_s^4}\begin{pmatrix}M/2 & 0 \\ 0 & M/2\end{pmatrix}^{-1}\mathbf{H}^T(\mathbf{Y}-\mathbf{Y_0}) \\ &= \frac{4(C_0+K_d)^2}{NK_dC_0}\begin{pmatrix}\sum_m y_m\cos\vartheta_m \\ \sum_m y_m\sin\vartheta_m\end{pmatrix} \simeq \frac{1}{\mu}\begin{pmatrix}z_1 \\ z_2\end{pmatrix},\end{aligned}$$

where in the last line we have noticed that $z_1 = \sum_n x_n\cos\varphi_n \simeq \sum_m y_m\cos\vartheta_m$ and $z_2 = \sum_n x_n\sin\varphi_n \simeq \sum_m y_m\sin\vartheta_m$ for sufficiently large M. By inverse mapping, we recover the MLE of $\mathbf{\Theta}$ given in Eq. (11). According to the asymptotic properties, $\widehat{\mathbf{\Theta}}_{\mathbf{mle}} = (\widehat{p},\widehat{\phi})^T \xrightarrow{d} \mathcal{N}\left(\mathbf{\Theta},\mathbf{I}(\mathbf{\Theta})^{-1}\right)$, with the Fisher information matrix:

$$\begin{aligned}\mathbf{I}(\mathbf{\Theta}) &= \left\langle\frac{\partial\ln P(\mathbf{Y}|\mathbf{\Theta})}{\partial\mathbf{\Theta}_i}\frac{\partial\ln P(\mathbf{Y}|\mathbf{\Theta})}{\partial\mathbf{\Theta}_j}\right\rangle \\ &\approx \begin{pmatrix}M\sigma_s^2/8 & 0 \\ 0 & M\sigma_s^2p^2/8\end{pmatrix} \\ &= \begin{pmatrix}\mu/2 & 0 \\ 0 & \mu p^2/2\end{pmatrix}. \tag{22}\end{aligned}$$

Then, we recover our main MLE results:

$$\widehat{p} = \mu^{-1}\sqrt{z_1^2+z_2^2} \xrightarrow{d} \mathcal{N}\left(p,\sigma_p^2=\frac{2}{\mu}\right), \tag{23}$$

$$\widehat{\phi} = \arctan(z_2/z_1) \xrightarrow{d} \mathcal{N}\left(\phi,\sigma_\phi^2=\frac{2}{\mu p^2}\right). \tag{24}$$

III.   GRADIENT SENSING FOR AN ELLIPTICAL CELL

Motile eukaryotic cells can become polarized and exhibit elongated shapes, which may affect the capacity of gradient sensing. In this section, we analyze a chemotactic cell with an elliptical geometry (Fig. 1). The ellipse can be described

in parametric form $(X(\omega), Y(\omega)) = (a\cos\omega, b\sin\omega)$, where a is the semi-major axis, b is semi-minor axis (thus $b \leq a$) and $0 \leqslant \omega \leqslant 2\pi$. With $\vartheta = 0$ measured from the major axis, the ellipse can also be described in polar coordinates by

$$r(\vartheta) = \frac{ab}{\sqrt{(a\sin\vartheta)^2 + (b\cos\vartheta)^2}} = \frac{b}{\sqrt{1 - (\varepsilon\cos\vartheta)^2}}, \tag{25}$$

where $\varepsilon \equiv \sqrt{1 - (b/a)^2}$ is the eccentricity of the ellipse and measures how much the geometry deviates from being circular. Obviously, ω and ϑ are related through

$$\tan\vartheta = \tan\left[\frac{Y(\omega)}{X(\omega)}\right] = \frac{b}{a}\tan\omega = \sqrt{1 - \varepsilon^2}\tan\omega, \tag{26}$$

$$r(\vartheta) = b/\sqrt{1 - \varepsilon^2\cos^2\vartheta} = \sqrt{X(\omega)^2 + Y(\omega)^2}$$

$$= a\sqrt{1 - \varepsilon^2\sin^2\omega} = r(\omega). \tag{27}$$

As before, we distribute the receptors uniformly along the cell circumference and divide the cell's circumference into M sectors with identical arc lengths (containing $N_s = N/M$ receptors). In the exponential gradient, the local concentration at the m_{th} sector with polar angle ϑ_m and radius r_m is

$$C_m = C_0 \exp\left[\frac{|\vec{\nabla}C|}{C_0} r_m \cos(\vartheta_m - \phi)\right]$$

$$= C_0 \exp\left[\frac{p}{2}\frac{r_m}{a}\cos(\vartheta_m(\omega_m) - \phi)\right]$$

$$= C_0 \exp\left[\frac{p}{2}\cos(\vartheta_m - \phi)\sqrt{1 - \varepsilon^2\sin^2\omega_m}\right], \tag{28}$$

where ϕ is the gradient direction measured from the major axis and $p \equiv 2a|\vec{\nabla}C|/C_0$ is the gradient steepness that quantifies the percentage concentration change across the major axis $2a$. Here, we have used that

$$r_m = r(\omega_m) = a\sqrt{1 - \varepsilon^2\sin^2\omega_m}$$

$$= r(\vartheta_m) = b/\sqrt{1 - \varepsilon^2\cos^2\vartheta_m}. \tag{29}$$

As before, the receptor occupancy number is $y_m = N_s C_m/(C_m + K_d) + \eta_m$, for $m = 1, ..., M$. For shallow gradients, we can expand y_m to leading order in p:

$$y_m \approx \frac{N_s C_0}{K_d + C_0} + \frac{N_s K_d C_0}{2(C_0 + K_d)^2}\frac{r_m}{a}p\cos(\vartheta_m - \phi) + \eta_m. \tag{30}$$

where the random component η_m is approximately normal: $\langle\eta_m\rangle = 0$ and $\langle\eta_m\eta_n\rangle = \delta_{mn}N_s C_m K_d/(C_m + K_d)^2 \approx \delta_{mn}N_s C_0 K_d/(C_0 + K_d)^2 = \delta_{mn}\sigma_s^2$. Thus, the likelihood function of $\mathbf{Y} = \{y_1, y_2, ..., y_M\}^T$ is $P(\mathbf{Y}|\mathbf{\Theta}) \approx (2\pi\sigma_s^2)^{-M/2}\exp[-\tilde{J}/(2\sigma_s^2)]$, where

$$\tilde{J} = \sum_{m=1}^{M}\left[y_m - \frac{N_s C_0}{C_0 + K_d} - \frac{\sigma_s^2 r_m}{2a}p\cos(\vartheta_m - \phi)\right]^2$$

$$= (\mathbf{Y} - \mathbf{Y_0} - \tilde{\mathbf{H}}\mathbf{\Theta}_\alpha)^T(\mathbf{Y} - \mathbf{Y_0} - \tilde{\mathbf{H}}\mathbf{\Theta}_\alpha), \tag{31}$$

with $\mathbf{\Theta}_\alpha = (\alpha_1, \alpha_2)^T = (p\cos\phi, p\sin\phi)^T$ and

$$\tilde{\mathbf{H}} = \frac{\sigma_s^2}{2a}\begin{pmatrix} r_1\cos\vartheta_1 & r_1\sin\vartheta_1 \\ r_2\cos\vartheta_2 & r_2\sin\vartheta_2 \\ ... & ... \\ r_M\cos\vartheta_M & r_M\sin\vartheta_M \end{pmatrix}. \tag{32}$$

Thus, the MLE of $\mathbf{\Theta}_\alpha$ is

$$\tilde{\mathbf{\Theta}}_{\alpha,\mathbf{mle}} = (\tilde{\alpha}_1, \tilde{\alpha}_2)^T = (\tilde{\mathbf{H}}^T\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^T(\mathbf{Y} - \mathbf{Y_0}). \tag{33}$$

Due to symmetry, one can directly find that

$$\sum_{m=1}^{M} r_m \cos \vartheta_m = \sum_{m=1}^{M} r(\vartheta_m) \cos \vartheta_m = 0, \tag{34}$$

$$\sum_{m=1}^{M} r_m \sin \vartheta_m = \sum_{m=1}^{M} r(\vartheta_m) \sin \vartheta_m = 0, \tag{35}$$

$$\sum_{m=1}^{M} r(\vartheta_m)^2 \cos \vartheta_m \sin \vartheta_m = 0. \tag{36}$$

For convenience, we define the following two functions which we will evaluate later on:

$$\Lambda_1(\varepsilon) \equiv \frac{1}{M} \sum_{m=1}^{M} \frac{r_m^2}{a^2} \cos^2 \vartheta_m = \frac{1}{M} \sum_{m=1}^{M} \frac{r(\omega_m)^2/a^2}{1 + \tan^2 \vartheta_m}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \frac{1 - \varepsilon^2 \sin^2 \omega_m}{1 + (1 - \varepsilon^2) \tan^2 \omega_m}, \tag{37}$$

$$\Lambda_2(\varepsilon) \equiv \frac{1}{M} \sum_{m=1}^{M} \frac{r_m^2}{a^2} \sin^2 \vartheta_m$$

$$= \frac{1}{M} \sum_{m=1}^{M} \frac{(1 - \varepsilon^2 \sin^2 \omega_m)(1 - \varepsilon^2) \tan^2 \omega_m}{1 + (1 - \varepsilon^2) \tan^2 \omega_m}. \tag{38}$$

Then, the MLE of $\boldsymbol{\Theta}_\alpha$ is given by

$$\begin{pmatrix} \widetilde{\alpha}_1 \\ \widetilde{\alpha}_2 \end{pmatrix} \approx \frac{4}{\sigma_s^4} \begin{pmatrix} M\Lambda_1(\varepsilon) & 0 \\ 0 & M\Lambda_2(\varepsilon) \end{pmatrix}^{-1} \widetilde{\mathbf{H}}^T (\mathbf{Y} - \mathbf{Y_0})$$

$$= \frac{1}{2\mu a} \begin{pmatrix} \Lambda_1^{-1} \sum_m y_m r_m \cos \vartheta_m \\ \Lambda_2^{-1} \sum_m y_m r_m \sin \vartheta_m \end{pmatrix}. \tag{39}$$

The expectations and variances of $\widetilde{\alpha}_{1,2}$ are

$$< \widetilde{\alpha}_1 > = \frac{\sigma_s^2 p}{4\mu\Lambda_1} \sum_m \frac{r_m^2}{a^2} \cos \vartheta_m \cos(\vartheta_m - \phi) = \alpha_1,$$

$$< \widetilde{\alpha}_2 > = \frac{\sigma_s^2 p}{4\mu\Lambda_2} \sum_m \frac{r_m^2}{a^2} \sin \vartheta_m \cos(\vartheta_m - \phi) = \alpha_2,$$

$$\sigma_{\widetilde{\alpha}_1}^2(\varepsilon) = \frac{\sigma_s^2}{4\mu^2\Lambda_1^2} \sum_m \frac{r_m^2}{a^2} \cos^2 \vartheta_m = \frac{1}{\mu\Lambda_1(\varepsilon)},$$

$$\sigma_{\widetilde{\alpha}_2}^2(\varepsilon) = \frac{\sigma_s^2}{4\mu^2\Lambda_2^2} \sum_m \frac{r_m^2}{a^2} \sin^2 \vartheta_m = \frac{1}{\mu\Lambda_2(\varepsilon)}.$$

One can also check that their covariance $\mathrm{Cov}[\widetilde{\alpha}_1, \widetilde{\alpha}_2] = 0$.

Next, we evaluate $\Lambda_{1,2}(\varepsilon)$. Each sector has an arc length equal to $\mathfrak{L}/M$, where $\mathfrak{L} = 4a\mathrm{E}(\varepsilon)$ is the circumference of the ellipse and $\mathrm{E}(\varepsilon) = \int_0^{\pi/2} \sqrt{1 - (\varepsilon \sin \omega)^2} d\omega$ is the complete elliptic integral of the second kind. The arc length of the ellipse between $\vartheta = 0$ and $\vartheta = \vartheta(\omega_0)$ is given by $s(\omega_0) = \int_0^{\omega_0} \sqrt{(a \sin \omega)^2 + (b \cos \omega)^2} d\omega = \int_0^{\omega_0} a\sqrt{1 - (\varepsilon \sin \omega)^2} d\omega$. Then, $ds = a\sqrt{1 - (\varepsilon \sin \omega)^2} d\omega$. For sufficiently large M, we can approximate:

$$\frac{1}{M} \sum_{m=1}^{M} g(\omega_m) \approx \int_0^{\mathfrak{L}} \frac{g(\omega)}{\mathfrak{L}} ds(\omega)$$

$$= \frac{1}{4\mathrm{E}(\varepsilon)} \int_0^{2\pi} g(\omega) \sqrt{1 - (\varepsilon \sin \omega)^2} d\omega. \tag{40}$$
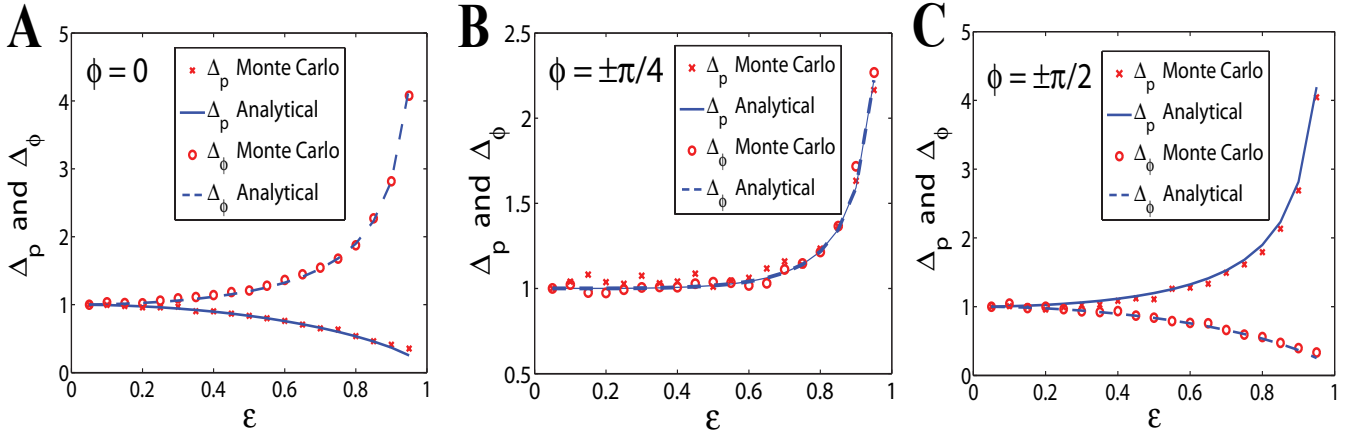
FIG. 2: (Color online). Δ_p and Δ_ϕ as a function of the eccentricity ε for different values of ϕ (lines: analytical expressions; symbols: Monte-Carlo simulations).

where $g(\cdot)$ represents a general integrable function. Therefore, we have

$$\begin{aligned}
\Lambda_1(\varepsilon) &\approx \frac{1}{4\mathrm{E}(\varepsilon)} \int_0^{2\pi} \frac{(1-\varepsilon^2 \sin^2 \omega)^{3/2}}{1+(1-\varepsilon^2)\tan^2 \omega} d\omega \\
&= \frac{(1+\varepsilon^2)E(\varepsilon) - (1-\varepsilon^2)K(\varepsilon)}{3\varepsilon^2 E(\varepsilon)},
\end{aligned} \tag{41}$$

where $\mathrm{K}(\varepsilon) = \int_0^{\pi/2} \frac{d\omega}{\sqrt{1-(\varepsilon \sin \omega)^2}}$ is the complete elliptic integral of the first kind. Similarly, one can find

$$\Lambda_2(\varepsilon) \approx \frac{(1-\varepsilon^2)\left[(-1+2\varepsilon^2)E(\varepsilon) + (1-\varepsilon^2)K(\varepsilon)\right]}{3\varepsilon^2 E(\varepsilon)}. \tag{42}$$

Since $\Lambda_{1,2}(\varepsilon = 0) = 1/2$, we have $\sigma^2_{\widetilde{\alpha}_{1,2}}(\varepsilon = 0) = 2/\mu$, consistent with Eq. (8). Therefore, $\widetilde{\alpha}_{1,2} \to \widehat{\alpha}_{1,2}$ as the cell shape tends to be circular. In the other extreme, we have $\lim_{\varepsilon \to 1} \Lambda_1(\varepsilon) = 2/3$ and $\lim_{\varepsilon \to 1} \Lambda_2(\varepsilon) = 0$.

The MLE of $\mathbf{\Theta} = (p, \phi)^T$ can be found directly,

$$\widetilde{\mathbf{\Theta}}_{\mathbf{mle}} = \begin{pmatrix} \widetilde{p} \\ \widetilde{\phi} \end{pmatrix} = \begin{pmatrix} \sqrt{\widetilde{\alpha}_1^2 + \widetilde{\alpha}_2^2} \\ \arctan(\widetilde{\alpha}_2/\widetilde{\alpha}_1) \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(\mathbf{\Theta}, \widetilde{\mathbf{I}}(\mathbf{\Theta})^{-1}\right), \tag{43}$$

where the last line is due to the asymptotic normality of MLE: as the sample size increases, the distribution of the MLE tends to the Gaussian distribution with mean Θ and covariance matrix equal to the inverse of the Fisher information matrix,

$$\begin{aligned}
\widetilde{\mathbf{I}}(\mathbf{\Theta})^{-1} &= \left\langle \frac{\partial \ln P(\mathbf{Y}|\mathbf{\Theta})}{\partial \mathbf{\Theta}_i} \frac{\partial \ln P(\mathbf{Y}|\mathbf{\Theta})}{\partial \mathbf{\Theta}_j} \right\rangle^{-1} \\
&\approx \frac{1}{\mu} \begin{pmatrix} \Xi_1(\varepsilon, \phi) & p \cdot \Xi_0(\varepsilon, \phi) \\ p \cdot \Xi_0(\varepsilon, \phi) & p^2 \cdot \Xi_2(\varepsilon, \phi) \end{pmatrix}^{-1},
\end{aligned} \tag{44}$$

where

$$\Xi_0(\varepsilon, \phi) \equiv \frac{1}{M} \sum_{m=1}^{M} \frac{r_m^2}{a^2} \cos(\vartheta_m - \phi)\sin(\vartheta_m - \phi), \tag{45}$$

$$\Xi_1(\varepsilon, \phi) \equiv \frac{1}{M} \sum_{m=1}^{M} \frac{r_m^2}{a^2} \cos^2(\vartheta_m - \phi), \tag{46}$$

$$\Xi_2(\varepsilon, \phi) \equiv \frac{1}{M} \sum_{m=1}^{M} \frac{r_m^2}{a^2} \sin^2(\vartheta_m - \phi). \tag{47}$$

Again, one can use Eq. (40) to calculate the above three functions. Here, we will skip all the intermediate technical steps and will only present the final results:

$$
\begin{aligned}
\sigma_{\widetilde{p}}^2(\varepsilon, \phi) &= \frac{\mu^{-1}\Xi_2(\varepsilon, \phi)}{\Xi_1(\varepsilon, \phi)\Xi_2(\varepsilon, \phi) - \Xi_0(\varepsilon, \phi)^2} \\
&= \frac{1}{\mu} \cdot \frac{A_1(\varepsilon)\left[A_2(\varepsilon, \phi) - A_3(\varepsilon, \phi)\right]}{A_6(\varepsilon) + A_7(\varepsilon) + A_8(\varepsilon)},
\end{aligned} \tag{48}
$$

$$
\begin{aligned}
\sigma_{\widetilde{\phi}}^2(\varepsilon, \phi) &= \frac{(\mu p^2)^{-1}\Xi_1(\varepsilon, \phi)}{\Xi_1(\varepsilon, \phi)\Xi_2(\varepsilon, \phi) - \Xi_0(\varepsilon, \phi)^2} \\
&= \frac{\sqrt{1 - \varepsilon^2}}{\mu p^2} \frac{A_1(\varepsilon)\left[A_4(\varepsilon, \phi) - A_5(\varepsilon, \phi)\right]}{A_6(\varepsilon) + A_7(\varepsilon) + A_8(\varepsilon)},
\end{aligned} \tag{49}
$$

where

$$
\begin{aligned}
A_1 &= 3\varepsilon^2 E(\varepsilon)/[2(1 - \varepsilon^2)], \\
A_2 &= 2E(\varepsilon)[(2\varepsilon^2 - \varepsilon^4) - (1 - \varepsilon^2 + \varepsilon^4)\cos(2\phi)], \\
A_3 &= (1 - \varepsilon^2)[\varepsilon^2 - (2 - \varepsilon^2)\cos(2\phi)]K(\varepsilon), \\
A_4 &= 2E(\varepsilon)[(2\varepsilon^2 - \varepsilon^4) + (1 - \varepsilon^2 + \varepsilon^4)\cos(2\phi)], \\
A_5 &= (1 - \varepsilon^2)[\varepsilon^2 + (2 - \varepsilon^2)\cos(2\phi)]K(\varepsilon), \\
A_6 &= (-1 + \varepsilon^2 + 2\varepsilon^4)E(\varepsilon)^2, \\
A_7 &= (2 - 3\varepsilon^2 + \varepsilon^4)E(\varepsilon)K(\varepsilon), \\
A_8 &= -(1 - \varepsilon^2)^2 K(\varepsilon)^2.
\end{aligned}
$$

It is easy to check that $\sigma_{\widetilde{p}}^2(0, \phi) = 2/\mu$ and $\sigma_{\widetilde{\phi}}^2(0, \phi) = 2/(\mu p^2)$, consistent with our results for the circular cell (Eqns. 13, 14). Also, we find that the covariance of $\widetilde{p}$ and $\widetilde{\phi}$ can be written as $\mathrm{Cov}[\widetilde{p}, \widetilde{\phi}] = A_0(\varepsilon)\sin(2\phi)$, which vanishes at $\phi = 0, \pm\pi, \pm\pi/2$. For the special case $\phi = 0$, Eq. (48) reduces to $\sigma_{\widetilde{p}}^2(\varepsilon, 0) = [\mu\Lambda_1(\varepsilon)]^{-1}$. As $b \to 0$ such that $\varepsilon \to 1$, the ellipse becomes a one-dimensional segment of length $2a$, which can measure the gradient steepness with an accuracy $\sigma_{\widetilde{p}}^2(\varepsilon \to 1, \phi = 0) = 3/(2\mu)$.

To determine how the eccentricity affects the accuracy of gradient estimates and to compare with our results for the circular cell, we fix the area of the elliptical cell, i.e., let $S \equiv \pi ab = \pi a^2/\sqrt{1 - \varepsilon^2}$ be constant. We also keep the gradient profile identical by fixing $p_0 \equiv |\vec{\nabla}C|/C_0$. Then, the gradient is proportional to the major axis: $p = 2ap_0 = p(a)$. These assumptions allow us to define the following ratios as a measure how the eccentricity will change the gradient sensing limits:

$$
\begin{aligned}
\Delta_p(\varepsilon, \phi) &\equiv \frac{\sigma_{\widetilde{p}}^2(\varepsilon, \phi)/p(a)^2}{\sigma_{\widetilde{p}}^2(0, \phi)/p(\sqrt{S/\pi})^2} = \frac{\mu S}{2\pi a^2}\sigma_{\widetilde{p}}^2(\varepsilon, \phi) \\
&= \frac{\sqrt{1 - \varepsilon^2}}{2} \frac{A_1(\varepsilon)\left[A_2(\varepsilon, \phi) - A_3(\varepsilon, \phi)\right]}{A_6(\varepsilon) + A_7(\varepsilon) + A_8(\varepsilon)} \\
&= \begin{cases} (2\Lambda_1)^{-1}\sqrt{1 - \varepsilon^2} & \text{for } \phi = 0, \\ (2\Lambda_2)^{-1}\sqrt{1 - \varepsilon^2} & \text{for } \phi = \pm\pi/2, \end{cases}
\end{aligned} \tag{50}
$$

$$
\begin{aligned}
\Delta_\phi(\varepsilon, \phi) &\equiv \frac{\sigma_{\widetilde{\phi}}^2(\varepsilon, \phi)}{\sigma_{\widetilde{\phi}}^2(0, \phi)} = \frac{\mu p^2}{2}\sigma_{\widetilde{\phi}}^2(\varepsilon, \phi) \\
&= \frac{\sqrt{1 - \varepsilon^2}}{2} \frac{A_1(\varepsilon)\left[A_4(\varepsilon, \phi) - A_5(\varepsilon, \phi)\right]}{A_6(\varepsilon) + A_7(\varepsilon) + A_8(\varepsilon)} \\
&= \begin{cases} (2\Lambda_2)^{-1}\sqrt{1 - \varepsilon^2} & \text{for } \phi = 0, \\ (2\Lambda_1)^{-1}\sqrt{1 - \varepsilon^2} & \text{for } \phi = \pm\pi/2, \end{cases}
\end{aligned} \tag{51}
$$

with $\Lambda_1(\varepsilon)$ and $\Lambda_2(\varepsilon)$ given by Eq. (41) and Eq. (42).

In Fig. 2 we have plotted, as solid and dashed lines, these ratios for different values of ϕ (the gradient direction measured from the major axis). In the same figure, we have also plotted as symbols the results of our Monte-Carlo simulations. In these simulations, the ellipse of a given eccentricity ε was divided into $M = 1000$ sectors with equal

arc lengths. Each sector contained $N_s = 40$ receptors and the receptor occupancy number for each sector, y_m, was assigned a Gaussian random variable with mean $N_s C_m/(C_m + K_d)$ and variance $N_s C_m K_d/(C_m + K_d)^2$. After plugging these random numbers into Eq. (39) and Eq. (43) for 10^4 realizations, we computed the numerical values for the sample variances of $\widetilde{p}$ and $\widetilde{\phi}$. As shown in the figure, the agreement between our analytical and numerical results is excellent.

In the case where the gradient direction points along the major axis of the elliptical cell ($\phi = 0$ and Fig. 2A), we see that Δ_p is a decreasing function of the eccentricity ε. This can be intuitively understood by realizing that for an elongated cell pointing in the gradient direction, the difference in concentration, and thus the difference in occupied receptors between the back and front sectors becomes large. Therefore, the more elongated the cell is along the gradient direction, the more accurately the cell is able to measure the gradient steepness. In contrast, Δ_ϕ increases with ε such that more elongated cells are less accurate in estimating the gradient direction than circular cells. When the external gradient is parallel to the cell's minor axis ($\phi = \pm\pi/2$ and Fig. 2C), we have exactly the opposite: Δ_p increases with ε while Δ_ϕ decreases with ε. Finally, for the intermediate case $\phi = \pm\pi/4$ (Fig. 2B), we have $\Delta_p = \Delta_\phi$ and they both increase with ε. Depending on the gradient angle ϕ, the ratios can be even non-monotonic with ε; for example, Δ_p at $\phi = \pi/6$ will first decrease and then increase with ε (data not shown).

It is noteworthy that gradient sensing generally fails in the line segment limit ($\varepsilon \to 1$), except for the special cases of steepness sensing at $\phi = 0, \pm\pi$ and direction sensing at $\phi = \pm\pi/2$. As a vector, the concentration gradient can be decomposed into two orthogonal components, one along the major axis of the cell $p_{major} = p\cos\phi$ and the other along the minor axis $p_{minor} = p\sin\phi$. When sensing the gradient magnitude (steepness), the line-segment cell is only able to detect the magnitude of p_{major}. But there is unlimited uncertainty in the other component p_{minor} (except for vanishing p_{minor} at $\phi = 0$ or $\pm\pi$). This explains why Δ_p always diverges as $\varepsilon \to 1$ except for $\phi = 0$ and $\pm\pi$. Similar arguments hold for directional sensing: a line-segment cell can only tell the direction of the gradient component perpendicular to itself (i.e., p_{minor}) but has unlimited uncertainty about p_{major} which disappears only for $\phi = \pm\pi/2$.

Our analytical and numerical results reveal that, for every gradient direction ϕ, a cell cannot improve the gradient steepness and direction estimates simultaneously by elongating its cell body. It is difficult to connect this result to the experimental measurements of the accuracy of chemotaxis. This is mostly done in the form of the chemotactic index (CI), defined as the distance traveled by the cell in the direction of the gradient divided by the total distance traveled. Clearly, the CI is a function of both the accuracy in gradient steepness and gradient direction. To determine the relative importance of these two parameters would require a precise knowledge of the gradient sensing and cell motility process, both of which are currently absent.

Even in the absence of such precise knowledge, however, we can conclude that elongated cell shapes can strongly affect the cell's accuracy of gradient sensing. In experiments, the cell shape varies depending on the precise conditions. Chemotactic neutrophils, for example, are found to have eccentricities up to 0.8 [29, 30]. For these values, we find that the variances of estimating gradient steepness and direction in the elliptical cells differ by a factor of 2 compared to the ones in circular cells. Furthermore, *Dictyostelium* cells can have much higher eccentricities ($\varepsilon \approx 0.95$) [31]. Such highly elliptical geometries are expected to have even more significant effects on the accuracy of gradient sensing. Thus, the sensing limits derived using a circular geometry are a good approximation only for cells that are not significantly elongated.

IV. DIRECTIONAL SENSING IN THE PRESENCE OF AN INTRACELLULAR BIAS

A large number of experiments have revealed that an external cue can induce spatial localization of several signaling proteins along the plasma membrane during eukaryotic directional sensing [7, 10, 32]. The membrane localization of these signaling molecules enables a cell to polarize and migrate toward the external chemical source. Samadani *et. al.* have quantitatively monitored the spatial and temporal localization of one of the key signaling component fused to GFP [33]. Their data provides evidence for inherent asymmetries in the intracellular signaling network of cells prior to stimulation. The magnitude of this asymmetry was found to vary significantly from cell to cell. This naturally raises the question: what is the accuracy of directional sensing in the presence of such an intracellular bias?

To address this question, we can use a nonuniform prior distribution $P(\phi)$ to represent this internal bias and apply the *maximum a posteriori* (MAP) estimation. Specifically, we assume that the prior distribution follows the *circular normal* (CN) distribution in directional statistics, $P(\phi) = \exp[\kappa_\varepsilon \cos(\phi - \phi_\varepsilon)]/(2\pi I_0(\kappa_\varepsilon))$, where ϕ_ε denotes the bias direction, κ_ε controls the magnitude of the bias, and $I_0(\cdot)$ is the modified Bessel function of order zero. We have plotted the CN distribution in Fig. 3A for various values of κ_ε.

The MAP estimate of ϕ is

$$
\begin{aligned}
\widehat{\phi}_{map} &= \arg\max_{\phi} P(\mathbf{Y}|\phi)P(\phi) \\
&= \arg\min_{\phi} \left[\frac{1}{2\sigma_s^2} J(\mathbf{Y}|\boldsymbol{\Theta}) - \kappa_\varepsilon \cos(\phi - \phi_\varepsilon) \right] \\
&= \arctan \left(\frac{\sum_m y_m \sin\vartheta_m + \gamma \sin\phi_\varepsilon}{\sum_m y_m \cos\vartheta_m + \gamma \cos\phi_\varepsilon} \right) \\
&= \arctan \left(\frac{z_1 + \gamma \sin\phi_\varepsilon}{z_2 + \gamma \cos\phi_\varepsilon} \right),
\end{aligned}
\tag{52}
$$

with $\gamma \equiv 2\kappa_\varepsilon/p$. This expression is similar to the maximal polarization angle derived from the geometric model in [33]. In the limit $\kappa_\varepsilon \to 0$, the prior distribution becomes uniform and the MAP estimate recovers the MLE of the gradient direction, i.e., $\widehat{\phi}_{map} \to \widehat{\phi}$. In the other extreme that $\kappa_\varepsilon \to \infty$, the prior distribution becomes a delta function and $\widehat{\phi}_{map} \to \phi_\varepsilon$. Let $\widetilde{z}_1 \equiv z_1 + \gamma \cos\phi_\varepsilon$ and $\widetilde{z}_2 \equiv z_2 + \gamma \sin\phi_\varepsilon$. Then, by Eq. (10), they are independent and both normal: $\widetilde{z}_1 \sim \mathcal{N}(\alpha_1 + \gamma \cos\phi_\varepsilon, \sigma^2)$ and $\widetilde{z}_2 \sim \mathcal{N}(\alpha_2 + \gamma \sin\phi_\varepsilon, \sigma^2)$. If we define $\widetilde{\mathbf{Z}} = \widetilde{z}_1 + i\widetilde{z}_2$, then $\widetilde{\mathbf{Z}}$ follows the complex Gaussian distribution and its phase variable is just the MAP estimator, i.e., $\widehat{\phi}_{map} = \arctan(\widetilde{z}_2/\widetilde{z}_1)$. As demonstrated in [25, 26], the phase of a complex Gaussian random variable has a distribution symmetric about the expected value of the phase. In our case, this implies that the MAP estimator $\widehat{\phi}_{map}$ is symmetrically distributed about its mean

$$
\begin{aligned}
\langle\widehat{\phi}_{map}\rangle &= \arctan \frac{\langle\widetilde{z}_2\rangle}{\langle\widetilde{z}_1\rangle} = \arctan \left(\frac{\mu p \cos\phi + \gamma \cos\phi_\varepsilon}{\mu p \sin\phi + \gamma \sin\phi_\varepsilon} \right) \\
&\approx \begin{cases} \phi_\varepsilon & \text{if } \kappa_\varepsilon \gg \kappa, \\ \phi & \text{if } \kappa_\varepsilon \ll \kappa. \end{cases}
\end{aligned}
\tag{53}
$$

One can see that the value of $\langle\widehat{\phi}_{map}\rangle$ is largely determined by the ratio $\gamma/(\mu p) = 2\kappa_\varepsilon/(\mu p^2) = \kappa_\varepsilon/\kappa$, where κ is the signal-to-noise ratio defined in Eq. (12).

In statistics, the bias function of an estimator is the difference between this estimator's expected value and the true value of the parameter being estimated. In our case, the bias function of the MAP estimator is

$$
b_{map}(\phi) \equiv \langle\widehat{\phi}_{map}\rangle - \phi = \arctan \left(\frac{\kappa \sin\phi + \kappa_\varepsilon \sin\phi_\varepsilon}{\kappa \cos\phi + \kappa_\varepsilon \cos\phi_\varepsilon} \right) - \phi.
\tag{54}
$$

Obviously, $\widehat{\phi}_{map}$ is a biased estimator of ϕ due to bias in $P(\phi)$. The squared estimation error of $\widehat{\phi}_{map}$ is

$$
\begin{aligned}
\sigma_{\phi,map}^2 &= \langle(\widehat{\phi}_{map} - \phi)^2\rangle \\
&= \langle(\widehat{\phi}_{map} - \langle\widehat{\phi}_{map}\rangle + \langle\widehat{\phi}_{map}\rangle - \phi)^2\rangle \\
&= \langle(\widehat{\phi}_{map} - \langle\widehat{\phi}_{map}\rangle)^2\rangle + b_{map}(\phi)^2.
\end{aligned}
\tag{55}
$$

According to the Cramér-Rao inequality for biased estimators [34] the first term is bounded by

$$
\langle(\widehat{\phi}_{map} - \langle\widehat{\phi}_{map}\rangle)^2\rangle \geq \left(1 + \frac{\partial b_{map}(\phi)}{\partial\phi} \right)^2 I_\phi^{-1},
\tag{56}
$$

where the Fisher information has been calculated before, $I_\phi = \left\langle (\partial_\phi \ln \mathcal{L})^2 \right\rangle = \sigma_\phi^{-2} = \mu p^2/2 = \kappa$. Therefore,

$$
\begin{aligned}
\sigma_{\phi,map}^2 &\geq \left(1 + \frac{\partial b_{map}(\phi)}{\partial\phi} \right)^2 I_\phi^{-1} + b_{map}(\phi)^2 \\
&= \kappa \left[\frac{\kappa + \kappa_\varepsilon \cos(\phi - \phi_\varepsilon)}{\kappa^2 + 2\kappa\kappa_\varepsilon \cos(\phi - \phi_\varepsilon) + \kappa_\varepsilon^2} \right]^2 \\
&\quad + \left[\arctan \left(\frac{\kappa \sin\phi + \kappa_\varepsilon \sin\phi_\varepsilon}{\kappa \cos\phi + \kappa_\varepsilon \cos\phi_\varepsilon} \right) - \phi \right]^2 \\
&= \begin{cases} \kappa/(\kappa + \kappa_\varepsilon)^2 & \text{if } \phi_\varepsilon = \phi, \\ \kappa/(\kappa - \kappa_\varepsilon)^2 & \text{if } \phi_\varepsilon = \phi \pm \pi. \end{cases}
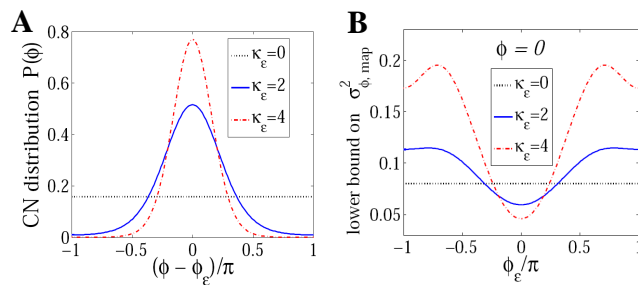\end{aligned}
\tag{57}
$$

FIG. 3: (Color online). A: The circular normal distribution for different values of κ_ε. B: The lower bound on the variance in gradient sensing $\sigma^2_{\phi,map}$ as a function of the direction of the bias ϕ_ε (relative to the gradient direction) for different values of the strength of the bias κ_ε. Parameter values are $N = 40,000$ and $C_0 = K_d$, resulting in a SNR of $\kappa = 12.5$.

In Fig. 3B we have plotted the lower bound of $\sigma_{\phi,map}$ as a function of the direction of the bias ϕ_ε for different values of the bias strength κ_ε. As expected, directional sensing is most accurate when the intracellular bias direction coincides with the extracellular gradient direction ($\phi_\varepsilon = \phi$) and becomes less accurate it differs significantly from the gradient direction. This is consistent with the experimental findings in [33] which showed that the GFP polarization is strongest at $\phi_\varepsilon = \phi$ and weaker at $\phi_\varepsilon = \phi \pm \pi$. Of course, we have chosen a fixed value for the parameter κ_ε, corresponding to a single cell. To deal with a population averaged as in typical experiments, one would have to assign a probability distribution to κ_ε (arising from cell individuality). This quenched randomness is expected to enlarge the cell-to-cell variability of directional sensing response, as observed in [33].

V. SUMMARY

In this paper we have used various concepts and techniques in estimation theory to investigate the physical limits of eukaryotic gradient sensing. We have derived explicit formulas for the variances of estimating both the gradient direction and steepness for an elliptical cell. Our theoretical and numerical results suggest that a cell cannot improve its sensing of both the gradient steepness and direction at the same time by simply elongating itself. We also show that highly eccentric cell shapes can significantly change the gradient sensing limits, which may be relevant in experimental observations for chemotactic eukaryotes like neutrophils and *Dictyostelium*. Finally, we examined how an intracellular bias may distort the cell's perception of external stimuli. As expected, the accuracy of gradient detection increases when the internal bias aligned with the external gradient but decreases when the direction of the internal bias is significantly different from the external gradient direction. Our approach is general and in principle can be extended to cases including non-uniformly distributed receptors or more complicated cell shapes.

VI. ACKNOWLEDGMENTS

[1] H. B. Barlow, Proc R Soc Lond B Biol Sci **212**, 1 (1981).
[2] F. Rieke and D. A. Baylor, Biophys J **75**, 1836 (1998).
[3] B. Houchmandzadeh, E. Wieschaus, and S. Leibler, Nature **415**, 798 (2002).
[4] D. Bray, M. D. Levin, and C. J. Morton-Firth, Nature **393**, 85 (1998).
[5] V. Sourjik and H. C. Berg, Proc Natl Acad Sci U S A **99**, 123 (2002).
[6] C. A. Parent and P. N. Devreotes, Science **284**, 765 (1999).
[7] P. J. V. Haastert and P. N. Devreotes, Nat Rev Mol Cell Biol **5**, 626 (2004).
[8] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice Hall, Upper Saddle River, 1993).
[9] C. Janetopoulos and R. A. Firtel, FEBS Lett **582**, 2075 (2008).
[10] W. J. Rappel and W. F. Loomis, Wiley Interdiscip Rev Syst Biol Med **1**, 141 (2009).
[11] M. Ueda, Y. Sako, T. Tanaka, P. N. Devreotes, and T. Yanagida, Science **294**, 864 (2001).

[12] L. Song, S. M. Nadkarni, H. U. Bödeker, C. Beta, A. Bae, C. Franck, W.-J. Rappel, W. F. Loomis, and E. Bodenschatz, Eur J Cell Biol **85**, 981 (2006).
[13] D. Fuller, W. Chen, M. Adler, A. Groisman, H. Levine, W. J. Rappel, and W. F. Loomis, PNAS **107**, 9656 (2010).
[14] H. C. Berg and E. M. Purcell, Biophysical Journal **20**, 193 (1977).
[15] W. Bialek and S. Setayeshgar, PNAS **102**, 10040 (2005).
[16] W. Bialek and S. Setayeshgar, Phys Rev Lett **100**, 258101 (2008).
[17] K. Wang, W. J. Rappel, R. Kerr, and H. Levine, Phys Rev E **75**, 061905 (2007).
[18] R. G. Endres and N. S. Wingreen, PNAS **105**, 15749 (2008).
[19] R. G. Endres and N. S. Wingreen, Phys Rev Lett **103**, 158101 (2009).
[20] W. J. Rappel and H. Levine, PNAS **105**, 19270 (2008).
[21] W. J. Rappel and H. Levine, Phys Rev Lett **100**, 228101 (2008).
[22] R. G. Endres and N. S. Wingreen, Prog Biophys Mol Biol **100**, 33 (2009).
[23] B. Hu, W. Chen, W. J. Rappel, and H. Levine, Phys Rev Lett **105**, 048104 (2010).
[24] B. Hu, W. Chen, H. Levine, and W. J. Rappel, Submitted.
[25] B. Hu, D. Fuller, W. F. Loomis, H. Levine, and W. J. Rappel, Phys Rev E **81**, 031906 (2010).
[26] H. Gudbjartsson and S. Patz, Magn Reson Med **34**, 910 (1995).
[27] P. J. van Haastert and M. Postma, Biophys J **93**, 1787 (2007).
[28] D. A. Lauffenburger and J. L. Linderman, *Receptors* (Oxford University Press, New York, 1996).
[29] M. D. Onsum, K. Wong, P. Herzmark, H. R. Bourne, and A. P. Arkin, Phys Biol **3**, 190 (2006).
[30] M. Nishio, K. Watanabe, J. Sasaki, C. Taya, S. Takasuga, R. Iizuka, T. Balla, M. Yamazaki, H. Watanabe, R. Itoh, et al., Nat Cell Biol **9**, 36 (2007).
[31] Monica Skoge, personal communication.
[32] A. J. Ridley, M. A. Schwartz, K. Burridge, R. A. Firtel, M. H. Ginsberg, G. Borisy, J. T. Parsons, and A. R. Horwitz, Science **302**, 1704 (2003).
[33] A. Samadani, J. Mettetal, and A. Oudenaarden, PNAS **103**, 11549 (2006).
[34] P. Dayan and L. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (The MIT Press, Boston, 2001).