Bo Hu · Wen Chen · Herbert Levine · Wouter-Jan Rappel ·

Quantifying information transmission in eukaryotic gradient sensing and chemotactic response

Abstract Eukaryotic cells are able to sense shallow chemical gradients by surface receptors and migrate toward chemoattractant sources. The accuracy of this chemotactic response relies on the ability of cells to infer gradients from the heterogeneous distribution of receptors bound by diffusing chemical molecules. Ultimately, the precision of gradient sensing is limited by the fluctuations of signaling components, including the stochastic receptor occupancy and noisy intracellular processing. Viewing the system as a Markovian communication channel, we apply techniques from information theory to derive upper bounds on the amount of information that can be reliably transmitted through a chemotactic cell. Specifically, we derive an expression for the mutual information between the gradient direction and the spatial distribution of bound receptors. We also compute the mutual information between the gradient direction and the motility direction using three different models for cell motion. Our results can be used to quantify the information loss during the various stages of directional sensing in eukaryotic chemotaxis.

Keywords Chemotaxis · Stochasticity · Mutual Information

1 Introduction

The directed movement of cells up or down a chemical gradient is known as chemotaxis. Although the underlying mechanisms are fundamentally different, both prokaryotic and eukaryotic cells employ chemotaxis as a way to direct cell motion. Bacteria direct their motion by measuring and comparing chemical concentrations over time [1]. Eukaryotic cells, on the other hand, are much slower and larger and are able to measure concentration differences in space. These eukaryotic cells, the focus of our study, plays an important role in a variety of biological processes, including neuronal patterning, wound healing, embryogenesis, and cancer metastasis [2–4].

The sensitivity of eukaryotic cells to gradients can be extremely high: both neutrophils and the social amoeba *Dictyostelium discoideum* cells can detect a 1-2% difference in concentration of the chemoattractant between the front and the back of the cell [5–7] and experiments with growth cones have claimed to exhibit axonal guidance in concentration differences as little as 0.1% [8]. Naturally, the question of how cells achieve such a high degree of sensitivity has attracted considerable attention. Clearly, chemotaxing cells are able to translate a shallow external gradient into a much larger internal asymmetry and this directional sensing capability has been the subject of numerous theoretical studies [9–15].

In eukaryotic cells, the first step in the chemotactic process consists of the binding of the chemoattractant to specific G-protein coupled receptors on the cell membrane. In the case of a chemoattractant gradient, this binding results in an asymmetric distribution of ligand-occupied receptors. These receptors then activate

Bo Hu, Wen Chen, Herbert Levine, and Wouter-Jan Rappel
Center for Theoretical Biological Physics and Department of Physics, University of California at San Diego, La Jolla, California 92093-0374, U.S.A.

multiple second-messenger pathways inside the cell, leading to asymmetric internal distributions of a multitude of signaling molecules. Eventually, these pathways drive the formation of actin-filled protrusions called pseudopodia. These pseudopodia are formed preferentially at the front, the side of highest chemoattractant concentration, and, together with a myosin-based trailing edge which pulls in the rear, results in directed cell movement. Many of the components responsible for translating the external chemoattractant gradient into cell motility are known and are conserved across species (for recent reviews, see [16–18]). The precise physical mechanism of this translation, however, remains poorly understood.

The binding of ligand molecules to chemoreceptors is an inherently noisy process and the question how noise influences cell motility has generated significant interest [7,19–28]. One way to study the effect of noise on the chemotactic process is to use information theoretic approaches [7,29]. We recently performed a theoretical investigation of the mutual information, a measure of the amount of information that two noisy variables share, between the input gradient direction and the resulting spatial distribution of ligand-bound receptors [7]. For shallow gradients, we were able to obtain approximate analytical expressions. Using a large experimental data set, we were also able to compute numerically the mutual information between the input gradient direction and the motility direction in the experiments. Comparing these two quantities allowed us to determine how much information was lost during intercellular processing.

Here, we extend our previous analysis and use information theoretic approaches to derive an explicit formula for the mutual information between the input gradient direction and the resulting distribution of ligand-bound receptors. This *external* mutual information reflects how the external receptor noise limits the gradient information acquisition at the cell membrane and provides an upper bound on the amount of information that can be reliably transmitted during gradient sensing at the receptor level. Furthermore, we propose and study several stochastic models that connect the external receptor signal to the output of chemotactic direction. These models allow us to calculate, analytically and/or numerically, the mutual information between the input source direction and the output chemotactic response angle. We will call this the *chemotactic* mutual information to distinguish it from the external mutual information. It quantifies the total information obtained by a chemotactic cell and will be at most equal to the external mutual information. In fact, by comparing this quantity to the external mutual information, we can determine how much information is dissipated due to intracellular fluctuations and nonlinear signal processing.

2 Results

2.1 Spatial distribution of stochastic ligand-receptor binding

Our model is shown in Fig. 1, along with the relevant notation and the various steps in the chemotactic process. We assume a circular cell with diameter L that is placed in a chemoattractant gradient with direction ϕ. A large number of receptors (typically, $N \approx 80,000$) is uniformly distributed on the cell surface, acting as the antennae for gradient sensing. 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 snapshot constitute a series of independent Bernoulli random variables, represented by

$$s_n = \begin{cases} 1, & \text{with probability} \quad P_n, \\ 0, & \text{with probability} \quad 1 - P_n, \end{cases} \qquad \text{for } n = 1, ..., N. \tag{1}$$

If the local concentration near the n_{th} receptor is C_n, then the probability of occupancy is $P_n = C_n/(C_n + K_d)$ for simple ligand-receptor kinetics with dissociation constant K_d. We divide the cell into M small sensory sectors such that each sector contains several hundreds of receptors exposing to almost the same local concentration. The gradient field at which the cell is situated can be fully described by the average local concentration C_{local}, the gradient steepness $p \equiv \frac{L}{C_{local}} \frac{\delta C}{\delta r}$ and the gradient direction ϕ. The steepness p reflects the percentage concentration change across the cell length L.

The local concentration at the m_{th} sector with angular position φ_m is given by $C_m = C_{local}[1 + \frac{p}{2}\cos(\varphi_m - \phi)]$. Neglecting any temporal correlation of receptors, we find that the number of occupied receptors in the m_{th} sector is $Y_m = E[Y_m] + \eta_m = N_s C_m/(C_m + K_d) + \eta_m$ for $m = 1, ..., M$, with $\langle \eta_m(t)\eta_n(s)\rangle \approx N_s C_m K_d \delta(t - s)\delta_{mn}/(C_m + K_d)^2$ [22] and where $N_s = N/M$. In this way, the receptor signal is decomposed into M independent but non-identical Gaussian random variables, denoted by $\mathbf{Y} = \{Y_1, Y_2, ..., Y_M\}^T$. Hereafter, boldface

Information Flow in Eukaryotic Chemotaxis
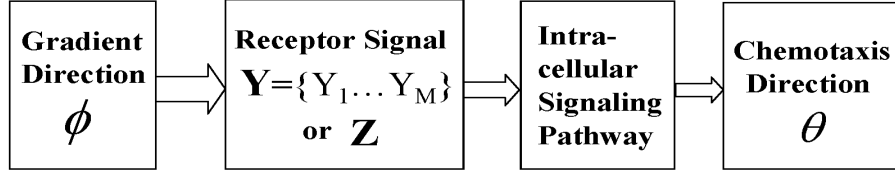
$$P(\phi) \rightarrow P(\mathbf{Y} \,|\, \phi) \rightarrow P(\mathbf{Z} \,|\, \phi) \rightarrow P(\theta \,|\, \phi)$$
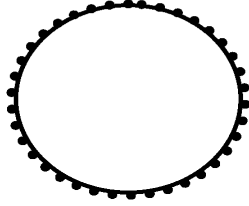
Receptor State	Probability	Description

$$S_n = \begin{cases} 1, & P_n, & \text{ligand-occupied} \\ 0, & 1 - P_n, & \text{unoccupied} \end{cases}$$

Fig. 1 Illustration of the information flow in eukaryotic chemotaxis. A chemical gradient with direction ϕ is first detected by receptors on the cell membrane, resulting in a spatially heterogeneous distribution of receptor occupancy, represented by $\mathbf{Y}$ or the sufficient statistic $\mathbf{Z}$. The noisy receptor signal is further transmitted through intracellular pathways, which convert the external asymmetry into an internal one. Ultimately, this leads to directed cell motion toward the chemical source, represented by the moving angle θ. We consider a two-dimensional circular cell with a large number of independent receptors uniformly located on the cell perimeter.

symbols denote vectors, and the superscript symbol ^{T} denotes the transpose. For small gradients ($p < 0.1$), we can expand Y_m in p, resulting in

$$Y_m = \frac{N_s C_{local}[1 + \frac{p}{2}\cos(\varphi_m - \phi)]}{K_d + C_{local} + \frac{p}{2}C_{local}\cos(\varphi_m - \phi)} + \eta_m \approx \frac{N_s C_{local}}{K_d + C_{local}} + \frac{N_s K_d C_{local}}{2(C_{local} + K_d)^2}p\cos(\varphi_m - \phi) + \mathscr{O}(p^2) + \eta_m,$$

(2)

To leading order in p, the covariance is written as $\langle \eta_m(t)\eta_n(s)\rangle \approx N_s K_d C_{local}\delta(t - s)\delta_{mn}/(C_{local} + K_d)^2 \equiv \sigma_s^2$ and is the same for all sectors. Thus, the receptor signal $\mathbf{Y} = \{Y_1, Y_2, ..., Y_M\}^T$ is a vector of independent Gaussian random variables with different means but approximately identical variance σ_s^2. Note that our approach uses the assumption that we can subdivide the membrane into independent sectors. This assumption is reasonable since there is no direct experimental evidence that individual receptors in chemotaxing eukaryotic cells are coupled. Also, as we will see below, our small gradient expansion is valid up to at least a gradient steepness of 20%.

2.2 The external mutual information

Receptors close to the gradient source are more likely to be occupied by ligands than those away from the source. This results in a heterogeneous distribution of the receptor occupancy on the cell surface, as reflected by Eq. 2. If all the intracellular processes are ignored then the best chemotactic decision the cell can make will be based on the spatially asymmetric distribution of this receptor occupancy. A natural question concerns how much information can be reliably transmitted in gradient sensing at the cell surface. This can be quantified by the mutual information between the spatial distribution of bound receptors $\mathbf{Y}$ and the gradient direction ϕ, which is chosen randomly from a preassigned distribution $P(\phi)$. By the Bayesian interpretation of probability, the prior distribution $P(\phi)$ expresses the cell's uncertainty about the gradient direction before taking any measurements into account.

Mutual information is one of the core concepts in information theory [30]. In the continuous case, the definition of mutual information between input (X) and output (Y), expressed in nats, is

$$I(X;Y) = \int_Y \int_X P(x,y) \ln \frac{P(x,y)}{P(x)P(y)} dx dy, \tag{3}$$

where $P(x,y)$ is the joint probability density function of X and Y, and $P(x)$ and $P(y)$ are the marginal probability density functions. By definition, we have $I(X;Y) = I(Y;X)$. Mutual information quantifies the mutual dependence of two random variables and is more general than the correlation coefficient which only measures the linear dependence of two variables.

Since in our model $\mathbf{Y}$ is a high-dimensional random vector, calculating the mutual information $I(\mathbf{Y};\phi)$ involves a difficult multidimensional integration which hampers direct analytical studies. This was shown in our recent study, where we also presented an expansion for shallow gradients [7]. A more tractable analytical expression can be found using a statistic $\mathbf{Z} \equiv \sum_{m=1}^{M} Y_m e^{i\varphi_m} = Z_1 + iZ_2$ which is *sufficient* for the gradient parameter ϕ (for more details, see the Appendix). In other words, the complex random variable $\mathbf{Z}$ contains the full information about the gradient direction and $\mathbf{Z}$ is just as informative as the full observations $\mathbf{Y}$. Thus, $I(\mathbf{Z};\phi) = I(\mathbf{Y};\phi)$ regardless of the input statistics $P(\phi)$. In the Appendix, we show that $\mathbf{Z}$, to first order of p, has a complex Gaussian distribution and that its real and imaginary parts are independently normal with identical variance $\sigma^2 \equiv \frac{1}{2} N K_d C_{local}/(C_{local} + K_d)^2$ but different means: $Z_1 \sim \mathcal{N}(\nu \cos\phi, \sigma^2)$ and $Z_2 \sim \mathcal{N}(\nu \sin\phi, \sigma^2)$, where $\nu = p\sigma^2/2$. In polar coordinates we can write $\mathbf{Z} = \rho e^{i\psi}$ where the magnitude variable $\rho \equiv \sqrt{Z_1^2 + Z_2^2}$ measures the degree of asymmetry in the receptor occupancy and where the phase variable $\psi \equiv \arctan(Z_2/Z_1)$ is found to be the optimal estimator of the gradient direction ϕ [31,32]. As shown in the Appendix, the magnitude ρ follows the Rice distribution, denoted by $P(\rho)$. Assuming that ϕ is chosen from a uniform distribution, i.e., $P(\phi) = 1/(2\pi)$, we can find an analytical expression of the mutual information $I(\mathbf{Z};\phi)$ between the gradient direction and $\mathbf{Z}$ (see Appendix). This expression is a monotone function of the *signal-to-noise ratio* (SNR), $\kappa \equiv \nu^2/\sigma^2 = Np^2 C_{local} K_d/(8(C_{local} + K_d)^2)$, and one can derive expressions in both the small and large κ limit:

$$I(\mathbf{Z};\phi) = \frac{\nu^2}{\sigma^2} - \left\langle \ln\left[I_0\left(\frac{\rho\nu}{\sigma^2}\right)\right]\right\rangle_{P(\rho)} \approx \begin{cases} \kappa/2 & \kappa \ll 1, \\ \ln\sqrt{2\pi\kappa/e} & \kappa \gg 1, \end{cases} \tag{4}$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind and order zero. Note that in the small SNR regime, the asymptotic expression of $I(\mathbf{Z};\phi)$ in Eq. 4 is identical to the approximate result for $I(\mathbf{Y};\phi)$ derived in our earlier study [7]. A more technical discussion of the various limits can be found in the appendix. In Fig. 2A, we plot $I(\mathbf{Z};\phi)$ as a function of the gradient steepness (blue dashed line), along with the numerically obtained values for $I(\mathbf{Y};\phi)$, using a Monte Carlo method detailed in Ref. [7]. As expected, the numerical values of $I(\mathbf{Y};\phi)$ lie exactly on the theoretical curve of $I(\mathbf{Z};\phi)$ given in Eq. 4. Note that the Monte Carlo method does not employ a small gradient expansion. Thus, the results of Fig. 2A demonstrate that the expansion we used in deriving our theoretical results is valid up to at least a gradient steepness of 20%. A plot of $I(\mathbf{Z};\phi)$ as a function of C_{local} (Fig. 2B) demonstrates that the mutual information is maximal when $C_{local} = K_d$.

By definition, the mutual information of two random variables measures their mutual dependence, or how much knowing one of these variables reduces the uncertainty about the other. Hence, the larger the mutual information $I(\mathbf{Y};\phi)$, the less uncertain the cell is about the gradient direction ϕ given the receptor signal $\mathbf{Y}$. In fact, for a Gaussian channel, one can prove that the *minimum mean-square error* (MMSE) in estimating the input given the output is related to the mutual information. It is given by twice the first-order derivative of the mutual information with respect to the SNR [33]. In our case, the channel is asymptotically Gaussian at large values of the SNR, and this elegant relationship suggests that the MMSE in estimating ϕ given the observation $\mathbf{Y}$ is equal to $1/\kappa$ or σ^2/ν^2. This is exactly the asymptotic variance of the optimal estimator ψ for the gradient parameter ϕ, i.e., $\psi \sim \mathcal{N}(\phi, 1/\kappa)$; see Appendix for more details. As an information measure, the external mutual information $I(\mathbf{Z};\phi)$ also sets an upper bound on the amount of information that could be achieved at the cell surface in a single snapshot. We can use it as a benchmark to compare with the information ultimately gained by the chemotactic cell, as to be examined in the next section.
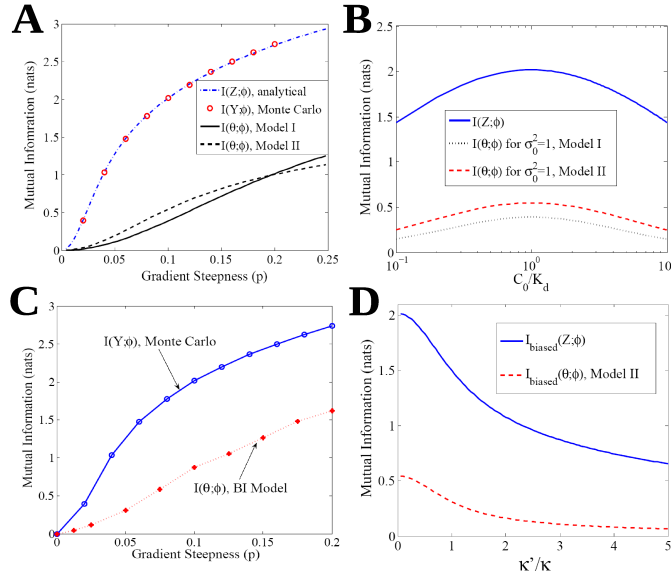
Fig. 2 The external and chemotactic mutual information for a cell with diameter $L = 10\mu m$, disassociation constant $K_d = 50nM$, number of receptors $N = 80000$, average local concentration $C_{local} = 50nM$, and added random motility noise $\sigma_0^2 = 1$. (A) The external mutual information $I(\mathbf{Z}; \phi)$ as a function of the gradient steepness p calculated using the analytically derived formula in Eq. 4 is plotted as a dashed line while the external mutual information obtained using a Monte Carlo algorithm is plotted using symbols. The chemotactic mutual information $I(\theta; \phi)$ for motility model I is shown as a solid black line and for motility and model II as a dashed line (B) Chemotactic mutual information $I(\theta; \phi)$ as a function of the average local concentration C_{local} for model I and model II using a fixed gradient steepness $p = 0.1$. Also plotted, as a solid blue line, is the external mutual information $I(\mathbf{Z}; \phi)$. (C) Chemotactic mutual information $I(\theta; \phi)$ as a function of the gradient steepness p for the BI model. The model parameter values are $k_a = 10s^{-1}$, $k_b = 3\mu m/s$, $k_{-a} = 0.2s^{-1}$, $k_{-b} = 0.2s^{-1}$, $k_i = 1000\mu m(s \cdot \text{molecule})^{-1}$, $D = 10\mu m^2/s$, and $D_m = 0.1\mu m^2/s$. As a comparison, the external mutual information $I(\mathbf{Z}; \phi)$ is plotted as a solid blue line. (D) External mutual information for biased input statistic, $I_{biased}(\mathbf{Z}; \phi)$, as a function of the prior knowledge about the gradient (represented by κ'/κ). The dashed line shows $I_{biased}(\theta; \phi)$ for model II using $p = 0.1$. Note that the parameter κ is determined by the other given parameters.

2.3 The chemotactic mutual information

The asymmetry of the receptor signal is amplified through a series of intracellular signaling events which eventually give rise to the chemotactic response (see Fig. 1). The receptor noise will propagate through the internal signaling networks which themselves are intrinsically noisy and may further interfere with the chemotactic decision-making. This noise is independent of the gradient sensing mechanism, but intrinsic to the cellular motility machinery and we will refer to it as the *random motility noise*. Due to this additional interference and due to possible non-linear signal processing, the amount of the gradient information that a chemotactic cell ultimately acquires could be much lower than that received at the cell surface. In our earlier study, we computed the information transmission during the entire chemotactic process using experimentally obtained cell tracks [7]. Here, we will investigate the mutual information $I(\theta; \phi)$ between the gradient direction ϕ and the response direction θ using three theoretical motility models. These models use the above computed spatial distribution of bound receptors as input and compute the resulting motility direction. In the first two models, the random motility noise is put in "by hand" while in the third model we explicitly model a proposed signaling pathway.

Model I. Since the complex random variable $\mathbf{Z}$ contains all the gradient information underlying the receptor signal $\mathbf{Y}$, we can use it as the input for a motility model. This leads to a model in which the cell is assumed to linearly transform the receptor signal $\mathbf{Z}$ to determine its direction. The effect of intracellular stochasticity can then be modeled by simply adding a noise term to $\mathbf{Z}$. For convenience, we write $\mathbf{Z}$ in the vector form: $\mathbf{Z} = (Z_1, Z_2)^T$. The model is defined as $\theta = \arctan(\widetilde{Z}_2/\widetilde{Z}_1)$ where $\widetilde{\mathbf{Z}} \equiv (\widetilde{Z}_1, \widetilde{Z}_2)^T = \beta \mathbf{Z} + \mathbf{\Sigma_0}$. Here, β is a constant scalar representing the signal amplification and $\mathbf{\Sigma_0}$ is a two-dimensional white random vector with zero mean and autocorrelation matrix $\langle \mathbf{\Sigma_0} \mathbf{\Sigma_0}^T \rangle = \sigma_0^2 \mathbf{I}$, representing random motility noise. Then, we have $\widetilde{Z}_1 \sim \mathscr{N}(\beta \nu \cos \phi, \beta^2 \sigma^2 + \sigma_0^2)$ and $\widetilde{Z}_2 \sim \mathscr{N}(\beta \nu \sin \phi, \beta^2 \sigma^2 + \sigma_0^2)$, which implies a new SNR,

$\widetilde{\kappa} \equiv \beta^2 v^2/(\beta^2\sigma^2 + \sigma_0^2)$. Thus, for large SNR, similar to the phase variable ψ in $\mathbf{Z}$, the response angle is asymptotically Gaussian, $\theta \sim \mathcal{N}(\phi, 1/\widetilde{\kappa})$. Again, using polar coordinates, we can write $\widetilde{\mathbf{Z}} = \widetilde{\rho}e^{i\theta}$ and we find the mutual information between gradient direction and motility direction is given by (see Appendix)

$$I(\theta;\phi) = I(\widetilde{\mathbf{Z}};\phi) = \widetilde{\kappa} - \left\langle \ln\left[I_0\left(\frac{\widetilde{\rho}\widetilde{\kappa}}{\beta v}\right)\right]\right\rangle_{P(\widetilde{\rho})} \approx \begin{cases} \widetilde{\kappa}/2 & \widetilde{\kappa} \ll 1, \\ \ln\sqrt{2\pi\widetilde{\kappa}/e} & \widetilde{\kappa} \gg 1. \end{cases} \tag{5}$$

For finite and positive β, $\widetilde{\kappa} < \kappa$ and hence $I(\theta;\phi) = I(\widetilde{\mathbf{Z}};\phi) < I(\mathbf{Z};\phi) = I(\mathbf{Y};\phi)$. As expected, the difference between $I(\theta;\phi)$ and $I(\mathbf{Y};\phi)$ in this linear model is enlarged when the intrinsic motility noise σ_0^2 increases or when the amplification parameter β decreases. If the intrinsic noise is negligible compared to the amplified receptor noise ($\sigma_0^2 \ll \beta^2\sigma^2$), we have $\widetilde{\kappa} \to \kappa$ and $I(\theta;\phi) \to I(\mathbf{Y};\phi)$. Thus, in this limit no additional information is lost in the internal steps. In Fig. 2A we have plotted $I(\theta;\phi)$ as a function of the gradient steepness (solid black line). For the particular value of the random motility noise chosen here, $I(\theta;\phi)$ is clearly reduced compared to $I(\mathbf{Z};\phi)$, reflecting the information loss due to noise. In addition, one can easily check that $\widetilde{\kappa}$, just as κ, has a maximum at $C_{local} = K_d$ as is shown in Fig. 2B. This can be understood by realizing that the applied intracellular linear transformation does not introduce an internal response threshold. Note that this simple model can be generalized easily. For example, we can replace the scalar β by a matrix (not necessarily symmetric) which provides a way to model any internal asymmetry of the cell.

Model II. Going beyond linear models, we can implement a more complicated transformation of $\mathbf{Z}$ to model the chemotactic response. Ideally, a chemotactic cell will try to align its movement with the estimated gradient direction ψ. The efficiency to adjust its direction may depend on the strength of the receptor signal (characterized by ρ) as well as on how responsive the cell is to that signal (parametrized by β). Based on symmetry requirements, we introduce a Langevin equation for the chemotactic response, $d\theta/dt = -\beta\rho\sin(\theta - \psi) + \eta_0$, where η_0 represents the random motility noise with $\langle\eta_0(t)\rangle = 0$ and $\langle\eta_0(t)\eta_0(s)\rangle = \sigma_0^2\delta(t-s)$ [25, 34].

In the small noise limit, we can expand this Langevin equation, resulting in

$$\begin{aligned} \frac{d\theta}{dt} &\approx -\beta v\sin(\theta - \phi) + \beta v\cos(\theta - \phi)\eta_\psi + \beta\sin(\theta - \phi)\eta_\rho + \eta_0, \\ &= -\beta v\sin(\theta - \phi) + \eta_{tot}, \end{aligned} \tag{6}$$

with the total noise given by

$$\langle\eta_{tot}^2\rangle = \beta^2 v^2(\sigma/v)^2\cos^2(\theta - \phi) + \beta^2\sigma^2\sin^2(\theta - \phi) + \sigma_0^2 = \beta^2\sigma^2 + \sigma_0^2 \equiv \sigma_{tot}^2. \tag{7}$$

Equivalently, this can be rewritten as a stochastic differential equation:

$$d\theta \approx -\beta v\sin(\theta - \phi)dt + \sigma_{tot}dW_t, \tag{8}$$

where W_t is the standard Wiener process or Brownian motion. The deterministic term βv in Eq. 8 can be interpreted as the mean restoring force that aligns the cell toward the true gradient direction ϕ. Solving the associated Fokker-Planck equation yields the stationary distribution of $\theta(t)$ given by

$$P(\theta|\phi) = \frac{\exp(\widehat{\kappa}\cos(\theta - \phi))}{2\pi I_0(\widehat{\kappa})}, \tag{9}$$

which is known as the *circular normal* (CN) distribution with modal direction ϕ and shape parameter $\widehat{\kappa} = 2\beta v/(\beta^2\sigma^2 + \sigma_0^2)$. In the limit of small $\widehat{\kappa}$, the CN density becomes the uniform distribution, while in the large $\widehat{\kappa}$ limit it approaches a Gaussian with variance $1/\widehat{\kappa}$. We have performed explicit Monte-Carlo simulations of the original Langevin equation ($d\theta/dt = -\beta\rho\sin(\theta - \psi) + \eta_0$) and have verified that the resulting directional distribution of θ agrees well with the CN distribution obtained from our approximate model (Eq. 8). We can compute (see Appendix) the mutual information $I(\theta;\phi)$ for this model as

$$I(\theta;\phi) = \widehat{\kappa}\frac{I_1(\widehat{\kappa})}{I_0(\widehat{\kappa})} - \ln I_0(\widehat{\kappa}) \approx \begin{cases} \widehat{\kappa}^2/4 - 3\widehat{\kappa}^4/64 + \mathcal{O}(\widehat{\kappa}^6) & \widehat{\kappa} \ll 1, \\ \ln\sqrt{2\pi\widehat{\kappa}/e} & \widehat{\kappa} \gg 1, \end{cases} \tag{10}$$

where $I_1(\cdot)$ is the first kind modified Bessel function of order one. In Fig. 2 we plot this mutual information as a function of the gradient steepness (dashed line, A) and of the background concentration (dashed line, B).

The mutual information is an increasing function of the gradient steepness while it reaches a maximum for $C_{local} = K_d$.

Model III. The last class of models we examine here explicitly incorporates the intracellular dynamics of second messenger pathways. Thus, instead of using ad-hoc additive noise terms as in Model I and II, we now represent the intracellular dynamics by a set of partial differential equations. We will focus here on one particular directional sensing model, the Balanced Inactivation (BI) model [10], although we have performed a similar analysis for other models (data not shown). In this model, receptor occupancy triggers the production of a membrane-bound activator A (read-out component) and a cytosolic diffusing inhibitor B at equal rates k_a. The diffusing species B can attach itself to the membrane at a rate k_b and become a membrane-bound species B_m. Near the membrane, it can irreversibly inactivate the read-out component A with rate k_i. Both A and B_m have a spontaneous degradation rates (k_{-a} and k_{-b}, respectively) which are assumed to be small compared with both the activation and the recombination rates. As the input for the BI model we use the stochastic receptor signal, $Y(\varphi_j)$ for $j = 1...M$ as computed above. These reactions can be represented by the following mathematical equations:

$$\frac{\partial A(\varphi_j)}{\partial t} = k_a Y(\varphi_j) - k_{-a}A - k_i A B_m + D_m \nabla_m^2 A, \qquad \text{at the membrane,}$$

$$\frac{\partial B_m(\varphi_j)}{\partial t} = k_b B - k_{-b}B_m - k_i A B_m + D_m \nabla_m^2 B_m, \qquad \text{at the membrane,} \qquad (11)$$

$$\frac{\partial B}{\partial t} = D\nabla^2 B, \qquad \text{in the cytosol,}$$

with a boundary condition for the outward pointing normal derivative of the cytosolic component:

$$D\frac{\partial B}{\partial n} = k_a Y(\varphi_j) - k_b B. \qquad (12)$$

Note that we have taken into account possible membrane diffusion of B_m, characterized by the membrane diffusion constant D_m.

We numerically solve this model for a 2D circular cell subdivided into M equal sized sectors and obtain the *steady-state* spatial distribution of the read-out component $A(\varphi_j)$ for each sector $j = 1,...,M$. Specifically, we generated 1000 independent realizations of the Gaussian random vector **Y** and used these as input to Eqns. 12. Importantly, to compute the chemotactic mutual information of the BI model we still need to relate the read-out component to motility. For this, we choose a spatial filter similar to the one we used before: $\widehat{\mathbf{Z}} \equiv \sum_{j=1}^{M} A(\varphi_j)e^{i\varphi_j} = \widehat{Z}_1 + i\widehat{Z}_2$ and $\theta \equiv \arctan(\widehat{Z}_2/\widehat{Z}_1)$. In this way, we are able to compute the output direction θ using the steady-state solution of A for each realization of **Y**. The resulting distribution of θ allows us to evaluate the mutual information $I(\theta;\phi)$ for the BI model, using either histogram estimation or kernel estimation [35,36]. The numerical result of $I(\theta;\phi)$ is plotted in Fig. 2C as a function of the gradient steepness. Again, as is the case for models I and II, the mutual information of the entire chemotactic process is reduced compared to the external mutual information $I(\mathbf{Z};\phi)$. Note, however, that receptor noise is the only stochasticity we have introduced into the BI model whereas models I and II have included extra fluctuations to account for noisy intracellular biochemical reactions. Since the BI model is nonlinear, the information reduction relative to $I(\mathbf{Z};\phi)$ can be attributed solely to the nonlinear signal processing by downstream second messengers. We should point out that a direct comparison between experimental values of the mutual information (as obtained in Ref. [7]) and values obtained from our models is difficult since our models do not contain an detailed motility module. For example, the mutual information in the experiments saturates as the gradient steepness is increased (see Fig. 3 in Ref. [7]). In contrast, the mutual information in our models is an increasing function of gradient steepness. This suggests that the our models do not contain one or more (unknown) mechanisms that limit the achievable mutual information.

We have also tested how the specific choice of biochemical parameters can affect the mutual information. For example, we found that the mutual information $I(\theta;\phi)$ only shows a slight tendency to increase within the range $D_m = 0$ - $100 \, \mu m^2/s$, as shown in Fig. 3A. This range of membrane diffusion constants encompasses the physiologically relevant range for membrane bound proteins. We should note however, that the mutual information is a monotonic increasing function of D_m. The reason for this is that for larger and larger values of D_m, the noise gets more and more suppressed. This comes at the expense of the amplitude in the signal (i.e., the difference between the maximum and minimum values of A), which becomes smaller and smaller. The mutual information, however, is strictly a function of the direction of the output and does not take this
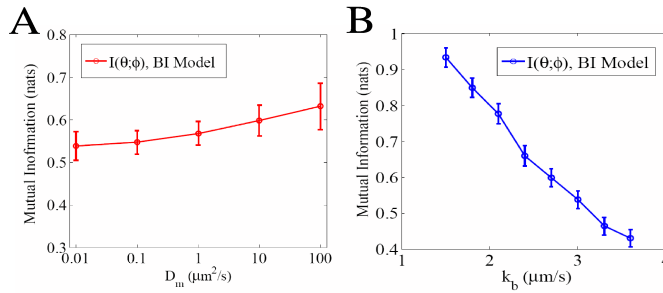
Fig. 3 (A) Mutual information $I(\theta;\phi)$ for the BI model as a function of the membrane diffusion coefficient D_m. The error bars represent the standard deviation of the Monte Carlo simulations. (B) $I(\theta;\phi)$ as a function of k_b for the BI model. In both plots, we have chosen the following parameter values: $L = 10\mu m$, $C_{local} = 20nM$, $K_d = 30nM$, $N = 80000$, $p = 0.075$, $k_a = 10s^{-1}$, $k_b = 3\mu m/s$ (left plot), $k_{-a} = 0.2s^{-1}$, $k_{-b} = 0.2s^{-1}$, $k_i = 1000\mu m(s \cdot \text{molecule})^{-1}$, $D = 10\mu m^2/s$, and $D_m = 0.1\mu m^2/s$ (right plot).

amplitude into account. A motility module that is downstream from the BI model will be dependent on this amplitude but is not part of our calculation. Other numerical tests demonstrated that the mutual information $I(\theta;\phi)$ is only sensitive to the parameter k_b, which denotes the rate that the diffusible species B can attach itself to the membrane and irreversibly become a membrane-bound species B_m. Larger values of this parameter correspond to higher levels of B_m and stronger inactivation of the read-out component A. Therefore, the shape of $I(\theta;\phi)$ as a function of the gradient steepness p can be most effectively tuned by choosing different values of k_b in the BI model.

2.4 The effect of a priori knowledge about the gradient

In the previous calculations, the prior distribution $P(\phi)$ has been assumed to be uniform, which can describe an unbiased cell that has no a priori knowledge about the gradient direction ϕ [29,37]. This may be the case when a cell is newly introduced into a gradient. However, long exposure to a gradient may bias the cell such that it expects the gradient to come predominantly from some direction ϕ'. To model this a priori knowledge we can use the circular normal distribution, i.e., $P(\phi) = \exp[\kappa'\cos(\phi - \phi')]/(2\pi I_0(\kappa'))$ where the parameter κ' controls the magnitude of the bias. In the limit $\kappa' \to 0$, the prior distribution is uniform ($P(\phi) = 1/(2\pi)$), whereas in the limit $\kappa' \to \infty$, it tends to be a Gaussian distribution with variance $1/\kappa'$. Without loss of generality, we will set $\phi' = 0$. Then the external mutual information is found to be (see Appendix)

$$I_{biased}(\mathbf{Z};\phi) = \kappa - \langle \ln[I_0(\rho\kappa/\nu)]\rangle_{P(\rho)} - \langle h(\kappa_\rho)\rangle_{P(\rho)} = I(\mathbf{Z};\phi) - \langle h(\kappa_\rho)\rangle_{P(\rho)}$$

$$\approx \begin{cases} \kappa\left(1 - [I_1(\kappa')/I_0(\kappa')]2\right)/2 & \kappa \ll 1, \\ \ln\sqrt{2\pi\kappa/e} - h(\kappa'') & \kappa \gg 1, \end{cases} \tag{13}$$

where $h(x) \equiv xI_1(x)/I_0(x) - \ln I_0(x)$, $\kappa_\rho \simeq \kappa'\nu\rho/(\kappa'\sigma 2 + \nu\rho)$ and $\kappa'' \simeq \kappa'\kappa/(\kappa' + \kappa)$. The function $h(x)$ is a monotone increasing function and is positive for all $x > 0$. Thus, the biased mutual information $I_{biased}(\mathbf{Z};\phi)$ is smaller than the mutual information $I(\mathbf{Z};\phi)$ for the unbiased case in Eq. 4 by the amount $\langle h(\kappa_\rho)\rangle$ which vanishes in the limit $\kappa' \to 0$. This is intuitively reasonable because a priori knowledge can help reduce the uncertainty (or entropy) of the input. The reduction in the mutual information can be seen in Fig. 2D where we plot $I_{bias}(\mathbf{Z};\phi)$ as a function of κ'/κ. A similar calculation for model II results in (see Appendix)

$$I_{biased}(\theta;\phi) = \widehat{\kappa}\frac{I_1(\widehat{\kappa})}{I_0(\widehat{\kappa})} - \ln\frac{I_0(\widehat{\kappa})}{I_0(\widehat{\kappa}')} - \widehat{\kappa}'\frac{I_1(\widehat{\kappa}')}{I_0(\widehat{\kappa}')} = I(\theta;\phi) - h(\widehat{\kappa}'), \tag{14}$$

where $\widehat{\kappa}' \simeq \widehat{\kappa}\kappa'/(\widehat{\kappa} + \kappa')$ and $\widehat{\kappa}$ has been defined in **Model II**. Again, the biased mutual information above is less than the unbiased result $I(\theta;\phi)$ in Eq. 10 and the difference $h(\widehat{\kappa}')$ is an increasing function of the bias parameter κ' (Fig. 2D). This implies that $I(\theta;\phi)$ is maximal when $P(\phi)$ is uniform (i.e. $\kappa' = 0$). We have

verified this prediction using the Blahut-Arimoto algorithm (see Appendix). Therefore, the mutual information in Eq. 10 under the assumption of the uniform prior distribution gives, by definition the *channel capacity* which represents an upper bound on the amount of information that can be reliably transmitted through the chemotactic information system. Similarly, the analytical result in Eq. 4 defines the channel capacity at the external, cell surface level.

3 Discussion

In this paper, we used the concept of mutual information to study the role of fluctuations in eukaryotic chemotaxis. There are several advantages to using the mutual information as a quantitative measure for chemotactic efficiency. In contrast to the commonly used chemotaxis index [6], it is possible to compute the mutual information at different stages of the chemotactic process. Also, it can tell us the minimum mean-squared error in estimating the input after observing the output, regardless of the input statistics. Finally, the channel capacity, defined as the mutual information maximized over all input distributions, gives the tightest upper bound on the amount of information that can be reliably transmitted over a communication channel.

We first computed the external mutual information for gradient sensing at the cell surface which is the first step in the overall chemotactic information processing. Therefore, this quantity represents the upper bound of information that can be transmitted to the intracellular decision-making system. We then proposed and analyzed three models that incorporate both the external receptor noise and random motility noise and computed their chemotactic mutual information. A comparison between the external and the chemotactic mutual information enabled us to determine how the internal signaling processes affect the chemotactic performance. For all three models, we find that the chemotactic mutual information is significantly reduced relative to the external mutual information. The information reduction in the linear model I is purely due to the addition of random motility noise. The third model (BI model with noisy receptor input) specifically models a second messenger pathway but does not incorporate any motility noise. Thus, the information reduction in the third model is caused by the nonlinear processing of the noisy receptor signal. Finally, the noisy receptor signal in model II is processed nonlinearly while this model contains additional random motility noise. These results together demonstrate that a significant amount of gradient information can be lost as a result of either intracellular motility noise or complex signal processing. It remains a challenge to further determine which factor matters more for specific experimental systems.

Our results are restricted to a snapshot of the receptor states, ignoring any auto-correlation of receptors. In reality, the receptors are correlated in time and this temporal correlation of signals changes the rate of information transmission at the receptor level. For observation times T that are comparable to or larger than the correlation time τ, the external mutual information Eq. 4 needs to be multiplied by a factor T/τ. For the external receptor binding process, this correlation time is determined by the diffusion and binding/unbinding of ligand molecules, i.e., $\tau = \tau_{\text{rec}} + \tau_{\text{diff}}$, where τ_{rec} is the time-scale of receptor-ligand reaction and τ_{diff} describes the average time to refresh the configuration by diffusion. For *Dictyostelium* cells, it is estimated that $\tau_{\text{rec}} \gg \tau_{\text{diff}}$ and $\tau_{\text{rec}} \sim 1s$, such that $\tau \sim 1s$. For the entire chemotactic process, the correlation time is most likely dominated by the lifetime of a pseudopod. This lifetime has been estimated to be approximately 10-20 seconds in *Dictyostelium* cells [38, 39].

In summary, we have used various tools from statistics and information theory to gain insights about the physical limits of gradient sensing and chemotactic efficiency. Our results may help derive a better understanding of the design principles of biological decision-making systems in general.

A Appendix

A.1 The Sufficient Statistic **Z** for Gradient Sensing

We first consider how to estimate the gradient direction ϕ from the receptor signal. Eq. 2 suggests that one can view $\mathbf{Y} = \{Y_1, Y_2, ..., Y_M\}^T$ as observations of a sinusoidal signal embedded in white Gaussian noise. It then becomes a problem that is solvable using the Maximum Likelihood Estimation (MLE). By Eq. 2, the probability density function (PDF) of $\mathbf{Y}$ conditional on the gradient direction ϕ reads (for small gradients $p < 0.1$),

$$P(\mathbf{Y}|\phi) \approx \frac{1}{(2\pi\sigma_s^2)^{M/2}} \exp\left[-\frac{1}{2\sigma_s^2}\sum_{m=1}^{M}\left(Y_m - \overline{Y} - \frac{\sigma_s^2}{2}p\cos(\varphi_m - \phi)\right)^2\right], \tag{15}$$

with $\overline{Y} \equiv N_s C_{local}/(C_{local}+K_d)$. This PDF can be factored as

$$P(\mathbf{Y}|\phi) \approx \frac{1}{(2\pi\sigma_s^2)^{M/2}} \exp\left[-\frac{1}{2\sigma_s^2}\sum_{m=1}^{M}\left(\frac{\sigma_s^4 p^2}{4}\cos^2(\varphi_m-\phi) - \sigma_s^2 p Z_1(\mathbf{Y})\cos\phi + \sigma_s^2 p Z_2(\mathbf{Y})\sin\phi\right)\right] \cdot \exp\left[-\frac{1}{2\sigma_s^2}\sum_{m=1}^{M}(Y_m-\overline{Y})^2\right] \quad (16)$$

where $Z_1(\mathbf{Y}) \equiv \sum_{m=1}^{M}(Y_m-\overline{Y})\cos\varphi_m = \sum_{m=1}^{M}Y_m\cos\varphi_m$ and $Z_2(\mathbf{Y}) \equiv \sum_{m=1}^{M}(Y_m-\overline{Y})\sin\varphi_m = \sum_{m=1}^{M}Y_m\sin\varphi_m$. According to the general Neyman-Fisher Theorem [40], $Z_1(\mathbf{Y})$ and $Z_2(\mathbf{Y})$ are jointly sufficient statistics for the estimation of ϕ. Therefore, we can introduce the spatial filter $\mathbf{Z} = \sum_{m=1}^{M}Y_m e^{i\varphi_m} = \sum_{m=1}^{M}Y_m\cos\varphi_m + i\sum_{m=1}^{M}Y_m\sin\varphi_m \equiv Z_1 + iZ_2$. This complex random variable itself is the sufficient statistic for the parameter ϕ.

We can evaluate $\mathbf{Z}$ by replacing the summation by integral when M is large, and expand the integrand around p for small gradients. Specifically, we obtain

$$\mathrm{E}[\mathbf{Z}] \approx M e^{i\phi}\int_0^{2\pi}\mathrm{E}[Y(\omega)]e^{i\omega}\frac{d\omega}{2\pi} = \int_0^{2\pi}\frac{N e^{i\phi}\left[1+\frac{p}{2}\cos\omega\right]e^{i\omega}}{1+K_d/C_{local}+\frac{p}{2}\cos\omega}\frac{d\omega}{2\pi} = \frac{pN C_{local}K_d}{4(C_{local}+K_d)^2}e^{i\phi} + \mathcal{O}(p^3), \quad (17)$$

$$\mathrm{Var}[Z_1] = \sum_{m=1}^{M}\sum_{n=1}^{M}\langle\eta_m\eta_n\rangle\cos\varphi_m\cos\varphi_n \approx \frac{M}{2\pi}\int_0^{2\pi}\frac{N C_m K_d\cos^2\varphi}{M(C_m+K_d)^2}d\varphi = \frac{N C_{local}K_d}{2(C_{local}+K_d)^2} + \mathcal{O}(p^2), \quad (18)$$

$$\mathrm{Var}[Z_2] = \sum_{m=1}^{M}\sum_{n=1}^{M}\langle\eta_m\eta_n\rangle\sin\varphi_m\sin\varphi_n \approx \mathrm{Var}[Z_1], \quad \mathrm{Cov}[Z_1,Z_2] = \sum_{m=1}^{M}\sum_{n=1}^{M}\langle\eta_m\eta_n\rangle\cos\varphi_m\sin\varphi_n = 0. \quad (19)$$

Therefore, Z_1 and Z_2 are independent Gaussian random variables with different means but approximately identical variance. Let

$$\nu = \frac{pN K_d C_{local}}{4(C_{local}+K_d)^2} \qquad \text{and} \qquad \sigma^2 = \frac{N C_{local}K_d}{2(C_{local}+K_d)^2}. \quad (20)$$

Then $\mathbf{Z} = Z_1 + iZ_2$ follows a complex Gaussian distribution to the first order of p,

$$P(Z_1,Z_2|\mathbf{X}) = \frac{1}{2\pi\sigma^2}\exp\left[-\frac{(Z_1-\nu\cos\phi)^2+(Z_2-\nu\sin\phi)^2}{2\sigma^2}\right], \quad (21)$$

In polar coordinates, we write $\mathbf{Z} = \rho e^{i\psi}$. The amplitude ρ measures the degree of asymmetry in the occupied receptor distribution, and the phase ψ is the MLE of the true gradient direction ϕ. Under the complex Gaussian density in Eq. 21, the amplitude ρ follows the Rice distribution:

$$P(\rho) = \frac{\rho}{\sigma^2}\exp\left[-\frac{\rho^2+\nu^2}{2\sigma^2}\right]I_0\left(\frac{\rho\nu}{\sigma^2}\right), \quad (22)$$

where I_0 is the modified Bessel function of the first kind and zeroth order. The raw moments for the Rice distribution are

$$\mu_k = \int\rho^k P(\rho)d\rho = \sigma^k 2^{k/2}\Gamma(1+k/2)L_{k/2}\left(-\frac{\nu^2}{2\sigma^2}\right), \quad (23)$$

where $\Gamma(\cdot)$ is the Gamma function and $L_{k/2}(\cdot)$ is the Laguerre polynomial. So, we have

$$\mathrm{E}[\rho] = \mu_1 = \sigma\sqrt{\frac{\pi}{2}}L_{1/2}\left(-\frac{\nu^2}{2\sigma^2}\right), \quad (24)$$

$$\mathrm{Var}[\rho] = \mu_2 - \mu_1^2 = 2\sigma^2 + \nu^2 - \frac{\pi\sigma^2}{2}L_{1/2}^2\left(-\frac{\nu^2}{2\sigma^2}\right). \quad (25)$$

For the phase variable, let $\widetilde{\psi} = \psi - \phi$ and $\xi = \frac{\nu^2}{2\sigma^2}$. Then, as in [41],

$$P(\widetilde{\psi}) = \frac{e^{-\xi}}{2\pi}\left[1 + \sqrt{\pi\xi}\cos\widetilde{\psi}\,e^{\xi\cos^2\widetilde{\psi}}(1+\mathrm{erf}(\xi^{\frac{1}{2}}\cos\widetilde{\psi}))\right], \quad (26)$$

$$\lim_{\xi\to 0}P(\widetilde{\psi}) = \frac{1}{2\pi} \quad \text{and} \quad \lim_{\xi\to\infty}P(\widetilde{\psi}) = \sqrt{\frac{\xi}{\pi}}\exp(-\xi\widetilde{\psi}^2). \quad (27)$$

Note that the probability $P(\widetilde{\psi})$ is symmetric about $\widetilde{\psi} = 0$, so the expectation is $\langle\widetilde{\psi}\rangle = 0$. This means that the MLE ψ is an unbiased estimator of the gradient direction ϕ, i.e. $\langle\psi\rangle = \phi = \arctan\frac{\langle Z_2\rangle}{\langle Z_1\rangle}$. Define the *signal to noise ratio* (SNR) as

$$\kappa \equiv \frac{\nu^2}{\sigma^2} = \frac{N p^2 C_{local}K_d}{8(C_{local}+K_d)^2}. \quad (28)$$

Numerically, it is found that when $\nu/\sigma > 3$, both ρ and ψ are approximately Gaussian [41]. In the white noise limit ($\tau \to 0$), one gets

$$\rho \approx \nu + \eta_\rho \quad \text{with} \quad \langle \eta_\rho(t)\eta_\rho(s) \rangle = \sigma^2 \delta(t-s) = \frac{NC_{local}K_d}{2(C_{local}+K_d)^2}\delta(t-s), \tag{29}$$

$$\psi \approx \phi + \eta_\psi \quad \text{with} \quad \langle \eta_\psi(t)\eta_\psi(s) \rangle = \frac{\sigma^2}{\nu^2}\delta(t-s) = \frac{8(C_{local}+K_d)^2}{NK_dC_{local}P^2}\delta(t-s) = \frac{1}{\kappa}\delta(t-s). \tag{30}$$

As an orthogonal transformation from the Cartesian coordinates, η_ρ is independent of η_ψ. For parameter values that represent *Dictyostelium* cells and for typical experimental conditions, we have verified that the above Gaussian approximation is excellent.

A.2 Calculation of the mutual information

Here, we derive expression for the various mutual information quantities in the main text. As mentioned above, we fix C_{local} and p for simplicity. For the distribution of the gradient direction ϕ, we consider two cases:

Case 1; unbiased cell: For a cell with no a priori knowledge of ϕ, we use the uniform distribution $P(\phi) = \frac{1}{2\pi}$. The mutual information of $\mathbf{Z}$ and ϕ is calculated as

$$\begin{aligned}
I(\mathbf{Z};\phi) &= H(\mathbf{Z}) - H(\mathbf{Z}|\phi) \\
&= -\int_{\mathbf{Z}} P(\mathbf{Z})\ln P(\mathbf{Z})d\mathbf{Z} + \int_{\mathbf{Z}} P(\mathbf{Z}|\phi)\ln P(\mathbf{Z}|\phi)d\mathbf{Z} \\
&= -\int_{\mathbf{Z}}\int_\phi P(\phi)P(\mathbf{Z}|\phi)d\phi \ln\left[\int_\phi P(\phi)P(\mathbf{Z}|\phi)d\phi\right]d\mathbf{Z} - \ln(2\pi e\sigma^2) \\
&= -\frac{1}{2\pi}\int_{\mathbf{Z}}\int_\phi P(\mathbf{Z}|\phi)d\phi \ln\left[\int_\phi P(\mathbf{Z}|\phi)d\phi\right]d\mathbf{Z} - \ln(e\sigma^2) \\
&= -\frac{1}{2\pi}\int_\rho\int_\psi \frac{P(\rho)}{\rho}\ln\left[\frac{P(\rho)}{\rho}\right]\rho d\rho d\psi - \ln(e\sigma^2) \\
&= \int_\rho P(\rho)\ln\rho d\rho - \int_\rho P(\rho)\ln P(\rho)d\rho - \ln(e\sigma^2) \\
&= \langle\ln\rho\rangle + H(\rho) - \ln(e\sigma^2) \\
&= \frac{\nu^2}{\sigma^2} - \left\langle\ln\left[I_0\left(\frac{\rho\nu}{\sigma^2}\right)\right]\right\rangle_{P(\rho)}.
\end{aligned} \tag{31}$$

In the above derivation, we have used the following results

$$\begin{aligned}
d\mathbf{Z} &= \rho d\rho d\psi, \\
\int_\phi P(\mathbf{Z}|\phi)d\phi &= \frac{1}{\sigma^2}\exp\left[-\frac{\rho^2+\nu^2}{2\sigma^2}\right]I_0\left(\frac{\rho\nu}{\sigma^2}\right) = \frac{P(\rho)}{\rho}, \\
\langle\ln\rho\rangle &= \int_\rho P(\rho)\ln\rho d\rho = \frac{1}{2}\Gamma\left(0, \frac{\nu^2}{2\sigma^2}\right) + \ln\nu, \\
H(\rho) &= \ln\sigma^2 - \langle\ln\rho\rangle + \int_\rho P(\rho)\frac{\rho^2+\nu^2}{2\sigma^2}d\rho - \int_\rho P(\rho)\ln\left[I_0\left(\frac{\rho\nu}{\sigma^2}\right)\right]d\rho, \\
&= \ln\sigma^2 - \langle\ln\rho\rangle + 1 + \frac{\nu^2}{\sigma^2} - \left\langle\ln\left[I_0\left(\frac{\rho\nu}{\sigma^2}\right)\right]\right\rangle_{P(\rho)}.
\end{aligned}$$

In the small SNR limit (i.e., $\kappa = \nu^2/\sigma^2 \ll 1$), we have $\ln I_0(\rho\nu/\sigma^2) \approx \rho^2\nu^2/(4\sigma^4)$. By Eq. 23, $\langle\ln I_0(\rho\nu/\sigma^2)\rangle_{P(\rho)} \approx \langle\rho^2\nu^2/(4\sigma^4)\rangle_{P(\rho)} = \nu^2(2\sigma^2+\nu^2)/(4\sigma^4)$. Thus, the approximate expression of the external mutual information for $\kappa \ll 1$ is $I(\mathbf{Z};\phi) \approx \nu^2/\sigma^2 - \nu^2(2\sigma^2+\nu^2)/(4\sigma^4) = \kappa/2 - \kappa^2/4 \approx \kappa/2 + \mathcal{O}(\kappa^2)$. Similarly, in the large SNR limit $\kappa \gg 1$, we found $I(\mathbf{Z};\phi) \approx \nu^2/\sigma^2 - \ln[I_0(\nu^2/\sigma^2)] - 1/2 = \kappa - \ln[I_0(\kappa)] - 1/2$. Since $I_0(\kappa) \approx e^\kappa/\sqrt{2\pi\kappa}$ as $\kappa \to \infty$, we can further simplify the expression: $I(\mathbf{Z};\phi) \approx \ln\sqrt{2\pi\kappa/e}$. In summary, the external mutual information is given by:

$$I(\mathbf{Z};\phi) = \frac{\nu^2}{\sigma^2} - \left\langle\ln\left[I_0\left(\frac{\rho\nu}{\sigma^2}\right)\right]\right\rangle_{P(\rho)} \approx \begin{cases} \kappa/2 & \kappa \ll 1, \\ \ln\sqrt{2\pi\kappa/e} & \kappa \gg 1. \end{cases} \tag{32}$$

The above equation indicates that $I(\mathbf{Z};\phi)$ is solely dependent on the SNR $= \nu^2/\sigma^2 \equiv \kappa$.

Note that the asymptotic result for small values of the signal-to-noise ratio coincides with the asymptotic result for a Gaussian channel with a normally distributed input [30]. In this case, the mutual information can be written as $\frac{1}{2}\log(1+\kappa)$ which is, in the limit of small *kappa*, simply $\kappa/2 - \kappa^2/4 + \mathcal{O}(\kappa^3)$. This similarity, however, is coincidental since our channel for small values

of κ is not Gaussian nor has a normally distributed input. For large values of κ, the channel becomes Gaussian (see Eq. 30) but the input remains uniformly distributed between 0 and 2π.

Next, we compute the chemotactic mutual information $I(\theta;\phi)$. For Model I, the chemotactic response angle is given by $\theta = \arctan(\widetilde{Z}_2/\widetilde{Z}_1)$ where $\widetilde{Z}_1 \sim \mathcal{N}(\beta v \cos\phi, \beta^2\sigma^2 + \sigma_0^2)$ and $\widetilde{Z}_2 \sim \mathcal{N}(\beta v \sin\phi, \beta^2\sigma^2 + \sigma_0^2)$. This implies a new SNR: $\widetilde{\kappa} \equiv \beta^2 v^2/(\beta^2\sigma^2 + \sigma_0^2)$. Thus, similar to the phase variable ψ in $\mathbf{Z}$, the response angle is asymptotically Gaussian, i.e., $\theta \sim \mathcal{N}(\phi, 1/\widetilde{\kappa})$. In polar coordinates, we can write $\widetilde{\mathbf{Z}} = \widetilde{Z}_1 + i\widetilde{Z}_2 = \widetilde{\rho}e^{i\theta}$ where again $\widetilde{\rho}$ follows the Rice distribution. Due to independence, the magnitude variable $\widetilde{\rho}$ does not tell us anything about the gradient direction ϕ. This means that, for the complex random variable $\widetilde{\mathbf{Z}}$, all the information about ϕ is contained in the phase variable θ. In other words, $I(\widetilde{\mathbf{Z}};\phi) = I(\theta;\phi)$. Similar to the calculation of $I(\mathbf{Z};\phi)$, we can easily derive the following

$$I(\widetilde{\mathbf{Z}};\phi) = \frac{\beta^2 v^2}{\beta^2\sigma^2 + \sigma_0^2} - \left\langle \ln\left[I_0\left(\frac{\widetilde{\rho}\beta v}{\beta^2\sigma^2 + \sigma_0^2}\right)\right]\right\rangle_{P(\widetilde{\rho})} \approx \begin{cases} \widetilde{\kappa}/2 & \widetilde{\kappa} \ll 1, \\ \ln\sqrt{2\pi\widetilde{\kappa}/e} & \widetilde{\kappa} \gg 1. \end{cases} \tag{33}$$

This proves Eq. 5.

For Model II, the response angle θ follows the *circular normal* (CN) distribution. Thus, we can directly find the chemotactic mutual information:

$$I(\theta;\phi) = \int_\theta \int_\phi P(\phi)P(\theta|\phi)\ln\frac{P(\theta|\phi)}{P(\theta)}d\theta d\phi = \ln 2\pi - H(\theta|\phi) = \widehat{\kappa}\frac{I_1(\widehat{\kappa})}{I_0(\widehat{\kappa})} - \ln I_0(\widehat{\kappa}). \tag{34}$$

where $P(\theta) = \int P(\theta|\phi)P(\phi)d\phi = \frac{1}{2\pi}$ and $\widehat{\kappa} \equiv 2\beta v/(\beta^2\sigma^2 + \sigma_0^2)$. In fact, $I(\theta;\phi) = h(\widehat{\kappa})$ is an increasing function of $\widehat{\kappa}$ and its limit behaviors are as follows:

$$I(\theta;\phi) \approx \begin{cases} \frac{1}{2}\ln(2\pi\widehat{\kappa}/e) & \widehat{\kappa} \gg 1, \\ \widehat{\kappa}^2/4 - 3\widehat{\kappa}^4/64 + \mathcal{O}(\widehat{\kappa}^6) & \widehat{\kappa} \ll 1. \end{cases} \tag{35}$$

Case 2; biased cell: We assume that $P(\phi) = \text{CN}(\phi', \kappa') = \exp[\kappa'\cos(\phi - \phi')]/(2\pi I_0(\kappa'))$, where the parameter κ' controls the degree of directional bias. Without loss of generality, we will set $\phi' = 0$. Then the external mutual information is,

$$\begin{aligned} I_{biased}(\mathbf{Z};\phi) &= -\int_\mathbf{Z}\int_\phi P(\mathbf{Z}|\phi)P(\phi)d\phi \ln\left[\int_\phi P(\mathbf{Z}|\phi)P(\phi)d\phi\right]d\mathbf{Z} - \ln(2\pi e\sigma^2) \\ &= -\int_\mathbf{Z}\left[\int_\phi P(\mathbf{Z}|\phi)\frac{\exp(\kappa'\cos\phi)}{2\pi I_0(\kappa')}d\phi\right]\ln\left[\int_\phi P(\mathbf{Z}|\phi)\frac{\exp(\kappa'\cos\phi)}{2\pi I_0(\kappa')}d\phi\right]d\mathbf{Z} - \ln(2\pi e\sigma^2) \\ &\approx -\int_\rho\int_\psi \frac{P(\rho)}{\rho}\frac{\exp(\kappa_\rho\cos\psi)}{2\pi I_0(\kappa_\rho)}\ln\left[\frac{P(\rho)}{\rho}\frac{\exp(\kappa_\rho\cos\psi)}{2\pi I_0(\kappa_\rho)}\right]\rho d\rho d\psi - \ln(2\pi e\sigma^2) \\ &= \int_\rho P(\rho)\ln\rho d\rho - \int_\rho P(\rho)\ln P(\rho)d\rho - \int_\rho P(\rho)\left[\kappa_\rho\frac{I_1(\kappa_\rho)}{I_0(\kappa_\rho)} - \ln I_0(\kappa_\rho)\right]d\rho - \ln(e\sigma^2) \\ &= \langle\ln\rho\rangle + H(\rho) - \ln(e\sigma^2) - \left\langle\kappa_\rho\frac{I_1(\kappa_\rho)}{I_0(\kappa_\rho)} - \ln I_0(\kappa_\rho)\right\rangle_{P(\rho)} \\ &= \frac{v^2}{\sigma^2} - \left\langle\ln\left[I_0\left(\frac{\rho v}{\sigma^2}\right)\right]\right\rangle_{P(\rho)} - \langle h(\kappa_\rho)\rangle_{P(\rho)}, \end{aligned} \tag{36}$$

where $\kappa_\rho(\rho)$ is defined by $\frac{I_1(\kappa_\rho)}{I_0(\kappa_\rho)} = \frac{I_1(\kappa')}{I_0(\kappa')}\frac{I_1(v\rho/\sigma^2)}{I_0(v\rho/\sigma^2)}$, or more conveniently $\kappa_\rho \simeq \kappa' v\rho/(\kappa'\sigma^2 + v\rho)$. The key result we have used is

$$\begin{aligned} P(\mathbf{Z}) &= \int_\phi P(\mathbf{Z}|\phi)P(\phi)d\phi \\ &= \int_\phi \frac{1}{2\pi\sigma^2}\exp\left[-\frac{(Z_1 - v\cos\phi)^2 + (Z_2 - v\sin\phi)^2}{2\sigma^2}\right]\frac{\exp(\kappa'\cos\phi)}{2\pi I_0(\kappa')}d\phi \\ &= \frac{1}{4\pi^2\sigma^2 I_0(\kappa')}\exp\left[-\frac{Z_1^2 + Z_2^2 + v^2}{2\sigma^2}\right]\int_\phi \exp\left[\kappa'\cos\phi + \frac{vZ_1}{\sigma^2}\cos\phi + \frac{vZ_1}{\sigma^2}\sin\phi\right]d\phi \\ &= \frac{1}{2\pi\sigma^2 I_0(\kappa')}\exp\left(-\frac{\rho^2 + v^2}{2\sigma^2}\right)I_0\left(\sqrt{\kappa'^2 + \frac{v^2\rho^2}{\sigma^4} + \frac{2\kappa'vZ_1}{\sigma^2}}\right) \\ &= \frac{P(\rho)}{\rho}\frac{1}{2\pi I_0(\kappa')I_0(v\rho/\sigma^2)}I_0\left(\sqrt{\kappa'^2 + \frac{v^2\rho^2}{\sigma^4} + \frac{2\kappa'v\rho\cos\psi}{\sigma^2}}\right) \\ &\approx \frac{P(\rho)}{\rho}\frac{\exp(\kappa_\rho\cos\psi)}{2\pi I_0(\kappa_\rho)}, \end{aligned} \tag{37}$$

where the last line is inspired by the convolution property of two CN distributions: Let φ_1 and φ_2 be independently distributed as $CN(\phi_1, \kappa_1)$ and $CN(\phi_2, \kappa_2)$, respectively. Then the probability density function of $\varphi = \varphi_1 + \varphi_2 \pmod{2\pi}$ is

$$P(\varphi = \varphi_1 + \varphi_2) = \frac{1}{4\pi^2 I_0(\kappa_1) I_0(\kappa_2)} \int \exp[\kappa_1 \cos(\vartheta - \phi_1) + \kappa_2 \cos(\varphi - \vartheta - \phi_2)] d\vartheta$$

$$= \frac{1}{2\pi I_0(\kappa_1) I_0(\kappa_2)} I_0 \left(\sqrt{\kappa_1^2 + \kappa_2^2 + 2\kappa_1 \kappa_2 \cos(\varphi - \phi_1 - \phi_2)} \right). \tag{38}$$

The convolution of two CN distributions is approximately a CN distribution [42], i.e., $P(\varphi) \approx CN(\phi_1 + \phi_2, \kappa_3)$ where κ_3 is the solution of $\frac{I_1(\kappa_3)}{I_0(\kappa_3)} = \frac{I_1(\kappa_1)}{I_0(\kappa_1)} \frac{I_1(\kappa_2)}{I_0(\kappa_2)}$. Through numerical testing, we find a more convenient expression $\kappa_3 = \kappa_1 \kappa_2 / (\kappa_1 + \kappa_2)$. In sum, the following approximation is excellent and useful:

$$\frac{1}{2\pi I_0(\kappa_1) I_0(\kappa_2)} I_0 \left(\sqrt{\kappa_1^2 + \kappa_2^2 + 2\kappa_1 \kappa_2 \cos(\varphi - \phi_1 - \phi_2)} \right) \approx \frac{\exp[\kappa_3 \cos(\varphi - \phi_1 - \phi_2)]}{2\pi I_0(\kappa_3)}. \tag{39}$$

Now considering the small SNR limit ($\kappa \ll 1$), we obtain

$$\frac{I_1(\nu\rho/\sigma^2)}{I_0(\nu\rho/\sigma^2)} \approx \frac{1}{2} \frac{\nu\rho}{\sigma^2} \quad \text{and therefore,} \quad \frac{\kappa_\rho}{2} \approx \frac{I_1(\kappa_\rho)}{I_0(\kappa_\rho)} = \frac{I_1(\kappa')}{I_0(\kappa')} \frac{I_1(\nu\rho/\sigma^2)}{I_0(\nu\rho/\sigma^2)} \approx \frac{I_1(\kappa')}{I_0(\kappa')} \frac{\nu\rho}{2\sigma^2}. \tag{40}$$

Thus, we have the following approximation at $\kappa \ll 1$,

$$\langle h(\kappa_\rho) \rangle_{P(\rho)} = \left\langle \kappa_\rho \frac{I_1(\kappa_\rho)}{I_0(\kappa_\rho)} - \ln I_0(\kappa_\rho) \right\rangle_{P(\rho)} \approx \left\langle \frac{\nu^2\rho^2}{2\sigma^4} \left[\frac{I_1(\kappa')}{I_0(\kappa')} \right]^2 - \ln \left[I_0 \left(\frac{I_1(\kappa')}{I_0(\kappa')} \frac{\nu\rho}{\sigma^2} \right) \right] \right\rangle_{P(\rho)}$$

$$\approx \left(\frac{\nu^2}{\sigma^2} + \frac{\nu^4}{2\sigma^4} \right) \left[\frac{I_1(\kappa')}{I_0(\kappa')} \right]^2 - \frac{\nu^2}{2\sigma^2} \left[\frac{I_1(\kappa')}{I_0(\kappa')} \right]^2$$

$$\approx \frac{\kappa}{2} \left[\frac{I_1(\kappa')}{I_0(\kappa')} \right]^2 + \mathcal{O}(\kappa^2). \tag{41}$$

In the large SNR limit $\kappa \gg 1$, the expectation and variance of ρ are asymptotically equal to ν and σ^2. So we can replace ρ by ν when evaluating $\langle h(\kappa_\rho) \rangle$. Define $\kappa'' = \kappa_\rho(\rho = \nu) = \kappa' \nu^2 / (\kappa' \sigma^2 + \nu^2) = \kappa' \kappa / (\kappa' + \kappa)$. Then,

$$\langle h(\kappa_\rho) \rangle_{P(\rho)} = \left\langle \kappa_\rho \frac{I_1(\kappa_\rho)}{I_0(\kappa_\rho)} - \ln I_0(\kappa_\rho) \right\rangle_{P(\rho)} \approx \kappa'' \frac{I_1(\kappa'')}{I_0(\kappa'')} - \ln I_0(\kappa'') = h(\kappa''). \tag{42}$$

and thus

$$I_{biased}(\mathbf{Z}; \phi) \approx \begin{cases} \kappa \left(1 - [I_1(\kappa')/I_0(\kappa')]^2 \right) / 2 & \kappa \ll 1, \\ \kappa - \ln[I_0(\kappa)] - 1/2 - \kappa'' I_1(\kappa'')/I_0(\kappa'') + \ln I_0(\kappa'') & \kappa \gg 1. \end{cases} \tag{43}$$

We can check the above approximation by considering the limit $\kappa' \to \infty$, under which we should have $I_{biased}(\mathbf{Z}; \phi) = 0$. Obviously, the approximate expression of $I_{biased}(\mathbf{Z}; \phi)$ at $\kappa \ll 1$ satisfies this limit property. Further, we notice that

$$\lim_{\kappa' \to \infty} \frac{I_1(\kappa_\rho)}{I_0(\kappa_\rho)} = \lim_{\kappa' \to \infty} \frac{I_1(\kappa')}{I_0(\kappa')} \frac{I_1(\nu\rho/\sigma^2)}{I_0(\nu\rho/\sigma^2)} = \frac{I_1(\nu\rho/\sigma^2)}{I_0(\nu\rho/\sigma^2)},$$

which means that $\kappa_\rho \to \nu\rho/\sigma^2$ and thus $\kappa'' \to \nu^2/\sigma^2 = \kappa$. Then, at the large SNR limit ($\kappa \gg 1$) and for $\kappa' \to \infty$, we have

$$\begin{aligned} I_{biased}(\mathbf{Z}; \phi) &\approx \kappa - \ln[I_0(\kappa)] - 1/2 - \kappa'' I_1(\kappa'')/I_0(\kappa'') + \ln I_0(\kappa'') \\ &\to \kappa - \ln[I_0(\kappa)] - 1/2 - \kappa I_1(\kappa)/I_0(\kappa) + \ln[I_0(\kappa)] \\ &\approx \kappa - 1/2 - \kappa(1 - \kappa/2) \\ &= 0, \end{aligned} \tag{44}$$

where we used that $I_1(x)/I_0(x) \approx 1 - \frac{1}{2x} + \mathcal{O}(\frac{1}{x^2})$ for large x. In fact, one can directly check that $I(\mathbf{Z}; \phi)$ vanishes at the limit $\kappa' \to \infty$ (such that $\kappa_\rho \to \nu\rho/\sigma^2$) since

$$\begin{aligned} I_{biased}(\mathbf{Z}; \phi) &\to \frac{\nu^2}{\sigma^2} - \left\langle \ln \left[I_0 \left(\frac{\rho\nu}{\sigma^2} \right) \right] \right\rangle_{P(\rho)} - \left\langle \frac{\nu\rho}{\sigma^2} \frac{I_1(\nu\rho/\sigma^2)}{I_0(\nu\rho/\sigma^2)} - \ln \left[I_0 \left(\frac{\rho\nu}{\sigma^2} \right) \right] \right\rangle_{P(\rho)} \\ &= \frac{\nu^2}{\sigma^2} - \int \frac{\nu\rho}{\sigma^2} \frac{I_1(\nu\rho/\sigma^2)}{I_0(\nu\rho/\sigma^2)} \frac{\rho}{\sigma^2} \exp \left[-\frac{\rho^2 + \nu^2}{2\sigma^2} \right] I_0 \left(\frac{\rho\nu}{\sigma^2} \right) d\rho \\ &= \frac{\nu^2}{\sigma^2} - \int \frac{\nu\rho^2}{\sigma^4} \exp \left[-\frac{\rho^2 + \nu^2}{2\sigma^2} \right] I_1 \left(\frac{\rho\nu}{\sigma^2} \right) d\rho \\ &= \frac{\nu^2}{\sigma^2} - \frac{\nu^2}{\sigma^2} \\ &= 0. \end{aligned} \tag{45}$$

Under the directional bias, the chemotactic mutual information for Model II is given by

$$
\begin{aligned}
I_{biased}(\theta;\phi) &= \int_\theta P(\theta|\phi)\ln P(\theta|\phi)d\theta - \int_\theta P(\theta)\ln P(\theta)d\theta d \\
&\approx \widehat{\kappa}\frac{I_1(\widehat{\kappa})}{I_0(\widehat{\kappa})} - \ln[2\pi I_0(\widehat{\kappa})] - \widehat{\kappa}'\frac{I_1(\widehat{\kappa}')}{I_0(\widehat{\kappa}')} - \ln[2\pi I_0(\widehat{\kappa}')] \\
&= \widehat{\kappa}\frac{I_1(\widehat{\kappa})}{I_0(\widehat{\kappa})} - \widehat{\kappa}'\frac{I_1(\widehat{\kappa}')}{I_0(\widehat{\kappa}')} + \ln\frac{I_0(\widehat{\kappa}')}{I_0(\widehat{\kappa})},
\end{aligned}
\tag{46}
$$

with $\widehat{\kappa}'$ defined through $\frac{I_1(\widehat{\kappa}')}{I_0(\widehat{\kappa}')} = \frac{I_1(\widehat{\kappa})}{I_0(\widehat{\kappa})}\frac{I_1(\kappa')}{I_0(\kappa')}$, or more conveniently, $\widehat{\kappa}' \simeq \widehat{\kappa}\kappa'/(\widehat{\kappa}+\kappa')$. The key calculation is

$$
\begin{aligned}
P(\theta) &= \int P(\theta|\phi)P(\phi)d\phi \\
&= \int \frac{\exp[\widehat{\kappa}\cos(\theta-\phi)]}{2\pi I_0(\widehat{\kappa})}\frac{\exp(\kappa'\cos\phi)}{2\pi I_0(\kappa')}d\phi \\
&= \frac{1}{2\pi I_0(\widehat{\kappa})I_0(\kappa')}I_0(\sqrt{\widehat{\kappa}^2+\kappa'^2+2\widehat{\kappa}\kappa'\cos\theta}) \approx \frac{\exp(\widehat{\kappa}'\cos\theta)}{2\pi I_0(\widehat{\kappa}')}.
\end{aligned}
\tag{47}
$$

A.3 Blahut-Arimoto Algorithm and Channel Capacity

Consider a simple communication channel $X \to Y$, where X and Y represent the input and output random variables, respectively. The maximum mutual information over all marginal input distribution $P(x)$ defines the channel capacity C, i.e.,

$$
C = \max_{P(x)} I(Y;X).
\tag{48}
$$

Blahut-Arimoto algorithm provides an efficient iterative method of determining the channel capacity [43,44]. Starting with an arbitrary marginal distribution $P(x)$, the algorithm updates $P(x)$ at each time by:

$$
P_{new}(x) = \frac{\exp\left[\int_Y P(y|x)\ln P(x|y)dy\right]}{\int_X \exp\left[\int_Y P(y|x)\ln P(x|y)dy\right]dx},
\tag{49}
$$

with $P(x|y) = P(y|x)P_{old}(x)/P(y)$. Iteration of the above process will converge to the optimal marginal distribution that maximizes $I(Y;X)$ and thus achieves the channel capacity. In our paper, the updating rule of Eq. 49 becomes

$$
P_{new}(\phi) = \frac{\exp\left[\int_\theta P(\theta|\phi)\ln P(\phi|\theta)d\theta\right]}{\int_\phi \exp\left[\int_\theta P(\theta|\phi)\ln P(\phi|\theta)d\theta\right]d\phi},
\tag{50}
$$

with $P(\phi|\theta) = P(\theta|\phi)P_{old}(\phi)/P(\theta)$. One can check that when $P_{old}(\phi) = \frac{1}{2\pi}$, the new marginal distribution $P_{new}(\phi)$ is also uniform by Eq. 50. First of all, $P(\theta) = \int_\phi P(\theta|\phi)P_{old}(\phi)d\phi = \frac{1}{2\pi}$, thus $P(\phi|\theta) = P(\theta|\phi)P_{old}(\phi)/P(\theta) = P(\theta|\phi)$. Next, we notice that

$$
\int_\theta P(\theta|\phi)\ln P(\phi|\theta)d\theta = \int_\theta P(\theta|\phi)\ln P(\theta|\phi)d\theta = -H(\theta|\phi) = \widehat{\kappa}I_1(\widehat{\kappa})/I_0(\widehat{\kappa}) - \ln[2\pi I_0(\widehat{\kappa})] = h(\widehat{\kappa}) - \ln(2\pi),
\tag{51}
$$

which is independent of ϕ. Hence,

$$
P_{new}(\phi) = \frac{\exp[h(\widehat{\kappa})-\ln(2\pi)]}{\int_\phi \exp[h(\widehat{\kappa})-\ln(2\pi)]d\phi} = \frac{1}{\int_\phi d\phi} = \frac{1}{2\pi}.
\tag{52}
$$

Therefore, the uniform source distribution must maximize $I(\theta;\phi)$ (and similarly $I(\mathbf{Z};\phi)$), implying that our mutual information results under the unbiased case give the channel capacity at different levels.

References

1. G.H. Wadhams, J.P. Armitage, Nat Rev Mol Cell Biol **5**(12), 1024 (2004)
2. R. Clark, *The molecular and cellular biology of wound repair* (Plenum, New York, 1996)
3. D. Dormann, C.J. Weijer, Curr Opin Genet Dev **13**(4), 358 (2003)
4. J. Condeelis, R.H. Singer, J.E. Segall, Annu Rev Cell Dev Biol **21**, 695 (2005)
5. S.H. Zigmond, J. Cell. Biol. **75**, 606 (1977)
6. L. Song, S.M. Nadkarni, H.U. Bödeker, C. Beta, A. Bae, C. Franck, W.J. Rappel, W.F. Loomis, E. Bodenschatz, Eur J Cell Biol **85**, 981 (2006)
7. D. Fuller, W. Chen, M. Adler, A. Groisman, H. Levine, W.J. Rappel, W.F. Loomis, PNAS **107**(21), 9656 (2010)
8. W.J. Rosoff, J.S. Urbach, M.A. Esrick, R.G. McAllister, L.J. Richards, G.J. Goodhill, Nat Neurosci **7**(6), 678 (2004)
9. W.J. Rappel, P.J. Thomas, H. Levine, W.F. Loomis, Biophys J **83**, 1361 (2002)
10. H. Levine, D.A. Kessler, W.J. Rappel, PNAS **103**(26), 9761 (2006)
11. A. Levchenko, P.A. Iglesias, Biophys J **82**, 50 (2002)
12. A. Narang, K.K. Subramanian, D.A. Lauffenburger, Ann. Biomed. Eng. **29**, 677 (2001)
13. H. Meinhardt, J. Cell. Sci. **112**, 2867 (1999)
14. M. Postma, P.J.M. van Haastert, Biophys. J. **81**, 1314 (2001)
15. C. Beta, G. Amselem, E. Bodenschatz, New J Phys **10**, 083015 (2008)
16. W.J. Rappel, W.F. Loomis, Wiley Interdiscip Rev Syst Biol Med **1**(1), 141 (2009)
17. C.Y. Chung, R.A. Firtel, J Muscle Res Cell Motil **23**(7-8), 773 (2002)
18. K.F. Swaney, C.H. Huang, P.N. Devreotes, Annu Rev Biophys **39**, 265 (2010)
19. R. Endres, N. Wingreen, PNAS **105**, 15749 (2008)
20. R.G. Endres, N.S. Wingreen, Phys Rev Lett **103**(15), 158101 (2009)
21. R.G. Endres, N.S. Wingreen, Prog Biophys Mol Biol **100**(1-3), 33 (2009)
22. K. Wang, W.J. Rappel, R. Kerr, H. Levine, Phys Rev E **75**, 061905 (2007)
23. W.J. Rappel, H. Levine, PNAS **105**(49), 19270 (2008)
24. W.J. Rappel, H. Levine, Phys Rev Lett **100**(22), 228101 (2008)
25. B. Hu, D. Fuller, W.F. Loomis, H. Levine, W.J. Rappel, Phys Rev E **81**(3), 031906 (2010)
26. W. Bialek, S. Setayeshgar, PNAS **102**(29), 10040 (2005)
27. D. Mortimer, J. Feldner, T. Vaughan, I. Vetter, Z. Pujic, W.J. Rosoff, K. Burrage, P. Dayan, L.J. Richards, G.J. Goodhill, PNAS **106**(25), 10296 (2009)
28. P.J. van Haastert, M. Postma, Biophys J **93**(5), 1787 (2007)
29. B.W. Andrews, P.A. Iglesias, PLoS Comput Biol **3**(8), e153 (2007)
30. T. Cover, J. Thomas, *Elements of Information Theory* (Wiley, New York, 1991)
31. B. Hu, W. Chen, W.J. Rappel, H. Levine, Phys Rev Lett **105**(4), 048104 (2010)
32. B. Hu, W. Chen, W.J. Rappel, H. Levine, Submitted (2011)
33. D. Guo, S. Shamai, S. Verdu, IEEE Trans. Inf. Theor. **51**, 1261 (2005)
34. M. Schienbein, K. Franke, H. Gruler, Phys Rev E **49**(6), 5462 (1994)
35. A.M. Fraser, H.L. Swinney, Phys Rev A **33**(2), 1134 (1986)
36. Y.I. Moon, B. Rajagopalan, U. Lall, Phys Rev E **52**(3), 2318 (1995)
37. A. Samadani, J. Mettetal, A. Oudenaarden, PNAS **103**, 11549 (2006)
38. L. Bosgraaf, P.J.V. Haastert, PLoS One **4**(4), e5253 (2009)
39. M. Skoge, M. Adler, A. Groisman, H. Levine, W.F. Loomis, W.J. Rappel, Integr Biol (Camb) (2010)
40. S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice Hall, Upper Saddle River, 1993)
41. H. Gudbjartsson, S. Patz, Magn Reson Med **34**(6), 910 (1995)
42. K. Mardia, *Statistics of Directional Data* (Academic Press, London, 1972)
43. R.E. Blahut, IEEE Trans. Inform. Theory **IT-18**, 460 (1972)
44. S. Arimoto, IEEE Trans. Inform. Theory **IT-18**, 14 (1972)