Role of spatial averaging in multicellular gradient sensing

Tyler Smith\textsuperscript{1,5}, Sean Fancher\textsuperscript{2,5}, Andre Levchenko\textsuperscript{3}, Ilya Nemenman\textsuperscript{4} and Andrew Mugler\textsuperscript{2,6}

\textsuperscript{1} Department of Physics, Emory University, Atlanta, GA 30322, USA
\textsuperscript{2} Department of Physics and Astronomy, Purdue University, West Lafayette, IN 47907, USA
\textsuperscript{3} Department of Biomedical Engineering and Yale Systems Biology Institute, Yale University, New Haven, CT 06520, USA
\textsuperscript{4} Departments of Physics and Biology, Emory University, Atlanta, GA 30322, USA
\textsuperscript{5} These authors contributed equally to this work.
\textsuperscript{6} Author to whom any correspondence should be addressed.
E-mail: amugler@purdue.edu

Keywords: gradient sensing, dimensionality, multicellular communication, cell and organoid morphology

Abstract

Gradient sensing underlies important biological processes including morphogenesis, polarization, and cell migration. The precision of gradient sensing increases with the length of a detector (a cell or group of cells) in the gradient direction, since a longer detector spans a larger range of concentration values. Intuition from studies of concentration sensing suggests that precision should also increase with detector length in the direction transverse to the gradient, since then spatial averaging should reduce the noise. However, here we show that, unlike for concentration sensing, the precision of gradient sensing decreases with transverse length for the simplest gradient sensing model, local excitation–global inhibition. The reason is that gradient sensing ultimately relies on a subtraction of measured concentration values. While spatial averaging indeed reduces the noise in these measurements, which increases precision, it also reduces the covariance between the measurements, which results in the net decrease in precision. We demonstrate how a recently introduced gradient sensing mechanism, regional excitation–global inhibition (REGI), overcomes this effect and recovers the benefit of transverse averaging. Using a REGI-based model, we compute the optimal two- and three-dimensional detector shapes, and argue that they are consistent with the shapes of naturally occurring gradient-sensing cell populations.

1. Introduction

Determining the strength and direction of a chemical concentration gradient is an essential task for a diverse array of biological processes. Gradient sensing underlies the polarization of single cells, the orientation and migration of cells and cell collectives, and the changes in tissue morphology that occur during embryogenesis and the subsequent development of an organism [1–12]. Experiments have shown that cells are remarkably precise gradient sensors [1, 12], and a large amount of effort has gone into understanding the mechanisms of, and the limits to, biological gradient sensing [2, 13–20].

At its core, gradient sensing requires the comparison of concentration measurements between the ’front’ and the ’back’ of a detector. Front and back here are defined with respect to the gradient direction, and the detector here is a single cell or a group of cells. If the front and back are more separated, then the concentration measurements are more different from each other, which improves the determination of the gradient. This implies that detectors that are longer in the gradient direction have a higher gradient sensing precision [14, 16–18]. This argument neglects the fact that information must be communicated between different parts of a detector, especially if the detector is multicellular. Recently we derived the limits to the precision of gradient sensing including communication, and we found that for a one-dimensional (1D) detector, the precision indeed increases with detector length, but then saturates due to the fact that communication introduces its own noise [12, 20]. Nonetheless, the precision of gradient sensing increases or saturates with the length of a 1D detector aligned with the gradient; it does not decrease.

Yet biological detectors are not 1D in general. Two-dimensional (2D) detectors include the quasi-cylindrical arrangement of cell nuclei during the early
stages of Drosophila development [21] and the planar arrangement of epithelial cell layers [7]. Three-dimensional (3D) detectors include single cells and the multicellular tips of growing epithelial ducts [22], as well as border cells exhibiting collective guidance in Drosophila [3]. This raises the question of what effect the dimensions transverse to the gradient direction have on the precision of gradient sensing.

Intuition about this question can be drawn from the similar task of sensing the value of a concentration (as opposed to sensing its difference between two points in space, i.e., the gradient). If the concentration profile is uniform in space, then the precision of concentration sensing benefits from increasing the detector length in any direction. The reason is that communication with other parts of the detector, or spatial averaging, does not change the mean of a particular measurement within the detector, but it does reduce the noise [13, 16, 17]. Even if the concentration profile is graded, but the goal is still concentration (rather than gradient) sensing, as in stripe formation in early Drosophila development, the precision still benefits from spatial averaging [23]. The benefit is especially clear in a direction transverse to the gradient direction: once again, spatial averaging in this direction does not change the mean of a particular measurement, but it does reduce the noise. These considerations, drawn from the problem of concentration sensing, suggest that the precision of gradient sensing should also increase with the length of a detector in a direction transverse to the gradient.

Here we investigate theoretically and computationally the precision of gradient sensing for 2D and 3D detectors. We start with one of the simplest models of gradient sensing, the local excitation–global inhibition (LEGI) model [15, 19]. This is an accepted basic model when gradient sensing is adaptive (that is, background concentration largely does not affect the gradient sensing). Surprisingly, in contrast to the case of concentration sensing, we find that the precision of gradient sensing decreases with the length of the detector in a direction transverse to the gradient direction. The reason is that gradient sensing fundamentally relies on a subtraction of concentration measurements, e.g., between the front and back of the detector. While spatial averaging reduces the intrinsic noise in these measurements, which increases precision, it also reduces the covariance between the measurements, which decreases precision. We demonstrate that the latter effect dominates, such that the net result is a decrease in precision with transverse detector size. Then we show that this decrease can actually be overcome by a gradient-sensing strategy that we recently introduced, termed regional excitation–global inhibition (REGI) [20]. We demonstrate that REGI retains a high covariance between measurements and restores the benefit of transverse averaging. Using a REGI-based model, we compute the optimal 2D and 3D detector shapes, which arise from an interplay of the effects of transverse averaging on both the signal and the noise of gradient detection. We argue that these shapes are consistent with the shapes of the multicellular tips of epithelial ducts, suggesting that this and other similarly shaped gradient-sensing systems benefit from spatial averaging in all dimensions.

2. Methods

As in previous work [12, 20], we consider the LEGI model of multicellular gradient sensing, which is a minimal, adaptive, spatially extended model of gradient sensing. We consider a signal concentration profile that varies linearly in a particular direction in 3D space, with concentration gradient \( c \) (figures 1(A) and (C)). In the \( n \)th cell, both a local molecular species \( X \) and a global molecular species \( Y \) are produced at a rate \( \beta \) and degraded at a rate \( \mu \). The production rate is also proportional to the number of signal molecules in the cell’s vicinity \( c_{n,a} \), where \( a \) is the cell diameter. Whereas the local species \( X \) is confined to each cell, the global species \( Y \) is exchanged between neighboring cells at a rate \( \gamma_y \) (figure 1(C); note that although all cells are producing \( X \) and \( Y \) molecules, we show as examples in figures 1(A) and (C) only those molecules originating from the rightmost, middle cell). Conceptually, \( X \) measures the local concentration of signal molecules, while \( Y \) represents their spatially averaged concentration. As in [12, 20] we consider the linear response regime, in which the dynamics of the local and global species satisfy the stochastic equations

\[
\frac{d x_n}{d t} = \beta (c_{n,a} a^3) - \mu x_n + \eta_n\]

\[
\frac{d y_n}{d t} = \beta (c_{n,a} a^3) - \mu y_n + \gamma_y \sum_{n' \in N_n} (y_{n'} - y_n) + \xi_n
\]

\[
= \beta (c_{n,a} a^3) - \mu \sum_{n'} M_{nm} y_{n'} + \xi_n.
\]

Here \( M_{nm} \equiv (1 + |N_n| \gamma_y / \mu) \delta_{nm} - (\gamma_y / \mu) \delta_{nm} \) is the connectivity matrix for the global species that accounts for degradation and molecule exchange. \( N_n \) and \( |N_n| \) denote the indices and the number of nearest neighbors of cell \( n \), respectively. The intrinsic noise terms \( \eta_n \) and \( \xi_n \) correspond to the Poissonian production, degradation, and exchange reactions [12]. The signal \( c_{n,a} \) also fluctuates, which introduces extrinsic noise. As described below, in this work we assume that these fluctuations are slow compared to the downstream signal processing, which is equivalent to assuming either slow diffusion of signaling molecules or instantaneous downstream processing, and leads to Poisson-distributed signal molecule counts [12].

In the LEGI paradigm, \( X \) excites a downstream species while \( Y \) inhibits it. If the cell is at the higher
edge of the gradient, then the local concentration (X) is higher than the spatial average (Y), and the excitation exceeds the inhibition. While such comparison of the excitation and the inhibition can be done by many different molecular mechanisms [19], we consider here the limit of shallow gradients, where the comparison is equivalent to subtracting Y from X [12]. This difference, $\Delta_n = x_n - y_n$, is the readout of the model. If $\Delta_n$ is positive, the $n$th cell is further up the gradient than average; if $\Delta_n$ is negative, the $n$th cell is further down the gradient than average. In this work, we always focus on the readout $\Delta_N$ of the cell highest up the gradient, which we denote as the Nth cell.

We assume that the cells do not average concentrations of the signal C and the messenger molecules X and Y over time (though generalizations with averaging are certainly possible [20]). Then the precision of gradient sensing is given by the square root of the instantaneous signal-to-noise ratio (SNR) for the readout, $\text{SNR}_\Delta = (\bar{\Delta}_N / \delta \Delta_N)^2$, where the mean and variance are given by [12]

$$\bar{\Delta}_N = \bar{x}_N - \bar{y}_N,$$

$$\bar{x}_N = G a^2 \bar{\xi}_N,$$

$$\bar{y}_N = G a^4 \sum_n K^y_n \bar{\xi}_{-n-},$$

and

$$\left(\delta \Delta_N\right)^2 = \left(\delta x_N\right)^2 + \left(\delta y_N\right)^2 - 2 \text{cov}(x_N, y_N),$$

$$\left(\delta x_N\right)^2 = \bar{x}_N + G a^2 \bar{\xi}_N,$$

$$\left(\delta y_N\right)^2 = \bar{y}_N + G a^4 \sum_n \left(\xi^y_n\right)^2 \bar{\xi}_{-n-},$$

$$\text{cov}(x_N, y_N) = G a^4 \sum_n K^x_n \bar{\xi}_n,$$

respectively. Here $K^y_n \equiv \langle M^y \rangle_{N-n-M}$ is the communication kernel, and $G \equiv \beta / \mu$ is the gain. The first terms in equations (7) and (8) correspond to intrinsic noise, while the second terms correspond to extrinsic noise and assume that the diffusion of the signal is slow.
[12]. Computing the precision for a given configuration of cells only requires inverting the connectivity matrix $M'$. That is, because our model is linear, both the means and (co)variances are obtained directly by matrix inversion, and no stochastic simulations are performed.

In the recently introduced REGI model [20], the local species $X$ is also exchanged among cells, but at a lower rate $\gamma < \gamma'$. Then equation (1) becomes analogous to equation (2), and equations (4), (7), and (9) are replaced by

$$x_N = G a^4 \sum_n K_n^2 \hat{c}_{N-n}$$

(10)

$$\langle x_N \rangle^2 = \hat{x}_N + G^2 a^4 \sum_n \langle K_n^2 \rangle \hat{c}_{N-n}$$

(11)

$$\text{cov}(x_N, y_N) = G^2 a^4 \sum_n K_n^2 \hat{c}_{N-n} \hat{c}_{N-n}$$

(12)

respectively, where $K_n^2 \equiv (M^2)_{N-n}^{-1}$ is the communication kernel for the local species, and $M_n^2 \equiv 1 + (1 + [\lambda/\mu])\delta_m \hat{e}_{m} - (\gamma / \mu) \delta_{m} \hat{e}_{m}$. Once more, computing the precision for a given configuration of cells in the REGI model only requires inverting the connectivity matrices $M^2$ and $M'$. While diffusion of $X$ decreases $x_N$ at the $N$th cell, and hence decreases the difference $\Delta_N$, it also averages $X$ over a larger volume, hence decreasing its noise. As shown in [20], under a broad range of conditions, the decrease in the noise dominates, and the overall precision of the REGI model is higher than that of LEGI.

3. Results

3.1. Concentration sensing precision increases with transverse detector size

Before investigating gradient sensing, we focus on the simpler problem of concentration sensing. In the LEGI model, both $X$ and $Y$ provide readouts of the local concentration, while their difference $\Delta$ provides a readout of the gradient. The concentration readout provided by $Y$ is spatially averaged, whereas the concentration readout provided by $X$ is not. Even if the signal profile is graded, $X$ and $Y$ are concentration readouts if viewed independently (with different spatial averaging), not gradient readouts. For example, during Drosophila development, the morphogen profiles are graded, but individual nuclei in the embryo measure (and threshold) the local concentration, possibly with some spatial averaging [21, 23–25].

How does the precision of concentration sensing depend on transverse detector size? To answer this question, we focus on the spatially averaged concentration readout $Y$. We consider a linear signal profile with gradient $g$ and compute the SNR of $Y$ in the $N$th cell, as we vary the number $M$ of rows of cells in a direction transverse to the gradient (figure 1(A)). We see in figure 1(B) (circles) that the precision of concentration sensing increases with $M$. The reason is that adding rows of cells transverse to the gradient allows for $Y$ molecules to be exchanged between rows (in addition to along each row). This does not change the mean $\hat{y}_N$ due to the translational symmetry in the transverse direction. However, it does reduce the variance, since the global species $Y$ is now averaged over more cells. The net effect is an increase in the SNR beyond what is allowed by longitudinal averaging.

We can elucidate the effect of spatial averaging more quantitatively by appealing to the expression for the variance in $Y$ (equation (8)). In a single row of cells, and in the limit of many cells ($N \gg 1$) and fast communication ($\gamma \ll \mu$), the kernel reduces to $K_N^Y \approx e^{-n/\mu} / n_j$, where $n_j \equiv \sqrt{\gamma / \mu}$ sets the effective length scale of communication [12]. Approximating the sum in equation (8) as an integral that extends to infinity (since $N$ is large) obtains $(\delta y_N^Y)^2 \approx \hat{y}_N + G^2 a^4 \hat{c}_N \hat{c}_N / (2n_j)$ for a single row. In the case of $M$ rows, $Y$ is averaged with the communication kernel $K^T$ over $M$ cells transverse to the gradient. This will not affect the intrinsic component of the variance (since the mean is unchanged), but the extrinsic component will be reduced according to $(\delta y_N^{(M)})^2 \approx \hat{y}_N + G^2 a^4 \hat{c}_N \hat{c}_N / (2n_j)$. We again make the exponential kernel approximation, this time normalizing over the finite domain of size $M$, giving $K_N^Y \approx e^{-y/\mu} / n_j$ in $\sum_{N=-M/2}^{N=M/2} e^{-y/\mu}$. Finally, again approximating the sums as integrals, we obtain

$$\left( \delta y_N^Y \right)^2 \approx \hat{y}_N + \frac{G^2 a^4 \hat{c}_N / n_j}{2n_j} \left( 1 - e^{-M/n_j} \right) \left( 1 - e^{-M/n_j} \right)^{1/2}.$$  

(13)

In deriving this approximation, we have neglected effects that the transverse edges have on the exponential shape of the kernel, as well as correlations between exchange reactions parallel and perpendicular to the gradient. Nonetheless, the SNR calculated using this approximation is compared with the numerical result in figure 1(B), and we see that the agreement is excellent. In the limit $M \ll n_j$, equation (13) simplifies to

$$\left( \delta y_N^Y \right)^2 \approx \hat{y}_N + \frac{G^2 a^4 \hat{c}_N / n_j}{2n_j}.$$  

(14)

We see that for a small number of rows, the averaging over rows is nearly uniform, and the extrinsic component of the variance is reduced by a factor of $M$, as expected.

3.2. Gradient sensing precision decreases with transverse detector size

We now turn our attention to gradient sensing. How does the precision of gradient sensing depend on transverse detector size? To answer this question for a linear signal profile, we compute the SNR of the gradient readout $\Delta_N$ as a function of the number $M$ of rows of cells in a direction transverse to the gradient (figure 1(C)). We see in figure 1(D) that the precision of gradient sensing decreases with $M$ (circles). This is
in contrast to the precision of concentration sensing, which increases with M (figure 1(B)).

To understand why the precision of gradient sensing decreases with M, we once again consider the mean and the variance of the readout. The mean $\Delta_N = \bar{x}_N - \bar{y}_N$ does not change with M because neither $\bar{x}_N$ nor $\bar{y}_N$ changes with M. However, the variance $(\delta \Delta_N)^2 = (\delta \bar{x}_N)^2 + (\delta \bar{y}_N)^2 - 2\text{cov}(\bar{x}_N, \bar{y}_N)$ changes with M due to two effects. First, the variance in the global species $(\delta \bar{y}_N)^2$ decreases with M due to spatial averaging, as discussed in the previous section. Second, the covariance $\text{cov}(\bar{x}_N, \bar{y}_N)$ also decreases with M because Y is exchanged with a larger number of cells, whereas X is not exchanged, so the two covary more weakly. The effects have opposite signs. To understand which effect dominates, we again appeal to analytic approximation. For a single row of cells, under the exponential kernel approximation, the covariance in equation (9) reduces to $\text{cov}^{\text{1D}}(x_N, y_N) \approx G^2a^3\bar{e}_{N-n/2}/n_y$. For M rows of cells, since only Y is exchanged, the covariance is reduced according to $\text{cov}^{\text{MD}}(x_N, y_N) = K_m^y - 0 \text{cov}^{\text{1D}}(x_N, y_N)$. Making the same approximation as above for $K_m^y$ of an exponential in a finite domain, we obtain

$$\begin{align*}
(\delta \bar{x}_N)^2 & \approx (\delta \bar{y}_N)^2 + \bar{y}_N + \frac{G^2a^3\bar{e}_{N-n/2}}{2n_y} - \frac{G^2a^3\bar{e}_N}{n_yM}, \\
(\delta \Delta_N)^2 & \approx (\delta \bar{x}_N)^2 + \bar{y}_N + \frac{G^2a^3\bar{e}_{N-n/2}}{2n_yM} - \frac{2G^2a^3\bar{e}_N}{n_yM}. \\
& \approx (\delta \bar{x}_N)^2 + \bar{y}_N - \frac{G^2a^3[4\bar{e}_N - \bar{e}_{N-n/2}]}{2n_yM}. 
\end{align*}$$

This SNR calculated using this approximation is compared with the numerical result in figure 3(D), and we see good agreement. In the limit $M \ll n_y$, equation (15) simplifies to

$$\begin{align*}
(\delta \Delta_N)^2 & \approx (\delta \bar{x}_N)^2 + \bar{y}_N + \frac{G^2a^3\bar{e}_{N-n/2}}{2n_yM} - \frac{2G^2a^3\bar{e}_N}{n_yM}, \\
& \approx (\delta \bar{x}_N)^2 + \bar{y}_N - \frac{G^2a^3[4\bar{e}_N - \bar{e}_{N-n/2}]}{2n_yM}. 
\end{align*}$$

Equation (16) shows that in this limit of near-uniform averaging, both (i) the extrinsic component of the variance in Y and (ii) the covariance are reduced by a factor of $M$, as expected. Furthermore, because the Nth cell is at the highest concentration, we have $\bar{e}_N > \bar{e}_{N-n/2}$ and we see that equation (17) is an increasing function of $M$. Thus, this limit elucidates the fact that the decrease of the covariance dominates over the decrease of the variance in Y, causing the variance of $\Delta_N$ to increase with M for all parameter values. Because the mean $\Delta_N$ does not change with M, we conclude that the precision of gradient sensing decreases with transverse detector size.

### 3.3. REGI mechanism recovers the benefit of transverse averaging

In the previous section we saw that the precision of gradient sensing using the LEGI model (local messenger X is not exchanged among the cells) decreases with the size of a detector in a direction transverse to the gradient, due to the fact that the covariance between the subtracted variables decreases with the transverse size. For the REGI model, exchange of the X molecules has an additional effect beyond increasing the sensing precision for a 1D line of cells [20]: it increases the covariance of X and Y, compared to the LEGI mechanism. Indeed, now both X and Y are downstream signals originating from some of the same external ligand molecules. Since the decrease of gradient sensing precision with transverse detector size is due to the loss of covariance (figure 1(D)), this raises the question of whether the REGI strategy can overcome this effect and allow gradient sensing precision to benefit from transverse averaging.

To answer this question, we once again consider a linear signal profile, and we compute the SNR of the gradient readout $\Delta_N$ under the REGI model (see methods), as a function of the number M of rows of cells in a direction transverse to the gradient (figure 2(A)). We see in figure 2(B) that for a sufficiently large value of $n_x \equiv \sqrt{\bar{e}_N}/\bar{\mu}$, which sets the lengthscale of spatial averaging for the local species, the precision of gradient sensing increases with M. This is in contrast to the case of LEGI, for which the precision decreases with M (figure 1(D) and black circles in figure 2(B)). Therefore, the recovery of covariance between X and Y in the REGI mechanism avoids the loss of gradient sensing precision and restores the benefit of transverse averaging.

To understand this effect quantitatively, we turn once more to analytic approximation. The variance in X (equation (11)) and Y (equation (5)) for M rows of cells will be approximated by expressions of the form of equation (13). The covariance (equation (12)) for a single row of cells under the exponential kernel approximation for both X and Y is $\text{cov}^{\text{1D}}(x_N, y_N) \approx G^2a^3\bar{e}_{N-n/2}/(n_x + n_y)$, where $n \equiv n_x n_y/(n_x + n_y)$. For M rows of cells the covariance is therefore $\text{cov}^{\text{MD}}(x_N, y_N) = \Sigma_{m=-M/2}^{M/2} K_m^x K_m^y \text{cov}^{\text{1D}}(x_N, y_N)$. Again approximating $K_m^x$ and $K_m^y$ as exponentials in a finite domain, we obtain

$$\begin{align*}
(\delta \Delta_N)^2 & \approx \bar{x}_N + \frac{G^2a^3\bar{e}_{N-n/2}}{2n_x} - \frac{G^2a^3\bar{e}_N}{n_xM} \\
& \approx \bar{y}_N + \frac{G^2a^3\bar{e}_{N-n/2}}{2n_y} - \frac{G^2a^3\bar{e}_N}{n_yM}.
\end{align*}$$

This approximation has assumed $[n_x, n_y] \gg 1$, and we find that for $n_y = 4$ the SNR calculated using this
This maximum is due to the previously understood tradeoff that is introduced when \( N_x \) increases; on the one hand the variance of \( X \) is reduced, which increases precision; on the other hand, the means of \( X \) and \( Y \) are more similar, which decreases the precision [20]. Here this tradeoff is modified by the additional benefit of increasing \( n_x \), namely that it increases the covariance of \( X \) and \( Y \) in the transverse direction, and thus further reduces the noise in gradient sensing.

3.4. Emergence of optimal detector shapes in two and three-dimensions

The emergence of an optimal number of transverse rows of cells, seen in the previous section, raises the more general question of whether there is an optimal detector shape for spatially extended gradient sensing. This question has relevance for both 2D and 3D multicellular geometries involved in gradient sensing. Is the optimal detector shape more ‘hairlike’, to maximize its extent in the gradient direction, or more ‘globular’, to exploit potential benefits of extending along the transverse direction?

To address this question, we perform a controlled optimization for both 2D and 3D multicellular geometries. For a fixed number of cells \( N = 50 \), we confine cells to an elliptical (2D) or ellipsoidal (3D) envelope, and compute the precision of gradient sensing as a function of the ellipse axis parameters (LEGI), as well as the ratio of averaging length scales \( n_x/n_y \) (REGI), exhaustively exploring substantial ranges of both. In addition to the extra shape parameter, there is one more important difference between the 2D and 3D cases: in the 2D case, we assume that every cell detects signal molecules, since we imagine that these molecules diffuse in the 3D bulk, while the cells form a sensory sheet exposed to the bulk. In contrast, in the 3D case, we assume that only the surface cells detect signal molecules, whereas cells that are blocked on all six sides by neighboring cells are ‘shielded’ and thus do not detect signal molecules (although all cells still communicate via molecule exchange).
optimal detector shapes determined by such exhaustive search for the REGI model are shown in figure 3(A), for 2D (top) and 3D (bottom).

To explain why these optimal shapes emerge, we present the precision of gradient sensing as a function of the control parameters. First we investigate the behavior of the LEGI model in 2D (figure 3(B)). The control parameter is \( N_g \), the (projected) number of cells in the gradient direction, which is set uniquely in 2D by the ratio of the ellipse axis parameters. Small \( N_g \to 1 \) corresponds to a chain of cells transverse to the gradient, while large \( N_g \to N \) corresponds to a chain of cells parallel to the gradient. The small ‘stair steps’ in the curves are due to the numerical task of fitting the discrete multicellular square lattice within the continuous elliptical envelope. We see that the precision vanishes at \( N_g = 1 \), as expected, since in our model a single cell cannot perform gradient detection. The precision is near maximal at \( N_g = N \). This trend is analogous to that seen for LEGI in figure 1(D), where here \( N/N_g \sim M \) is the analog of the number of transverse rows. However, unlike in figure 1(D), we see in figure 3(B) that there is a weak optimum at an intermediate value of \( N_g \). This is due to a difference between the protocols of adding rows of cells (figure 1(D)) and reshaping a fixed number of cells (figure 3(B)). Adding rows does not change \( \Delta_N \). In contrast, as seen in the inset of figure 3(B), reshaping changes \( \Delta_N \). The reason is that elliptical configurations (like figure 3(A), top) are not translationally symmetric in the transverse direction. In particular, a large density of cells in the middle of the configuration is a sink for molecules of Y. This decreases the mean number of Y in the rightmost cell, \( \bar{y}_N \), which weakly increases the signal \( \Delta_N = \bar{x}_N - \bar{y}_N \) at intermediate values of \( N_g \) (figure 3(B) inset), and therefore increases the precision (figure 3(B)). Finally, we see that the precision increases with the gain \( G \), as expected, and that the increase saturates with \( G \), since then the variance of \( X \) and \( Y \) is dominated entirely by extrinsic, and not intrinsic, noise (see methods).
Next we investigate the behavior of REGI in 2D (figure 3(C)). Once again the control parameter is $N_y$. Additionally, at every $N_y$ we optimize the local species’ averaging length scale $n_x$ (generally we find an optimal value between $\sim 0.1 n_y$ and $\sim 0.5 n_y$, see figure 3 caption). We see in figure 3(C) that the trend of precision versus $N_y$ is similar to that of the LEGI model (figure 3(C)), but with two key differences. First, the precision is higher for REGI than for LEGI. This is due to regional averaging reducing the variance of the local species, as was known previously for the 1D model [20]. Second, the optimum in the precision as a function of $N_y$ is more pronounced for REGI than for LEGI. This is because the region surrounding the optimum corresponds to near-circular ellipses, where considerable transverse averaging occurs. As shown in the previous section, transverse averaging increases precision in the REGI model. Overall, the optimal structure (figure 3(A), top) is closer to a ‘globular’ circle than a ‘hairlike’ chain (compare locations of the optima to the dashed vertical line in figure 3(C), which corresponds to a perfect circle). Therefore, we see that optimal gradient sensing by a 2D structure benefits from an elliptical shape in which transverse averaging occurs.

Finally, we investigate the behavior of REGI in 3D (figure 3(D)). Here there are two control parameters: the number of cells in the gradient direction $N_y$ and the asymmetry of the ellipsoid in the two directions transverse to the gradient. Generally we find that the optimal shape at a fixed $N_y$ displays symmetry in the two transverse directions, and therefore we impose this symmetry explicitly and focus on the control parameter $N_y$. As before, at every $N_y$ we optimize the local species’ averaging length scale $n_x$. Importantly, in the 3D geometry, we find that the optimal value at every $N_y$ is $n_y^* = 0$, corresponding to no averaging of the local species (an effective LEGI model). This is due to the shielding of internal cells: since internal cells do not detect signal molecules, averaging of the local species would dramatically reduce the mean local readout, making it far less than the actual local signal value at the edge cell. This would severely reduce the mean $\tilde{\Delta}_N$, and thus the precision. The dependence of precision on $N_y$ is shown in figure 3(D). The additional jaggedness is again due to the incommensurate nature of the cubic cell lattice with the smooth ellipsoidal envelope, here amplified due to the additional dimension. We see in figure 3(D) that there is again an optimum. In fact, it is much more pronounced than in 2D: the overall value of the precision is ten-fold higher than in 2D. This is again due to the shielding of internal cells: the global species $Y$ is averaged among internal cells that do not produce it, which sharply decreases $\tilde{\bar{n}}_N$, and thereby increases $\tilde{\Delta}_N$ and thus the precision.

Overall, the optimal structure is very ‘globular’ (figure 3(A), bottom). Indeed, it is almost a sphere (compare the optima to the dashed vertical line in figure 3(D)). We conclude that, due to the combined effects of spatial averaging and shielding, the optimal 3D detector of linear gradients extends significantly in all three spatial dimensions. The effects of shielding are expanded upon in the Discussion.

4. Discussion

We have investigated theoretically and computationally the ways in which the precision of spatially extended, multi-component gradient sensing is affected by detector geometry. Using a minimal model of adaptive gradient sensing (LEGI), we have found that, unlike for concentration sensing, the precision of gradient sensing decreases with the size of the detector in a direction transverse to the gradient. This is due to the competing effects of noise reduction and a reduction of the covariance between concentrations subtracted to estimate the gradient. We have demonstrated that a simple modification of LEGI (REGI) restores the covariance and recovers the benefit of transverse averaging for gradient sensing. The result is that the optimal detectors in 2D and 3D are more globular than hairlike.

Our study elucidates several important roles of spatial averaging in gradient sensing. First, there is spatial averaging along the gradient. In both LEGI and REGI, the global species $Y$ is averaged along the gradient. For a linear signal profile, this averaging both increases the signal $\tilde{\Delta}^2$, and decreases the noise $(\delta \Delta)^2$. Therefore, it is optimal for $Y$ to be averaged along the gradient to as large an extent as possible. Second, in the REGI model, the local species $X$ is also averaged along the gradient. This decreases the signal but also decreases the noise [20]. Therefore, there is often an optimal ratio $n_x/n_y$ of the spatial extents of the averaging. Third, there is spatial averaging transverse to the gradient. In the LEGI model, only $Y$ is averaged transverse to the gradient. In a translationally symmetric geometry, this does not change the signal, but it changes the noise by both decreasing the variance of $Y$ and decreasing the covariance between $X$ and $Y$. These have opposite effects on the precision. For LEGI, the latter dominates, decreasing the precision. Therefore, transverse averaging is detrimental for gradient sensing. However, in the REGI model, $X$ is also averaged transverse to the gradient. Once again, in a translationally symmetric geometry, this does not change the signal with respect to REGI in 1D, but it decreases the noise, both by further reducing the variance in $X$ and by restoring a larger covariance between $X$ and $Y$. Therefore, transverse averaging is beneficial for REGI-type gradient sensing. These roles of spatial averaging are modified in geometries without translational symmetry as we discussed above. However, the net result remains the same: the optimal 2D and 3D REGI-type

---

8 Note that this particular effect of shielding will result in the value of $\Delta_N$ being positive in every edge cell, instead of only the edge cells at the high end of the gradient. The sensory outcomes are still biased, but are less adaptive, similar to ‘tug-of-war’ chemotaxis mechanisms that have been proposed [26].
gradient detectors are globular, benefitting from extensive spatial averaging in the transverse directions.

Our study also reveals the effects of shielding of signal from the inner cells in a 3D geometry. Shielding amplifies the effect of spatial averaging, since the measurements performed by edge cells, which detect signal, are averaged with those of their interior neighbors, which do not detect signal. This amplification increases the signal in a particular edge cell, but makes the system less adaptive, since every edge cell has an above-average readout. With shielding, a more appropriate measure of the sensory outcome might therefore be the difference in readouts between cells up and down the gradient, e.g. $\Delta_N - \Delta_i$. This measure is likely to depend nontrivially on internal and geometric parameters such as $n_x$ and $M$, and will likely result in a nontrivial optimal local averaging length scale, $n_x^* \neq 0$. Another possibility is that gain $G$ should be different in equations (4) and (5), compensating for the two messenger molecules averaging over different numbers of neighbors that do not detect the ligand. We leave both of these interesting explorations for future investigations.

In this work, we have emphasized the distinction between (i) concentration sensing within a graded concentration profile and (ii) gradient sensing. For example, in Drosophila development, individual nuclei in the embryo measure (and are thought to threshold) the local concentration, even though the morphogen gradient is graded [21, 23–25]. This is an example of concentration sensing. In contrast, gradient sensing, as explored here, is the task of obtaining an internal readout of the difference in local signal concentrations at two or more different points in space. In other words, unlike concentration sensing, gradient sensing determines the direction in which the concentration changes, and it allows subsequent directional polarization of the sensor. This definition of gradient sensing, by construction, is adaptive: the readout does not depend on the background concentration. Systems that respond adaptively and directionally to chemical gradients, such as ameba [27] and epithelial cell groups [12], are performing gradient sensing. Because concentration sensing and gradient sensing are distinct, it may not be so surprising that transverse averaging has very different effects on them: the precision of concentration sensing increases with the transverse size, whereas the precision of LEG gradient sensing decreases with the transverse size (figure 1).

How do our results compare to experimental systems? A well-studied example of a natural-gradient-sensing system is the growth factor-directed expansion of mammary epithelial ducts [10, 22]. Gradient sensing in this system has been shown to be multicellular and adaptive [12]. In vivo, the extension is led by an ‘end bud’ of cells at the duct tip. These tips can form either long hairlike structures or coalesce into nearly spherical globules, as was observed in organotypic studies with different chemical and genetic perturbations [12]. Long hairs could act as ‘feelers’ for the duct, sampling a long swath of the environment in the gradient direction. However, our analysis predicts that such hairlike morphologies are suboptimal, and the globular bud shape, as in figure 3(A), would produce a better precision. In agreement with the prediction, the end buds in wild-type mice are nearly spherical, and the globule is often wider than the duct itself [10]. Similarly, neither chemotaxing ameba [27] and neutrophils [2], nor growing neurons [1] form very thin hairlike protrusions to facilitate sensing. Instead they keep the aspect ratio of the gradient sensing part of the protrusions closer to one, again supporting our findings. Further, in Drosophila border cell migration, another example of directional collective cell behavior, groups of cells travel as a sphere in a confined space, where it would have been easier to travel as a chain [5]. All of these examples provide indirect evidence that transverse averaging is used in multiple biological contexts. While direct tests of effects of transverse averaging have not been done, they are certainly possible. Indeed, as mentioned above, different perturbations to organotypic epithelial cultures result in them assuming different geometric shapes [12]. Thus it should be possible to measure the accuracy of sensing (and the subsequent organoid polarization) as a function of the shape. Such experiments would allow direct testing of our main prediction that transverse averaging leads to more accurate directional sensing outcomes, especially in REGI-type models.

Acknowledgments

TS was supported by the Woodruff Fellowship at Emory University. SF and AM were supported in part by Simons Foundation grant 376198. AL was supported in part by NSF grant PoLS-1410545. IN was supported in part by NSF grant PoLS-1410978.

References

[1] Rosoff W J, Urbach J S, Esrick M A, Mcallister R G, Richards L J and Goodhill G J 2004 Nat. Neurosci. 7 678–82
[2] Onsum M D, Wong K, Herzmark P, Bourne H R and Arkin A P 2006 Phys. Biol. 3 190–9
[3] Song L, Nadkarni S M, Bodeker H U, Beta C, Bae A, Franck C, Rappel W J, Loomis W F and Bodenschatz E 2006 Eur. J. Cell. Biol. 85 981–9
[4] Sternlicht M D, Kourous-Mehr H, Lu P and Werb Z 2006 Differentiation 74 365–81
[5] Bianco A, Poukkula M, Cliffe A, Mathieu J, Luque C, Fulga T and Barth P 2007 Nature 448 362–5
[6] Friedl P and Gilmour D 2009 Nat. Rev. Mol. Cell. Biol. 10 445–37
[7] Mani M, Goyal S, Irvine K D and Shraiman B I 2013 Proc. Natl Acad. Sci. 110 20420–5
[8] Donà E et al 2013 Nature 503 285–9
[9] Pochapska S M and Montell D J 2014 Annu. Rev. Genet. 48 295–318
[10] Huebner R J and Ewald A J 2014 Sem. Cell Dev. Biol. 31 124–31
[11] Malet-Engra G, Yu W, Oldani A, Rey-Barroso J, Gov N S, Scita G and Dupré L 2015 *Curr. Biol.* **25** 242–50
[12] Ellison D et al 2016 *Proc. Natl Acad. Sci. USA* **13** E679–88
[13] Berg H C and Purcell E M 1977 *Biophys. J.* **20** 193–219
[14] Goodhill G J and Urbach J S 1999 *J. Neurobiol.* **41** 230–41
[15] Levchenko A and Iglesias P A 2002 *Biophys. J.* **82** 50–63
[16] Endres R G and Wingreen N S 2008 *Proc. Natl Acad. Sci. USA* **105** 15749–54
[17] Endres R G and Wingreen N S 2009 *Prog. Biophys. Mol. Biol.* **100** 33–9
[18] Hu B, Chen W, Rappel W J and Levine H 2010 *Phys. Rev. Lett.* **105** 48104
[19] Jilkine A and Edelstein-Keshet L 2011 *PLoS. Comput. Biol.* **7** e1001121
[20] Mugler A, Levchenko A and Nemenman I 2016 *Proc. Natl Acad. Sci. USA* **13** E689–95
[21] Gregor T, Tank D W, Wieschaus E F and Bialek W 2007 *Cell* **130** 153–64
[22] Ewald A J, Brenot A, Duong M, Chan B S and Werb Z 2008 *Dev. Cell* **14** 570–81
[23] Erdmann T, Howard M and Ten Wolde P R 2009 *Phys. Rev. Lett.* **103** 238101
[24] Lander A D 2011 *Cell* **144** 955–69
[25] Sokolowski T R and Tkacik G 2015 *Phys. Rev. E* **91** 062710
[26] Camley B A, Zimmermann J, Levine H and Rappel W J 2016 *Phys. Rev. Lett.* **116** 098101
[27] Swaney K, Huang C and Devreotes P 2010 *Annu. Rev. Biophys.* **39** 265–89