The ability to sense the environment is heterogeneously distributed in cell populations

Andrew Goetz[1*], Hoda Akl[1*], and Purushottam Dixit[1,2,3,4,5,6]

[1]Department of Physics, University of Florida, Gainesville FL
[2]Department of Chemical Engineering, University of Florida, Gainesville FL
[3]Genetics Institute, University of Florida, Gainesville FL
[4]Cancer center, University of Florida, Gainesville FL
[5]Institute for fundamental theory, Department of Physics, University of Florida, Gainesville FL
[6]Corresponding author: pdixit@ufl.edu
[*]Contributed equally

Abstract

Channel capacity of signaling networks quantifies their fidelity in sensing extracellular inputs. Low estimates of channel capacities for several mammalian signaling networks suggest that cells can barely detect the presence/absence of environmental signals. However, given the extensive heterogeneity in cell states, we hypothesize that the sensing ability itself varies from cell to cell in a cell state dependent manner. In this work, we present an information theoretic framework to quantify the distribution of sensing abilities from single cell data. Using data on two mammalian pathways, we show that sensing abilities are widely distributed in the population and most cells achieve better resolution of inputs than what is implied by traditional cell state agnostic estimates. We verify these predictions using live cell imaging data on the IGFR/FoxO pathway. Importantly, we identify cell state variables that correlate with cells' sensing abilities. This information theoretic framework will significantly improve our understanding of how cells sense in their environment.

Introduction

In cell populations, there is a significant overlap in cellular responses to environmental stimuli of differing strengths. This raises a fundamental question[1]: do signaling networks in cells relay accurate information about their environment to take appropriate action? And if not, where along the signal transduction pathway is the information lost[2,3]? Mutual information (MI) quantifies the information content in a cellular output about extracellular inputs. For an input u (e.g., concentration of a ligand) and an output x (e.g., intracellular species concentration[4–6] or a cellular phenotype[7,8]), the MI is defined as[9]

$$I = \sum_{x,u} p(x|u)p(u) \log_2 \frac{p(x|u)}{\sum_{u'} p(x|u')p(u')}. \tag{1}$$

Experimental single cell methods such as flow cytometry[10], immunofluorescence[10], mass spectrometry[11], or live cell imaging[12] allow us to estimate response histograms $p(x|u)$ across several inputs. Using these distributions, we can estimate the maximum of the MI (the channel capacity, CC) by optimizing Eq. 1 over input distributions $p(u)$. The CC quantifies fidelity of signal transduction[13]. For example, a CC of 1 bit implies that that the cells can *at best* distinguish between two input levels, with higher CCs indicating that cells can resolve multiple input states. Importantly, CC can be used to identify bottlenecks in signaling[3,14].

The CC has been estimated for input-output relationships in several mammalian signaling networks[4,6,15–18]. When the output is defined as levels of a single protein a fixed time, the CC was found to be surprisingly low, $\sim 1 - 1.5$ bits. These estimates have been improved by considering multidimensional outputs[5] or time varying inputs[19]. While these modifications led to somewhat higher CC estimates, the overall conclusion that cells know little about their environment remains well established. In contrast, significantly higher CC estimates have been found when the output is defined as an average over the cell population[6], suggesting that the only way to overcome low sensing fidelity is collective response.

These previous calculations estimated one channel capacity for all cells in a population, implicitly assuming that individual cells have similar sensing capabilities. We now know that cells exhibit extensive heterogeneity in cell state variables such as abundances of key signaling proteins[11,20,21], gene expression levels[21], epigenetic status[22], etc. These variables can remain constant throughout cells' entire life cycle and beyond[23]. Heterogeneity in cell state variables leads to a variability in response to extracellular cues, including chemotherapeutic drugs[23,24], mitogens[25], hormones[26], and chemotactic signals[7,8]. Therefore, it

is quite reasonable to expect that cells have variable abilities to sense their environment[27,28]. And that the differences in cells' abilities are explained by differences in cell states.

There is currently no conceptual framework or computational algorithms to estimate the distribution of sensing abilities in cell populations. To that end, we introduce an information theoretic quantity CeeMI: **Ce**ll sta**te** dependent **M**utual **I**nformation. We show that using typically collected single cell data and a computational model of the underlying signaling network, we can estimate the distribution $p_{\mathrm{CeeMI}}(I)$ of single cell signaling fidelities (single cell mutual information values). We also show that we can identify cell state variables that make some cells better and others worse at sensing their environment.

Using an illustrative example, we show that in heterogeneous cell populations, the traditional cell state agnostic estimate of mutual information is significantly lower than the mutual information of signaling networks in typical cells. Next, using previously collected experimental data, we estimate $p_{\mathrm{CeeMI}}(I)$ for two important mammalian signaling pathways; the Epidermal growth factor (EGF)/EGF receptor pathway[29] and the Insulin like growth factor (IGF)/Forkhead Box protein O (FoxO) pathway[29], we show that while the cell state agnostic CC estimates for both pathways are ~ 1 bit, most individual cells are predicted to be significantly better at resolving different inputs. Using live cell imaging data for the IGF/FoxO pathway, we show that our estimate of variability in sensing abilities matches closely with experimental estimates. Importantly, for this pathway, we also verify our prediction that specific cell state variables dictate cells' sensing abilities. We believe that CeeMI will be an important tool to explore variability in cellular sensing abilities and in identifying what makes some cells better and others worse in sensing their environment.

Results

Conditional mutual information models single cells as cell state dependent channels

Consider a cell population where cells are characterized by state variables $\boldsymbol{\theta}$**.** These include abundances of signaling proteins and enzymes, epigenetic states, etc. We assume that cell states are temporally stable[30], that is, $\boldsymbol{\theta}$ remains constant over a time scale that is much longer than typical fluctuations in environmental inputs. We assume that these variables are distributed in the population according to a distribution $p(\boldsymbol{\theta})$. If x denotes an output of choice (e.g., phosphorylation levels of some protein at a specific time) and u denotes the input (e.g., ligand concentration), the experimentally measured response distribution $p(x|u)$ can be decomposed as:

$$p(x|u) = \int p(x|u, \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{2}$$

where $p(x|u, \boldsymbol{\theta})$ is the distribution of the output x conditioned on the input u *and* cell state variables $\boldsymbol{\theta}$. In most cases, $p(x|u, \boldsymbol{\theta})$ may not be experimentally accessible. However, it is conceptually well defined.

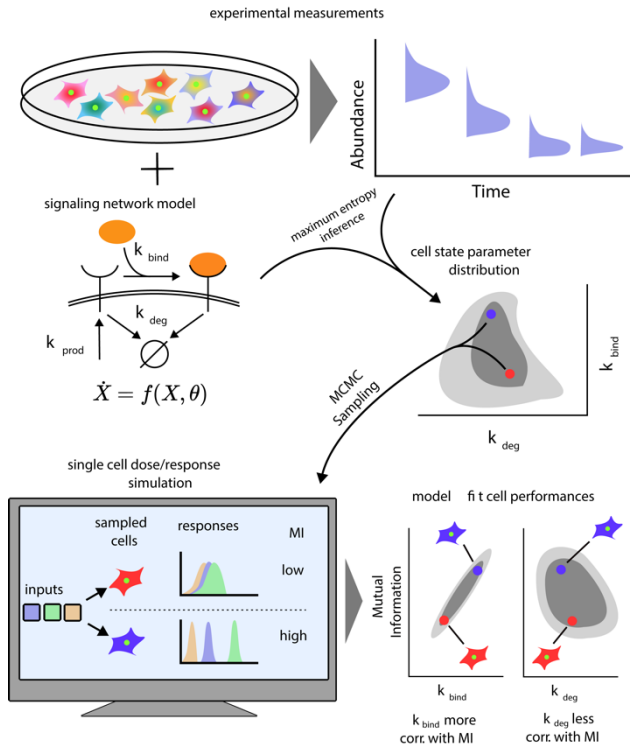

Figure 1. A schematic of our computational approach. (top) single cell data across different input conditions and time points are integrated with a stochastic model of a signaling network using a previous developed maximum entropy approach leading to a distribution over signaling network parameters $p(\boldsymbol{\theta})$ (middle). (bottom) In silico cells are generated using the inferred parameter distribution and cell-state specific mutual information $I(\boldsymbol{\theta})$ and population distribution of cell performances $p_{\mathrm{CeeMI}}(I)$ is estimated. The model also evaluates the correlation between cells' performance and biochemical parameters.

Using $p(x|u, \boldsymbol{\theta})$, we can define the cell state dependent mutual information $I(\boldsymbol{\theta})$ for a fixed input distribution $p(u)$ analogously to Eq. 1:

$$I(\boldsymbol{\theta}) = \sum_{x,u} p(x|u, \boldsymbol{\theta}) p(u) \log_2 \frac{p(x|u, \boldsymbol{\theta})}{\sum_{u'} p(x|u', \boldsymbol{\theta}) p(u')}. \tag{3}$$

$I(\boldsymbol{\theta})$ quantifies individual cells' ability to sense their environment as a function of the cell state parameters $\boldsymbol{\theta}$. The distribution $p_{\mathrm{CeeMI}}(I)$ of single cell sensing abilities is

$$p_{\mathrm{CeeMI}}(I) = \int p(\boldsymbol{\theta})\, \delta\big(I - I(\boldsymbol{\theta})\big) d\boldsymbol{\theta} \tag{4}$$

where $\delta(x)$ is the Dirac delta function. We can also compute the joint distribution between the single cell mutual information and any cell state variable of interest χ (e.g., cell surface receptor levels):

$$p_{\mathrm{CeeMI}}(I,\chi) = \int p(\boldsymbol{\theta})\delta\big(I - I(\boldsymbol{\theta})\big)\delta(\chi - \chi(\boldsymbol{\theta}))d\boldsymbol{\theta} \tag{5}$$

where $\chi(\boldsymbol{\theta})$ is the value of the biochemical parameter when cell state parameters are fixed at $\boldsymbol{\theta}$. The distributions defined in Eq. 5 quantify the interdependency between a cell's signaling fidelity $I(\boldsymbol{\theta})$ and cell specific biochemical parameters $\chi(\boldsymbol{\theta})$. As we will see below, the distributions in Eq. 4 and Eq. 5 can be experimentally verified when appropriate measurements are available. Finally, we define the population average of the cell state dependent mutual information:

$$I_{\mathrm{Cee}} = \int I(\boldsymbol{\theta})\, p(\boldsymbol{\theta})d\boldsymbol{\theta}. \tag{6}$$

In information theory, I_{Cee} is known as the conditional mutual information[9] between the input u and the output x conditioned on $\boldsymbol{\theta}$. From these definitions of it can be shown that $I_{\mathrm{Cee}} \geq I$ (SI Section 1). That is, if cell state variables remain approximately constant, at least some cells in the population have better sensing abilities compared to what is implied by the cell state agnostic mutual information (Eq. 1). Since I_{Cee} depends on the input distribution $p(u)$,we can find an optimal input distribution $p(u)$ that maximizes I_{Cee} (SI Section 1). Going forward, unless an input distribution is specified, the distributions $p_{\mathrm{CeeMI}}(I)$ and $p_{\mathrm{CeeMI}}(I,\chi)$ are discussed in the context of this optimal input distribution.

Maximum entropy inference can estimate $p_{\mathrm{CeeMI}}(I)$

Estimation of $p_{\mathrm{CeeMI}}(I)$ requires decomposing the experimentally observed response $p(x|u)$ into cell specific output distributions $p(x|u,\boldsymbol{\theta})$ and the distribution of cell state variables $p(\boldsymbol{\theta})$ (Eq. 3 and 4). This problem is difficult to solve given that neither of these distributions are accessible in typical experiments. For many signaling networks, stochastic biochemical models can be constructed to obtain the cell specific output distribution $p(x|u,\boldsymbol{\theta})$. Here, $\boldsymbol{\theta}$ comprise protein concentrations and effective rates of enzymatic reactions and serve as a proxy for cell state variables. Given the experimentally measured cell state averaged response $p(x|u)$ and the model predicted cell specific output distribution $p(x|u,\boldsymbol{\theta})$, we need a computational method to infer $p(\boldsymbol{\theta})$ (see Eq. 2). The problem of inference of parameter heterogeneity is a non-trivial inverse problem[31]. We use our previously develop maximum entropy-based approach[32] to infer $p(\boldsymbol{\theta})$. This way, we can estimate $p_{\mathrm{CeeMI}}(I)$ using experimentally obtained cell state agnostic response $p(x|u)$ and a stochastic biochemical model of the underlying signaling network.

Cell state agnostic mutual information underestimates cells' ability to sense their environment

To illustrate the effect of heterogeneity of cell state variables on the cell state agnostic estimate of mutual information (which we call cell state agnostic mutual information or I_{CSA} from now onwards), we consider a simple stochastic biochemical network of a receptor-ligand system. Details of the model and the calculations presented below can be found in SI Section 2. Briefly, extracellular ligand L binds to cell surface receptors R. Steady state levels of the ligand bound receptor B is considered the output. The signaling network is parametrized by several cell state variables $\boldsymbol{\theta}$ such as receptor levels, rates of binding/unbinding, etc. For simplicity, we assume that only one variable, steady state receptor level R_0 in the absence of the ligand, varies in the population. Calculations with variability in other parameters are presented in SI Section 2.

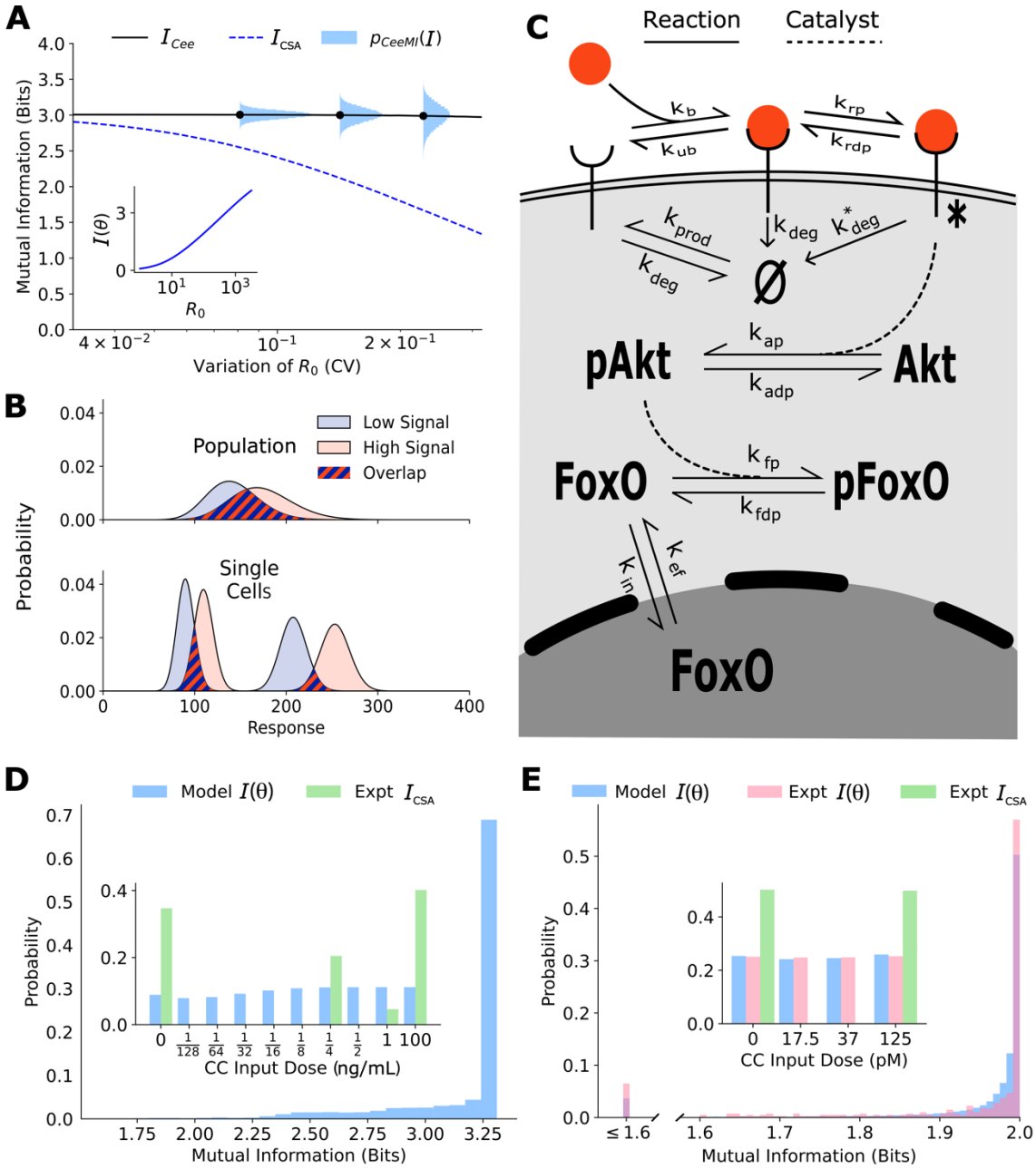

Figure 2. A. The distribution of single cell sensing abilities (horizontal blue histograms) and its average plotted as a function of the coefficient of variation of the distribution of one cell state variable, the cell surface receptor number The dashed blue lines show the traditional cell state averaged mutual information (Eq. 1). The inset shows the dependence between cell state specific mutual information and cell state variable The input distribution is assumed to be a gamma distribution. **B.** A schematic showing the effect of heterogeneity in cell states on population level response. Even when individual cells have little overlap in their responses to extracellular signal (bottom), the population level responses could have significant overlap (top), leading to a low mutual information between cell state averaged response and the input. **C.** A combined schematic of the two growth factor pathways. Extracellular growth factor ligand (red circle) binds to cell surface receptors which are shuttled to and from the plasma membrane continuously. Ligand bound receptors are phosphorylated and activate Akt. Phosphorylated Akt leads to phosphorylation of FoxO which effectively shuttles it out of the nucleus. **D**. The estimated distribution of single cell mutual information values for the EGF/EGFR pathway. The inset shows the input distribution corresponding to the maximum of the average of

(blue), along with the input distribution corresponding to the channel capacity of I_{CSA} (green). **E**. Same as **D** for the IGF/FoxO pathway. We additionally show the experimentally estimated $p_{\mathrm{CeeMI}}(I)$ (pink).

In this population, cells' response $p(B|L, R_0)$ is distributed as a Poisson distribution whose mean is proportional to the cell state variable R_0 (SI Section 2). That is, when all other parameters are fixed, a higher R_0 leads to lower noise. To calculate cell state dependent mutual information (Eq. 3), we assume that $p(L)$ is a gamma distribution. As expected, $I(R_0)$ (Eq. 3) between the output B and the input L increases monotonically with R_0 (inset in Fig. 2A). Moreover, given that R_0 varies in the population (assumed to be a gamma distribution as well), the sensing ability varies in the population as well (Fig. 2A). Notably, the average I_{Cee} of $I(R_0)$ remains relatively robust to variation in R_0. At the same time, the traditional estimate I_{CSA}, which is the mutual information between the input L and the cell state agnostic population response $p(B|L) = \int p(B|L, R_0)p(R_0)dR_0$ (Eq. 1 and Eq. 2), decreases as the population heterogeneity in R_0 increases. Importantly, I_{CSA} is significantly lower than the sensing ability of most cells in the population (Fig. 2A). This is because the overlap in the population response distributions is significantly larger than single cell response distributions (Fig. 2B), which arises due to the variability in cell state variables. This simple example illustrates that the traditional mutual information estimates may severely underestimate cells' ability to resolve inputs, especially when parameters are heterogeneously distributed. Moreover, it is crucial to explicitly account for heterogeneity in cell state variables when estimating fidelity of cellular communication channels.

Experimental verification of $p_{\mathrm{CeeMI}}(I)$ using growth factor signaling networks

In real cell populations, state variables that govern signaling dynamics such as protein levels[20,26] (receptors, kinases, dephosphatases, etc.), endocytosis rates[33], ligand binding rates[34], etc. are known to vary from cell to cell. Therefore, we expect the sensing abilities to be widely distributed as well. To experimentally verify the computational prediction of the distribution $p_{\mathrm{CeeMI}}(I)$ of sensing abilities, we need an experimental system that allows us to approximate the cell state specific response distribution $p(x|u, \boldsymbol{\theta})$. The IGF/FoxO pathway is an ideal system for these explorations for several reasons. First, following IGF stimulation, the transcription factor FoxO is pulled out of the nucleus[29]. GFP-tagged variant of FoxO can be used to detect the dynamics of nuclear levels of FoxO at the single cell level[28]. Second, nuclear FoxO levels reach an approximate steady state within 30-45 minutes of constant IGF stimulation with FoxO levels decreasing with increasing IGF dose[28]. As a result, an approximate cell state specific distribution $p(x|u, \boldsymbol{\theta})$ of steady state levels of FoxO can be obtained by stimulating the same cell with successive IGF levels[28]. Finally, the biochemical interactions in the IGF/FoxO are well studied[29] (Fig. 2C),

214 allowing us to a build stochastic biochemical model based on previous computational efforts[35] that fits
215 the single cell data accurately. Another system where $p_{\mathrm{CeeMI}}(I)$ can in principle be verified is the
216 EGF/EGFR pathway[29] (Fig. 2C). Here too, abundance of cell surface EGFR levels can be tracked in live
217 cells following EGF stimulation[36,37], allowing us to obtain cell-state specific response distribution
218 $p(x|u,\boldsymbol{\theta})$. Below, we show estimates of $p_{\mathrm{CeeMI}}(I)$ for both pathways and an experimental verification of
219 our estimates for the IGF/FoxO pathway where live-cell imaging data was previously collected[28].
220
221 The details of the calculations presented below can be found in SI Sections 3 and 4. Briefly, we first
222 constructed stochastic biochemical models for the two pathways based on known interactions[29] (Fig. 2C)
223 and previous models[35,38]. The output for the EGFR pathway was defined as the steady state levels of cell
224 surface EGF receptors and the output for the IGF/FoxO pathway was defined as the steady state nuclear
225 levels of the transcription factor FoxO. Using previously collected single cell data on the two
226 pathways[28,32] and our maximum entropy-based framework[32], we estimated the distribution over
227 parameters $p(\boldsymbol{\theta})$ for the two pathways (SI Section 3). Using these inferred distributions, and the model-
228 predicted cell state specific response distribution $p(x|u,\boldsymbol{\theta})$, we can compute $p_{\mathrm{CeeMI}}(I)$ for any specified
229 input distribution $p(u)$. We choose the support of the input distribution as the ligand concentrations
230 used in the experimental setup. The estimates shown in Fig. 2D and Fig. 2E show $p_{\mathrm{CeeMI}}(I)$
231 corresponding to the input distribution that maximizes the conditional mutual information I_{Cee} (see Eq.
232 6).
233
234 Similar to the illustrative example (Fig. 2A), there is a wide distribution of single cell sensing fidelities in
235 real populations (Fig. 2D and Fig. 2E). Moreover, most cells are better sensors compared to what is
236 indicated by the maximum of I_{CSA} which was estimated to be ~ 1 bit for both pathways. Indeed, as seen
237 in the insets of Fig. 2D and Fig. 2E, the input distribution corresponding to the maximum of I_{CSA} is
238 concentrated on the lowest and the highest input for both pathways, indicating that cells may be able to
239 detect only two input levels. In contrast, the input distribution corresponding to the maximum of I_{Cee} is
240 close to uniform, suggesting that individual cells can in fact resolve different ligand levels[28].
241
242 To verify our computational estimate of $p_{\mathrm{CeeMI}}(I)$ for the IGF/FoxO pathway, we reanalyzed previously
243 collected data[28] wherein several cells were stimulated using successive IGF levels. The details of our
244 calculations can be found in SI Section 4. Briefly, the cells reached an approximate steady state within 60
245 minutes of each stimulation and nuclear FoxO levels measured between 60 to 90 minutes were used to
246 approximate an experimental cell state specific response distribution $p(x|u,\boldsymbol{\theta})$. The distribution

$p_{\mathrm{CeeMI}}(I)$ was then obtained by maximizing the average mutual information I_{Cee} (averaged over all cells) with respect to the input distribution. As seen in Fig. 2D and Fig. 2E, the experimentally evaluated single cell signaling fidelities match closely with our computational estimates. Moreover, as predicted using our computational analysis, individual cells in the population were significantly better at sensing than what is implied by the maximum of I_{CSA}. Indeed, the distribution of steady state FoxO levels were found to be well-resolved at the single cell level as well (Fig. 2E). Live cell imaging data for the EGFR pathway was not available and we leave it to future studies to validate our predictions.

Our calculations show that real cell populations comprise cells that have differing sensing fidelities, individual cells are significantly better at sensing their environment than what traditional estimates would indicate, and importantly, the CeeMI approach can accurately estimate the variation in signaling performances using readily collected experimental data and stochastic biochemical models.

CeeMI identifies biochemical parameters that determine cells' ability to sense their environment

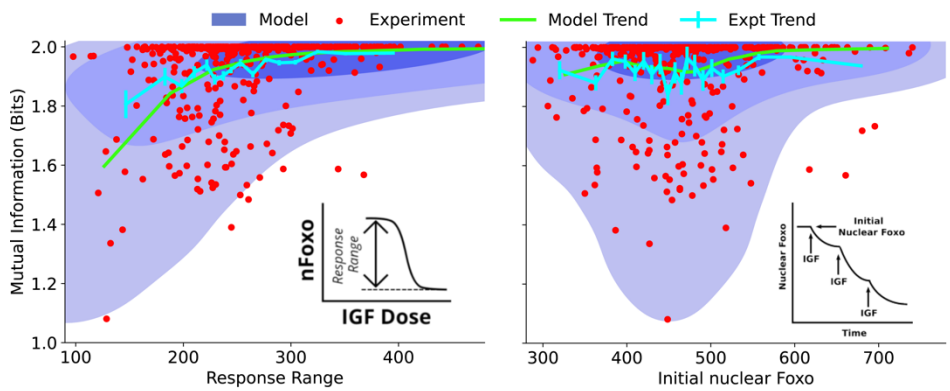

Figure 3. Dependence on cell state dependent mutual information on biochemical parameters. (left) The joint distribution $p_{\mathrm{CeeMI}}(I,\chi)$ of cell state specific mutual information and biochemical parameter χ chosen to be the single cell response range of nuclear FoxO levels (x-axis, see inset for a cartoon). The shaded blue regions are model predictions, and the green line is the model average. The darker shades represent higher probabilities. The red dots represent experimental cells. The cyan line represents experimental averages. (right) same as (left) with biochemical parameter χ chosen to be steady state nuclear Foxo levels in the absence of stimulation. The contours represent 1% to 10%, 10% to 50%, and 50% to 100% of the total probability mass (from faint to dark shading).

We expect that cells' ability to sense their environment depends on their state variables. Our approach allows us to quantify the joint distribution $p_{\mathrm{CeeMI}}(I,\chi)$ of single cell signaling fidelity and biochemical state variables that take part in the signaling network. To test whether we can identify variables that are predictive of cellular fidelity, we estimated the joint distribution $p_{\mathrm{CeeMI}}(I,\chi)$ for two variables that were experimentally accessible, response range of nuclear FoxO (Fig. 3 left, see inset) and total nuclear FoxO levels prior to IGF stimulation (Fig. 3 right, see inset). In both figures, the shaded regions show

computational estimates of the joint probability distributions, and the red circles represent real cells. The green and the cyan trend lines represent computational and experimental binned averages.

One may expect that higher total nuclear FoxO levels could result in lower noise and therefore better sensing abilities. However, we find that total nuclear FoxO levels only weakly correlate with cell state dependent mutual information (correlation coefficient $r = 0.16$ for computational estimates and $r = 0.04$ for experimental data). In contrast, cell state dependent mutual information depended strongly on the dynamic range of the response (correlation coefficient $r = 0.53$ for computational estimates and $r = 0.29$ for experimental data). Importantly, the model captured the observation that cells with a small response range had a variable sensing abilities while cells with a large response range all performed well in resolving extracellular IGF levels. Surprisingly, the total nuclear FoxO levels only weakly correlated with the cell specific mutual information. In SI Section 5, we show the predicted joint distributions $p_{\mathrm{CeeMI}}(I, \chi)$ for several other biochemical variables that can potentially govern single cells' response to extracellular IGF stimuli. This way, CeeMI can be used to systematically identify cell state variables that differentiate between good sensors and bad sensors of the environmental stimuli.

Discussion

Cell populations are characterized by heterogeneity in cell state variables that is responsible for important phenotypic differences including sensitivity to drugs[23,24], response to chemotactic signals[7,8], and following proliferation cues[25]. Therefore, it is reasonable to expect that cells' ability to sense their environment depends on their state and is therefore variable across cells in a population. To quantify this heterogeneity, here, we developed a novel information theoretic framework that takes as input easily measurable single cell data and models of signaling networks to quantify the distribution $p_{\mathrm{CeeMI}}(I)$ of single cell sensing abilities. Our framework also quantifies the joint distribution $p_{\mathrm{CeeMI}}(I, \chi)$ of cell specific sensing ability and any biochemical cell state variable. Importantly, using two growth factor pathways, we showed that individual cells in real cell populations were much better at sensing their environment compared to what is implied by the traditional estimate of channel capacities of signaling networks.

Our framework will be useful in identifying bottlenecks in signal transduction. Many cellular phenotypes such as chemotaxis[7,8] and cell proliferation[39] exhibit a weak correlation between cellular outputs (e.g., directional alignment with chemical gradients) with the input (e.g., gradient strength), resulting in a low

channel capacity even for individual cells. In such cases, it is important to understand where exactly along the information transduction pathway is the information about the gradient is lost. If traditional calculations are pursued, for example, for movement of mammalian cells under EGF gradients[7,8], one may conclude that the information loss likely happens right at the receptor level (Fig. 2D). In contrast, CeeMI will allow us to disentangle the effect of cell state heterogeneity and noisy cellular response to precisely pinpoint intracellular signaling nodes that are responsible for signal corruption.

A limitation of the current approach is the assumption that cell state variables remain constant over the time scale of typical environmental fluctuations. While many cell state variables remain constant over the life spans of cells and beyond[23,40], it is likely that state changes may occur within the lifespan of a cell[41]. These dynamical changes can be accommodated easily in the framework. Here, instead of fixing the cell state variables $\boldsymbol{\theta}$, we can treat them as initial conditions and propagate them stochastically with their own dynamics. In the limit of very fast dynamics where individual cells rapidly transition cell states, our framework will agree with traditional cell state agnostic estimates of the channel capacity. In contrast, if the cell state dynamics are slow, our framework highlights differences between cells in the population.

In summary, we showed that like other phenotypes, the ability to sense the environment is itself heterogeneously distributed in a cell population. Moreover, we also showed that mammalian cells appear to be significantly better at sensing their environment than what traditional mutual information calculation suggests. Finally, we showed that we could identify cell state variables that made some cells better sensors compared to others. We believe that CeeMI will be an important framework in quantifying fidelity of input/output relationships in heterogeneous cell populations.

Acknowledgments

AG, HA, and PD are supported by NIGMS grant R35GM142547. The authors would like to thank Andrew Levchenko for useful discussions.

References

1. Levchenko, A. & Nemenman, I. Cellular noise and information transmission. *Curr. Opin. Biotechnol.* **28**, 156–164 (2014).
2. ten Wolde, P. R., Becker, N. B., Ouldridge, T. E. & Mugler, A. Fundamental Limits to Cellular Sensing. *J. Stat. Phys.* **162**, 1395–1424 (2016).

3. Rhee, A., Cheong, R. & Levchenko, A. The application of information theory to biochemical signaling systems. *Phys. Biol.* **9**, 045011 (2012).
4. Cheong, R., Rhee, A., Wang, C. J., Nemenman, I. & Levchenko, A. Information Transduction Capacity of Noisy Biochemical Signaling Networks. *Science* **334**, 354–358 (2011).
5. Selimkhanov, J. *et al.* Accurate information transmission through dynamic biochemical signaling networks. *Science* **346**, 1370–1373 (2014).
6. Suderman, R., Bachman, J. A., Smith, A., Sorger, P. K. & Deeds, E. J. Fundamental trade-offs between information flow in single cells and cellular populations. *Proc. Natl. Acad. Sci.* **114**, 5755–5760 (2017).
7. Varennes, J., Moon, H., Saha, S., Mugler, A. & Han, B. Physical constraints on accuracy and persistence during breast cancer cell chemotaxis. *PLOS Comput. Biol.* **15**, e1006961 (2019).
8. Moon, H., Saha, S., Mugler, A. & Han, B. Signal processing capacity of the cellular sensory machinery regulates the accuracy of chemotaxis under complex cues. *iScience* **24**, 103242 (2021).
9. Cover, T. M. & Thomas, J. A. *Elements of information theory*. (Wiley-Interscience, 2006).
10. Wu, M. & Singh, A. K. Single-cell protein analysis. *Curr. Opin. Biotechnol.* **23**, 83–88 (2012).
11. Slavov, N. Single-cell protein analysis by mass spectrometry. *Curr. Opin. Chem. Biol.* **60**, 1–9 (2021).
12. Lemon, W. C. & McDole, K. Live-cell imaging in the era of too many microscopes. *Curr. Opin. Cell Biol.* **66**, 34–42 (2020).
13. Tang, Y. & Hoffmann, A. Quantifying information of intracellular signaling: progress with machine learning. *Rep. Prog. Phys.* **85**, 086602 (2022).
14. Hansen, A. S. & O'Shea, E. K. Limits on information transduction through amplitude and frequency regulation of transcription factor activity. *eLife* **4**, e06559 (2015).
15. Uda, S. *et al.* Robustness and Compensation of Information Transmission of Signaling Pathways. *Science* **341**, 558–561 (2013).
16. Frick, C. L., Yarka, C., Nunns, H. & Goentoro, L. Sensing relative signal in the Tgf-β/Smad pathway. *Proc. Natl. Acad. Sci.* **114**, (2017).
17. Garner, K. L. *et al.* Information Transfer in Gonadotropin-releasing Hormone (GnRH) Signaling. *J. Biol. Chem.* **291**, 2246–2259 (2016).
18. Billing, U. *et al.* Robustness and Information Transfer within IL-6-induced JAK/STAT Signalling. *Commun. Biol.* **2**, 27 (2019).
19. Lee, J. B., Caywood, L. M., Lo, J. Y., Levering, N. & Keung, A. J. Mapping the dynamic transfer functions of eukaryotic gene regulation. *Cell Syst.* **12**, 1079-1093.e6 (2021).
20. Meyer, R. *et al.* Heterogeneous kinetics of AKT signaling in individual cells are accounted for by variable protein concentration. *Front. Physiol.* **3**, (2012).
21. Frei, A. P. *et al.* Highly multiplexed simultaneous detection of RNAs and proteins in single cells. *Nat. Methods* **13**, 269–275 (2016).
22. Bonasio, R., Tu, S. & Reinberg, D. Molecular Signals of Epigenetic States. *Science* **330**, 612–616 (2010).
23. Shaffer, S. M. *et al.* Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature* **546**, 431–435 (2017).
24. Spencer, S. L., Gaudet, S., Albeck, J. G., Burke, J. M. & Sorger, P. K. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **459**, 428–432 (2009).
25. Matson, J. P. & Cook, J. G. Cell cycle proliferation decisions: the impact of single cell analyses. *FEBS J.* **284**, 362–375 (2017).
26. Norris, D. *et al.* Signaling Heterogeneity is Defined by Pathway Architecture and Intercellular Variability in Protein Expression. *iScience* **24**, 102118 (2021).
27. Keshelava, A. *et al.* High capacity in G protein-coupled receptor signaling. *Nat. Commun.* **9**, 876 (2018).

392 28. Gross, S. M., Dane, M. A., Bucher, E. & Heiser, L. M. Individual Cells Can Resolve Variations in
393 Stimulus Intensity along the IGF-PI3K-AKT Signaling Axis. *Cell Syst.* **9**, 580-588.e4 (2019).
394 29. *Signal transduction: principles, pathways, and processes*. (Cold Spring Harbor Laboratory Press,
395 2014).
396 30. Trapnell, C. Defining cell types and states with single-cell genomics. *Genome Res.* **25**, 1491–1498
397 (2015).
398 31. Banks, H. T., Kenz, Z. R. & Thompson, W. C. A review of selected techniques in inverse problem
399 nonparametric probability distribution estimation. *J. Inverse Ill-Posed Probl.* **20**, (2012).
400 32. Dixit, P. D., Lyashenko, E., Niepel, M. & Vitkup, D. Maximum Entropy Framework for Predictive
401 Inference of Cell Population Heterogeneity and Responses in Signaling Networks. *Cell Syst.* **10**, 204-
402 212.e8 (2020).
403 33. Kallenberger, S. M. *et al.* Correlated receptor transport processes buffer single-cell heterogeneity.
404 *PLOS Comput. Biol.* **13**, e1005779 (2017).
405 34. Chung, J. C., Sciaky, N. & Gross, D. J. Heterogeneity of epidermal growth factor binding kinetics on
406 individual cells. *Biophys. J.* **73**, 1089–1102 (1997).
407 35. Wimmer, R. J. *et al.* Mathematical modeling reveals modulation of both nuclear influx and efflux of
408 Foxo1 by the IGF-I/PI3K/Akt pathway in skeletal muscle fibers. *Am. J. Physiol.-Cell Physiol.* **306**,
409 C570–C584 (2014).
410 36. Liu, L., Shi, H., Chen, X. & Wang, Z. Regulation of EGF-Stimulated EGF Receptor Endocytosis During
411 M Phase. *Traffic* **12**, 201–217 (2011).
412 37. Fortian, A. & Sorkin, A. Live cell fluorescence imaging reveals high stoichiometry of Grb2 binding to
413 the EGF receptor sustained during endocytosis. *J. Cell Sci.* jcs.137786 (2013) doi:10.1242/jcs.137786.
414 38. Lyashenko, E. *et al.* Receptor-based mechanism of relative sensing and cell memory in mammalian
415 signaling networks. *eLife* **9**, e50342 (2020).
416 39. Gross, S. M. & Rotwein, P. Unraveling Growth Factor Signaling and Cell Cycle Progression in
417 Individual Fibroblasts. *J. Biol. Chem.* **291**, 14628–14638 (2016).
418 40. Filipczyk, A. *et al.* Network plasticity of pluripotency transcription factors in embryonic stem cells.
419 *Nat. Cell Biol.* **17**, 1235–1246 (2015).
420 41. MacLean, A. L., Hong, T. & Nie, Q. Exploring intermediate cell states through the lens of single cells.
421 *Curr. Opin. Syst. Biol.* **9**, 32–41 (2018).
422