Poor sensing maximises microbial fitness when few out of many signals are sensed

Age J. Tjalma[a], **Robert Planqué**[b], **and Frank J. Bruggeman**[a,1]

[a]Systems Biology Lab, AIMMS, VU University, De Boelelaan 1085, 1081 HV, Amsterdam, The Netherlands; [b]Dept. of Mathematics, VU University, De Boelelaan 1111, 1081 HV, Amsterdam, The Netherlands

This manuscript was compiled on April 8, 2020

An open problem in biology is to understand when particular adaptation strategies of microorganisms are selected during evolution. They range from random, bet-hedging strategies to deterministic, responsive strategies, relying on signalling circuits. We present an evolutionary model that integrates basic statistical physics of molecular circuits with fitness maximisation and information theory. This model provides an explanation for a puzzling observation on responsive strategies: the accuracy with which signalling networks track external signals seems remarkably low. Single cells often distinguish only between 2 to 4 concentration ranges, corresponding to 1 or 2 bits of mutual information between signal and response. Why did evolution lead to such low-fidelity signalling systems? Our theory offers an explanation by taking a novel perspective. It considers the fitness benefit of all signals, including those that are not sensed. We introduce a new concept, 'latent information', which captures the mutual information between all non-sensed signals and the optimal response. The theory predicts that it is often evolutionarily optimal to transduce sensed signals noisily when latent information is present. It indicates that fitness can indeed be maximal when the mutual information extracted from sensed signals is not maximal, but rather has a low value of about 1 or 2 bits. Cells likely do not sense all signals because of the fitness cost of expressing idle signalling systems that consume limited biosynthetic resources. Our theory illustrates that as the total available information about the optimal behaviour decreases, the cell should trust the available information less, and gamble more.

Cellular decision making | Phenotypic adaptation strategies | Bet-hedging | Cellular signalling | Information Theory | Mutual information | Biochemical noise

DRAFT

Evolutionary theories about the fitness of phenotypic strategies (1–6) have made successful predictions of the outcomes of natural selection in fluctuating environments (3, 4, 7, 8). In those theories, the (geometric) fitness (4–6) is generally maximised. This fitness measure equals the logarithm of the fold change in the number of organisms divided by the number of environmental periods, or the total time. Organisms that produced most offspring have maximal geometric fitness. These theories evaluate the fitness contributions of phenotypic strategies, ranging from random, bet-hedging strategies to deterministic, responsive strategies.

Bet-hedging strategies (14) are independent of environmental conditions. Cells can switch randomly between alternative phenotypic states. This ensures the existence of subpopulations that might be maladapted to the current condition, but are prepared for future conditions. For example, slow-growing, stress-tolerant cells (so-called persister cells) and fast-growing, stress-sensitive cells have been discovered in microbial cell populations (9–12). The fast growing cells largely determine the current fitness, while the persister cells are 'insurance policies', guaranteeing survival when conditions suddenly become harsh and extinction-threatening. Together, they maximise the geometric fitness of the population in conditions that fluctuate between benign and adverse conditions.

Close-to-deterministic adaptation strategies, on the other hand, also exist. Cells then perceive and transmit a signal to infer the state of the environment and adapt their phenotype accordingly. The omnipresence of two-component signalling circuits across microorganisms (13) highlights the importance of this mode of adaptation. However, inevitable molecular stochasticity in signalling circuits can cause random mismatches between the phenotype and the environmental state, leading to fitness losses.

Since the entire range of phenotypic behaviours is found amongst microorganisms, and because they are great experimental systems for physiology and evolution, microbiology is the perfect playground for testing predictions and improving evolutionary theories. We therefore focus our theory on microbial phenotypic adaptation.

The question under which conditions bet-hedging adaptation strategies are more evolutionarily successful than sensing strategies has a long history in evolutionary biology. Bet-hedging has been predicted to be evolutionarily advantageous under at least two conditions (14): In slowly-changing, mild environments where sensing machinery would rarely provide an evolutionary benefit for long periods of time, and in environments that change quickly into extinction-threatening states where responsive adaptation would be too slow (1–4, 6, 14). In all other cases, responsive signalling strategies should be favoured.

What remains poorly understood is what the optimal accuracy of signalling systems should be. In particular, experiments indicate that single cells have a remarkably low capacity for accurate tracking of environmental signals (15–19) far below

Significance Statement

Surprisingly, microorganisms appear to sense only very few environmental signals (such as nutrients and stresses) compared to the number of conditions they can encounter. Even worse, the signals they do sense are transduced at low fidelity. We study the accuracy of sensing in the situation where multiple signals determine the optimal response, but only few signals are actually sensed. We show that it is in fact to be expected that sensing in the presence of latent information should be imprecise, and that signalling circuits should underperform.

AJT performed the research. AJT, RP and FJB wrote the paper.

[1]To whom correspondence should be addressed. E-mail: f.j.bruggeman@vu.nl

theoretical predictions: cells should be able to measure much more accurately than they actually do. Intuitively, one would expect that more accurate signalling improves evolutionary success as it reduces maladaptation. Accordingly, many theories are based on maximisation of the mutual information between signal and response (20–22). However, in reality cells only appear to sense few external signals and in a noisy regime, leading to stochastic responses (23–26). Moreover, experiments report poor signalling capabilities of single cells; no more than 2 bits of mutual information between an external signal and an internal response have been found (16–19). Thus, either reliable signal transduction is not that important for evolutionary success, or it is very hard for cells to achieve reliable signal transduction.

Reliable signal transduction might be difficult to achieve due to the inevitable randomness of molecular systems (27). Increased expression of signalling proteins can make them more reliable (28), but requires a fitness-reducing diversion of resources from growth processes. Only if signalling activity leads to a net fitness increase one would expect signalling systems to evolve. Therefore, expression of idle signalling systems is likely fitness reducing in the absence of the signal and, moreover, genes encoding signalling machinery may then eventually accumulate random mutations that reduce signalling performance. In addition, many signalling systems rely on signal receptors in the membrane, and their expression is at the expense of nutrient importers; this reduces fitness too. These considerations may explain why cells express so few signalling systems; *E. coli* can grow, for instance, on hundreds of carbon sources, but it only has a handful one- and two-component sensing systems dedicated to nutrient sensing. However, they do not explain why exploited signalling systems display such remarkably low mutual information values between their signal and their response. Answering this question is the goal of this paper.

In information-theoretic studies on cellular signalling, maximisation of mutual information between a signal and a response variable is considered beneficial for cells. We argue that this is not necessarily so. Imagine, for example, a flat relationship between the fitness value and the response variable versus one that is sharply peaked. In the former case, infinite mutual information between signal and response variable would not improve fitness, while it would in the latter case. More importantly, another limitation is that most studies consider only one signal and one response variable. It is, however, likely that cells integrate different signals to induce a single response. Imagine a fitness landscape as a function of several response variables that contains ridges and peaks. In such a landscape, mutual information maximisation with respect to one signal and one response variable is insufficient. Instead, phenotypic adaptation strategies should be evaluated on the basis of fitness maximisation—considering all signals. Importantly, also 'latent signals' – those that are not sensed by the cell, but that are fitness enhancing if they would be sensed – should be considered.

Results

The geometric fitness model. We aim to investigate fitness maximisation of an isogenic population in the presence of fluctuating signals. All single cells sense a signal s that induces the expression of a response protein with concentration p (figure 1A). In addition to the sensed signal, 'latent' signals, $s_{l1}, ..., s_{ln}$, exist that are not sensed by cells. They are the elements of the vector $\mathbf{s}_l$. All signals, both sensed and latent, determine the optimal protein concentration $p_o(s, \mathbf{s}_l)$ that maximises fitness.

We use the following fitness function, which reaches its maximal value $f_o(s, \mathbf{s}_l)$ at $p = p_o$,

$$f(s,\mathbf{s}_l, p, p_o) = f_o(s, \mathbf{s}_l) \exp\left\{-\frac{1}{2}\frac{(p-p_o)^2}{\omega^2}\right\}. \quad [1]$$

The fitness reduction, due to a deviation of the protein concentration p from its optimal value p_o is determined by the width of the "fitness-landscape" ω (figures 1A and 1C). The interpretation of the fitness equation is explained in figure 1C.

To keep the model analytically tractable, we assume in addition that: (1) the optimum p_o is given by the deterministic function $q(s, \mathbf{s}_l)$, which is a nonlinear function of s plus a weighted sum of all latent signals $\mathbf{s}_l$ such that

$$p_o = q(s, \mathbf{s}_l) = g(s) + \sum_{j=1}^{n} \beta_j s_{lj}; \quad [2]$$

(2) all latent signals in $\mathbf{s}_l$ are independent, i.e., they do not covary, and are normally distributed.

Given these assumptions, following the methodology explained in figure 1B, we obtain for the fitness of a population of isogenic cells sensing one signal, averaged over all latent signals (see the appendix for the general derivation with multiple sensed signals),

$$G(s) = \ln\left[f_o(s, \mathbf{s}_l)\right] - \frac{1}{2}\ln\left[\frac{\omega^2+\tau^2+\alpha^2\sigma^2}{\omega^2}\right] - \frac{1}{2}\frac{\int(\alpha s - g(s))^2 h(s)ds + \sum_{j=1}^{n}\beta_j^2\langle\delta^2 s_{lj}\rangle}{\omega^2+\tau^2+\alpha^2\sigma^2}. \quad [3]$$

The first term $\ln[f_o(s, \mathbf{s}_l)]$ equals the maximal geometric fitness, we will denote it as $\ln[f_o]$ from here on. The last two terms are fitness costs. The first captures the cost of noisy signalling and quantifies the fitness reduction due to protein fluctuations that lead to an average deviation from the optimal concentration. The total variance of the protein fluctuations equals $\tau^2 + \alpha^2\sigma^2$. The second cost term captures a distance from optimality due to two suboptimal (deterministic) effects: i. the relationship $\langle p|s\rangle = \alpha s$ might deviate from the optimal relation $g(s)$ and ii. not sensing of latent signals that are informative about the value of $q(s, \mathbf{s}_l)$, leads to a fitness loss too, in terms of the variances of these latent signals.

The last cost term in equation 3 contains $\int(\alpha s - g(s))^2 h(s)ds$, which is reminiscent of a mean squared error (MSE) of an estimator. Its occurrence suggests that evolution can be interpreted as minimising the distance of a cell's behaviour to its optimal behaviour. It is debatable whether this term can be made small by natural selection. For this to occur, the dependency $g(s)$ should not become a too complex function, as otherwise cells would not be able to approximate it by a steady-state input-output relation of a molecular circuit.

Considering that we aim to investigate optimal sensing cells, we assume that $g(s) \approx \alpha s$. Therefore, the equation for

DRAFT

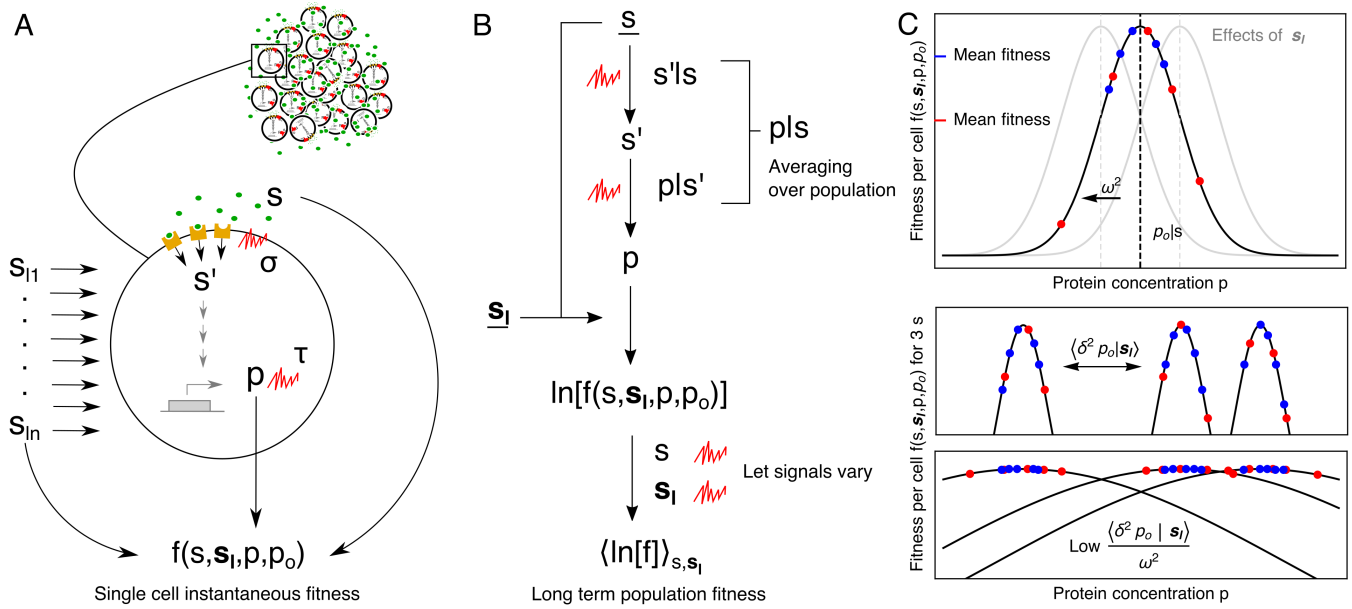

Fig. 1. Biological situation, methodological workflow, and intuition for the fitness function. A. Evolution acts on an (isogenic) population. An individual cell senses an environmental signal s and estimates its concentration to be s'. In this process noise is added, with variance σ^2. The signal is transduced and eventually leads to protein concentration p, which contains noise on its own; with variance τ^2. The concentration p leads to a certain fitness of the cell depending on how far it is from the optimum p_o following the fitness function $f(s, \mathbf{s}_l, p, p_o)$. The optimal protein concentration p_o is a function of the sensed signal s and all latent signals $s_{l1}, \ldots, s_{ln}$. **B.** Under a fixed environment we use $s'|s \sim \mathcal{N}(s, \sigma^2)$ and $p|s' \sim \mathcal{N}(\alpha s', \tau^2)$ to obtain $p|s \sim \mathcal{N}(\alpha s, \alpha^2\sigma^2 + \tau^2)$, as shown in the appendix. The average fitness of the phenotypically heterogeneous population at one environment can now be determined and depends on s. Taking the logarithm and averaging over all signals (i.e., all environments) leads to the geometric mean of the fitness. **C.** The fitness of one cell is given by: $f(s, \mathbf{s}_l, p, p_o) = f_o(s, \mathbf{s}_l) \exp\left\{-\frac{1}{2}\frac{(p-p_o)^2}{\omega^2}\right\}$ where ω determines the width of the fitness function, and $f_o(s, \mathbf{s}_l)$ determines the fitness (growth rate) in the optimum. The optimal protein concentration p_o is determined by the signal s (black curve), but is also influenced by the latent signals $\mathbf{s}_l$ (gray curves). Considering two hypothetical populations of blue and red cells, where the blue cells are genetically superior in transducing signals, we see how the mean fitness of the population in one environment depends on the accuracy of estimating the optimum p_o. However, depending on the latent signals $\mathbf{s}_l$, the optimum and thus the mean fitness can still change. The middle plot shows how, under fixed $\mathbf{s}_l$, the variance in optima determines how far the peaks of the fitness curves lie apart. The bottom plot illustrates how inaccurate signal transduction will have fewer fitness consequences when the variance in optima is small relative to the width of the fitness curve.

DRAFT

the geometric mean fitness of sensing cells in the presence of multiple signals, both latent and observed, reduces to

$$G = \ln[f_o] - \frac{1}{2}\ln\left[\frac{\omega^2 + \tau^2 + \alpha^2\sigma^2}{\omega^2}\right] - \frac{1}{2}\frac{\sum_{j=1}^{n}\beta_j^2\langle\delta^2 s_{lj}\rangle}{\omega^2 + \tau^2 + \alpha^2\sigma^2}. \quad [4]$$

The final term in equation 4 is the fitness cost of varying latent signals. It quantifies how much fitness-influencing uncertainty remains about the environment, caused by the influence of all latent signals $\mathbf{s}_l$ on the optimum p_o. This fitness cost is dependent on the width of the fitness function and the internal variance, and can thus be reduced by increased protein fluctuations, i.e., by increasing $\tau^2 + \alpha^2\sigma^2$. This would lead to bet-hedging or 'noisy sensing' strategies, in agreement with previous results (1–3). Such strategies can be optimal, which is shown by differentiating equation 4 with respect to the internal variance and solving for its optimum value. Division by ω^2 then gives,

$$\left(\frac{\tau^2 + \alpha^2\sigma^2}{\omega^2}\right)^{opt} = \frac{\sum_{i=1}^{n}\beta_j^2\langle\delta^2 s_{lj}\rangle}{\omega^2} - 1; \quad [5]$$

indicating that the relative internal variance should only exceed zero when the relative uncertainty in the environment exceeds 1 (figure 2A). In this case noisy signalling pays off due uncertainty in the environment. The geometric mean fitness of a cell is more sensitive to the relative uncertainty about its environment than to its own relative internal variance (figure 2A).

Thus, an optimal sensing cell, which does not sense all signals that are informative about the optimal behaviour, will still have to be noisy to overcome fitness variation due to the influences of all latent, non-sensed signals. Such a cell would therefore not follow a deterministic, pure sensing strategy, but allow for some bet-hedging behaviour, indicating that the mutual information between s and p should not always be maximised. This is an important insight from our theory that we will explore further.

Sensing vs non-sensing cells. The geometric mean fitness of a non-sensing population can also be derived from our general framework, outlined in the appendix. The variance induced by sensing machinery σ^2 drops out of the equation, as non-sensing cells lack this machinery. However, since all signals are now latent, an extra variance term appears in the second cost term:

$$G_{ns} = \ln[f_o] - \frac{1}{2}\ln\left[\frac{\omega^2 + \tau^2}{\omega^2}\right] - \frac{1}{2}\frac{\langle\delta^2 g(s)\rangle + \sum_{j=1}^{n}\beta_j^2\langle\delta^2 s_{lj}\rangle}{\omega^2 + \tau^2}. \quad [6]$$

This extra variance term, $\langle\delta^2 g(s)\rangle$, is the variance in the optimum due to the signal s, which is the signal of interest in the sensing population. Since we assumed above $g(s) \approx \alpha s$, we substitute $\langle\delta^2 g(s)\rangle$ with $\alpha^2\langle\delta^2 s\rangle$. The optimal, normalised

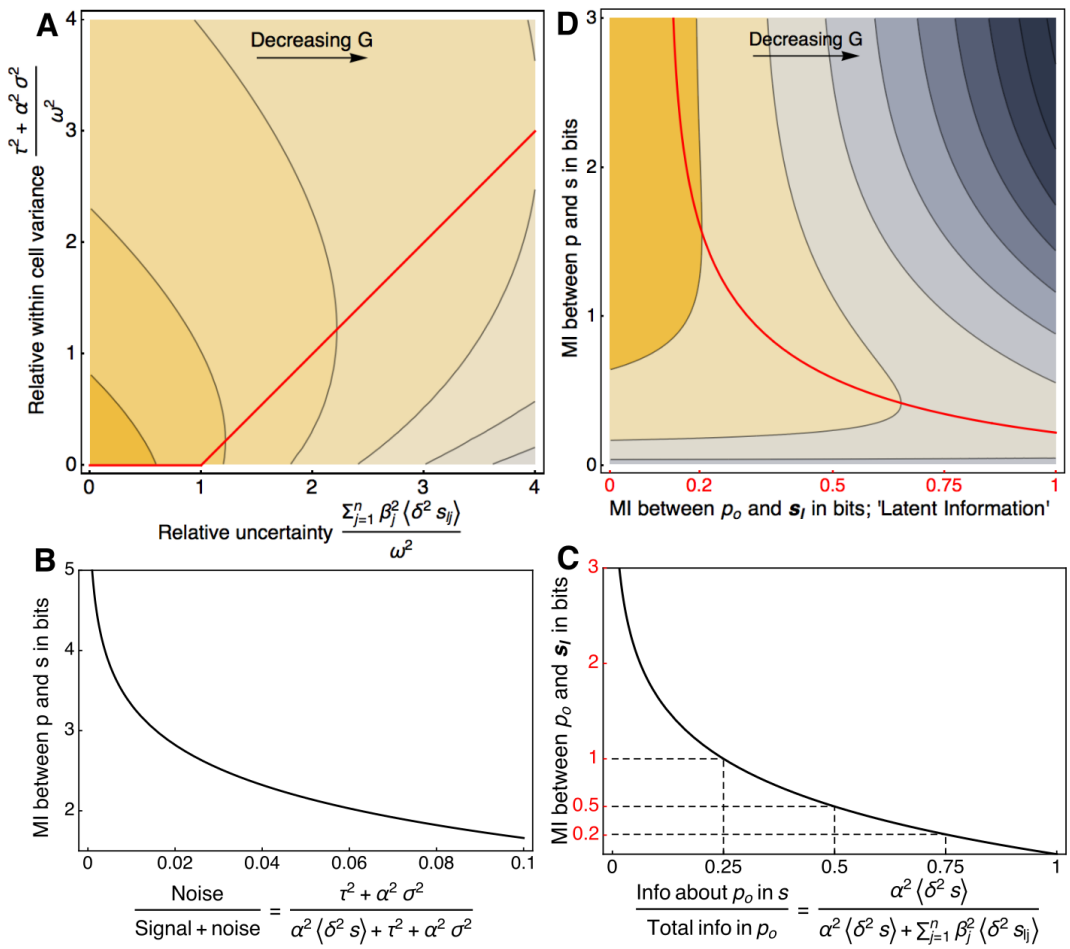

Fig. 2. Mutual information, latent information and fitness of sensing cells. A. Contour plot of the geometric mean fitness (G) as a function of the relative within-cell variance and relative uncertainty in the environment. The colour indicates G as a percentage of the maximum (given by $\ln[f_o]$), here it ranges from 90% (bottom left) to 30% (bottom right). The red line shows the optimal relative internal variance for a given relative uncertainty. **B.** Mutual information (MI) between p and s as a function of the fraction of internal variance (i.e., noise) over all variance in p. This gives an indication of the range of the MI between p and s given a certain noise level. Note that the x-axis ranges from 0 to 0.1. **C.** MI between p_o and $\mathbf{s}_l$ as a function of the propagated variance in p_o from s over the total variance in p_o. Since we are considering normal distributions, variance is proportional to mutual information. **D.** Contour plot of G as a function of the MI between the sensed signal s and the cell's response p and the MI between the optimum p_o and all latent signals $\mathbf{s}_l$ (i.e., the latent information). The colouring indicates G as a percentage of the maximum; ranging from 75% (top left) to -150% (top right), the red line is the optimal MI between p and s, and $\rho = 5$. The x-axis covers a much smaller range than the y-axis, it corresponds to the red y-axis in C.

internal variance for non-sensing cells is now given by,

$$\left(\frac{\tau^2}{\omega^2}\right)^{opt} = \frac{\alpha^2 \langle \delta^2 s \rangle + \sum_{j=1}^{n} \beta_j^2 \langle \delta^2 s_{lj} \rangle}{\omega^2} - 1. \qquad [7]$$

This result is similar to the results of Bull (1) and Haccou & Iwasa (2). The total variance in the environment, defined in those papers as a single parameter, equals $\alpha^2 \langle \delta^2 s \rangle + \sum_{j=1}^{n} \beta_j^2 \langle \delta^2 s_{lj} \rangle$ in our extended formalism. Cells should introduce noise when the relevant variance in the environment exceeds the width of the cellular fitness function.

In the appendix we evaluate the fitness difference between optimally sensing and optimally non-sensing cells, i.e., $G^{opt} - G_{ns}^{opt}$. There we show (in generalised form) that $G^{opt} - G_{ns}^{opt}$ is minimally 0, which only occurs when $\alpha^2 \langle \delta^2 s \rangle = 0$ or $\omega^2 \to \infty$. These limits are in agreement with biological intuition, since the only reason not to sense a signal, when sensing does not come at any cost, is when it contains no information about the optimum or it has no fitness consequence.

Mutual and latent information. In order to interpret the geometric mean fitness of sensing cells from an information theoretic perspective we express it in terms of mutual information (MI). First we define the MI between the signal s that is being sensed and the response of the cell p. This way of using MI is very common in the evaluation of the accuracy of signal transduction. We find for the MI, in bits, between the sensed signal s and the internal response p (see appendix) the following relation

$$I(p; s) = -\frac{1}{2} \log_2 \left[\frac{\tau^2 + \alpha^2 \sigma^2}{\alpha^2 \langle \delta^2 s \rangle + \tau^2 + \alpha^2 \sigma^2} \right]. \qquad [8]$$

The argument of the logarithm is the fraction of all variance in p that is caused by internal variance (i.e., noise). The MI depends on this fraction as shown in figure 2B, which indicates that when it decreases below 10%, the relevant range of MI is approximately 1-5 bits. The relevant range of MI as a function of the internal noise depends on both the type of signal transduction and the choice of the input distribution.

For instance, when two modes of a bimodal input distribution are perfectly separated and equally likely, the MI increases with 1 bit, effectively doubling the number of perceived states.

We will now introduce the crucial information measure for the fitness of sensing cells. In contrast to what is commonly done, we define an MI term between the optimum and the latent signals. From the cell's perspective, this measure can be interpreted as 'latent information',

$$I(p_o; \mathbf{s}_l) = -\frac{1}{2}\log_2\left[\frac{\alpha^2\langle\delta^2 s\rangle}{\alpha^2\langle\delta^2 s\rangle + \sum_{j=1}^{n}\beta_j^2\langle\delta^2 s_{lj}\rangle}\right]. \quad [9]$$

The argument of the logarithm captures the fraction of information on the optimum that the cell retrieved (figure 2C).

The relevant range of the latent information can be very low. For instance, when the sensed signal s contains 25% of the total information about the optimum, the latent information is 1 bit. When s contains 75% of all information on the optimum, the latent information is only 0.2 bit (figure 2C). An informative reference point to keep in mind is that at 0.5 bit latent information, s provides 50% of all information on the optimum p_o.

Equation 9 indicates that the cell can reduce latent information by sensing more signals. By sensing an extra signal its variance is added to the numerator in equation 9, decreasing the latent information. We note that latent information is not reduced by sensing an already sensed signal more accurately.

Maximal mutual information is not always optimal. We can write the geometric mean fitness of sensing cells (equation 4) in terms of the two mutual information measures. Before doing so, we define $\rho = \frac{\alpha^2\langle\delta^2 s\rangle}{\omega^2}$, the normalised variance in optima caused by the perceived signal s. The geometric mean fitness of a sensing population of isogenic cells equals

$$G = \ln[f_o] - \frac{1}{2}\ln\left[1 + \frac{\rho}{4^{I(p;s)}-1}\right] - \frac{1}{2}\frac{\rho\left(4^{I(p_o;\mathbf{s}_l)}-1\right)}{1+\frac{\rho}{4^{I(p;s)}-1}}. \quad [10]$$

When we study G as a function of the mutual and latent information, given a certain value of ρ, we observe that the fitness decreases sharply in the direction of increasing latent information (figure 2D, note the low range on the x-axis).

Differentiating equation 10 with respect to the MI between p and s and setting it to 0 gives the optimal MI:

$$I(p;s)^{opt} = \frac{1}{2}\log_2\left[1 + \frac{\rho}{\rho\left(4^{I(p_o;\mathbf{s}_l)}-1\right)-1}\right] \quad [11]$$

The optimal MI, $I(p;s)^{opt}$, given a certain amount of latent information, $I(p_o;\mathbf{s}_l)$, is plotted in red in figure 2D. The fact that there is an optimum for the MI means that cells should not maximise this term. This is exactly what a great part of previous research has focused on. The optimum is valid as long as $4^{I(p_o;\mathbf{s}_l)} - 1 > \frac{1}{\rho}$, which shows that as ρ increases there are lower values of $I(p_o;\mathbf{s}_l)$ for which an optimal number of bits MI between p and s exists. When this optimum does not exist, fitness is maximised as $I(p;s)$ approaches infinity, the fitness in this limit is given by:

$$\lim_{I(p;s)\to\infty} G = \ln[f_o] - \frac{1}{2}\rho\left(4^{I(p_o;\mathbf{s}_l)}-1\right) \quad [12]$$

This is the maximal fitness in the 'pure sensing' regime, where more MI always leads to higher fitness.

Distributed information leads to 'noisy sensing' of individual signals. An important conclusion can be drawn from figure 2D. It shows that for each given amount of latent information, there exists an optimal value of mutual information. This optimum converges to infinity as the latent information becomes very small.

As discussed above, the only way for cells to reduce the latent information is to sense more signals. The red line in figure 2D shows that the mutual information should increase when the latent information decreases. This concerns not just the mutual information of p and s, but also of the newly sensed signal (see appendix). This suggests that the total mutual information should increase as the cell reduces its latent information, but we do not know the optimal mutual information per signal. Within our framework, the only reason to maximise the mutual information with one signal is that this signal contains almost all information about the optimal response of the cell.

High mutual information between one signal and a response seems rarely necessary. By distinguishing between the regimes where cells should maximise mutual information and where there exists an optimum, we can plot which strategy should be chosen for certain combinations of ρ and the latent information (figure 3A).

Pure sensing is generally only preferred given low latent information. Another option seen in 3A, is when ρ becomes small. This is can be caused by two phenomena: i. $\alpha^2\langle\delta^2 s\rangle \to 0$, in this scenario the latent information will automatically approach infinity however (see equation 9), and ii. $\omega^2 \to \infty$, now there is no consequence to missing out on the effects of the latent signals, pure sensing is now optimal because we do not consider enzyme costs to sensing, and the entire relevant fitness landscape is convex (which means averaging over a region can more often lead to a lower result than choosing one point on the landscape). How low the latent information should become in order for pure sensing to become favourable depends on the value of ρ.

Evolution will drive the cell to reduced latent information values (figure 3B). However, cells can only achieve this by sensing more signals. In our model the cell would therefore start sensing all relevant signals, ending up in the pure sensing regime. In addition, as long as the variance in optima caused by the sensed signal ($\rho > 0$) remains, sensing is always preferred over bet-hedging. In reality however, cells will only sense signals that contain a sufficiently large and frequent enough fitness benefit, likely creating a lower bound on the latent information. As discussed, sensing machinery with low fitness benefit might be lost via selection and genetic drift; in figure 3B we see that this could happen at high latent information.

When one signal contains almost all information about the optimum, the mutual information should be maximised (figure 3C). We note, however, that even for high ρ the fitness curve saturates quickly, i.e., the amount of fitness gained per bit of mutual information quickly becomes very low. This is a consequence of two factors: firstly, mutual information is an exponential measure, where the number of bits MI leads to 2^{MI} perceived states; secondly, the usage of normal distributions leads to 4^{MI} instead of 2^{MI}, as can be seen in equation 10. In figure 3D we see how, when the latent information increases, the previously discussed optimum appears in the fitness curve. The exact latent information value when this occurs depends

DRAFT

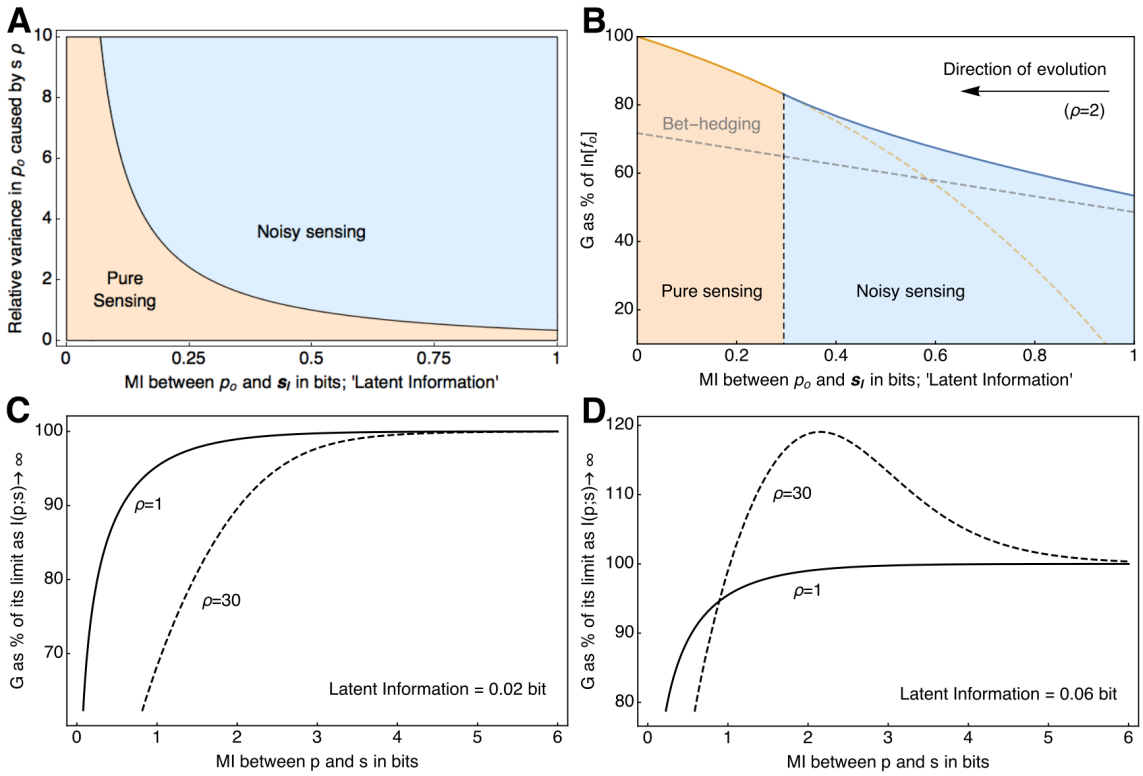

Fig. 3. Comparison of pure and noisy sensing strategies A. Optimal strategies given a certain relative variance in the optimum p_o caused by s (ρ) and a certain number of bits of mutual information (MI) between p_o and all latent signals $\mathbf{s}_l$ (i.e., the latent information). In this plot the MI between p and s is optimal in each point, so ∞ in the pure sensing regime and $\frac{1}{2}\log_2\left[1+\frac{\rho}{\rho\,(4^{I(p_o;\mathbf{s}_l)}-1)-1}\right]$ in the noisy sensing regime. **B.** The geometric mean fitness G as a percentage of $\ln[f_o]$ (the maximal obtainable fitness) as a function of the MI between p_o and $\mathbf{s}_l$. Again the MI between p and s is optimal in each point. The dashed gray line shows the fitness of the optimal bet-hedging strategy in this situation. The direction of evolution is towards higher fitness, note that this can only be achieved by sensing more signals, as the total amount of information that is contained in one signal about the optimum is a given. **C.** G as a percentage of $\lim_{I(p;s)\to\infty} G$, as a function of the MI between p an s. The curve saturates quickly, even when ρ is increased drastically. The latent information (mutual information of p_o and $\mathbf{s}_l$) is 0.02 bit. **D.** G as a percentage of $\lim_{I(p;s)\to\infty} G$, as a function of the MI between p an s. As the latent information increases from 0.02 (in C) to 0.06 bit an optimum appears in the line where $\rho = 30$, meaning that in this situation cells should not maximise the MI between the sensed signal s and their response p.

on the value of ρ.

Discussion

Our theory indicates that a signal should be transduced as accurately as possible only when it contains nearly all information about the optimal response. When information about the optimal response is partially contained in signals that are not sensed, it is evolutionarily optimal to perceive the sensed signals noisily. When the ability to sense one signal is lost, the remaining sensed signals should not be perceived more, but *less* accurately. This is because, as the total available information about the optimal behaviour decreases, the cell should trust the available information less, and gamble more.

Maximisation of the mutual information between the input and the output of a signalling circuit has received a lot of theoretical attention. One of its advantages is that it can be applied without the use of a mechanistic, stochastic model of the signalling network. In one application, only three functions are needed: the input/output relation, the dependency of the output noise on the input, and the probability distributions of input values, i.e., of environmental states (20). Maximisation of mutual information then allows for the prediction of one of those three functions from the two others. It is often rationalised by saying that it is a requirement for fitness maximisation. Our theory has shed doubt on this argument.

How precise the signalling machinery of a single microbe should track environmental signals depends on how important those signals are for its fitness. If latent signals exist that would improve fitness if they would be sensed, then the accuracy with which sensed signals are transduced should be low. Thus, our theory predicts that the optimal mutual information between the sensed signal and the cell's response should generally not be maximal for fitness maximisation when latent signals occur.

It is likely that latent signals occur. Firstly, because microorganisms display only a handful of signalling systems – in particular in the light of the huge number of nutrients that they can grow on. Secondly, expressing idle signalling systems reduces immediate fitness, because of biosynthetic resource consumption by non-growth promoting processes. Thirdly, idle, scanning signalling systems are evolutionary unstable; selection and drift would randomly mutate those unused systems. Natural selection for maximal geometric fitness in the presence of latent signals leads to optimal mutual information values that are not maximal.

Our approach does however have limitations, most of which are caused by our aim to create an analytically tractable model that has general applicability. We have used non-truncated normal distributions, such that distributions of compound

concentrations go to minus infinity. We assumed all variances to be independent of fluctuations in compound concentrations. Also, using normal distributions might by overly simplistic. Lastly, assumptions of linearity might not always be in agreement with experimental data. However, by keeping our model analytically tractable, we obtained general qualitative insights into the fitness effects of phenotypic adaptation mechanisms by considering a minimal model. When particular systems are of interest, models that incorporate mechanisms in more detail could provide additional, system-specific insights.

It has become clear in this work that the concept of mutual information should be used with care when it comes to quantifying cell performance. Firstly, biological performance should not be measured in terms of bits, but in terms of the fitness consequences of these bits. Secondly, it is difficult to distinguish characteristics of the input distribution from characteristics of the transduction system. This aspect of mutual information can lead to a certain number of bits being perceived as low, but for the distribution of inputs that is being considered it might well be that it is close to all information that is contained in the input. This could be resolved by normalizing with the entropy of the input distribution. Thirdly, we show that the commonly used mutual information measure, the mutual information between a signal and a cell's response, is not the relevant measure when considering its fitness consequences. In conclusion, this does not mean that mutual information cannot be used in cell biology; it only means that one has to look at what is relevant for fitness. We have shown that in cellular adaptation there is a very relevant mutual information term, which is the mutual information between all latent signals and the optimum, the latent information. Thus, it is not about how well the cell can perceive one aspect of the environment, but about how the whole environment determines the optimal behaviour. Using this perspective, it became clear that only signals that contain nearly all information about the optimal behaviour should be sensed as accurately as possible.

Our model implies that cells should improve their fitness by sensing as many informative signals as are available. In reality, we see that cells only sense a few signals. Apart from the possibility of flux-based regulation (e.g., catabolite repression (29, 30)), this is likely caused by the fact that to obtain specific information on individual signals, a cell needs separate signalling machinery for each of these signals. Whether or not a cell evolves sensing machinery for a particular signal depends on the fitness gain that can be achieved by sensing this individual signal. When this fitness gain is higher than the fitness cost of having the sensing machinery, the machinery should evolve. Even when the fitness cost of having sensing machinery is very low, sensing machinery for signals with low or infrequent fitness gain will be lost due to genetic drift. So the signals that are being sensed by microbes must have exceeded the 'fitness threshold', having sensing machinery for these signals is evolutionarily beneficial. The accuracy with which these signals should be sensed depends on the cumulative information in all signals that did not reach the fitness threshold, and are thus not being sensed, i.e., the latent information. This leads to the somewhat counter-intuitive conclusion that when the ability to sense one signal is lost, the other sensed signals should be perceived less accurately – not more.

This work contributes to a better understanding of optimal phenotypic adaptation strategies and of the use of information theoretic concepts. By considering multiple signals in the light of their fitness consequence we were able to show that not the mutual information between one signal and one response is what is crucial to cells, but that the latent information is what ultimately determines evolutionary success.

ACKNOWLEDGMENTS. We would like to thank Riccardo Muolo and Daan de Groot for their time and valuable comments during our weekly discussions, and Bram van de Putte, Tom Clement and Joanne Preuter for their support and insights over many cups of coffee.

1. J. Bull, Evolution of phenoptypic variance, Evolution 41 (1987) 303–315.
2. P. Haccou, Y. Iwasa, Optimal mixed strategies in stochastic environments, Theoretical Population Biology 47 (1995) 212–243.
3. L. Wolf, O. K. Silander, E. van Nimwegen, Expression noise facilitates the evolution of gene regulation, eLife 4 (2015) 1–48.
4. E. Kussell, S. Leibler, Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments, Science 309 (2005) 2075–2078.
5. B. E. Sæther, S. Engen, The concept of fitness in fluctuating environments, Trends in Ecology and Evolution 30 (2015) 273–281.
6. O. Rivoire, S. Leibler, The value of information for populations in varying environments, Journal of Statistical Physics 142 (2011) 1124–1166.
7. M. Acar, J. T. Mettetal, A. van Oudenaarden, Stochastic switching as a survival strategy in fluctuating environments, Nature genetics 40 (2008).
8. J. M. Pedraza, D. A. Garcia, M. F. Pérez-Ortiz, Noise, Information and Fitness in Changing Environments, Frontiers in Physics 6 (2018) 1–12.
9. J. Bigger, Treatment of staphylococcal infections with penicillin by intermittent sterilisation, The Lancet 244 (1944) 497–500.
10. K. Lewis, Persister cells, dormancy and infectious disease, Nature Reviews Microbiology 5 (2007) 48.
11. N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, S. Leibler, Bacterial persistence as a phenotypic switch, Science 305 (2004) 1622–1625.
12. A. Harms, E. Maisonneuve, K. Gerdes, Mechanisms of bacterial persistence during stress and antibiotic exposure, Science 354 (2016).
13. A. M. Stock, V. L. Robinson, P. N. Goudreau, Two-component signal transduction, Annual review of biochemistry 69 (2000) 183–215.
14. T. M. Norman, N. D. Lord, J. Paulsson, R. Losick, Stochastic Switching of Cell Fate in Microbes, Annual Review of Microbiology 69 (2015) 381–403.
15. C. G. Bowsher, P. S. Swain, Environmental sensing, information transfer, and cellular decision-making, Current Opinion in Biotechnology 28 (2014) 149–155.
16. R. Cheong, A. Rhee, C. J. Wang, I. Nemenman, A. Levchenko, Information transduction capacity of noisy biochemical signaling networks, Science (2011) 1204553.
17. M. Coppey, A. N. Boettiger, A. M. Berezhkovskii, S. Y. Shvartsman, Nuclear trapping shapes the terminal gradient in the drosophila embryo, Current Biology 18 (2008) 915–919.
18. S. Uda, T. H. Saito, T. Kudo, T. Kokaji, T. Tsuchiya, H. Kubota, Y. Komori, Y.-i. Ozaki, S. Kuroda, Robustness and compensation of information transmission of signaling pathways, Science 341 (2013) 558–561.
19. M. Voliotis, R. M. Perrett, C. McWilliams, C. A. McArdle, C. G. Bowsher, Information transfer by leaky, heterogeneous, protein kinase signaling systems, Proceedings of the National Academy of Sciences 111 (2014) E326–E333.
20. W. Bialek, Biophysics: searching for principles, Princeton University Press, Woodstock, Oxfordshire, 2012.
21. L. A. Rhee Alex, Cheong Raymond, The application of information theory to biochemical signaling systems, Physical Biology 9 (2013).
22. A. Levchenko, I. Nemenman, Cellular noise and information transmission, Current Opinion in Biotechnology 28 (2014) 156–164.
23. T. J. Perkins, P. S. Swain, Strategies for cellular decision-making, Molecular Systems Biology 5 (2009).
24. O. Feinerman, J. Veiga, J. R. Dorfman, R. N. Germain, G. Altan-Bonnet, Variability and robustness in t cell activation from regulated heterogeneity in protein levels, Science 321 (2008) 1081–1084.
25. N. Geva-Zatorsky, N. Rosenfeld, S. Itzkovitz, R. Milo, A. Sigal, E. Dekel, T. Yarnitzky, Y. Liron, P. Polak, G. Lahav, et al., Oscillations and variability in the p53 system, Molecular systems biology 2 (2006).
26. A. Sigal, R. Milo, A. Cohen, N. Geva-Zatorsky, Y. Klein, Y. Liron, N. Rosenfeld, T. Danon, N. Perzov, U. Alon, Variability and memory of protein levels in human cells, Nature 444 (2006) 643.
27. A. Eldar, M. B. Elowitz, Functional roles for noise in genetic circuits, Nature 467 (2010) 167.
28. J. Paulsson, Summing up the noise in gene networks, Nature 427 (2004) 415.
29. B. Görke, J. Stülke, Carbon catabolite repression in bacteria: many ways to make the most out of nutrients, Nature Reviews Microbiology 6 (2008) 613.
30. V. Chubukov, L. Gerosa, K. Kochanowski, U. Sauer, Coordination of microbial metabolism, Nature Reviews Microbiology 12 (2014) 327–340.

DRAFT