

1 First-principles prediction of the information processing 2 capacity of a simple genetic circuit

3 Manuel Razo-Mejia¹ and Rob Phillips^{1, 2, 3, *}

4 ¹*Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA*

5 ²*Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA*

6 ³*Department of Applied Physics, California Institute of Technology, Pasadena, CA 91125, USA*

7 ^{*}*Correspondence: phillips@pboc.caltech.edu*

8 Abstract

9
10 Given the stochastic nature of gene expression, genetically identical cells exposed to the same
11 environmental inputs will produce different outputs. This heterogeneity has consequences for how
12 cells are able to survive in changing environments. Recent work has explored the use of information
13 theory as a framework to understand the accuracy with which cells can ascertain the state of their
14 surroundings. Yet the predictive power of these approaches is limited and has not been rigorously
15 tested using precision measurements. To that end, we generate a minimal model for a simple genetic
16 circuit in which all parameter values for the model come from independently published data sets.
17 We then predict the information processing capacity of the genetic circuit for a suite of biophysical
18 parameters such as protein copy number and protein-DNA affinity. We compare these parameter-
19 free predictions with an experimental determination of the information processing capacity of *E.*
20 *coli* cells, and find that our minimal model accurately captures the experimental data.

21 As living organisms thrive in some given environment, they are faced with constant changes in their
22 surroundings. From abiotic conditions such as temperature fluctuations or changes in osmotic pressure,
23 to biological interactions such as cell-to-cell communication in a tissue or in a bacterial biofilm, living
24 organisms of all types sense and respond to external signals. Fig. 1(A) shows a schematic of this
25 process for a bacterial cell sensing a concentration of an extracellular chemical. At the molecular
26 level where signal transduction unfolds mechanistically, there are physical constraints on the accuracy
27 and precision of these responses given by intrinsic stochastic fluctuations [1]. This means that two
28 genetically identical cells exposed to the same stimulus will not have an identical response [2].

29 The implication of this biological noise is that cells do not have an infinite resolution to distinguish
30 signals and, as a consequence, there is a one-to-many mapping between inputs and outputs. Further-
31 more, given the limited number of possible outputs, there are overlapping responses between different
32 inputs. In that sense, one might think of cells performing a Bayesian inference of the state of the
33 environment given their phenotypic response, as schematized in Fig. 1(B). The question then becomes
34 how to analyze this probabilistic rather than deterministic relationship between inputs and outputs?
35 The abstract answer to this question was worked out in 1948 by Claude Shannon who, in his seminal
36 work, founded the field of information theory [3]. Shannon developed a general framework for how
37 to analyze information transmission through noisy communication channels. In his work, Shannon
38 showed that the only quantity that satisfies simple conditions of how a metric for information should
39 behave, was of the same functional form as the thermodynamic entropy – thereby christening his met-
40 ric the information entropy [4]. He also gave a definition, based on this information entropy, for the
41 relationship between inputs and outputs known as the mutual information. The mutual information
42 $I(p; c)$ between input c and output p , given by

$$I(p; c) = \sum_c P(c) \sum_p P(p | c) \log_2 \frac{P(p | c)}{P(p)}, \quad (1)$$

43 quantifies how much we learn about the state of the input c given that we get to observe the output
44 p .

45 It is natural to conceive of scenarios in which living organisms that can better resolve signals might
46 have an evolutionary benefit, making it more likely that their offspring will have a fitness advantage [5].
47 In recent years there has been a growing interest in understanding the theoretical limits on cellular
48 information processing [6, 7], and in quantifying how close evolution has pushed cellular signaling
49 pathways to these theoretical limits [8–10]. While these studies have treated the signaling pathway
50 as a “black box” explicitly ignoring all the molecular interactions taking place in them, other studies
51 have explored the role that molecular players and regulatory architectures have on these information
52 processing tasks [11–17]. Despite the great advances in our understanding of the information processing
53 capabilities of molecular mechanisms, the field still lacks a rigorous experimental test of these ideas
54 with precision measurements on a simple system tractable both theoretically and experimentally.

55 Over the last decade the dialogue between theory and experiments in gene regulation has led to
56 predictive power not only over the mean, but the noise in gene expression as a function of relevant
57 parameters such as regulatory protein copy numbers, affinity of these proteins to the DNA promoter,
58 as well as the extracellular concentrations of inducer molecules [18–21]. These models based on
59 equilibrium and non-equilibrium statistical physics have reached a predictive accuracy level such that
60 for simple cases it is now possible to design input-output functions [22, 23]. This opens the possibility
61 to exploit these predictive models to tackle the question of how much information genetic circuits
62 can process. The question lays at the heart of understanding the precision of the cellular response to
63 environmental signals. Fig. 1(C) schematizes a scenario in which two bacterial strains respond with
64 different levels of precision to three possible environmental states, i.e. inducer concentrations. The
65 overlap between the three different responses is what precisely determines the resolution with which
66 cells can distinguish different inputs. This is analogous to how for an imaging system the point spread
67 function limits the ability to resolve two light emitting point sources.

68 In this work we follow the same philosophy of theory-experiment dialogue used to determine model
69 parameters to predict from first principles the effect that biophysical parameters such as transcription
70 factor copy number and protein-DNA affinity have on the information processing capacity of a simple
71 genetic circuit. Specifically, to predict the mutual information between an extracellular chemical signal
72 (input c) and the corresponding cellular response in the form of protein expression (output p) (Eq. 1) we
73 must compute the input-output function $P(p | c)$. To do so, we use a master-equation-based model to
74 construct the protein copy number distribution as a function of an extracellular inducer concentration
75 for different combinations of transcription factor copy numbers and binding sites. Having these input-
76 output distributions allows us to compute the mutual information between inputs and outputs $I(c; p)$
77 for any arbitrary input distribution $P(c)$. We opt to compute the channel capacity, i.e. the maximum
78 information that can be processed by this gene regulatory architecture, defined as Eq. 1 maximized
79 over all possible input distributions $P(c)$. By doing so we can examine the physical limits of what cells
80 can do in terms of information processing by harboring these genetic circuits. All parameters used for
81 our model were inferred from a series of studies that span several experimental techniques [19, 24–26],
82 allowing us to perform parameter-free predictions of this information processing capacity [27].

83 These predictions are then contrasted with experimental data, where the channel capacity is in-
84 ferred from single-cell fluorescence distributions taken at different concentrations of inducer for cells
85 with previously characterized biophysical parameters [19, 26]. We find that our parameter-free pre-
86 dictions closely match the experiments. In this sense we demonstrate how our minimal model can

87 be used to quantify the resolution with which cells can resolve the environmental state with no free
 88 parameters.

89 The remainder of the paper is organized as follows. In Section 1.1 we define the minimal theoretical
 90 model and parameter inference for a simple repression genetic circuit. Section 1.2 discusses how
 91 all parameters for the minimal model are determined from published datasets that explore different
 92 aspects of the simple repression motif. Section 1.3 computes the moments of the mRNA and protein
 93 distributions from this minimal model. In Section 1.4 we explore the consequences of variability in
 94 gene copy number during the cell cycle. In this section we compare experimental and theoretical
 95 quantities related to the moments of the distribution. Specifically the predictions for the fold-change
 96 in gene expression (mean expression relative to an unregulated promoter) and the gene expression
 97 noise (standard deviation over mean). Section 1.5 follows with reconstruction of the full mRNA and
 98 protein distribution from the moments using the maximum entropy principle. Finally Section 1.6 uses
 99 the distributions from Section 1.5 to compute the maximum amount of information that the genetic
 100 circuit can process. Here we again contrast our zero-parameter fit predictions with experimental
 101 inferences of the channel capacity.

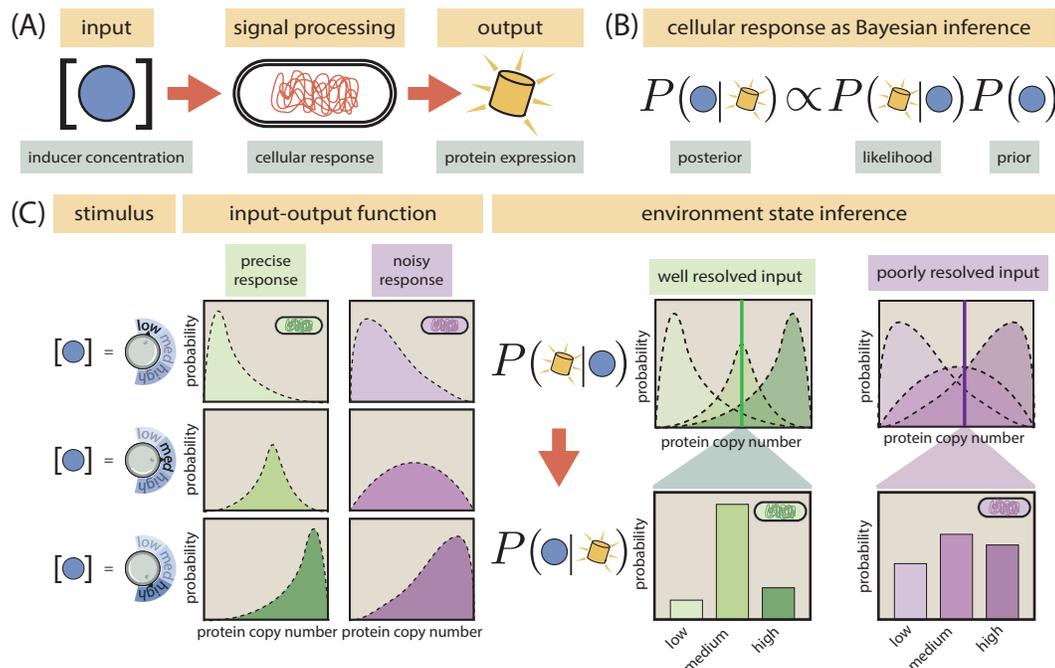


Figure 1. Cellular signaling systems sense the environment with different degrees of precision.

(A) Schematic representation of a cell as a noisy communication channel. From an environmental input (inducer molecule concentration) to a phenotypic output (protein expression level), cellular signaling systems can be modeled as noisy communication channels. (B) We treat cellular response to an external stimuli as a Bayesian inference of the state of the environment. As the phenotype (protein level) serves as the internal representation of the environmental state (inducer concentration), the probability of a cell being in a specific environment given this internal representation $P(c | p)$ is a function of the probability of the response given that environmental state $P(p | c)$. (C) The precision of the inference of the environmental state depends on how well can cells resolve different inputs. For three different levels of input (left panel) the green strain responds more precisely than the purple strain since the output distributions overlap less (middle panel). This allows the green strain to make a more precise inference of the environmental state given a phenotypic response (right panel).

1 Results

1.1 Minimal model of transcriptional regulation

We begin by defining the simple repression genetic circuit to be used throughout this work. As a tractable circuit for which we have control over the parameters both theoretically and experimentally we chose the so-called simple repression motif, a common regulatory scheme among prokaryotes [28]. This circuit consists of a single promoter with an RNA-polymerase (RNAP) binding site and a single binding site for a transcriptional repressor [19]. The regulation due to the repressor occurs via exclusion of the RNAP from its binding site when the repressor is bound, decreasing the likelihood of having a transcription event. As with many important macromolecules, we consider the repressor to be allosteric, meaning that it can exist in two conformations, one in which the repressor is able to bind to the specific binding site (active state) and one in which it cannot bind the specific binding site (inactive state). The environmental signaling occurs via passive import of an extracellular inducer that binds the repressor, shifting the equilibrium between the two conformations of the repressor [26]. In previous publications we have extensively characterized the mean response of this circuit under different conditions using equilibrium based models [27]. In this work we build upon these models to characterize the full distribution of gene expression with parameters such as repressor copy number and its affinity for the DNA being systematically varied.

Given the discrete nature of molecular species copy numbers inside cells, chemical master equations have emerged as a useful tool to model the inherent probability distribution of these counts [29]. In Fig. 2(A) we show the minimal model and the necessary set of parameters needed to predict mRNA and protein distributions. Specifically, we assume a three-state model where the promoter can be found 1) In a transcriptionally active state (A state), 2) in a transcriptionally inactive state without the repressor bound (I state) and 3) with the repressor bound (R state). We do not assume that the transition between the active state A and the inactive state I happens due to RNAP binding to the promoter. The transcriptional initiation kinetics involve several more steps than simple binding [30]. We coarse-grain all these steps into an effective “on” and “off” states for the promoter consistent with experiments demonstrating the bursty nature of gene expression in *E. coli* [18]. These three states generate a system of coupled differential equations for each of the three state distributions $P_A(m, p; t)$, $P_I(m, p; t)$ and $P_R(m, p; t)$, where m and p are the mRNA and protein count per cell, respectively and t is the time. Given the rates shown in Fig. 2(A) we define the system of ODEs for a specific m and p . For the transcriptionally active state we have

$$\begin{aligned} \frac{dP_A(m, p)}{dt} = & - \overbrace{k_{\text{off}}^{(p)} P_A(m, p)}^{A \rightarrow I} + \overbrace{k_{\text{on}}^{(p)} P_I(m, p)}^{I \rightarrow A} \\ & + \overbrace{r_m P_A(m-1, p)}^{m-1 \rightarrow m} - \overbrace{r_m P_A(m, p)}^{m \rightarrow m+1} + \overbrace{\gamma_m (m+1) P_A(m+1, p)}^{m+1 \rightarrow m} - \overbrace{\gamma_m m P_A(m, p)}^{m \rightarrow m-1} \\ & + \overbrace{r_p m P_A(m, p-1)}^{p-1 \rightarrow p} - \overbrace{r_p m P_A(m, p)}^{p \rightarrow p+1} + \overbrace{\gamma_p (p+1) P_A(m, p+1)}^{p+1 \rightarrow p} - \overbrace{\gamma_p p P_A(m, p)}^{p \rightarrow p-1}. \end{aligned} \quad (2)$$

133 For the transcriptionally inactive state I we have

$$\begin{aligned}
 \frac{dP_I(m, p)}{dt} = & \overbrace{k_{\text{off}}^{(p)} P_A(m, p)}^{A \rightarrow I} - \overbrace{k_{\text{on}}^{(p)} P_I(m, p)}^{I \rightarrow A} + \overbrace{k_{\text{off}}^{(r)} P_R(m, p)}^{R \rightarrow I} - \overbrace{k_{\text{on}}^{(r)} P_I(m, p)}^{I \rightarrow R} \\
 & + \overbrace{\gamma_m(m+1) P_I(m+1, p)}^{m+1 \rightarrow m} - \overbrace{\gamma_m m P_I(m, p)}^{m \rightarrow m-1} \\
 & + \overbrace{r_p m P_I(m, p-1)}^{p-1 \rightarrow p} - \overbrace{r_p m P_I(m, p)}^{p \rightarrow p+1} + \overbrace{\gamma_p(p+1) P_I(m, p+1)}^{p+1 \rightarrow p} - \overbrace{\gamma_p p P_I(m, p)}^{p \rightarrow p-1}.
 \end{aligned} \tag{3}$$

134 And finally, for the repressor bound state R we have

$$\begin{aligned}
 \frac{dP_R(m, p)}{dt} = & - \overbrace{k_{\text{off}}^{(r)} P_R(m, p)}^{R \rightarrow I} + \overbrace{k_{\text{on}}^{(r)} P_I(m, p)}^{I \rightarrow R} \\
 & + \overbrace{\gamma_m(m+1) P_R(m+1, p)}^{m+1 \rightarrow m} - \overbrace{\gamma_m m P_R(m, p)}^{m \rightarrow m-1} \\
 & + \overbrace{r_p m P_R(m, p-1)}^{p-1 \rightarrow p} - \overbrace{r_p m P_R(m, p)}^{p \rightarrow p+1} + \overbrace{\gamma_p(p+1) P_R(m, p+1)}^{p+1 \rightarrow p} - \overbrace{\gamma_p p P_R(m, p)}^{p \rightarrow p-1}.
 \end{aligned} \tag{4}$$

135 As we will discuss later in Section 1.4 the protein degradation term γ_p is set to zero since we do
 136 not consider protein degradation as a Poisson process, but rather we explicitly implement binomial
 137 partitioning as the cells grow and divide.

138 It is convenient to rewrite these equations in a compact matrix notation [29]. For this we define
 139 the vector $\mathbf{P}(m, p)$ as

$$\mathbf{P}(m, p) = (P_A(m, p), P_I(m, p), P_R(m, p))^T, \tag{5}$$

140 where T is the transpose. By defining the matrices \mathbf{K} to contain the promoter state transitions, \mathbf{R}_m
 141 and $\mathbf{\Gamma}_m$ to contain the mRNA production and degradation terms, respectively, and \mathbf{R}_p and $\mathbf{\Gamma}_p$
 142 to contain the protein production and degradation terms, respectively, the system of ODEs can then be
 143 written as (See Appendix S1 for full definition of these matrices)

$$\begin{aligned}
 \frac{d\mathbf{P}(m, p)}{dt} = & (\mathbf{K} - \mathbf{R}_m - m\mathbf{\Gamma}_m - m\mathbf{R}_p - p\mathbf{\Gamma}_p) \mathbf{P}(m, p) \\
 & + \mathbf{R}_m \mathbf{P}(m-1, p) + (m+1)\mathbf{\Gamma}_m \mathbf{P}(m+1, p) \\
 & + m\mathbf{R}_p \mathbf{P}(m, p-1) + (p+1)\mathbf{\Gamma}_p \mathbf{P}(m, p+1).
 \end{aligned} \tag{6}$$

144 1.2 Inferring parameters from published data sets

145 A decade of research in our group has characterized the simple repression motif with an ever
 146 expanding array of predictions and corresponding experiments to uncover the physics of this genetic
 147 circuit [27]. In doing so we have come to understand the mean response of a single promoter in the
 148 presence of varying levels of repressor copy numbers and repressor-DNA affinities [19], due to the effect
 149 that competing binding sites and multiple promoter copies impose [25], and in recent work, assisted by

150 the Monod-Wyman-Changeux (MWC) model, we expanded the scope to the allosteric nature of the
151 repressor [26]. All of these studies have exploited the simplicity and predictive power of equilibrium
152 approximations to these non-equilibrium systems [31]. We have also used a similar kinetic model to
153 the one depicted in Fig. 2(A) to study the noise in mRNA copy number [24]. As a test case of the
154 depth of our theoretical understanding of the so-called “hydrogen atom” of transcriptional regulation
155 we combine all of the studies mentioned above to inform the parameter values of the model presented
156 in Fig. 2(A). Fig. 2(B) schematizes the data sets and experimental techniques used to measure gene
157 expression along with the parameters that can be inferred from them.

158 Appendix S2 expands on the details of how the inference was performed for each of the parameters.
159 Briefly the promoter activation and inactivation rates $k_{\text{on}}^{(p)}$ and $k_{\text{off}}^{(p)}$, as well as the transcription rate
160 r_m were obtained in units of the mRNA degradation rate γ_m by fitting a two-state promoter model
161 (no state R from Fig. 2(A)) [32] to mRNA FISH data of an unregulated promoter (no repressor
162 present in the cell) [24]. The repressor on rate is assumed to be of the form $k_{\text{on}}^{(r)} = k_o[R]$ where k_o
163 is a diffusion-limited on rate and $[R]$ is the concentration of active repressor in the cell [24]. This
164 concentration of active repressor is at the same time determined by the mean repressor copy number
165 in the cell, and the fraction of repressors in the active state. Existing estimates of the transition rates
166 between conformations of allosteric molecules set them at the microsecond scale [33]. By considering
167 this to be representative for our repressor of interest, the separation of time-scales between the rapid
168 conformational changes of the repressor and the slower downstream processes such as the open-complex
169 formation processes allow us to model the probability of the repressor being in the active state as an
170 equilibrium MWC process. The parameters of the MWC model K_A , K_I and $\Delta\epsilon_{AI}$ were previously
171 characterized from video-microscopy and flow-cytometry data [26]. For the repressor off rate $k_{\text{off}}^{(r)}$ we
172 take advantage of the fact that the mean mRNA copy number as derived from the model in Fig. 2(A)
173 cast in the language of rates is of the same functional form as the equilibrium model cast in the
174 language of binding energies [34]. Therefore the value of the repressor-DNA binding energy $\Delta\epsilon_r$
175 constrains the value of the repressor off rate $k_{\text{off}}^{(r)}$. These constraints on the rates allow us to make
176 self-consistent predictions under both, the equilibrium and the kinetic framework.

177 1.3 Computing the moments of the mRNA and protein distributions

178 Solving chemical master equations represent a challenge that is still an active area of research.
179 An alternative approach is to find schemes to approximate the distribution. One such scheme, the
180 maximum entropy principle, makes use of the moments of the distribution to approximate the full
181 distribution. In this section we will demonstrate an iterative algorithm to compute the mRNA and
182 protein distribution moments.

183 Our simple repression kinetic model depicted in Fig. 2(A) consists of an infinite system of ODEs
184 for each possible pair m, p . To compute any moment of the distribution we define a vector

$$\langle \mathbf{m}^x \mathbf{p}^y \rangle \equiv (\langle m^x p^y \rangle_A, \langle m^x p^y \rangle_I, \langle m^x p^y \rangle_R)^T, \quad (7)$$

185 where $\langle m^x p^y \rangle_S$ is the expected value of $m^x p^y$ in state $S \in \{A, I, R\}$ for $x, y \in \mathbb{N}$. In other words, just
186 as we defined the vector $\mathbf{P}(m, p)$, here we define a vector to collect the expected value of each of the
187 promoter states. By definition any of these moments $\langle m^x p^y \rangle_S$ are computed as

$$\langle m^x p^y \rangle_S \equiv \sum_{m=0}^{\infty} \sum_{p=0}^{\infty} m^x p^y P_S(m, p). \quad (8)$$

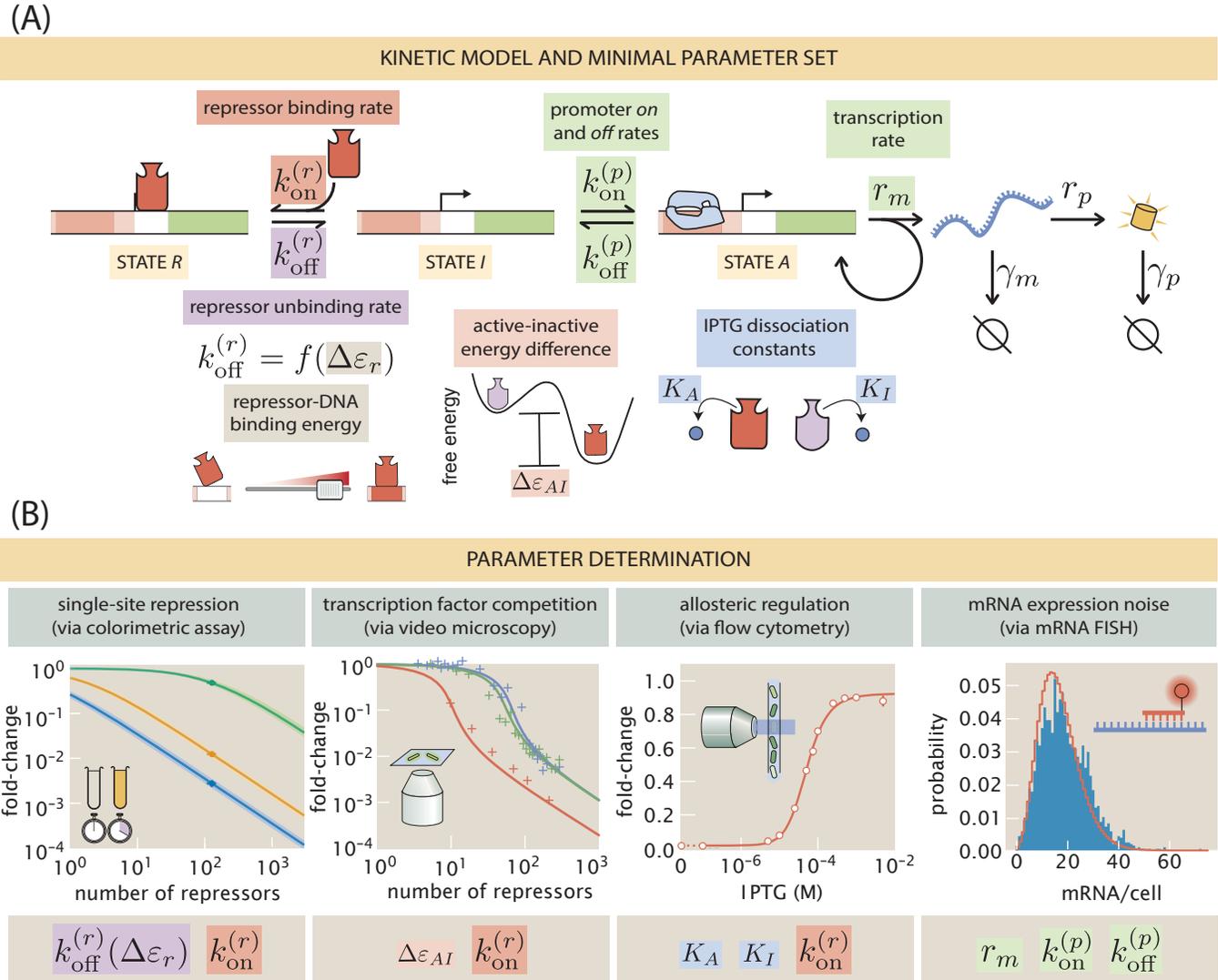


Figure 2. Minimal kinetic model of transcriptional regulation for a simple repression architecture.

(A) Three-state promoter stochastic model of transcriptional regulation by a repressor. The regulation by the repressor occurs via exclusion of the transcription initiation machinery, not allowing the promoter to transition to the transcriptionally active state. All parameters highlighted with colored boxes were determined from published datasets based on the same genetic circuit. (B) Data sets used to infer the parameter values. From left to right Garcia & Phillips [19] is used to determine $k_{\text{off}}^{(r)}$ and $k_{\text{on}}^{(r)}$, Brewster et al. [25] is used to determine $\Delta\epsilon_{AI}$ and $k_{\text{on}}^{(r)}$, Razo-Mejia et al. [26] is used to determine K_A , K_I , and $k_{\text{on}}^{(r)}$ and Jones et al. is used to determine r_m , $k_{\text{on}}^{(p)}$, and $k_{\text{off}}^{(p)}$.

188

Summing over all possible m and p values in Eq. 6 results in a ODE for any moment of the

189 distribution of the form (See Appendix S3 for full derivation)

$$\begin{aligned} \frac{d\langle \mathbf{m}^x \mathbf{p}^y \rangle}{dt} &= \mathbf{K} \langle \mathbf{m}^x \mathbf{p}^y \rangle \\ &+ \mathbf{R}_m \langle \mathbf{p}^y [(\mathbf{m} + \mathbf{1})^x - \mathbf{m}^x] \rangle + \mathbf{\Gamma}_m \langle \mathbf{m} \mathbf{p}^y [(\mathbf{m} - \mathbf{1})^x - \mathbf{m}^x] \rangle \\ &+ \mathbf{R}_p \langle \mathbf{m}^{(x+1)} [(\mathbf{p} + \mathbf{1})^y - \mathbf{p}^y] \rangle + \mathbf{\Gamma}_p \langle \mathbf{m}^x \mathbf{p} [(\mathbf{p} - \mathbf{1})^y - \mathbf{p}^y] \rangle. \end{aligned} \quad (9)$$

190 Given that all transitions in our stochastic model are first order reactions, Eq. 9 has no moment-
 191 closure problem [13]. What this means is that the dynamical equation for a given moment only depends
 192 on lower moments (See Appendix S3 for full proof). This feature of our model implies, for example,
 193 that the second moment of the protein distribution $\langle p^2 \rangle$ depends only on the first two moments of the
 194 mRNA distribution $\langle m \rangle$, and $\langle m^2 \rangle$, the first protein moment $\langle p \rangle$ and the cross-correlation term $\langle mp \rangle$.
 195 We can therefore define $\boldsymbol{\mu}^{(x,y)}$ to be a vector containing all moments up to $\langle \mathbf{m}^x \mathbf{p}^y \rangle$ for all promoter
 196 states. This is

$$\boldsymbol{\mu}^{(x,y)} = [\langle \mathbf{m}^0 \mathbf{p}^0 \rangle, \langle \mathbf{m}^1 \mathbf{p}^0 \rangle, \dots, \langle \mathbf{m}^x \mathbf{p}^y \rangle]^T. \quad (10)$$

197 Explicitly for the three-state promoter model depicted in Fig. 2(A) this vector takes the form

$$\boldsymbol{\mu}^{(x,y)} = [\langle m^0 p^0 \rangle_A, \langle m^0 p^0 \rangle_I, \langle m^0 p^0 \rangle_R, \dots, \langle m^x p^y \rangle_A, \langle m^x p^y \rangle_I, \langle m^x p^y \rangle_R]^T. \quad (11)$$

198 Given this definition we can compute the general moment dynamics as

$$\frac{d\boldsymbol{\mu}^{(x,y)}}{dt} = \mathbf{A} \boldsymbol{\mu}^{(x,y)}, \quad (12)$$

199 where \mathbf{A} is a square matrix that contains all the numeric coefficients that relate each of the moments.
 200 We can then use Eq. 9 to build matrix \mathbf{A} by iteratively substituting values for the exponents x and y
 201 up to a specified value. In the next section, we will use Eq. 12 to numerically integrate the dynamical
 202 equations for our moments of interest as cells progress through the cell cycle.

203 1.4 Accounting for cell-cycle dependent variability in gene dosage

204 As cells progress through the cell cycle, the genome has to be replicated to guarantee that each
 205 daughter cell receives a copy of the genetic material. This replication of the genome implies that
 206 cells spend part of the cell cycle with multiple copies of each gene depending on the cellular growth
 207 rate and the relative position of the gene with respect to the replication origin [35]. Genes closer to
 208 the replication origin spend a larger fraction of the cell cycle with multiple copies compared to genes
 209 closer to the replication termination site [35]. Fig. 3(A) depicts a schematic of this process where
 210 the replication origin (*oriC*) and the relevant locus for our experimental measurements (*galK*) are
 211 highlighted.

212 Since this change in gene copy number has been shown to have an effect on cell-to-cell variability in
 213 gene expression [24, 36], we now extend our minimal model to account for these changes in gene copy
 214 number during the cell cycle. We reason that the only difference between the single-copy state and the
 215 two-copies states of the promoter is a doubling of the mRNA production rate r_m . In particular the
 216 promoter activation and inactivation rates $k_{\text{on}}^{(p)}$ and $k_{\text{off}}^{(p)}$ and the mRNA production rate r_m inferred in

217 Section 1.1 assume that cells spend a fraction f of the cell cycle with one copy of the promoter (mRNA
218 production rate r_m) and a fraction $(1 - f)$ of the cell cycle with two copies of the promoter (mRNA
219 production rate $2r_m$). This inference was performed considering that at each cell state the mRNA
220 level immediately reaches the steady state value for the corresponding mRNA production rate. This
221 assumption is justified since the timescale to reach this steady state depends only on the degradation
222 rate γ_m , which for the mRNA is much shorter (≈ 3 min) than the length of the cell cycle (100 min
223 for our experimental conditions) [37]. Appendix S2 shows that a model accounting for this gene copy
224 number variability is able to capture the experimental data from single molecule mRNA counts of an
225 unregulated (constitutively expressed) promoter.

226 Given that the protein degradation rate γ_p in our model is set by the cell division time, we do
227 not expect that the protein count will reach the corresponding steady state value for each stage in
228 the cell cycle. In other words, cells do not spend long enough with two copies of the promoter for the
229 protein level to reach the steady state value corresponding to a transcription rate of $2r_m$. We therefore
230 use the dynamical equations developed in Section 1.3 to numerically integrate the time trajectory of
231 the moments of the distribution with the corresponding parameters for each phase of the cell cycle.
232 Fig. 3(B) shows an example corresponding to the mean mRNA level (upper panel) and the mean
233 protein level (lower panel) for the case of the unregulated promoter. Given that we inferred the
234 promoter rates parameters considering that mRNA reaches steady state at each stage, we see that the
235 numerical integration of the equations is consistent with the assumption of having the mRNA reach
236 a stable value at each stage (See Fig. 3(B) upper panel). On the other hand, the mean protein level
237 does not reach a steady state at either of the cellular stages. Nevertheless it is interesting to observe
238 that after a couple of cell cycles the trajectory from cycle to cycle follows a repetitive pattern (See
239 Fig. 3(B) lower panel). Previously we have experimentally observe this repetitive pattern by tracking
240 the expression level over time with video microscopy as shown in Fig. 18 of [27].

241 To test the effects of including this gene copy number variability in our model we now compare
242 the predictions of the model with experimental data. Specifically as detailed in Methods we obtained
243 single-cell fluorescence values of different *E. coli* strains under twelve different inducer concentrations.
244 The strains imaged spanned three orders of magnitude in repressor copy number and three distinct
245 repressor-DNA affinities. Since growth was asynchronous, we reason that cells were randomly sampled
246 at all stages of the cell cycle. Therefore when computing statistics from the data such as the mean
247 fluorescence value, in reality we are averaging over the cell cycle. In other words, as depicted in
248 Fig. 3(B) quantities such as the mean protein copy number change over time, i.e. $\langle p \rangle \equiv \langle p(t) \rangle$. This
249 means that computing the mean of a population of unsynchronized cells is equivalent to averaging
250 this time dependent mean protein copy number over the span of the cell cycle. Mathematically this
251 is expressed as

$$\langle p \rangle_c = \int_{t_o}^{t_d} \langle p(t) \rangle P(t) dt, \quad (13)$$

252 where $\langle p \rangle_c$ represents the average protein copy number over a cell cycle, t_o represents the start of the
253 cell cycle, t_d represents the time of cell division, and $P(t)$ represents the probability of any cell being
254 at time $t \in [t_o, t_d]$ of their cell cycle. We do not consider cells uniformly distributed along the cell
255 cycle since it is known that cells follow an exponential distribution, having more younger than older
256 cells at any time point [38]. All computations hereafter are therefore done by applying an averaging
257 like the one in Eq. 13 for the span of a cell cycle. We remind the reader that these time averages are
258 done under a fixed environmental state. It is the trajectory of cells over cell cycles under a constant

259 environment what we need to account for.

260 Fig. 3(C) compares zero-parameter fit predictions (lines) with experimentally determined quanti-
261 ties (points). The upper row shows the non-dimensional quantity known as the fold-change in gene
262 expression [19]. This fold-change is defined as the relative mean gene expression level with respect to
263 an unregulated promoter. For protein this is

$$\text{fold-change} = \frac{\langle p(R \neq 0) \rangle_c}{\langle p(R = 0) \rangle_c}, \quad (14)$$

264 where $\langle p(R \neq 0) \rangle_c$ represents the mean protein count for cells with non-zero repressor copy number
265 count R over the entire cell cycle, and $\langle p(R = 0) \rangle_c$ represents the equivalent for a strain with no
266 repressors present. The experimental points were determined from the fluorescent intensities of cells
267 with varying repressor copy number and a $\Delta lacI$ strain with no repressor gene present (See Methods
268 for further details). The fold-change in gene expression has previously served as a metric to test
269 the validity of equilibrium-based models [34]. We note that the curves shown in the upper panel of
270 Fig. 3(C) are consistent with the predictions from equilibrium models [26] despite being generated
271 from a clearly non-equilibrium process as shown in Fig. 3(B). The kinetic model from Fig. 2(A) goes
272 beyond the equilibrium picture to generate predictions for moments of the distribution other than the
273 mean mRNA or mean protein count. To test this extended predictive power the lower row of Fig. 3(C)
274 shows the noise in gene expression defined as the standard deviation over the mean protein count.
275 The good correspondence between the zero-parameter fit theoretical predictions and the experimental
276 data is only achieved when considering the gene copy number variability introduced in this section.
277 (See Appendix S4 for comparison when this variability is not included).

278 1.5 Maximum Entropy approximation

279 Having numerically computed the moments of the mRNA and protein distributions as cells progress
280 through the cell cycle we now proceed to make an approximating reconstruction of the full distributions
281 given this limited information. As hinted in Section 1.3 the maximum entropy principle, first proposed
282 by E.T. Jaynes in 1957, approximates the entire distribution by maximizing the Shannon entropy
283 subject to constraints given by the values of the moments of the distribution, among other quantities
284 [39]. This procedure leads to a probability distribution P_H of the form (See Appendix S5 for full
285 derivation)

$$P_H(m, p) = \frac{1}{\mathcal{Z}} \exp \left(- \sum_{(x,y)} \lambda_{(x,y)} m^x p^y \right), \quad (15)$$

286 where $\lambda_{(x,y)}$ is the Lagrange multiplier associated with the constraint set by the moment $\langle m^x p^y \rangle$, and
287 \mathcal{Z} is a normalization constant. The more moments $\langle m^x p^y \rangle$ included as constraints, the more accurate
288 the approximation resulting from Eq. 15 becomes.

289 The computational challenge then becomes a minimization routine in which the values for the
290 Lagrange multipliers $\lambda_{(x,y)}$ that are consistent with the constraints set by the moments values $\langle m^x p^y \rangle$
291 need to be found. Appendix S5 details our implementation of a robust algorithm to find such values.
292 Fig. 4 shows example predicted protein distributions reconstructed using the first six moments of
293 the protein distribution for a suite of different biophysical parameters and environmental inducer
294 concentrations. As repressor-DNA binding affinity (columns in Fig. 4) and repressor copy number
295 (rows in Fig. 4) are varied, the responses to different signals (i.e. inducer concentrations) overlap to

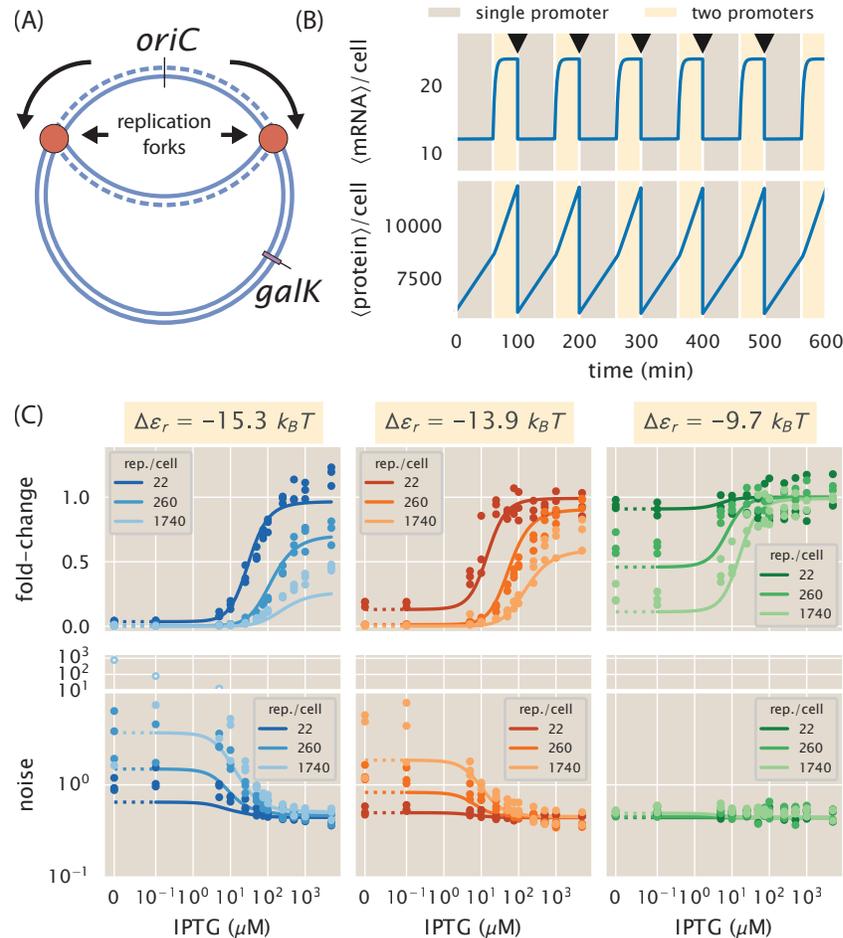


Figure 3. Accounting for gene copy number variability during the cell cycle. (A) Schematic of a replicating bacterial genome. As cells progress through the cell cycle the genome is replicated, duplicating gene copies for a fraction of the cell cycle. *oriC* indicates the replication origin, and *galK* indicates the locus at which the reporter construct was integrated. (B) mean mRNA (upper panel) and mean protein (lower panel) dynamics. Cells spend a fraction of the cell cycle with a single copy of the promoter (light brown) and the rest of the cell cycle with two copies (light yellow). Black arrows indicate time of cell division. (C) Zero parameter-fit predictions (lines) and experimental data (circles) of the gene expression fold-change (upper row) and noise (lower row) for repressor binding sites with different affinities (different columns) and different repressor copy numbers per cell (different lines on each panel). Dotted lines indicate linear scale while solid lines indicate logarithmic scale. White dots on the lower row are plotted on a different scale for visual clarity.

296 varying degrees. For example the upper right corner frame with a weak binding site ($\Delta\epsilon_r = -9.7 k_B T$)
 297 and a low repressor copy number (22 repressors per cell) has virtually identical distributions regardless
 298 of the input inducer concentration. This means that cells with this set of parameters cannot resolve
 299 any difference in the concentration of the signal. As the number of repressors is increased, the degree
 300 of overlap between distributions decreases, allowing cells to better resolve the value of the signal input.
 301 On the opposite extreme the lower left panel shows a strong binding site ($\Delta\epsilon_r = -15.3 k_B T$) and
 302 a high repressor copy number (1740 repressors per cell). This parameter combination shows overlap
 303 between distributions since the high degree of repression skews all distributions towards lower copy
 304 numbers, giving again little ability for the cells to resolve the inputs. In Appendix S5 we show the

305 comparison of these predicted distributions with the experimental single-cell fluorescence distributions.
 306 In the following section we formalize the notion of how well cells can resolve different inputs from an
 307 information theoretic perspective via the channel capacity.

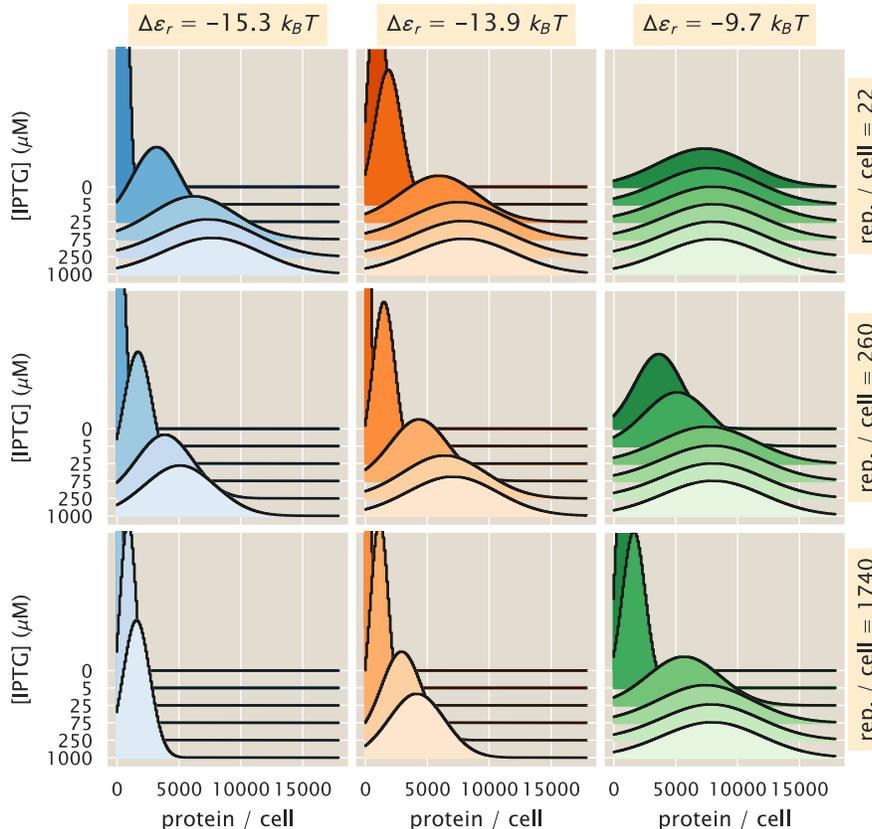


Figure 4. Maximum entropy protein distributions for varying physical parameters. Predicted protein distributions under different inducer (IPTG) concentrations for different combinations of repressor-DNA affinities (columns) and repressor copy numbers (rows). The first six moments of the protein distribution used to constrain the maximum entropy approximation were computed by integrating Eq. 9 as cells progressed through the cell cycle as described in Section 1.4.

308 1.6 Theoretical prediction of the channel capacity

309 As a useful measure of the ability of the genetic circuit to allow the cell to infer the environmental
 310 state, i.e. the inducer concentration, we turn to the channel capacity. The channel capacity is defined
 311 as the mutual information between input and output, maximized over all possible input distributions.
 312 Putting this into mathematical terms we define c as the inducer concentration. $P(c)$ represents
 313 the distribution of inducer and $P(p | c)$ the distribution of protein counts given a fixed inducer
 314 concentration - effectively the distributions shown in Fig. 4. The channel capacity is then given by

$$C \equiv \max_{P(c)} I(p; c), \quad (16)$$

315 where $I(p; c)$, the mutual information between protein count and inducer concentration is given by
 316 Eq. 1.

317 If used as a metric of how reliably a signaling system can infer the state of the external signal,
318 the channel capacity, when measured in bits, is commonly interpreted as the logarithm of the number
319 of states that the signaling system can properly resolve. For example, a signaling system with a
320 channel capacity of C bits is interpreted as being able to resolve 2^C states, though channel capacities
321 with fractional values are allowed. As a result, we prefer the Bayesian interpretation that the mutual
322 information, and as a consequence the channel capacity, quantifies the improvement in the inference
323 of the input when considering the output compared to just using the prior distribution of the input
324 by itself for prediction [13, 40]. Under this interpretation a channel capacity of a fractional bit still
325 quantifies an improvement of the ability of the signaling system to infer the value of the extracellular
326 signal compared to having no sensing system at all.

327 Computing the channel capacity as defined in Eq. 16 implies optimizing over an infinite space of
328 possible distributions $P(c)$. For special cases in which the noise is small compared to the dynamic
329 range, approximate analytical equations have been derived [16]. But given the high cell-to-cell variabil-
330 ity that our model predicts, the conditions of the so-called small noise approximation are not satisfied.
331 We therefore appeal to a numerical solution known as the Blahut-Arimoto algorithm [41]. This algo-
332 rithm, starting on any (discrete) distribution $P(c)$, converges to the distribution at channel capacity.
333 Fig. 5(A) shows zero-parameter fit predictions of the channel capacity as a function of the number
334 of repressors for different repressor-DNA affinities (solid lines). These predictions are contrasted with
335 experimental determinations of the channel capacity as inferred from single-cell fluorescence intensity
336 distributions taken over 12 different concentrations of inducer. Briefly, from single-cell fluorescent
337 measurements we can approximate the input-output distribution $P(p | c)$. Once these conditional
338 distributions are fixed, the task of finding the input distribution at channel capacity become a compu-
339 tational minimization routine that can be undertaken using conjugate gradient or similar algorithms.
340 For the particular case of the channel capacity on a system with a discrete number of inputs and
341 outputs the Blahut-Arimoto algorithm is built in such a way that it guarantees the convergence to-
342 wards the optimal input distribution (See Appendix S6 for further details). Fig. 5(B) shows example
343 input-output functions for different values of the channel capacity. This illustrates that having access
344 to no information (zero channel capacity) is a consequence of having overlapping input-output func-
345 tions (lower panel). On the other hand, the more separated the input-output distributions are (upper
346 panel) the higher the channel capacity can be.

347 Fig. 5(A) has interesting features that are worth highlighting. On one extreme for cells with no
348 transcription factors there is no information processing potential as this simple genetic circuit would
349 be constitutively expressed regardless of the environmental state. As cells increase the transcription
350 factor copy number, the channel capacity increases until it reaches a maximum to then fall back down
351 at high repressor copy number since the promoter would be permanently repressed. The steepness
352 of the increment in channel capacity as well as the height of the maximum expression highly depend
353 on the repressor-DNA affinity. For strong binding sites (blue curve in Fig. 5(A)) there is a rapid
354 increment in the channel capacity, but the maximum value reached is smaller compared to a weaker
355 binding site (orange curve in Fig. 5(A)).

356 Discussion

357 Building on Shannon's formulation of information theory, there have been significant efforts using
358 this theoretical framework to understand the information processing capabilities of biological systems,
359 and the evolutionary consequences for organisms harboring signal transduction systems [1, 5, 8, 42–

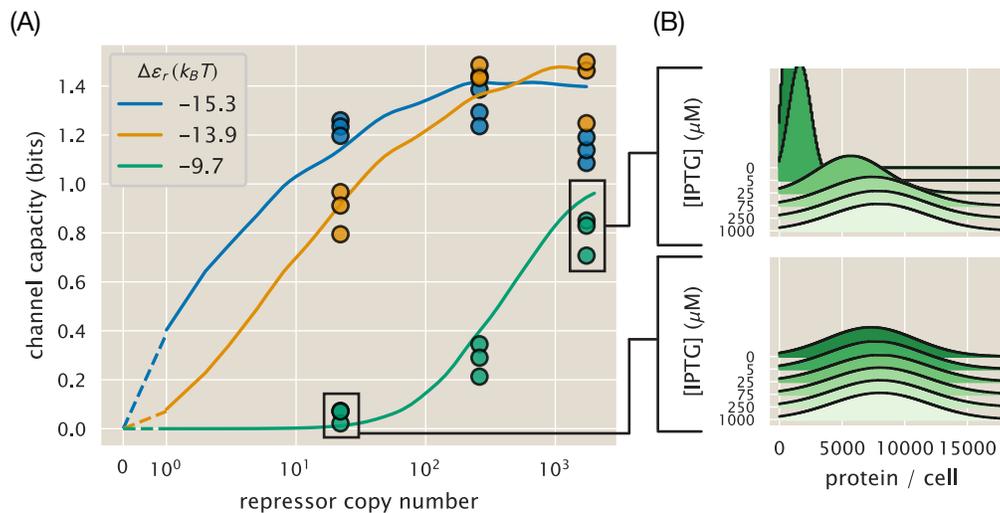


Figure 5. Comparison of theoretical and experimental channel capacity. (A) Channel capacity as inferred using the Blahut-Arimoto algorithm [41] for varying number of repressors and repressor-DNA affinities. All inferences were performed using 12 IPTG concentrations as detailed in the Methods. Lines represent zero-parameter fit predictions done with the maximum entropy distributions as those shown in Fig. 4. Points represent inferences made from single cell fluorescence distributions (See Appendix S6 for further details). Solid lines indicate plot in logarithmic scale, while dashed line indicates linear scale. (B) Example input-output functions of opposite limits of channel capacity. Lower panel illustrates that zero channel capacity indicates that all distributions overlap. Upper panel illustrates that as the channel capacity increases, the separation between distributions increases as well.

360 44]. Recently, with the mechanistic dissection of molecular signaling pathways significant progress has
 361 been made on the question of the physical limits of cellular detection and the role that features such
 362 as feedback loops play in this task [6, 13, 15, 45, 46]. But the field still lacks a rigorous experimental
 363 test of these ideas with precision measurements on a system that is tractable both experimentally and
 364 theoretically.

365 In this paper we take advantage of the recent progress on the quantitative modeling of input-output
 366 functions of genetic circuits to build a minimal model of the so-called simple repression motif [27]. By
 367 combining a series of studies on this circuit spanning diverse experimental methods for measuring gene
 368 expression under a myriad of different conditions, we infer all parameter values of our model - allowing
 369 us to generate parameter-free predictions for processes related to information processing. Some of the
 370 model parameters for our kinetic formulation of the input-output function are informed by inferences
 371 made from equilibrium models. We use the fact that if both, kinetic and thermodynamic languages
 372 describe the same system, the predictions must be self-consistent. In other words, if the equilibrium
 373 model can only make statements about the mean mRNA and mean protein copy number because
 374 of the way these models are constructed, those predictions must be equivalent to what the kinetic
 375 model has to say about these same quantities. This condition therefore constrains the values that the
 376 kinetic rates in the model can take. To test whether or not the equilibrium picture can reproduce
 377 the predictions made by the kinetic model we compare the experimental and theoretical fold-change
 378 in protein copy number for a suite of biophysical parameters and environmental conditions. The
 379 agreement between theory and experiment demonstrates that these two frameworks can indeed make
 380 consistent predictions.

381 The kinetic treatment of the system brings with it increasing predictive power compared to the
382 equilibrium picture. Under the kinetic formulation, the predictions are not limited only to the mean
383 but to any moment of the mRNA and protein distribution. We first test these novel predictions by
384 comparing the noise in protein copy number (standard deviation / mean) with experimental data.
385 Since the model is able to accurately predict the noise in protein count we extended our analysis to
386 infer entire protein distributions at different input signal concentrations by using the maximum entropy
387 principle. What this means is that we compute moments of the protein distribution, and then use
388 these moments to build an approximation to the full distribution. These predicted distributions are
389 then compared with experimental single-cell distributions as shown in Appendix S5. The agreement
390 between our predictions and the experimental data at the full protein distribution means that we can
391 use our model to predict the information processing capacity of the genetic circuit.

392 By maximizing the mutual information between input signal concentration and output protein
393 distribution over all possible input distributions we predict the channel capacity for a suite of biophys-
394 ical parameters such as varying repressor protein copy number and repressor-DNA binding affinity.
395 We compare these theoretical channel capacity predictions with experimental determinations, finding
396 that our minimal model is able to predict with no free parameters this quantity. In principle since
397 our predicted input-output distributions were in close agreement with experimental data we could
398 have chosen any arbitrary input distribution $P(c)$ and compute the mutual information between input
399 and outputs. The relevance of the channel capacity comes from its interpretation as a metric of the
400 limits of how precise the inference that cells can make about what the state of the environment is
401 given this simple genetic circuit. Our model makes non-trivial predictions such as the existence of
402 an optimal repressor copy number for a given repressor-DNA binding energy (See Fig. 5). We note
403 that this differs from previous theoretical results since this optimal combination does not come from
404 adding a cost term for the regulation [15]. This is a consequence of the parameters inferred in [26] for
405 the allosteric repressor never allowing all repressors to go into the inactive (non-DNA binding) state.
406 That means that even at saturating concentrations of inducer, as the number of repressors increases,
407 a significant number of them are still able to bind to the promoter. This causes all of the input-output
408 functions to be biased towards low expression levels, decreasing the amount of information that the
409 circuit is able to process.

410 It is important to highlight the limitations of the work presented here. As first reported in [26],
411 our model fails to capture the steepness of the fold-change induction curve for the weakest repressor
412 binding site (See Fig. 3(B)). This systematic deviation for weak binding sites remains an unresolved
413 problem that deserves further investigation. Also the minimal model in Fig. 2(A), despite being
414 widely used, is an oversimplification of the physical picture of how the transcriptional machinery
415 works. The coarse-graining of all the kinetic steps involved in the transcription initiation into two
416 effective promoter states - active and inactive - ignores potential kinetic regulatory mechanisms of
417 intermediary states [47]. Furthermore it has been argued that despite the fact that the mRNA count
418 distribution does not follow a Poisson distribution, this effect could be caused by unknown factors not
419 at the level of transcriptional regulation [48].

420 The findings of this work open the opportunity to accurately test intriguing ideas that connect
421 Shannon's metric of how accurately a signaling system can infer the state of the environment, with
422 Darwinian fitness [5]. Beautiful work along these lines has been done in the context of the developmen-
423 tal program of the early *Drosophila* embryo [8, 10]. These studies demonstrated that the input-output
424 function of the pair-rule genes works at channel capacity, suggesting that selection has acted on these

425 signaling pathways, pushing them to operate at the limit of what the physics of these systems allows.
426 Our system differs from the early embryo in the sense that we have a tunable circuit with variable
427 amounts of information processing capabilities. Furthermore, compared with the fly embryo in which
428 the organism tunes both the input and output distributions over evolutionary time, we have exper-
429 imental control of the distribution of inputs that the cells are exposed to. What this means is that
430 instead of seeing the final result of the evolutionary process, we can set different environmental chal-
431 lenges, and track over time the evolution of the population. These experiments could shed light into
432 the suggestive hypothesis of information bits as a metric on which natural selection acts. We see this
433 exciting direction as part of the overall effort in quantitative biology of predicting evolution [49].

434 2 Materials and Methods

435 2.1 *E. coli* strains

436 All strains used in this study were originally made for [26]. We chose a subset of three repressor copy
437 numbers that span 3 orders of magnitude. We refer the reader to [26] for detail on the construction
438 of these strains. Briefly the strains have a construct consisting of the *lacUV5* promoter, one of
439 three possible binding sites for the *lac* repressor (O1, O2, and O3) controlling the expression of a
440 YFP reporter gene. This construct is integrated into the genome at the *galK* locus. The number of
441 repressors per cell is varied by changing the ribosomal binding site controlling the translation of the *lac*
442 repressor gene. The repressor constructs were integrated in the *ybcN* locus. Finally all strains used in
443 this work constitutively express an mCherry reporter from a *loc* copy number plasmid. This serves as
444 a volume marker that facilitates the segmentation of the cells when processing the microscopy images.

445 2.2 Growth conditions

446 For all experiments cultures were initiated from a 50% glycerol frozen stock at -80°C. Three strains -
447 autofluorescence (*auto*), $\Delta lacI$ (Δ), and a strain with a known binding site and repressor copy number
448 (*R*) - were inoculated into individual tubes with 2 mL of Lysogeny Broth (LB Miller Powder, BD
449 Medical) with 20 $\mu\text{g}/\text{mL}$ of chloramphenicol and 30 $\mu\text{g}/\text{mL}$ of kanamycin. These cultures were grown
450 overnight at 37°C and rapid agitation to reach saturation. The saturated cultures were diluted 1:1000
451 into 500 μL of M9 minimal media (M9 5X Salts, Sigma-Aldrich M6030; 2 mM magnesium sulfate,
452 Mallinckrodt Chemicals 6066-04; 100 mM calcium chloride, Fisher Chemicals C79-500) supplemented
453 with 0.5% (w/v) glucose on a 2 mL 96-deep-well plate. The *R* strain was diluted into 12 different
454 wells with minimal media, each with a different IPTG concentration (0 μM , 0.1 μM , 5 μM , 10 μM ,
455 25 μM , 50 μM , 75 μM , 100 μM , 250 μM , 500 μM , 1000 μM , 5000 μM) while the *auto* and Δ strains
456 were diluted into two wells (0 μM , 5000 μM). Each of the IPTG concentration came from a single
457 preparation stock kept in 100-fold concentrated aliquots. The 96 well plate was then incubated at
458 37°C with rapid agitation for 8 hours before imaging.

459 2.3 Microscopy imaging procedure

460 The microscopy pipeline used for this work followed exactly the steps from [26]. Briefly, twelve 2%
461 agarose (Life Technologies UltraPure Agarose, Cat.No. 16500100) gels were made out of M9 media (or
462 PBS buffer) with the corresponding IPTG concentration (See growth conditions) and placed between
463 two glass coverslips for them to solidify after microwaving.

464 After the 8 hour incubation in minimal media 1 μL of a 1:10 dilution of the cultures into fresh
465 media or PBS buffer was placed into small squares (roughly 10 mm \times 10 mm) of the different agarose
466 gels. A total of 16 agarose squares - 12 concentrations of IPTG for the *R* strain, 2 concentrations for
467 the Δ and 2 for the *auto* strain - were mounted into a single glass-bottom dish (Ted Pella Wilco Dish,
468 Cat. No. 14027-20) that was sealed with parafilm.

469 All imaging was done on an inverted fluorescent microscope (Nikon Ti-Eclipse) with custom-
470 built laser illumination system. The YFP fluorescence (quantitative reporter) was imaged with a
471 CrystaLaser 514 nm excitation laser coupled with a laser-optimized (Semrock Cat. No. LF514-C-000)
472 emission filter. All strains, including the *auto* strain included a constitutively expressed mCherry
473 protein to aid for the segmentation. Therefore for each image 3 channels YFP, mCherry, and phase
474 contrast were acquired.

475 On average 30 images with roughly 20 cells per condition were taken. 25 images of a fluorescent
476 slide and 25 images of the camera background noise were taken every time in order to flatten the
477 illumination. The image processing pipeline for this work is exactly the same as [26].

478 Acknowledgements

479 The authors would like to acknowledge Griffin Chure, Muir Morrison, and Nathan Belliveau for
480 fruitful discussions. We would like to also thank William Bialek, Emanuel Flores, Hernan Garcia,
481 Alejandro Granados, Jane Kondev, Sarah Marzen, Porfirio Quintero, Alvaro Sanchez, Rachel Taub-
482 man, Gašper Tkačik, Catherine Triandafillou, and Ned Wingreen for useful advice and discussion. We
483 thank Rob Brewster for providing the raw mRNA FISH data for inferences, Heun-Jin Lee for support
484 with the quantitative microscopy, and David Drabold for advice on the maximum entropy inferences.
485 This work was supported by La Fondation Pierre-Gilles de Gennes, the Rosen Center at Caltech, and
486 the NIH 1R35 GM118043 (MIRA).

487 References

- 488 [1] Ilya Nemenman. “Information theory and adaptation”. In: *Quantitative Biology: From Molecular*
489 *to Cellular Systems*. Vol. 30322. Taylor and Francis, 2010, pp. 1–12.
- 490 [2] Avigdor Eldar and Michael B Elowitz. “Functional roles for noise in genetic circuits”. In: *Nature*
491 467.7312 (2010), pp. 167–173.
- 492 [3] C. E. Shannon. “A Mathematical Theory of Communication”. In: *Bell System Technical Journal*
493 27.3 (1948), pp. 379–423.
- 494 [4] David JC MacKay. *Information theory, inference and learning algorithms*. Cambridge university
495 press, 2003.
- 496 [5] Samuel F. Taylor, Naftali Tishby, and William Bialek. “Information and fitness”. In: *ArXiv*
497 (2007).
- 498 [6] William Bialek and Sima Setayeshgar. “Physical limits to biochemical signaling”. In: *Proceedings*
499 *of the National Academy of Sciences* 102.29 (2005), pp. 10040–10045.
- 500 [7] Thomas Gregor, David W. Tank, Eric F. Wieschaus, and William Bialek. “Probing the Limits
501 to Positional Information”. In: *Cell* 130.1 (2007), pp. 153–164.

- 502 [8] Gasper Tkacik, Curtis G Callan, and William Bialek. “Information flow and optimization in
503 transcriptional regulation”. In: *Proceedings of the National Academy of Sciences* 105.34 (2008),
504 pp. 12265–12270.
- 505 [9] Julien O Dubuis, Gasper Tkacik, Eric F Wieschaus, Thomas Gregor, and William Bialek. “Posi-
506 tional information, in bits”. In: *Proceedings of the National Academy of Sciences* 110.41 (2013),
507 pp. 16301–16308.
- 508 [10] Mariela D. Petkova, Gašper Tkačik, William Bialek, Eric F. Wieschaus, and Thomas Gregor.
509 “Optimal Decoding of Cellular Identities in a Genetic Network”. In: *Cell* 176.4 (2019), 844–
510 855.e15.
- 511 [11] Georg Rieckh and Gašper Tkačik. “Noise and Information Transmission in Promoters with
512 Multiple Internal States”. In: *Biophysical Journal* 106.5 (2014), pp. 1194–1204.
- 513 [12] Etay Ziv, Ilya Nemenman, and Chris H. Wiggins. “Optimal Signal Processing in Small Stochastic
514 Biochemical Networks”. In: *PLoS ONE* 2.10 (2007). Ed. by Gustavo Stolovitzky, e1077.
- 515 [13] Margaritis Voliotis, Rebecca M Perrett, Chris McWilliams, Craig a McArdle, and Clive G Bow-
516 sher. “Information transfer by leaky, heterogeneous, protein kinase signaling systems”. In: *Pro-
517 ceedings of the National Academy of Sciences* 111.3 (2014), E326–E333.
- 518 [14] Filipe Tostevin and Pieter Rein ten Wolde. “Mutual Information between Input and Output
519 Trajectories of Biochemical Networks”. In: *Physical Review Letters* 102.21 (2009), p. 218101.
- 520 [15] Gašper Tkačik and Aleksandra M Walczak. “Information transmission in genetic regulatory
521 networks: a review”. In: *Journal of Physics: Condensed Matter* 23.15 (2011), p. 153102.
- 522 [16] Gašper Tkačik, Curtis G Callan, and William Bialek. “Information capacity of genetic regulatory
523 elements”. In: *Physical Review E* 78.1 (2008), p. 011910.
- 524 [17] Omar P Tabbaa and C Jayaprakash. “Mutual information and the fidelity of response of gene
525 regulatory models”. In: *Physical Biology* 11.4 (2014), p. 046004.
- 526 [18] Ido Golding, Johan Paulsson, Scott M Zawilski, and Edward C Cox. “Real-Time Kinetics of
527 Gene Activity in Individual Bacteria”. In: *Cell* 123.6 (2005), pp. 1025–1036.
- 528 [19] Hernan G Garcia and Rob Phillips. “Quantitative dissection of the simple repression input-
529 output function”. In: *Proceedings of the National Academy of Sciences* 108.29 (2011), pp. 12173–
530 12178.
- 531 [20] Jose M G Vilar and Leonor Saiz. “Reliable Prediction of Complex Phenotypes from a Modular
532 Design in Free Energy Space: An Extensive Exploration of the lac Operon”. In: *ACS Synthetic
533 Biology* 2.10 (2013), pp. 576–586.
- 534 [21] Heng Xu, Leonardo A. Sepúlveda, Lauren Figard, Anna Marie Sokac, and Ido Golding. “Com-
535 bining protein and mRNA quantification to decipher transcriptional regulation”. In: *Nature
536 Methods* 12.8 (2015), pp. 739–742.
- 537 [22] Robert C Brewster, Daniel L Jones, and Rob Phillips. “Tuning Promoter Strength through
538 RNA Polymerase Binding Site Design in *Escherichia coli*”. In: *PLoS Computational Biology*
539 8.12 (2012). Ed. by Erik van Nimwegen, e1002811.
- 540 [23] Stephanie L. Barnes, Nathan M. Belliveau, William T. Ireland, Justin B. Kinney, and Rob
541 Phillips. “Mapping DNA sequence to transcription factor binding energy in vivo”. In: *PLoS
542 Computational Biology* 15.2 (2019). Ed. by Gary D. Stormo, e1006226.

- 543 [24] Daniel L Jones, Robert C Brewster, and Rob Phillips. “Promoter architecture dictates cell-to-cell
544 variability in gene expression”. In: *Science* 346.6216 (2014), pp. 1533–1536.
- 545 [25] Robert C. Brewster, Franz M. Weinert, Hernan G. Garcia, Dan Song, Mattias Rydenfelt, and
546 Rob Phillips. “The Transcription Factor Titration Effect Dictates Level of Gene Expression”.
547 In: *Cell* 156.6 (2014), pp. 1312–1323.
- 548 [26] Manuel Razo-Mejia, Stephanie L Barnes, Nathan M Belliveau, Griffin Chure, Tal Einav, Mitchell
549 Lewis, and Rob Phillips. “Tuning Transcriptional Regulation through Signaling: A Predictive
550 Theory of Allosteric Induction”. In: *Cell Systems* 6.4 (2018), 456–469.e10.
- 551 [27] Rob Phillips, Nathan M Belliveau, Hernan Garcia, Manuel Razo, and Clarissa Scholes. “Figure
552 1 Theory Meets Figure 2 Experiments in the Study of Gene Expression”. In: *Annual Review of*
553 *Biophysics* (2019), pp. 1–46.
- 554 [28] Mattias Rydenfelt, Hernan G Garcia, Robert Sidney Cox, and Rob Phillips. “The Influence
555 of Promoter Architectures and Regulatory Motifs on Gene Expression in *Escherichia coli*”. In:
556 *PLoS ONE* 9.12 (2014). Ed. by Jordi Garcia-Ojalvo, e114347.
- 557 [29] Alvaro Sanchez, Sandeep Choubey, and Jane Kondev. “Stochastic models of transcription: From
558 single molecules to single cells”. In: *Methods* 62.1 (2013), pp. 13–25.
- 559 [30] Douglas F Browning and Stephen J. W. Busby. “The regulation of bacterial transcription initi-
560 ation”. In: *Nature Reviews Microbiology* 2.1 (2004), pp. 57–65.
- 561 [31] Nicolas E Buchler, Ulrich Gerland, and Terence Hwa. “On schemes of combinatorial transcription
562 logic”. In: *Proceedings of the National Academy of Sciences* 100.9 (2003), pp. 5136–5141.
- 563 [32] J. Peccoud and B. Ycart. “Markovian Modeling of Gene-Product Synthesis”. In: *Theoretical*
564 *Population Biology* 48.2 (1995), pp. 222–234.
- 565 [33] Qiang Cui and Martin Karplus. “Allostery and cooperativity revisited”. In: *Protein Science* 17.8
566 (2008), pp. 1295–1307.
- 567 [34] Rob Phillips. “Napoleon Is in Equilibrium”. In: *Annual Review of Condensed Matter Physics*
568 6.1 (2015), pp. 85–111.
- 569 [35] H. D. P. P. Bremer and P. P. Dennis. “Modulation of chemical composition and other parameters
570 of the cell by growth rate”. In: *Escherichia coli and Salmonella: cellular and molecular biology*
571 2.2 (1996), pp. 1553–69.
- 572 [36] Joseph R. Peterson, John A. Cole, Jingyi Fei, Taekjip Ha, and Zaida A. Luthey-Schulten. “Effects
573 of DNA replication on mRNA noise”. In: *Proceedings of the National Academy of Sciences* 112.52
574 (2015), pp. 15886–15891.
- 575 [37] Hengjiang Dong and C.G. Kurland. “Ribosome Mutants with Altered Accuracy Translate with
576 Reduced Processivity”. In: *Journal of Molecular Biology* 248.3 (1995), pp. 551–561.
- 577 [38] E. O. POWELL. “Growth Rate and Generation Time of Bacteria, with Special Reference to
578 Continuous Culture”. In: *Journal of General Microbiology* 15.3 (1956), pp. 492–511.
- 579 [39] E. T. Jaynes. “Information Theory and Statistical Mechanics”. In: *Physical Review* 106.4 (1957),
580 pp. 620–630.
- 581 [40] Clive G. Bowsher and Peter S. Swain. “Environmental sensing, information transfer, and cellular
582 decision-making”. In: *Current Opinion in Biotechnology* 28 (2014), pp. 149–155.

- 583 [41] R. Blahut. “Computation of channel capacity and rate-distortion functions”. In: *IEEE Trans-*
584 *actions on Information Theory* 18.4 (1972), pp. 460–473.
- 585 [42] C.T. Bergstrom and Michael Lachmann. “Shannon information and biological fitness”. In: *In-*
586 *formation theory workshop. IEEE, 2004*. IEEE, 2004, pp. 50–54.
- 587 [43] Daniel Polani. “Information: Currency of life?” In: *HFSP Journal* 3.5 (2009), pp. 307–316.
- 588 [44] Olivier Rivoire and Stanislas Leibler. “The Value of Information for Populations in Varying
589 Environments”. In: *Journal of Statistical Physics* 142.6 (2011), pp. 1124–1166.
- 590 [45] Eric Libby, Theodore J Perkins, and Peter S Swain. “Noisy information processing through
591 transcriptional regulation”. In: *Proceedings of the National Academy of Sciences* 104.17 (2007),
592 pp. 7151–7156.
- 593 [46] Alex Rhee, Raymond Cheong, and Andre Levchenko. “The application of information theory to
594 biochemical signaling systems”. In: *Physical Biology* 9.4 (2012), p. 045011.
- 595 [47] Clarissa Scholes, Angela H. DePace, and Álvaro Sánchez. “Combinatorial Gene Regulation
596 through Kinetic Control of the Transcription Cycle”. In: *Cell Systems* 4.1 (2017), 97–108.e9.
- 597 [48] Sandeep Choubey, Jane Kondev, and Alvaro Sanchez. “Distribution of Initiation Times Reveals
598 Mechanisms of Transcriptional Regulation in Single Cells”. In: *Biophysical Journal* 114.9 (2018),
599 pp. 2072–2082.
- 600 [49] Michael Lässig, Ville Mustonen, and Aleksandra M. Walczak. “Predicting evolution”. In: *Nature*
601 *Ecology & Evolution* 1.3 (2017), p. 0077.