

RESEARCH ARTICLE

The Impact of Different Sources of Fluctuations on Mutual Information in Biochemical Networks

Michael Chevalier^{1*}, Ophelia Venturelli², Hana El-Samad^{1*}

1 Department of Biochemistry and Biophysics, California Institute for Quantitative Biosciences, University of California San Francisco, San Francisco, California, United States of America, **2** Department of Bioengineering, University of California, Berkeley, Berkeley, California, United States of America

* Michael.Chevalier@ucsf.edu (MC); Hana.El-Samad@ucsf.edu (HE)



CrossMark
click for updates

OPEN ACCESS

Citation: Chevalier M, Venturelli O, El-Samad H (2015) The Impact of Different Sources of Fluctuations on Mutual Information in Biochemical Networks. *PLoS Comput Biol* 11(10): e1004462. doi:10.1371/journal.pcbi.1004462

Editor: Feilim Mac Gabhann, Johns Hopkins University, UNITED STATES

Received: April 28, 2014

Accepted: July 20, 2015

Published: October 20, 2015

Copyright: © 2015 Chevalier et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the National Institute of General Medical Sciences (<http://www.nigms.nih.gov/Pages/default.aspx>), grant P50 GM081879, and the Paul Allen Family Foundation (<http://www.pgafamilyfoundation.org/>), 2013 Allen Distinguished Investigator award. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Stochastic fluctuations in signaling and gene expression limit the ability of cells to sense the state of their environment, transfer this information along cellular pathways, and respond to it with high precision. Mutual information is now often used to quantify the fidelity with which information is transmitted along a cellular pathway. Mutual information calculations from experimental data have mostly generated low values, suggesting that cells might have relatively low signal transmission fidelity. In this work, we demonstrate that mutual information calculations might be artificially lowered by cell-to-cell variability in both initial conditions and slowly fluctuating global factors across the population. We carry out our analysis computationally using a simple signaling pathway and demonstrate that in the presence of slow global fluctuations, every cell might have its own high information transmission capacity but that population averaging underestimates this value. We also construct a simple synthetic transcriptional network and demonstrate using experimental measurements coupled to computational modeling that its operation is dominated by slow global variability, and hence that its mutual information is underestimated by a population averaged calculation.

Author Summary

This work demonstrates how different sources of variability within biochemical networks impact the interpretation of information transmission. These sources are the intrinsic noise generated within the pathway of a single cell, variability due to initial conditions and/or global parameters across the population. A theoretical analysis of a simple signaling pathway and experimental exploration of a synthetic circuit are used to discuss the contributions of these sources of variability to information transmission using mutual information as a metric.

Introduction

To survive in challenging conditions, cells need to detect, transduce, and process signals from their environment. A cell's ability to precisely process environmental signals is limited by intrinsic fluctuations and variability of its cellular processes. This variability takes root in the stochastic nature of biochemical reactions. For a given pathway, this includes the stochastic steps involved in transcription and translation [1–4] as well as diffusion-reactions, dissociations, allosteric changes, and degradation of biological molecules. A signal propagates across cellular networks through molecules undergoing these various reactions, and gets distorted and altered by their probabilistic nature. Therefore, metrics for quantifying the limits of faithful information propagation (signaling fidelity) in biological pathways are crucial for understanding their information processing and transduction capabilities.

Mutual information [5] (MI) is a natural metric for characterizing information transmission between the inputs of a stochastic network and its nodes. MI quantifies the level of precision with which a given node(s) in a network estimates and responds to an input(s) by accounting for both the mean and variability in the response. Recent studies have used MI to characterize information transmission between environmental inputs and transcription factors in a number of genetic circuits [6–10]. In these studies, steady-state MI was computed for a variety of *in silico* networks to assess their stationary response as a function of input dose. More recently, these ideas were extended to optimize time-dependent MI in delay circuits with binary inputs, and MI was used to discuss maximally informative network topologies in these contexts [11]. In addition, time dependent MI calculations were used to obtain fundamental limits on the suppression of molecular fluctuations for different network topologies [12].

Several experimental studies have also used MI to assess signaling fidelity. MI was used as a metric to argue that negative feedback enables dose-response alignment and enhances information transmission in the pheromone pathway in yeast [13]. Similarly, MI was used to estimate time-dependent information transfer in tumor necrosis factor (TNF) signaling, and to assess transmission bottlenecks in this system [14]. Recently, robustness and compensation of information transmission in different pathways and pharmacological perturbations were attempted in PC12 cells using similar measurements [15]. These experimental studies relied on driving isogenic cell populations with various inputs, and then calculating the mutual information based on the overall variability in the population response. Such calculations mostly found low MI values, suggesting that cellular pathways might have on average low information transmission capacity. In this work, we argue that these calculations often under-estimate MI of a pathway in a single cell, since they do not account for 1) *variability in initial conditions* and 2) *variability that is extrinsic to the pathway*. The overall effect of these two sources of variability is that any single cell has a quantitatively distinct input-output relationship [3, 4, 16] and that calculations that take this into account are needed for more accurate estimation of MI from experimental data. By assuming that extrinsic variability manifests as cell-to-cell differences in a global parameter, such as translation capacity, we demonstrate in a simple *in silico* circuit that mixing cells with different parameters sets (and/or different initial conditions) reduces the value of the computed MI. We also argue this point experimentally by building a simple synthetic circuit that exhibits strong extrinsic variability, and then demonstrating with the help of computational modeling that single cells within the population have a larger mutual information than that exhibited by the averaged population. These results indicate that cells might possess higher capacity for information transmission than previously appreciated.

Results

To compute mutual information in a given biological network, we apply simple step functions [14] of the appropriate environmental input to N populations of the same isogenic cells. The step function is mathematically defined as $x(t) = 0$ for $t < 0$ and $x(t) = X_+$ for $t \geq 0$, where X_+ is a constant within a given population. For each of the N populations, X_+ is sampled from a discrete uniform distribution, $p_u(x_+)$, over the range of interest (0 to X_{max}). The uniform distribution represents an unbiased distribution (other than the choice of X_{max}) that has been routinely applied to steady-state mutual information calculations [6]. Experimentally, one can implement this scheme by growing replicas of the same culture in an N -well plate and stimulating each well with a different step function as defined above (Fig 1a). For a given population n and sampled input amplitude $X_+(n)$, the stochastic time-dependent response of measurable proteins $y = [y_1 \dots y_m]$ of the population at t will be $p(y, t|X_+(n))$. For a general x_+ between 0 and X_{max} , we interpolate the central moments of adjacent sampled distributions to construct $p(y, t|x_+)$. The time-dependent mutual information is then given by

$$\begin{aligned} I(x(t); y, t) &= \sum_{x(t)} \sum_y p(y, t|x(t)) p(x(t)) \log_2 \frac{p(y, t|x(t))}{p(y, t)} \\ &= \sum_{x_+} \sum_y p(y, t|x_+) p_u(x_+) \log_2 \frac{p(y, t|x_+)}{p(y, t)} \\ &= I(x_+; y, t) \end{aligned} \tag{1}$$

The value of N (that is the number of experiments) can be chosen based on well established methods [14] to approximate the MI in Eq (1) (see Materials and Methods for further details).

We first illustrate MI calculations using an *in silico* model of a simple signaling cascade (Fig 1b). Here the input $X_n(t)$ causes the transformation of the inactive molecule Y_1^* to its active form Y_1 . Y_1 could be a receptor or transcription factor responsive to the given input. Y_1 in turn activates transcription of Y_2 , whose mRNA is denoted as M_{y_2} (chemical equations are detailed in Materials and Methods with the parameter values listed in Table 1). The uniform input distribution is between 0 nM and $X_{max} = 250$ nM. We found that for this system and its corresponding parameters, $N \geq 20$ was a conservatively large number of experiments to approximate MI.

Contribution of initial condition variability to time-dependent mutual information

We first assumed that this circuit is isolated from the rest of the cell, and that any stochasticity it exhibits is only the result of its chemical reactions (intrinsic variability). When this system is unstimulated ($t \leq 0$), its molecular species assume a joint steady-state distribution, $p(y_1^*, y_1, M_{y_2}, y_2, t \leq 0)$. As an example, we show the marginal distribution of Y_1^* in Fig 1c. This distribution represents the range of initial conditions in Y_1^* that a population containing this network would exhibit before any input is applied.

We will first compute the MI of the network while ignoring this initial distribution of states, assuming that all cells in the population start from the same initial condition (for state S1, this is the mean of the initial joint distribution, see Fig 1c) which we refer to as a homogeneous initial condition. This could be thought about as the mutual information of one cell in that population. We plot the time-dependent mutual information between the input and the different species of the circuit: $I(x_+; y_1, t|S1)$, $I(x_+; M_{y_2}, t|S1)$, $I(x_+; y_2, t|S1)$ (Fig 1d). The MI from the input to y_1 , $I(x_+; y_1, t|S1)$, has rapid dynamics, peaking initially and decaying with time to a

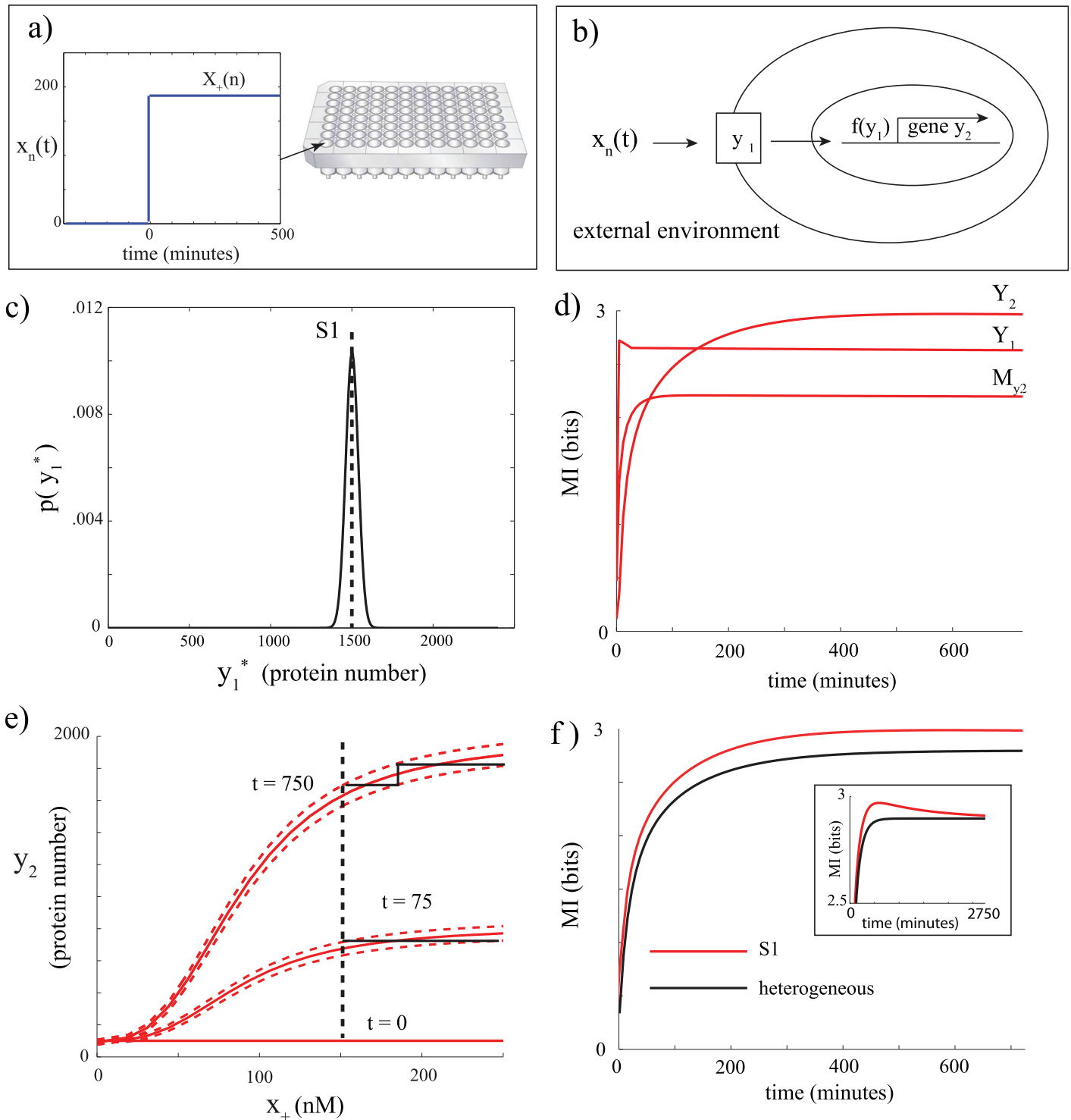


Fig 1. Time-dependent mutual information transmission in a simple biochemical circuit. a) Every well contains the same type of isogenic cell population where each well receives a different step input sampled from a uniform distribution. The resulting time series data is used to compute time-dependent mutual information. b) Schematic of pathway. c) Initial distribution of Y_1^* before any input is given, reflecting cell-cell variability in initial conditions. In this case, a single parameter set is used and the distribution is the result of intrinsic variability of the circuit. S1 denotes the mean of the distribution. d) $I(x_+, y_1, t|S1)$, $I(x_+, M_{y_2}, t|S1)$, and $I(x_+, y_2, t|S1)$. e) Time-dependent dose-response relationships at $t = 0, 75$ and $t = 750$ minutes between y_2 and the input. The vertical black dashed line is where $x_+ = 150$ nM. f) Mutual information between y_2 and the input computed for heterogeneous initial conditions (full distribution in panel c, black) and for homogeneous initial conditions (S1, red). Inset: same data but with time plotted to 2750 minutes and the MI plotted between 2.5 to 3 bits.

doi:10.1371/journal.pcbi.1004462.g001

Table 1. Parameters for circuit in numerical example.

parameter:	β_{y_1} (mol s ⁻¹)	$\gamma_{y_1} = \gamma_{y_1}$ (s ⁻¹)	θ_x (mol ⁻¹ s ⁻¹)	θ_{y_1} (s ⁻¹)	n_1	y_{1_0} (mol)
value:	.01	.0000067	.001	1.5	3	85
parameter:	$\beta_{m_2}^*$ (mol s ⁻¹)	β_{m_2} (mol s ⁻¹)	γ_{m_2} (s ⁻¹)	β_{y_2} (s ⁻¹)	γ_{y_2} (s ⁻¹)	
value:	.0083	.1583	.00083	.0014	.00014	

doi:10.1371/journal.pcbi.1004462.t001

steady-state. The initial peak in this MI is due solely to the activation and inactivation of y_1 , while the subsequent decrease to steady-state is due to the fluctuations in the synthesis and degradation of y_1 . By contrast, the MI from the input to M_{y_2} ($I(x_+; M_{y_2}, t|S1)$) is slower and on the order of tens of minutes, while that of the protein y_2 ($I(x_+; y_2, t|S1)$) is on the order of hours. This is not unexpected, as the mutual information signals for each species follow the causality of the circuit where y_2 shows the largest delay.

The increase of $I(x_+; y_2, t|S1)$ as a function of time has an intuitive explanation in terms of y_2 dynamics. To visualize this, we plot $\bar{y}_2(n, t)$, the mean of y_2 as well as $\bar{y}_2(n, t) \pm \sigma_{y_2}(n, t)$ versus $X_+(n)$ for $t = 0, 75$, and 750 minutes (Fig 1e, red lines), where $\sigma_{y_2}(n, t)$ is the standard deviation in y_2 . We will refer to these plots as the *time-dependent dose response relationships*. For $t = 750$, more values of x_+ are resolvable from measurement of y_2 than at time $t = 75$. For example, for $x_+ > 150$, steps in the dose response curves constrained between the standard deviations (Fig 1e, black lines for $t = 75$ and 750 minutes) approximate how well a measurement in y_2 can infer the value of x_+ . At time 750 , about 2 steps are resolvable allowing for two distinct ranges of x_+ to be inferred. While for $t = 75$ minutes, only one distinct range of x_+ is inferrable. The larger the number of resolved states, the higher the value of the mutual information.

When mutual information is calculated between the input and a given node over the entire time duration of the signals, the mutual information between the input and each successive node has an upper bound equal to that of the prior node. This is known as the data processing inequality [5]. However, since we are evaluating the time-dependent MI at a given time t , the instantaneous value y_2 can have more information about the input than y_1 . Indeed, at t approximately greater than 150 minutes, we find that the MI $I(x_+; y_2, t|S1)$ is greater than $I(x_+; y_1, t|S1)$ or $I(x_+; M_{y_2}, t|S1)$ (Fig 1d). This is because for the particular parameter set used in this example, the noise propagated from y_1 onto y_2 is averaged out, and the only variability in y_2 stems from its own production and degradation. As a result, $I(x_+; y_2, t|S1)$ can be modulated to be higher or lower than $I(x_+; y_1, t|S1)$ by changing the rates of y_2 production and degradation [17]. On the other hand, increasing the number of y_1 molecules would increase its mutual information as this would reduce the noise in the y_1 signal. Therefore, the mutual information at each node of this pathway can be modulated through choice of kinetic parameters. Similar observations that filtering can improve time-dependent MI between success nodes have been discussed in the context of other types of pathways [18].

Next, we examined mutual information while accounting for the fact that cells assume a distribution of initial states across the population upon receiving the input stimulus. We do so by incorporating the pre-stimulus steady state initial joint distribution into the MI calculations. This variability in initial conditions transiently reduces the MI (Fig 1f). At steady-state, the mutual information curves computed for a single or a distribution of initial states eventually converge onto each other at approximately $t = 2750$ minutes (Fig 1f, inset). This convergence at longer times occurs because a population in which every cell assumes the same exact initial conditions will eventually produce a heterogeneous distribution of states due to the intrinsic

stochasticity of the biochemical reactions. For the values of parameters used in this example, the convergence of the two MI curves proceeds very slowly. Therefore, even when only intrinsic fluctuations are present, with no extrinsic contributions to variability, and for a given distribution of initial conditions, a single cell still transiently assumes, on average, a higher time-dependent mutual information than the whole population. In our case, this difference is very modest.

Time-dependent mutual information transmission with global parameter variation

Thus far, in our MI calculations, we have only accounted for variability in initial conditions given a single parameter set for the pathway. More realistically, any given pathway in a cell is subjected to variability through coupling to other cellular activities. This is known as extrinsic noise to distinguish it from intrinsic noise generated by the pathway itself. There are many extrinsic sources of variability that cellular pathways experience. For example, different cells may contain different numbers of polymerases or ribosomes, and hence have different capacities for transcription and translation [3]. This extrinsic variability can be accounted for in many ways, the simplest is to assume that the transcription or translation rate constants themselves can assume different values in different cells across the population.

To demonstrate the contribution of extrinsic variability to MI calculations, we consider a simple case where cells in the population have different translation rates. To do so, we add a stochastic global variable, G , which affects the protein creation rates such that $\hat{\beta}_{y_1} = \beta_{y_1} G/\bar{G}$ and $\hat{\beta}_{y_2} = \beta_{y_2} G/\bar{G}$, where β_{y_1} and β_{y_2} are the nominal values for the parameters used above. In this way, the protein creation rates keep their mean value, but fluctuate because of their coupling to G . For this example, G follows a memoryless birth/death process such that the mean of G is $\bar{G} = \beta_g/\gamma_g$ (β_g, γ_g are the birth and death rates). It follows that G has a coefficient of variation given by $\eta_g = 1/\sqrt{\bar{G}}$.

First, setting $\bar{G} = 50$, we chose $\beta_g = 1.5 \times 10^{-6}$ mol-s⁻¹ and $\gamma_g = 3 \times 10^{-8}$ s⁻¹. These values establish a stationary distribution of states, which we use as an initial distribution for the MI calculations. Fluctuations in the translation rate induce extra variability in the pathway components (compare the initial distributions of Y_1^* in Fig 2a to Fig 1c). As a result, mutual information calculations with this added extrinsic variability (and using the population distribution of initial conditions) show that $I(x_+, y_2, t)$ is now drastically reduced compared to the case when a single parameter set is used to represent the lack of global variability (compare black line in Fig 2b with value in Fig 1d). Here also, as expected, MI calculations from a single initial state corresponding to parameters $G = \bar{G}$ (state S1), $G = \bar{G} - \sqrt{\bar{G}}$ (state S2) and $G = \bar{G} + \sqrt{\bar{G}}$ (state S3), generate high transient values (red (S1), blue (S2) and green (S3) curves in Fig 2b). This discrepancy between single cell and population MI is further highlighted by examining the time-dependent dose response relationship between y_2 and x_+ at $t = 750$ (Fig 2d (full population) and Fig 2c (S1, S2, and S3)). Again, the sub-populations generated from S1, S2, S3 each have little variability (high mutual information) relative to the full population.

While constraining $\bar{G} = 50$, we investigated the time-dependent MI for different values of β_g and γ_g . Our original choice of $\gamma_g = 3 \times 10^{-8}$ forces G , and hence the translation rates $\hat{\beta}_{y_1}$ and $\hat{\beta}_{y_2}$, to fluctuate very slowly. Therefore, the convergence of the MI values computed from a single initial condition versus the full distribution also proceeds slowly. As γ_g increases, this convergence proceeds faster (Fig 2e). Therefore, these results indicate that the mutual information of a pathway can be severely underestimated by population-based measurements if the

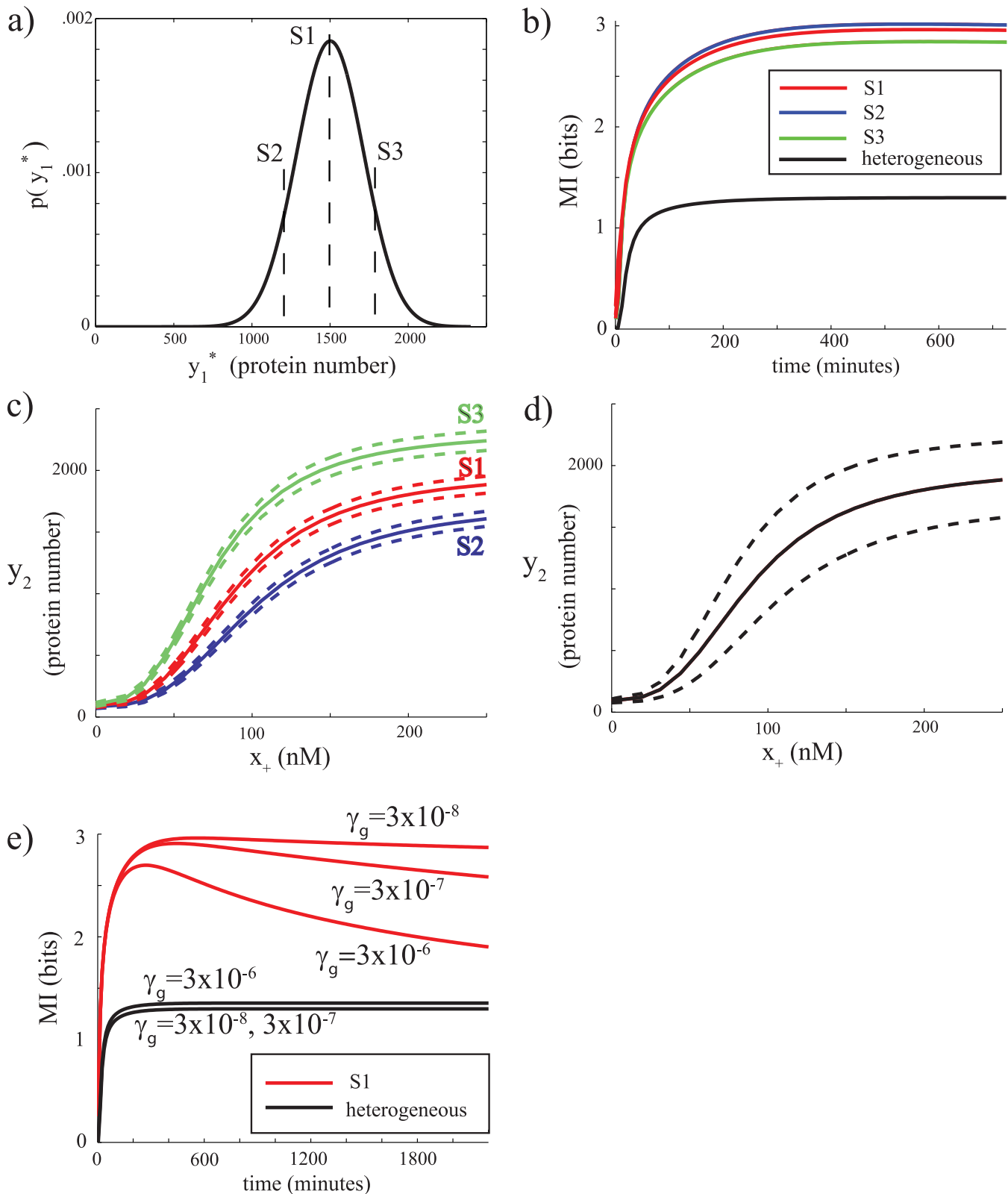


Fig 2. Global variability has large impact on mutual information. a) Initial distribution of Y_1^* before any input is given, reflecting cell-cell variability in initial conditions. In this case, the protein synthesis parameter is stochastic, reflecting a globally varying source of noise. S1 denotes the mean of the distribution, while S2 and S3 denote cells that are one standard deviation away from the mean. b) Time-dependent mutual information computed for heterogeneous initial

conditions (the whole distribution in panel a, black), or homogeneous initial conditions (S1: red, S2: blue, and S3: green) c) Time-dependent dose-response relationships between y_2 and input at $t = 750$ minutes for a population starting from homogeneous initial conditions corresponding to S_1 , S_2 , and S_3 . d) Time-dependent dose-response relationship between y_2 and input for heterogeneous initial conditions corresponding to full distribution in panel a. Calculations for panels a-d correspond to a slowly fluctuating global variable with $\gamma_g = 3 \times 10^{-8}$. e) Mutual information for different timescales of the fluctuations in the global variable.

doi:10.1371/journal.pcbi.1004462.g002

pathway is subjected to global fluctuations that proceed on a slower timescale than the pathway itself.

Probing the mutual information of a simple synthetic circuit

Next, we sought to probe the major determinants of mutual information for a simple synthetic transcriptional circuit (Fig 3a). In this circuit, a constitutively expressed transcription factor Y_1^* interacts with a small molecule X_+ , leading to the activation of the transcription factor. The active transcription factor Y_1 translocates into the nucleus and activates expression of a gene Y_2 .

In our implementation, Y_1^* is an estradiol (input X_+) responsive chimeric transcription regulator (TR) consisting of three fused elements: an activation domain (from MSN2), a lipid-binding domain (from the human estradiol receptor, hER-LBD), and a DNA binding domain (from GAL4). When estradiol binds to the LBD, the activated TR Y_1 translocates to the nucleus and controls the expression of promoters containing Gal4-binding-sites. Therefore, the protein Y_2 (in this case a fluorescent protein) is produced from a Gal4-responsive GAL10 promoter (See [Materials and Methods](#) for more details). At the same time, Y_1^* is produced from an altered version of the promoter of the alcohol dehydrogenase 1 gene (ADH1). We constructed two strains for measurement purposes. Strain 1 contains the circuit in addition to two copies of the GAL10 promoter, one driving YFP and the other driving mCherry. Strain 2 contains the circuit, but this time with two copies of the ADH1 promoter, one driving the production of Y_1^* and the other driving the production of YFP (which we will refer to as Y_{1r}). The same strain also contains a GAL10 promoter driving the production of the mCherry (Y_2) protein. These strains were useful for two reasons. First, we wanted to establish how mutual information computations depend on the ability to simultaneously measure different quantities in a circuit (e.g. Y_1 and Y_2 versus Y_2 alone). Given that this necessitates the use of two fluorescent proteins, in this case YFP and mCherry, we wanted to ascertain that the results we obtain are qualitatively independent of the choice of fluorophores, given that mCherry has lower dynamic range than YFP with higher background fluorescence and hence increased noise at low concentrations.

For each strain, we subjected 12 exponentially growing populations (wells) of cells cultured in non-repressive media to input concentrations of estradiol (x_+) log-sampled between 0 and 100 nM. The 12 measurement points sufficiently sampled the dose response relationships. The number of cells measured from each well was greater than 3000, ensuring good statistics for approximating the MI [14]. All cultures were started from zero estradiol concentrations. Samples were taken at $t = 0, 65, 165, 330$ and 580 minutes. Fig 3b shows the time-dependent dose-response relationships of estradiol versus y_2 (in this case YFP, strain 1) for these timepoints, where fluorescence values were normalized with respect to side scatter in order to minimize the effects of cellular volume and shape dependent differences.

The dose-response relationships of y_2 normalized by their respective maximum mean values (Fig 3d) exhibit an interesting trend: for the last 3 time points, the traces for the mean and variability are very similar to each other. The only outlier to this trend is the time point at $t = 65$ minutes after stimulation (Fig 3d). For this timepoint, fluorescence is weak and strongly

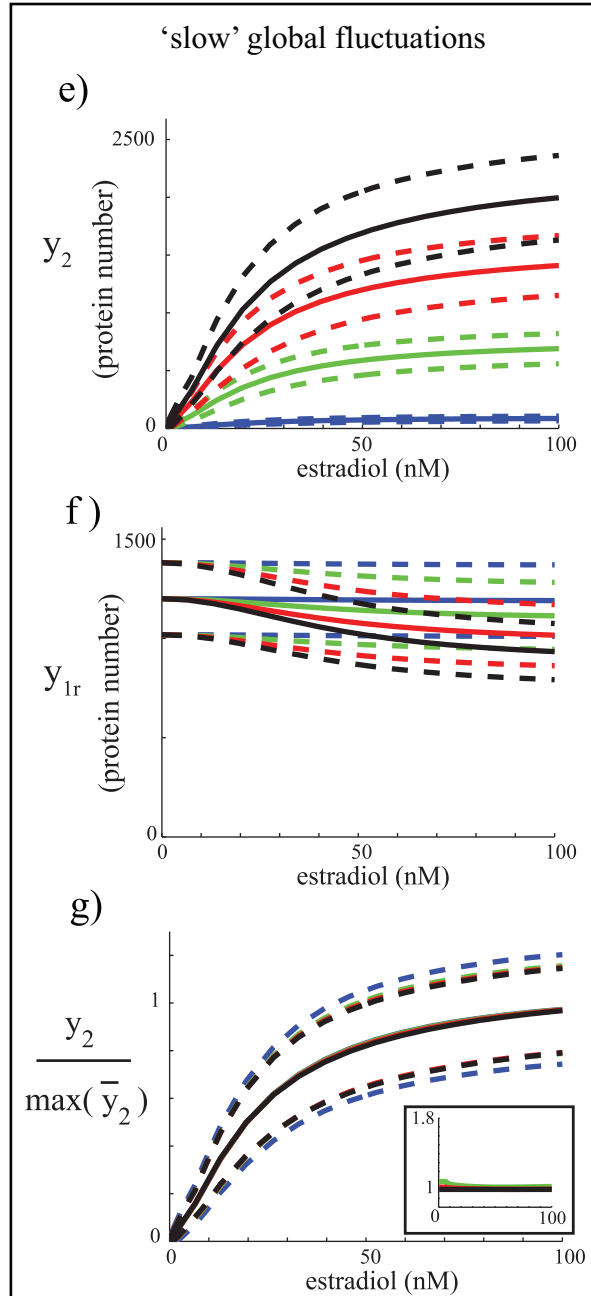
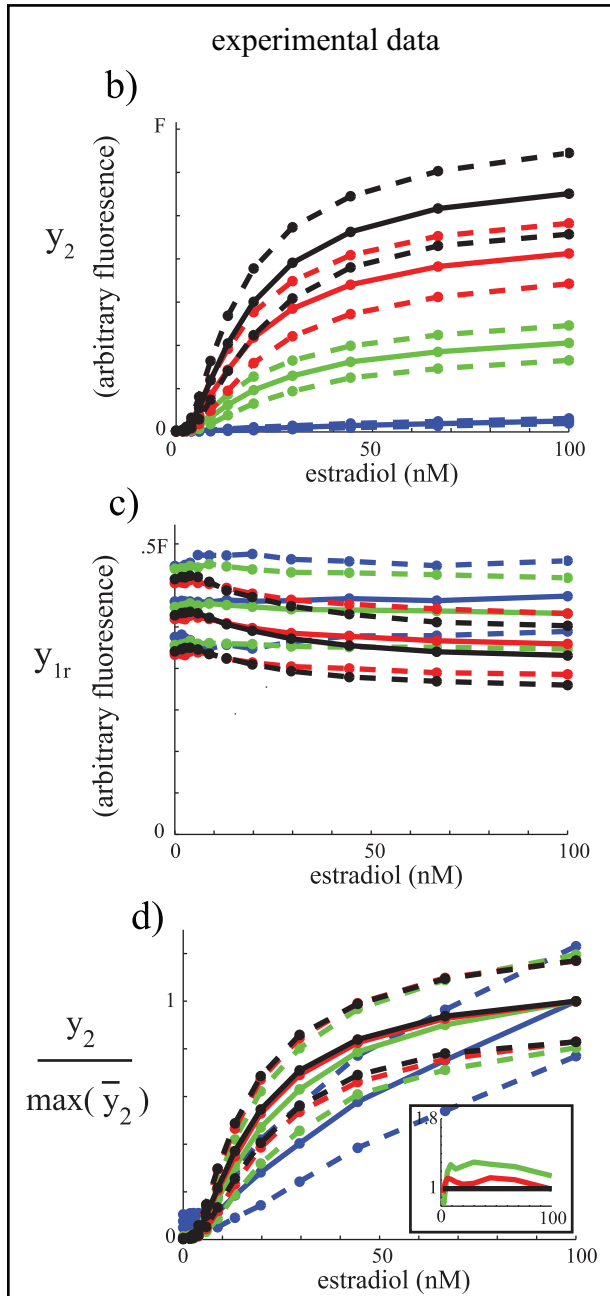
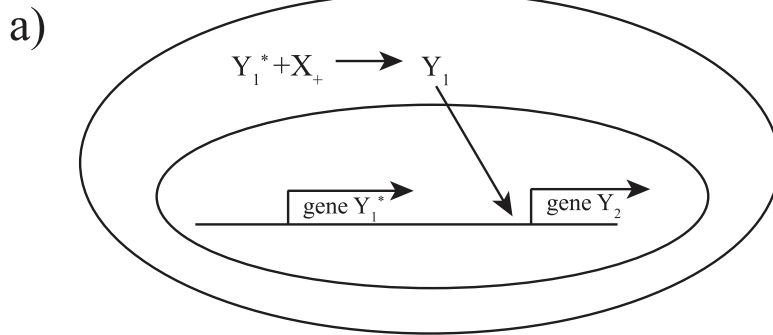


Fig 3. Variability in a transcriptional synthetic circuit is dominated by slowly fluctuating global variable. a) Schematic of the simple synthetic circuit. b-g) Dose-response relationships for $t = 65$ (blue), 165 (green), 330 (red), and 580 (black) minutes. Insets in d and g represent the ratio of the noise at time t to that at time $t = 580$ where for green: $t = 165$, red: $t = 330$, and black: $t = 580$. Noise is defined as standard deviation over the mean. b) Experimental y_2 data. c) Experimental y_{1r} data. d) Normalized experimental y_2 data. Each curve is normalized by its maximum mean value in panel b. e) y_2 data generated by a computational model of the circuit with slow fluctuations in protein synthesis rates. f) y_{1r} data generated by a computational model of the circuit with slow fluctuations in protein synthesis rates. g) Normalized y_2 data generated by a computational model of the circuit with slow fluctuations in protein synthesis rates. Each curve is normalized by its maximum mean value in panel e.

doi:10.1371/journal.pcbi.1004462.g003

overlaps with autofluorescence and folding delays, and therefore the true signal cannot be accurately estimated. Autofluorescence and folding delay also contributes, albeit less dramatically, to the measurement at the $t = 165$ minutes timepoint (Fig 3d). The mCherry measurements (strain 1 or strain 2) generated the same trend (S1a and S1b Fig) albeit with a noisier outcome than YFP due to the limited dynamic range of mCherry. As a consequence, the y_2 (YFP) and y_{1r} (YFP) experimental measurements from the two strains can be used in combination for comparison of modeling with data. The fact that variability in the y_2 data irrespective of the fluorescent protein does not decrease with increasing mean values suggests that dominant fluctuations are unlikely to be intrinsic to the pathway.

Fig 3c plots measurements of y_{1r} (YFP, strain 2). Unexpectedly, despite the common assumption that the ADH1 promoter has constitutive and constant expression, we found that it exhibits a modest dependence on estradiol. We do not know the root of this dependence, but it is likely to reflect the influence of the circuit itself on the metabolic state of the cell, hence affecting ADH1 promoter activity. Overall, the growth rate of these strains is independent of estradiol for concentrations under 100 nM over the duration of the experiment (S1c and S1d Fig), and therefore this effect can be compensated for in the mutual information calculations. It is worth noting here that we are making the assumption that despite the fact that YFP (Y_{1r}) and Y_1^* are different proteins sharing only the same transcription rate (since both are driven by the ADH1 promoter), they share the same dominant noise characteristics. This would be the case if their intrinsic noise, which can be different, is insignificant compared to a dominant source of extrinsic noise affecting both. Next, we present data and modeling demonstrating that, indeed, noise in both Y_1 and Y_2 is most likely dominated by the same extrinsic global component.

Since the measured distributions are approximately gaussian for the majority of estradiol concentrations (S2 Fig) and the synthetic circuit (Fig 3a) follows the same basic chemical equations as the simple pathway we have studied in Fig 1b, we used this already established model to computationally explore different noise scenarios (see Materials and Methods for a more technical justification of the model). Specifically, we simulated the model (parameter values listed in Table 2) with both intrinsic variability and added global extrinsic parameter variability as sources of stochasticity. The data we collected are in fluorescent units, therefore we set our model to arbitrarily yield maximum y_2 protein expression levels of about 2500 molecules, likely

Table 2. Parameters for model of synthetic circuit (global model).

parameter:	β_{m_1} (mol s ⁻¹)	γ_{m_1} (s ⁻¹)	β_{y_1} (s ⁻¹)	$\gamma_{y_1} = \gamma_{y_1}$	θ_x	θ_{y_1}	n_1	y_{1_0}	$\beta_{m_2}^*$
value:	.667	.00067	.0000139	.0000463	.002	1.5	1.5	85	0
parameter:	β_{m_2}	γ_{m_2}	β_{y_2}	$\gamma_{y_2} = \gamma_{y_{1r}}$	α_x	x_e (mol)	n_x		
value:	.083	.00083	.0014	.0000463	.35	45	2		

doi:10.1371/journal.pcbi.1004462.t002

Table 3. Parameters for model of synthetic circuit (intrinsic model).

parameter:	β_{m_1}	γ_{m_1}	$\beta_{y_1^*}$	$\gamma_{y_1^*} = \gamma_{y_1}$	θ_x	θ_{y_1}	n_1	y_{1_0}	$\beta_{m_2}^*$	β_{m_2}	γ_{m_2}
value:	.667	.00067	.0000019	.0000463	.002	1.5	1.5	1	0	.083	.00083
parameter:	β_{y_2}	$\gamma_{y_2} = \gamma_{y_{1r}}$	α_x	x_e	n_x						
value:	.0014	.0000463	.35	45	2						

doi:10.1371/journal.pcbi.1004462.t003

an underestimation of the actual system. However, this choice constitutes a scaling factor and does not affect any of our results. We also accounted for the estradiol dependence of ADH1 (Fig 3c) by adding to the model a term depicting the modest estradiol dependent repression of this promoter.

For global parameter variability, we again chose to focus on the parameters affecting protein expression. We potentially could model the global parameter variability with cell-to-cell heterogeneity in the protein degradation rates. However, given that our experimental data does not measure the expression of genes involved in either of these processes, i.e. no way to experimentally distinguish the source(s) of global parameter variability, we chose to model global variability in the protein creation rates. Following the same procedure as in the previous section, we added a stochastic global parameter, G , which affects the protein creation rates for Y_1^* and Y_2 such that $\hat{\beta}_{y_1^*} = \beta_{y_1^*} G / \bar{G}$ and $\hat{\beta}_{y_2} = \beta_{y_2} G / \bar{G}$. The noise in the experimental Y_{1r} data is approximately .155, therefore modeling Y_1 and Y_{1r} using Poissonian statistics sets the mean of the global noise variable to $\bar{G} \approx 42$. We first assumed that global parameter fluctuations are slow relative to the circuit timescales ($\gamma_g = 3 \times 10^{-6}$). Simulating the model with this slow global source of fluctuations (SGF model, (Fig 3e–3g)) generated profiles for normalized y_2 (Fig 3g) that recapitulated the highly similar variance envelopes of the experimental time-dependent dose responses (Fig 3d). This behavior was a characteristic feature of the model for any $\gamma_g < = 3 \times 10^{-6}$. By contrast, as the global fluctuating variable assumes a faster timescale ($\gamma_g > 3 \times 10^{-6}$), the variability envelopes in the normalized time-dependent dose response of y_2 started to diverge from each other (S3a and S3b Fig, $\gamma_g = 3 \times 10^{-4}$). As expected, the system modeled with intrinsic variability only ($\bar{G} = \infty$, parameter values listed in Table 3) shows a normalized time-dependent dose response in which variability decreases as a function of time as the protein levels increase (S3c and S3d Fig). Given the data in Fig 3b, if the fluctuations were purely intrinsic, the ratio of the standard deviation to the mean between times 165 and 580 minutes would decrease by a factor of approximately 1.7. This is a change we should be able to detect in our data since for the number of cells sampled, the error in estimating the means and standard deviations in the dose response relationships are .5 percent and 2 percent, respectively. However, as previously discussed, the experimental data shows that this ratio is relatively invariant for the last 3 time points (Fig 3d, inset) while increasing for the both the fast global fluctuations model (S3b Fig, inset) and the intrinsic variability model (S3d Fig, inset). Our argument is further strengthened by the fact that in order to capture the noise observed in y_{1r} with the intrinsic variability model for the first timepoint, we had to set the y_{1r} mean copy number in the model to an unrealistically low value for a strong promoter such as ADH1 (approximately 40 proteins), further indicating that variability is unlikely to be intrinsic. The results for the SGF model for $\gamma_g < = 3 \times 10^{-6}$ are not an artifact of the estradiol dependence of ADH1 since an SGF model without this effect yields indistinguishable results (S3e and S3f Fig). We therefore conclude that the dominant source of variability in this synthetic circuit is likely to be due to a globally slow fluctuating variable. This is consistent with previous results, which also indicated that global parameters play a dominant role in cell to cell variability and

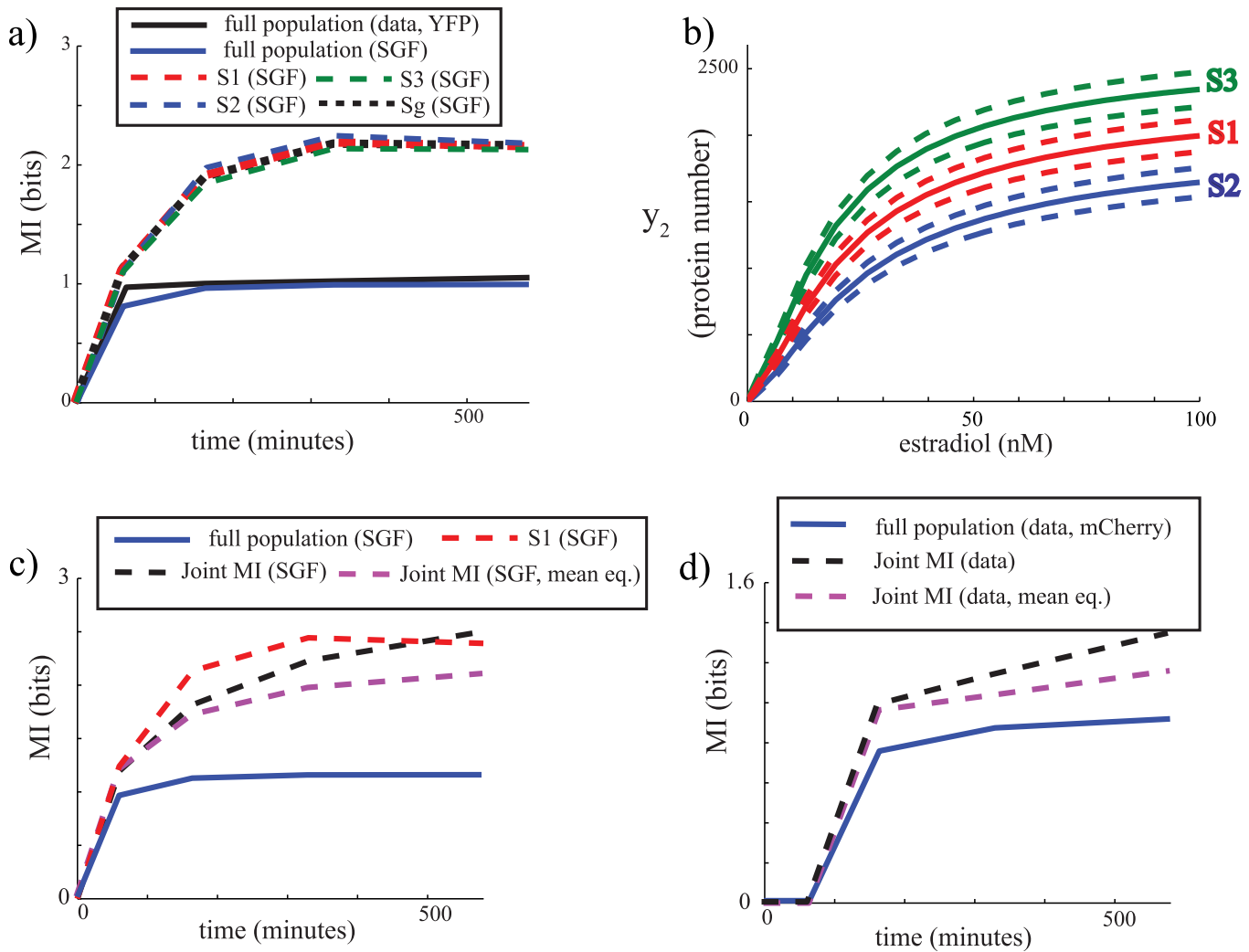


Fig 4. Mutual information modeling predictions and experimental measurements of the transcriptional synthetic circuit. a) Mutual information $I(x_+, y_2, t)$ from experimental data (black solid, YFP in Strain 1), $I(x_+, y_2, t)$ from the SGF model (blue solid), and $I(x_+, y_2, t|S_i)$ from the SGF model conditioned on S1 (red dashed), S2 (blue dashed), S3 (green dashed), and Sg (black dashed). b) Dose response of y_2 as a function of estradiol at time 580 minutes for the SGF model with initial condition S1, S2, and S3. c) Mutual information $I(x_+, [y_{1r}, y_2], t)$ based on joint measurement of y_{1r} and y_2 computed for the SGF model (black dashed), same calculations for equalized mean of y_{1r} (magenta dashed). Also shown for comparison are the mutual informations $I(x_+, y_2, t)$ (SGF, blue solid) and $I(x_+, y_2, t|S1)$ (SGF, red dashed). d) Measured joint mutual information $I(x_+, [y_{1r}, y_2], t)$ (black dashed) including the mean equalized case (magenta dashed). The measured mutual information $I(x_+, y_2, t)$ (mCherry, Strain 2) is also shown for comparison (blue solid).

doi:10.1371/journal.pcbi.1004462.g004

that these parameters exhibit fluctuations at a slower timescale than fluctuations of processes involved in gene expression [4].

In terms of mutual information, the fact that the normalized time-dependent dose responses coincide in terms of their variability (Fig 3d) implies that the experimentally computed mutual information $I(x_+, y_2, t)$ at $t = 165, 330$ and 580 minutes should be similar. This is indeed the case (Fig 4a, solid black). Importantly, $I(x_+, y_2, t)$ peaks and reaches a plateau at approximately 1 bit, at an earlier time than when y_2 reaches its steady-state. This further lends credence to the idea that the variability in the population is dominated by global parameter variability. Gratifyingly, the model with ‘slow’ global parameter fluctuations (with $\gamma_g = 3 \times 10^{-6}$) also captures the

time-dependent mutual information seen in the data without any further parameter tuning (Fig 4a, solid blue).

Since slow global fluctuations seem to dominate in this circuit, our analysis above indicates that the population mutual information might be under-estimating the fidelity of a single cell. To illustrate this point, we used the model to computationally isolate and compute the mutual information $I(x_+; y_2, t|S_i)$ for single cells S1, S2 and S3 as defined in the computational example above. These calculations yield a substantially higher MI value than the population MI for the time span simulated (Fig 4a). Evidently, and as explained above, the MI for S1, S2, or S3 will eventually converge back to the whole population MI, but here it will do so on a much slower timescale than that of the system. For example, the dose response and distribution of y_2 at time 580 minutes when the system is started from S1, S2 and S3 (Fig 4b) still shows tighter variability than that of the full population. Allowing for intrinsic variability in the initial conditions, i.e. starting with cells with $G = \bar{G}$ at time zero (state S_g) yields a similar MI value to that of $I(x_+; y_2, t|S1)$ for $\gamma_g = 3 \times 10^{-6}$ (Fig 4a).

Finally, we explored how simultaneous measurement of y_{1r} and y_2 affects mutual information calculations. Calculations using the model indicate that as expected, knowledge of y_{1r} improves the estimate of mutual information. For a slow globally fluctuating variable ($\gamma_g = 3 \times 10^{-6}$), the joint mutual information $I(x_+; [y_{1r}, y_2], t)$ is larger than $I(x_+; y_2, t)$. It can be shown that $I(x_+; [y_{1r}, y_2], t) = I(x_+; y_{1r}, t) + E[I(x_+; y_2, t|y_{1r})]$ where $E[I(x_+; y_2, t|y_{1r})]$ is the expected value of $I(x_+; y_2, t|y_{1r})$. Since the influence of estradiol on y_{1r} adds (albeit very slightly (Fig 3c (data) and Fig 3f (model))) to the mutual information, i.e. $I(x_+; y_{1r}, t) > 0$, we normalized for this effect. To do so, at a given time, we set the y_{1r} mean at each estradiol value to the value of the y_{1r} mean at zero estradiol while adjusting the variance to preserve the noise in y_{1r} at each estradiol value. Importantly, this operation does not affect correlation between y_2 and y_{1r} at each estradiol value, but enforces $I(x_+; y_{1r}, t) = 0$. We confirm that this does not change our conclusions that knowledge of y_{1r} improves the estimate of mutual information (compare Fig 4c blue, dashed black and dashed magenta).

For comparison, we can carry out mutual information from the data obtained using Strain 2 in which both y_{1r} and y_2 are measured. In this strain, y_2 is the fluorescent protein mCherry which has a limited dynamic range. Importantly, at the highest estradiol values and peak mCherry signal (time 580 minutes), the measured correlation between y_{1r} and y_2 (greater than .79) is less than ten percent below the model predictions. Even at these signal levels the noise in the mCherry signal still deteriorates the correlation. For decreasing values of estradiol the correlations become increasingly inaccurate. Therefore, the values of the MI cannot be quantitatively compared to the model which was fitted to YFP data. However, the qualitative trend of increased MI due to measurement of y_{1r} relative to computing the MI with no knowledge of y_{1r} should also hold. This is indeed seen to be the case (Fig 4d). This insight is in agreement with recent work [19] that studied mutual information in the RAS/ERK pathway. Nuclear ERK (erk_{nuc}) was used as a readout of pathway information transmission. The MI at time t between this readout and the input x was conditioned for single cell ERK levels, using measurement of total ERK (erk_{tot}). It was also shown that $I(x; [erk_{tot}, erk_{nuc}], t)$ is greater than $I(x; erk_{nuc}, t)$. Therefore, simultaneous measurements of different cellular variables improve estimates of mutual information capabilities of single cells.

Discussion

In this work, we illustrated how variability in initial conditions across a population, as well as slow-fluctuating extrinsic (global) variables can generate low values for the population mutual information in response to an input. We also demonstrated that when subpopulations of cells

that have similar parameters or initial states are isolated, their mutual information values are transiently much higher than those of the whole population. These findings are important in light of the fact that many previous studies have found that extrinsic variability is a substantial contributor to pathway fluctuations. Indeed, our own experimental data using a synthetic circuit also implicated extrinsic fluctuations as a major source of variability. As a result, cells in a population cannot be considered to be the same noisy channel for mutual information calculations. Rather, each cell is a different noisy channel possessing its own parameters. Recent work [20] using light-inducible input signals [21, 22] to a mammalian RAS/MAPK pathway observed that different isogenic single cells have quantitatively different dose-response relationships. Interestingly, for the RAS/MAPK mammalian system, the dose-response relationships were repeatable for hours within a given single cell [20], suggesting slow global parameters that affect that pathway for that duration.

A direct assessment of mutual information requires repeated time-resolved measurements in single cells. Another strategy to better approximate mutual information is to simultaneously measure a large number of interconnected variables, including global states. This might be increasingly feasible with breakthrough technologies such as mass-cytometry (a.k.a. CyTOF) [23] as well as improvements in fluorescent reporter technologies. In the mean time, however, we have demonstrated that computational modeling, especially with respect to the patterns of time-dependent variability, can generate valuable insights into whether intrinsic or extrinsic fluctuations dominate variability in a circuit. These results produce a more accurate quantification of mutual information, and therefore promise to generate a more realistic assessment of signaling fidelity in cellular circuits.

Our results and those from [20] support a view in which individual cells have distinct transfer functions over relevant signaling timescales, and have superior signaling fidelity (> 1.5 bits) than estimated from pooled measurements of a population. From this perspective, it could be the case that a diversity of high fidelity but different single cell signaling transfer functions across the population is a beneficial trait. However, some situations might arise where variability in population signal transmission capacity is not desirable. In this case, cells might use strategies such as negative feedback to constrain this variability. In either case, cells might also capitalize on the integration of signals from many pathways that respond to a given input(s) in order to generate a desired population response. In this view, each such pathway will add to the mutual information of the desired cellular output (e.g. level or activity of a transcription factor), allowing the population to further circumvent in this way any information fidelity bottleneck. Researchers of the subject are likely to encounter both situations, and perhaps a revised form of population mutual information might be needed to quantify these effects, along with the formulation of new information theoretic metrics. As an example, for any given input $x(t)$, the mutual information $I(y_2; y_1^*, t | x(t))$ gives us a sense of the diversity (or spread) in responses in y_2 given the cell-to-cell variability encoded in y_1^* . S4 Fig shows the results of this metric applied to the simple signaling cascade (Fig 1b) for different input step function amplitudes x_+ and for different times. We envision these kind of metrics to reflect the different subpopulations with similar parameters within a given population and to serve as a potential tool to quantify how cell-to-cell variability across a population might change in structure due to various time-dependent inputs.

Finally, most studies to date have focused on variability in populations of non-communicating cells. Information fidelity in cells that communicate, for example through quorum sensing for bacterial communities or cell-to-cell mechanical coupling for tissues, is still largely unstudied. How cell-to-cell communication modulates global variability and variability in initial conditions across a population, and hence mutual information of cellular pathways, is a topic that

should be explored in order to determine whether and when multicellularity offers a beneficial strategy in terms of signaling fidelity.

Materials and Methods

Approximating mutual information with N experiments

Because we are using a finite number of experiments, the input distribution $p_u(x_+)$ is sampled with N discrete points. In practice, these points are spaced to accurately sample the input-output transfer function $p(y, t|x_+)$ for x_+ ranging from 0 to X_{max} . The time-dependent mutual information is then calculated with this data. For values of x_+ between the sampled values, $p(y, t|x_+)$ is approximated by linearly interpolating the moments of the adjacent sampled distributions. Since the distributions generated by systems in this paper are approximately gaussian (and approximately negative binomial at very low x_+ for the synthetic circuit data), only the means and covariances are required. A larger number of experiments (N) generates a more accurate approximation of mutual information. However, we observed that convergence to accurate MI values does not increase monotonically with N for the logarithmic sampling of the doses response that we have adopted. Rather, convergence proceeds exponentially, followed by marginal gains in accuracy as N increases. Therefore, for every N , we examine the last three sample number values, N , $N-1$ and $N-2$. Given their measured convergence rates, we can extrapolate an upper bound on the MI at an infinite number samples. We choose N whose calculated MI at N is within 1 percent of the extrapolated upper bound.

Chemical equations for the simple *in silico* network

The chemical equations for the circuit in [Fig 1b](#) are



where $f_1(y_1) = \beta_{m_2}^* + \frac{\beta_{m_2} y_1^{n_1}}{\gamma_{y_1}^{n_1} + y_1^{n_1}}$. The propensities of the reactions appear above the reaction

arrows. The system is a simple cascade of reactions where the input X activates Y_1 , and subsequently the Y_1 -dependent transcription of Y_2 . The parameter values are tabulated in [Table 1](#).

Here the mean total number of Y_1 molecules, active and inactive, is $\beta_{y_1^*} / \gamma_{y_1^*} = 1500$. The max mean numbers for Y_2 mRNA and Y_2 protein are $\beta_{m_2} / \gamma_{m_2} = 200$ and $\frac{\beta_{m_2} \beta_{y_2}}{\gamma_{m_2} \beta_{y_2}} = 2000$, respectively. This system has only a single stationary solution. This allows us to approximate and efficiently calculate the master equation with a local affine assumption using the first two moments Eqs (6) and (7) taken from [24].

Inclusion of global parameter variability within chemical equations. In this work, we have assumed that the stochastic global parameter, G , manifests itself in variation in translation rates. To incorporate G we define our new protein creations rates $\hat{\beta}_{y_1^*} = \beta_{y_1^*} G / \bar{G}$ and $\hat{\beta}_{y_2} = \beta_{y_2} G / \bar{G}$. Here, $\beta_{y_1^*}$ and β_{y_2} are the nominal values in Eq (2). For simplicity we define G to follow a memoryless birth/death process through the reaction equations



where $\bar{G} = \beta_g / \gamma_g$ and the noise in G , η_g , is given as the standard deviation over the mean: $\eta_g = 1 / \sqrt{\bar{G}}$.

Chemical equations for the synthetic circuit. The chemical equations for the synthetic circuit in Fig 3a are the same as the simple circuit except that the production of Y_1^* now involves an mRNA step, which does not directly affect any of our results. Also, we have added a YFP reporter of Y_1^* that has a half-life of 6 hours which we set the transcription factor itself to be the same. We define the estradiol dependence in Y_1^* mRNA as $\beta_{m_1^*}(x) = \beta_{m_1}(1 - \alpha_x) + \alpha_x x_e^{n_x} / (x_e^{n_x} + x^{n_x})$. The parameters for the circuit are given in Table 2 (global fluctuation model) and Table 3 (intrinsic fluctuation model).

Computation of first two moments using affine assumption

The formulation that we assume in our model and data consists of a system of well-stirred chemical reactions with N molecular species. For some environmental input $X(t)$, we define the pathway state $Y(t)$ to denote the vector whose integer elements $Y_i(t)$ are the number of molecules of the i th species at time t . If there are M elementary chemical reactions that can occur among these N species, then we associate with each reaction r_j ($j = 1, \dots, M$) a non-negative propensity function defined such that $a_j(Y(t)) \tau + o(\tau)$ is the probability that reaction r_j will happen in the next small time interval $[t, t + \tau]$, as $\tau \rightarrow 0$. The polynomial form of the propensities $a_j(y)$ may be derived from fundamental principles under certain assumptions [25]. The occurrence of a reaction r_j leads to a change of $v_j \in \mathbf{Z}^N$ (the set of nonnegative integers) for the state Y . v_j is therefore a stoichiometric vector that reflects the integer change in reactant species due to a reaction r_j .

This set of well-stirred chemical reactions can be represented by the joint probability density function $P(y, t | X(t))$ which describes the probability of the system being in state y at time t , given the environmental signal $X(t)$. The evolution of $P(y, t | X(t))$ is given by

$$\frac{\partial P(y, t | X(t))}{\partial t} = \sum_{j=1}^M [a_j(y - v_j) P(y - v_j, t | X(t)) - a_j(y) P(y, t | X(t))] \tag{4}$$

Eq (4) is the so-called chemical master equation (CME) [26, 27].

To approximate the CME with moment equations, we approximate the propensity function $a_j(y)$ with a locally affine Taylor series expansion [24] about the mean of the distribution, $z(t)$, to get

$$a_j(y) \approx a_j(z(t)) + \sum_{i=1}^N \frac{\partial a_j(y)}{\partial y_i} \Big|_{y=z(t)} [y_i - z_i(t)] \tag{5}$$

Table 4. Strains used in this study.

Strain name	Genotype
S-1	W303 MATa leu2::LEU2-P _{GAL10} Venus, trp1, ura3::URA3-P _{ADH1} Gal4DBD-ER, ade2::ADE2, his3::HIS3-P _{GAL10} mCherry, GAL4::HPH
S-2	W303 MATa leu2, trp1::TRP1-P _{ADH1} Venus, ura3::URA3-P _{ADH1} Gal4DBD-ER, ade2::ADE2, his3::HIS3-P _{GAL10} mCherry, GAL4::HPH

doi:10.1371/journal.pcbi.1004462.t004

From the time dependent mean equation for the k th species is

$$\frac{\partial z_k(t)}{\partial t} = \sum_{j=1}^K v_{jk} a_j(z(t)) \quad (6)$$

and the time dependent covariance equation for the k th and k' th species is

$$\frac{\partial C_{kk'}(t)}{\partial t} = \sum_{j=1}^K \left(v_{jk} \sum_i \frac{\partial a_j(y)}{\partial y_i} \Big|_{y=z(t)} C_{ik'} + v_{j k'} \sum_{i'} \frac{\partial a_j(y)}{\partial y_{i'}} \Big|_{y=z(t)} C_{k i'} + v_{jk} v_{j k'} a_j(z(t)) \right) \quad (7)$$

The calculation of the mutual information requires probability distributions. Given that we solve the first two moments, we constrain our distributions to be either a negative binomial distribution or a normal distribution. For cases when $\mu_k < 3\sqrt{C_{kk}}$, we apply the negative binomial distribution since it only requires the first two moments and is non-negative. The negative binomial is very close to a normal distribution for $\mu_k > 3 * \sqrt{C_{kk}}$ and we therefore apply the normal distribution in these regions. The value of 3 used is heuristic, but the tail of the normal distribution at negative values is negligible at this point. For linear transcriptional systems, the negative binomial is a natural steady state solution [28] which was our motivation for applying it. Importantly, our data never violated any constraints required by the negative binomial distribution, for example, $\mu_k \leq C_{kk}$. Note that the negative binomial distribution is only required for our modeling of the synthetic circuit data. Our theoretical example in the first half of the paper has large enough basal levels at zero input which always keeps it in the normal distribution regime. As a demonstration of the validity of the moment approach, S5 Fig shows very good agreement in the distributions derived from stochastic simulations (SSA) and the moment equations for the synthetic circuit model ($\gamma_g = 3 \times 10^{-6}$, initial condition S1).

Synthetic circuit constructs

Strains. All plasmids used in this study were derived from a set of yeast single integration vectors containing selectable markers and targeting sequences for the LEU2, HIS3, TRP1 and URA3 loci. These vectors were linearized by digestion with PmeI and transformed using standard yeast transformation techniques. All strains were derived from haploid W303a and were deleted for GAL4 to eliminate competition between the endogenous Gal4p and the previously described estradiol-inducible Gal4 chimera (Gal4DBD-ER) for binding to the GAL10 promoter [29]. The sequences for the ADH1 and GAL10 promoters were 658 and 646 bp upstream from the start codons for these genes, respectively. The genotypes for these strains are listed in Table 4.

Growth conditions and flow cytometry. Cells were grown in YPD at 30°C. Prior to the experiment, cells were grown and maintained in exponential phase (optical density <.15) for approximately twenty-four hours in the absence of estradiol and then diluted to an optical density (OD600) of 0.05 at the beginning of the experiment and periodically diluted to stay under

0.15. Using a deep-well 96-well plate (Thermo), estradiol concentrations decreased from columns one to eleven in logarithmically spaced points according to the following equation: $100(2/3)^{c-1}$ nM where c denotes the column number. Well 12 did not receive estradiol. Estradiol did not change the growth rate of the cell population. Two replicates were performed. Measurements were taken at $t = 0, 65, 165, 330,$ and 580 minutes.

Fluorescence measurements were performed on a LSRII analyzer (BD Biosciences). A blue (488 nm) laser was used to excite YFP (Venus) and a green (561 nm) laser was used to excite RFP (mCherry). Emission was detected using a 530/30-nm bandpass filter for Venus (Chroma) and a 610/20 bandpass filter for mCherry (Chroma). Greater than three thousand cells were collected for each measurement. Flow cytometry data was analyzed in MATLAB (Mathworks).

Mutual information of multivariate measurements

Here we discuss how multi-variate MI measurements relates to MI measurements from particular initial conditions: We start with the distribution $p(y_m, y_s, t|x(t))$ where y_m are the dynamic cellular pathway/network signals, y_s are the slowly fluctuating pathway component quantities relative to the timescale of a given experiment, and $x(t)$ is the input signal(s). The time dependent mutual information is

$$\begin{aligned}
 I(x(t); [y_m, y_s], t) &= \sum_{x(t)} \sum_{y_s} \sum_{y_m} p(y_m, y_s, t|x(t)) p(x(t)) \log_2 \frac{p(y_m, y_s, t|x(t))}{p(y_m, y_s, t)} \\
 &= \sum_{x(t)} \sum_{y_s} \sum_{y_m} p(y_m, t|y_s, x(t)) p(y_s, t|x(t)) p(x(t)) \\
 &\quad \log_2 \frac{p(y_m, t|y_s, x(t)) p(y_s, t|x(t))}{p(y_m, t|y_s) p(y_s, t)}
 \end{aligned} \tag{8}$$

where the second line is simply a chain-rule representation. In addition to the assumption that the quantities of y_s are fluctuating extremely slowly, we will also impose that the quantities in y_s are independent of $x(t)$. This results in $p(y_s, t|x(t)) = p(y_s, t) \approx p(y_s)$. The time dependent MI is approximated as

$$\begin{aligned}
 I(x(t); [y_m, y_s], t) &\approx \sum_{x(t)} \sum_{y_s} \sum_{y_m} p(y_m, t|y_s, x(t)) p(y_s) p(x(t)) \log_2 \frac{p(y_m, t|y_s, x(t)) p(y_s)}{p(y_m, t|y_s) p(y_s)} \\
 &= \sum_{y_s} p(y_s) \sum_{x(t)} \sum_{y_m} p(y_m, t|y_s, x(t)) p(y_s) p(x(t)) \log_2 \frac{p(y_m, t|y_s, x(t))}{p(y_m, t|y_s)} \\
 &= \sum_{y_s} p(y_s) I(x(t); y_m, t|y_s) \\
 &= E[I(x(t); y_m, t|y_s)]
 \end{aligned} \tag{9}$$

Finally, we can examine the mutual information between y_s and y_m for a given input signal(s) $x(t)$ using the formula

$$I(y_m; y_s, t|x(t)) = \sum_{y_s} \sum_{y_m} p(y_m, t|y_s, x(t)) p(y_s, t|x(t)) \log_2 \frac{p(y_m, t|y_s, x(t))}{p(y_m, t|x(t))} \tag{10}$$

Supporting Information

S1 Fig. Comparison of Y_2 mCherry measurements show that Strains 1 and 2 are equivalent in Y_2 output. Strains 1 and 2 have similar growth rates that are independent of estradiol concentrations less than 100 nM. a-b) Dose-response relationships for $t = 65$ (blue), 165 (green), 330 (red), and 580 (black) minutes. The top panel is the normalized experimental y_2 data. Each curve is normalized by its maximum mean value in the corresponding lower panel. The lower panel is the un-normalized experimental y_2 data. a) Experimental y_2 (mCherry) data for Strain 1. b) Experimental y_2 (mCherry) data for Strain 2. c-d) Single well growth curves accounting for dilution of 1/2 at $t = 165$ and 1/2 at $t = 330$ (single wells (red), 0 nM estradiol well (blue), 100 nM estradiol well (green), average (black)). The flow cytometer sample volume changed at $t = 165$ minutes, hence the jump in cell count. There was also pipetting error from the deep well plate to the sample plate at $t = 0$ and 65 minutes resulting in error in cell count across wells. For $t = 165, 330$ and 580 the pipetting error was minimized. c) Growth curves for Strain 1. d) Growth curves for Strain 2.
(EPS)

S2 Fig. The majority of peak normalized dose-dependent distributions of Y_2 (YFP) for estradiol doses from 2.6 to 44.4 nM are close to gaussian. The distributions presented in ascending mean values are for ascending estradiol nM values of 2.60, 5.85, 8.78, 13.2, 19.8, 29.6, and 44.4, respectively. a) Peak normalized distributions of Y_2 (YFP) at 580 minutes. b) Close-up of peak normalized distributions at 580 minutes for low doses of estradiol. c) Peak normalized distributions of Y_2 (YFP) at 165 minutes. d) Close-up of peak normalized distributions at 165 minutes for low doses of estradiol.
(EPS)

S3 Fig. Comparison of alternative computational models for the transcriptional synthetic circuit demonstrate that only the slow global fluctuations model can recapitulate the experimental data. a-f) Dose-response relationships for $t = 65$ (blue), 165 (green), 330 (red), and 580 (black) minutes for alternative computational models. Insets in b, d and f represent the ratio of the noise at time t to that at time $t = 580$ where for green: $t = 165$, red: $t = 330$, and black: $t = 580$. Noise is defined as standard deviation over the mean. a)-b) Fast global fluctuations model. a) $y_{1,r}$ data. b) Normalized y_2 data. Each curve is normalized by its maximum mean value. c)-d) Intrinsic fluctuations model. c) $y_{1,r}$ data. d) Normalized y_2 data. Each curve is normalized by its maximum mean value. e)-f) Slow global fluctuations model with no estradiol dependent effects on $y_{1,r}$ production. e) $y_{1,r}$ data. f) Normalized y_2 data. Each curve is normalized by its maximum mean value.
(EPS)

S4 Fig. Time snapshots of mutual information, $I(y_2; y_{1,r}^*, t | x(t))$. The mutual information is applied to the simple pathway model with slow global model fluctuations for $\gamma_g = 3 \times 10^{-8}$ and $\bar{G} = 50$. Here $x(t)$ is a simple step function with amplitude x_+ . Therefore, the range of x_+ represents a range of different $x(t)$. Times are $t = 65$ (blue), 165 (green), 330 (red), and 750 (black) minutes.
(EPS)

S5 Fig. Comparison of SSA and moment equations for the synthetic circuit model ($\gamma_g = 3 \times 10^{-6}$ and initial condition S1). a) Plot of dose response from moment equations at time 580 minutes. b) Plots of the distribution $p(y_2, t = 580 | x(t)) / \max(p(y_2, t = 580 | x(t)))$ derived from the SSA and moment equations for select estradiol values up to 8.5 nM. c) Plots of the distribution $p(g, t = 580 | x(t)) / \max(p(g, t = 580 | x(t)))$ derived from the SSA and moment

equations. Analytical steady-state distribution shows that the moment equations and SSA results have evolved the global variable distribution $p(g, t = 580|x(t))$ about halfway to steady state.

(EPS)

Acknowledgments

We would like to thank members of the El-Samad lab, especially Rajat Bhatnagar, Kieran Mace, and Ben Heineke for critical reading of the manuscript. We would also like to thank Lucien Bogar for help in strain construction and flow cytometry measurements.

Author Contributions

Conceived and designed the experiments: MC OV. Performed the experiments: MC OV. Analyzed the data: MC OV HE. Contributed reagents/materials/analysis tools: MC OV. Wrote the paper: MC OV HE.

References

1. McAdams HH, Arkin A. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences*. 1997; 94(3):814–819. doi: [10.1073/pnas.94.3.814](https://doi.org/10.1073/pnas.94.3.814)
2. Kepler TB, Elston TC. Stochasticity in Transcriptional Regulation: Origins, Consequences, and Mathematical Representations. *Biophysical Journal*. 2001; 81(6):3116–3136. doi: [10.1016/S0006-3495\(01\)75949-8](https://doi.org/10.1016/S0006-3495(01)75949-8) PMID: [11720979](https://pubmed.ncbi.nlm.nih.gov/11720979/)
3. Swain PS, Elowitz MB, Siggia ED. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*. 2002; 99(20):12795–12800. doi: [10.1073/pnas.162041399](https://doi.org/10.1073/pnas.162041399)
4. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic Gene Expression in a Single Cell. *Science*. 2002; 297(5584):1183–1186. doi: [10.1126/science.1070919](https://doi.org/10.1126/science.1070919) PMID: [12183631](https://pubmed.ncbi.nlm.nih.gov/12183631/)
5. Shannon CE, Weaver W. *The Mathematical Theory of Communication*. Urbana, Illinois: University of Illinois Press; 1949.
6. Ziv E, Nemenman I, Wiggins C. Optimal Signal Processing in Small Stochastic Biochemical Networks. *PLoS ONE*. 2007; 2(10):e1077. doi: [10.1371/journal.pone.0001077](https://doi.org/10.1371/journal.pone.0001077) PMID: [17957259](https://pubmed.ncbi.nlm.nih.gov/17957259/)
7. Tkačik G, Callan CG, Bialek W. Information capacity of genetic regulatory elements. *Phys Rev E*. 2008 Jul; 78:011910. doi: [10.1103/PhysRevE.78.011910](https://doi.org/10.1103/PhysRevE.78.011910)
8. Tkačik G, Callan CG, Bialek W. Information flow and optimization in transcriptional regulation. *Proceedings of the National Academy of Sciences*. 2008; 105(34):12265–12270. doi: [10.1073/pnas.0806077105](https://doi.org/10.1073/pnas.0806077105)
9. Tkačik G, Walczak AM, Bialek W. Optimizing information flow in small genetic networks. *Phys Rev E*. 2009 Sep; 80:031920. doi: [10.1103/PhysRevE.80.031920](https://doi.org/10.1103/PhysRevE.80.031920)
10. Walczak AM, Tkačik G, Bialek W. Optimizing information flow in small genetic networks. ii. Feed-forward interactions. *Phys Rev E*. 2010 Apr; 81:041905. doi: [10.1103/PhysRevE.81.041905](https://doi.org/10.1103/PhysRevE.81.041905)
11. Mancini F, Wiggins CH, Marsili M, Walczak AM. Time-dependent information transmission in a model regulatory circuit. *Phys Rev E*. 2013 Aug; 88:022708. doi: [10.1103/PhysRevE.88.022708](https://doi.org/10.1103/PhysRevE.88.022708)
12. Lestas I, Vinnicombe G, Paulsson J. Fundamental limits on the suppression of molecular fluctuations. *Nature*. 2010; 467:174–178. doi: [10.1038/nature09333](https://doi.org/10.1038/nature09333) PMID: [20829788](https://pubmed.ncbi.nlm.nih.gov/20829788/)
13. Yu RC, Pesce CG, Colman-Lerner A, Lok L, Pincus D, Serra E, et al. Negative feedback that improves information transmission in yeast signalling. *Nature*. 2008; 456:755–761. doi: [10.1038/nature07513](https://doi.org/10.1038/nature07513) PMID: [19079053](https://pubmed.ncbi.nlm.nih.gov/19079053/)
14. Cheong R, Rhee A, Wang CJ, Nemenman I, Levchenko A. Information Transduction Capacity of Noisy Biochemical Signaling Networks. *Science*. 2011; 334(6054):354–358. doi: [10.1126/science.1204553](https://doi.org/10.1126/science.1204553) PMID: [21921160](https://pubmed.ncbi.nlm.nih.gov/21921160/)
15. Uda S, Saito TH, Kudo T, Kokaji T, Tsuchiya T, Kubota H, et al. Robustness and Compensation of Information Transmission of Signaling Pathways. *Science*. 2013; 341(6145):558–561. doi: [10.1126/science.1234511](https://doi.org/10.1126/science.1234511) PMID: [23908238](https://pubmed.ncbi.nlm.nih.gov/23908238/)

16. Colman-Lerner A, Gordon A, Serra E, Chin T, Resnekov O, Endy D, et al. Regulated cell-to-cell variation in a cell-fate decision system. *Nature*. 2005; 437:699–706. doi: [10.1038/nature03998](https://doi.org/10.1038/nature03998) PMID: [16170311](https://pubmed.ncbi.nlm.nih.gov/16170311/)
17. Paulsson J. Models of stochastic gene expression. *Physics of Life Reviews*. 2005; 2(2):157–175. doi: [10.1016/j.plrev.2005.03.003](https://doi.org/10.1016/j.plrev.2005.03.003)
18. Mugler A, Walczak AM, Wiggins CH. Information-Optimal Transcriptional Response to Oscillatory Driving. *Phys Rev Lett*. 2010 Jul; 105:058101. doi: [10.1103/PhysRevLett.105.058101](https://doi.org/10.1103/PhysRevLett.105.058101) PMID: [20867954](https://pubmed.ncbi.nlm.nih.gov/20867954/)
19. Voliotis M, Perrett RM, McWilliams C, McArdle CA, Bowsher CG. Information transfer by leaky, heterogeneous, protein kinase signaling systems. *Proceedings of the National Academy of Sciences*. 2014; 111(3):E326–E333. doi: [10.1073/pnas.1314446111](https://doi.org/10.1073/pnas.1314446111)
20. Toettcher J, Weiner O, Lim W. Using Optogenetics to Interrogate the Dynamic Control of Signal Transmission by the Ras/Erk Module. *Cell*. 2013; 155(6):1422–1434. doi: [10.1016/j.cell.2013.11.004](https://doi.org/10.1016/j.cell.2013.11.004) PMID: [24315106](https://pubmed.ncbi.nlm.nih.gov/24315106/)
21. Toettcher JE, Gong D, Lim WA, Weiner OD. Light-based feedback for controlling intracellular signaling dynamics. *Nat Meth*. 2011; 8:837–839. doi: [10.1038/nmeth.1700](https://doi.org/10.1038/nmeth.1700)
22. Toettcher JE, Gong D, Lim WA, Weiner OD. Chapter seventeen—Light Control of Plasma Membrane Recruitment Using the Phy-PIF System. In: Voigt C, editor. *Synthetic Biology, Part A*. vol. 497 of *Methods in Enzymology*. Academic Press; 2011. p. 409–423. Available from: <http://www.sciencedirect.com/science/article/pii/B9780123850751000172>
23. Bodenmiller B, Zunder ER, Finck R, Chen TJ, Savig ES, Bruggner RV, et al. Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nat Biotech*. 2012; 30:858–867. doi: [10.1038/nbt.2317](https://doi.org/10.1038/nbt.2317)
24. Engblom S. Computing the moments of high dimensional solutions of the master equation. *Applied Mathematics and Computation*. 2006; 180(2):498–515. doi: [10.1016/j.amc.2005.12.032](https://doi.org/10.1016/j.amc.2005.12.032)
25. Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*. 1976; 22(4):403–434. doi: [10.1016/0021-9991\(76\)90041-3](https://doi.org/10.1016/0021-9991(76)90041-3)
26. McQuarrie D. Stochastic approach to chemical kinetics. *Journal of Applied Probability*. 1967; 4(4):413–478. doi: [10.2307/3212214](https://doi.org/10.2307/3212214)
27. Gillespie DT. A rigorous derivation of the chemical master equation. *Physica A: Statistical Mechanics and its Applications*. 1992; 188(1-3):404–425. doi: [10.1016/0378-4371\(92\)90283-V](https://doi.org/10.1016/0378-4371(92)90283-V)
28. Shahrezaei V, Swain PS. Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences*. 2008; 105(45):17256–17261. doi: [10.1073/pnas.0803850105](https://doi.org/10.1073/pnas.0803850105)
29. Stewart-Ornstein J, Weissman JS, El-Samad H. Cellular Noise Regulons Underlie Fluctuations in *Saccharomyces cerevisiae*. *Molecular Cell*. 2012; 45(4):483–493. doi: [10.1016/j.molcel.2011.11.035](https://doi.org/10.1016/j.molcel.2011.11.035) PMID: [22365828](https://pubmed.ncbi.nlm.nih.gov/22365828/)