

Competing constraints shape the non-equilibrium limits of cellular decision making

Nicholas C. Lammers^a, Avi I. Flamholz^b, and Hernan G. Garcia^{a,c,d,e,f,1}

^aBiophysics Graduate Group, University of California at Berkeley, Berkeley, California; ^bPhysics Department, California Institute of Technology, Pasadena, California; ^cDepartment of Physics, University of California at Berkeley, Berkeley, California; ^dInstitute for Quantitative Biosciences-QB3, University of California at Berkeley, Berkeley, California; ^eDepartment of Molecular and Cell Biology, University of California at Berkeley, Berkeley, California; ^fChan Zuckerberg Biohub, San Francisco, California

1 **Gene regulation is central to cellular function. Yet, despite decades of work, we lack quantitative models that**
2 **can predict how transcriptional control emerges from molecular interactions at the gene locus. Thermodynamic**
3 **models of transcription, which assume that gene circuits operate at equilibrium, have previously been employed**
4 **with considerable success in the context of bacterial systems. However, the presence of ATP-dependent**
5 **processes within the eukaryotic transcriptional cycle suggests that equilibrium models may be insufficient to capture**
6 **how eukaryotic gene circuits sense and respond to input transcription factor concentrations. Here, we employ**
7 **simple kinetic models of transcription to investigate how energy dissipation within the transcriptional cycle**
8 **impacts the rate at which genes transmit information and drive cellular decisions. We find that biologically plausible**
9 **levels of energy input can lead to significant gains in how rapidly gene loci transmit information, but discover**
10 **that the regulatory mechanisms underlying these gains change depending on the level of interference from**
11 **non-cognate activator binding. When interference is low, information is maximized by harnessing energy to push**
12 **the sensitivity of the transcriptional response to input transcription factors beyond its equilibrium limits. Conversely,**
13 **when interference is high, conditions favor genes that harness energy to increase transcriptional specificity**
14 **by proofreading activator identity. Our analysis further reveals that equilibrium gene regulatory mechanisms break**
15 **down as transcriptional interference increases, suggesting that energy dissipation may be indispensable in systems**
16 **where non-cognate factor interference is sufficiently large.**

17 **Gene Regulation | Decision Theory | Non-equilibrium**
18 **Transcriptional Dynamics | Cell Signaling**

19 **Correspondence: hggarcia@berkeley.edu (HGG)**

20 Introduction

21 Throughout biology, systems must make accurate decisions under time constraints using noisy molecular machinery. Eukaryotic gene regulation exemplifies this challenge: genes must read out input concentrations of transcription factor proteins and respond by producing appropriate levels of gene product (mRNA and eventually protein) in order to drive downstream cellular decisions. Interestingly, the gene activity underlying cellular decision-making is often subject to large amounts of noise. Indeed, experiments across a wide range of organisms have revealed that eukaryotic transcrip-

22 tion is highly stochastic, occurring in episodic bursts (Bothma et al., 2014; Tantale et al., 2016; Nicolas et al., 2017; Lionnet and Wu, 2021)—periods of activity interspersed with periods of transcriptional silence—that unfold over timescales ranging from minutes to hours (Lammers et al., 2020). Because of this stochasticity, the transcription rate is a noisy reflection of transcription factor concentration. Over time, the accumulation of gene products tends to average out this noise, but biological processes must operate under time constraints: cells in developing fruit fly embryos have only minutes to determine their developmental fates (et al. Alberts B, Johnson A, Lewis J, 2002; Desponds et al., 2020), antigen recognition in T-cells unfolds over a single day (Obst, 2015), and cells in adult tissues are constrained by mRNA half-lives that range from minutes to days (Pérez-Ortín et al., 2013).

23 A key question, therefore, is how the molecular architecture of gene loci—the number and identity of biochemical steps in the transcriptional cycle and the reaction rates connecting these steps—dictates the amount of time needed for bursty gene expression to drive accurate cellular decisions. In particular, while it is widely accepted that processes within the eukaryotic transcriptional cycle consume biochemical energy (Coulon et al., 2013; Wong and Gunawardena, 2020), we do not yet know what non-equilibrium should “look like” in the context of transcriptional systems. Indeed, it remains challenging not only to predict unambiguous signatures of energy expenditure that can be detected experimentally (Hammar et al., 2014; Park et al., 2019; Eck et al., 2020), but also to establish how energy consumption can be harnessed to improve gene regulatory performance in the first place (Zoller et al., 2021).

24 Here, we use concepts from information theory and statistical physics as a lens to investigate how energy dissipation impacts the timescale on which gene circuits can drive cellular decisions. We consider a simple binary choice scenario wherein a cell must decide, as rapidly as possible, whether it is subjected to a high (c_1) or low (c_0) concentration of a transcriptional activator based on the transcriptional output of a gene locus. The basis for this decision is the gene’s input-output function (Figure 1A and B), which emerges from microscopic interactions between input activator molecules and their target gene loci (Figure 1C) that induce differences in the output dynamics of transcriptional burst-

ing (Figure 1D) for high and low activator concentrations. In turn, these differences in burst dynamics drive different rates of mRNA accumulation (Figure 1E). Because each ON/OFF fluctuation is stochastic, the resulting gene expression levels are noisy, and the cell must wait some time T before it is possible to accurately distinguish between c_1 and c_0 . Our central question in this work is whether energy dissipation within the molecular processes driving transcription allows gene loci to decrease the decision time, T , and, if so, how this performance gain manifests in terms of measurable features of the transcriptional input-output function.

There are multiple ways in which energy dissipation could alter the input-output behavior of a gene locus to improve cellular decision-making. As illustrated in Figure 1A and B, non-equilibrium processes could increase sensitivity to differences in input transcription factor concentration (“sharpness”) or suppress transcriptional noise (“precision”). Since our model assumes that, in addition to the concentration of the cognate activator, C , the gene locus is subject to some level of non-cognate factors, W , energy dissipation could also buffer against interference from off-target activation (“specificity”).

Recent works have begun to uncover a complex space of tradeoffs among these three aspects of transcriptional performance both at and away from thermodynamic equilibrium. A recent study found that systems operating at thermodynamic equilibrium suffer from strict tradeoffs between transcriptional specificity and transcriptional precision, but this tradeoff can be overcome by gene circuits that spend energy to enhance specificity through a scheme reminiscent of classical kinetic proofreading (Shelansky and Boeger, 2020; Ninio, 1975; Hopfield, 1974). Likewise, a separate study demonstrated that energy dissipation can enhance transcriptional sharpness (Estrada et al., 2016). Interestingly, while energy can increase sharpness and specificity *separately*, another study found that non-equilibrium levels of specificity come at the cost of sub-optimal sharpness (Grah et al., 2020). The authors also found that energy dissipation tends to *decrease* transcriptional precision, although this conclusion likely hinges on the study’s modeling assumptions (Grah et al., 2020). Despite this progress, it remains unclear how these non-equilibrium gains and tradeoffs ultimately impact how effectively gene circuits can harness differences in transcription factor concentrations to drive cellular decisions.

In this work, we identify a key quantity, the rate of information transmission from input transcription factor concentrations to output transcription rates as the quantitative link between energy-dependent changes in the transcriptional input-output function (Figure 1B) and the speed at which gene loci drive accurate biological decisions (Figure 1E) (Siggia and Vergassola, 2013; Deponds et al., 2020). We use this rate as a lens to examine the interplay between energy dissipation and cellular decision-making. We consider model gene cir-

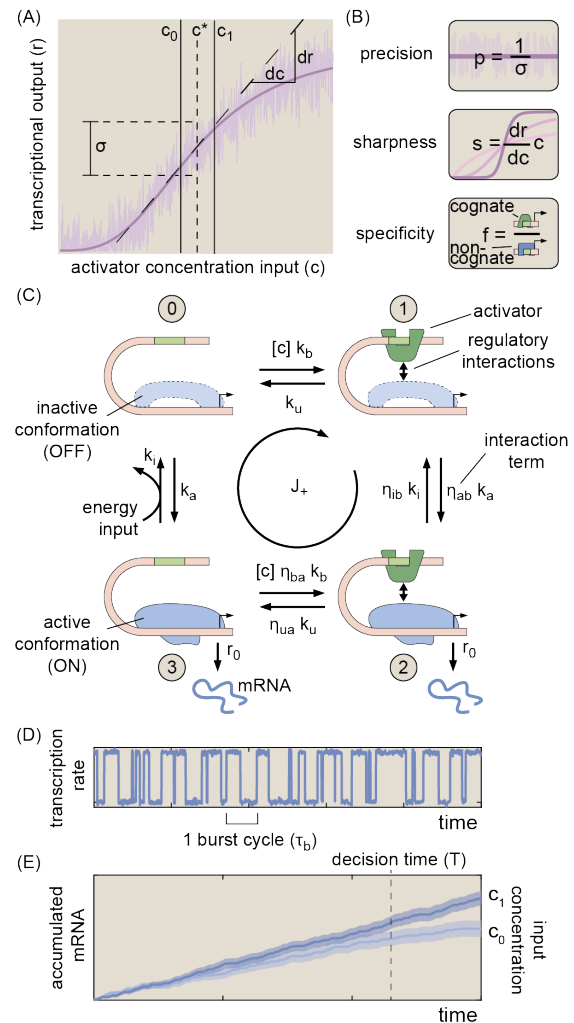


Figure 1. Three regulatory features shaping transcriptional information transmission. (A) Gene regulatory input-output function illustrating the basic biological problem considered in this work. Here, a cell must distinguish between two activator concentrations, c_0 and c_1 , based on the transcriptional output of a gene locus (purple curve). (B) We examine how three regulatory features of the transcriptional input-output function—sharpness, precision, and specificity—combine to dictate the rate at which the transcriptional output drives biological decisions. (C) Four-state MWC-like model (Phillips and Orme, 2020) of transcription used as the foundation of our investigations. Here, a single activator (green square) may bind to a specific site at the gene locus, and mRNA production occurs when the gene locus switches to its active (ON) conformation. A hypothetical energy input is depicted along the rate from state 3 to state 0. In practice, our framework permits non-equilibrium driving to occur along any of the eight transition rates in the model. (D) Simulated burst dynamics for one realization of the model shown in (C). Activator binding drives different burst dynamics at loci exposed to high and low activator concentrations. The burst cycle time is defined as the average time required to complete one ON \rightarrow OFF \rightarrow ON cycle and sets the timescale over which biological decisions unfold. (E) Illustrative simulation results for accumulated mRNA levels driven by c_1 and c_0 . Solid lines show trajectories for a single locus, and shaded regions indicate the standard deviation of levels taken across 100 simulated trajectories. The vertical dashed line indicates the “decision time,” when the expected mRNA levels driven by c_1 and c_0 are sufficiently different to permit an accurate decision about the input activator concentration.

cuits with varying numbers of activator binding sites. We also examine models with different numbers of molecular steps in the activation pathway, since transcriptional activation is also thought to require multiple molecular steps beyond activator binding itself, such as the localization of key transcription factors to the gene locus (No-gales et al., 2017).

We demonstrate that energy dissipation increases the rate at which genes can drive cellular decisions for all models considered. Moreover, the presence of multiple activation steps enables gene loci to more effectively harness energy to increase information transmission. At the level of the transcriptional input-output function (Figure 1A), while energy input can drive increases in all three regulatory features considered (sharpness, precision, and specificity; Figure 1B), genes cannot realize these non-equilibrium gains simultaneously. In particular, we show that the upper limit of information transmission is defined by a shifting tradeoff between sharpness and specificity. When the relative concentration of wrong-to-right activator species is small (e.g., in the fruit fly embryo), non-equilibrium gene circuits that maximize sharpness drive the fastest decisions. However, when the ratio of non-cognate to cognate activator concentrations is larger than the intrinsic difference in their binding affinities (e.g., in mammalian cells), gene circuits must instead prioritize transcriptional specificity.

In closing, we identify hallmarks of non-equilibrium gene regulation that may be amenable to experimental detection. We use our model to illustrate how simple point mutations in activator binding sites can lead to robust signatures of non-equilibrium regulatory processes. Additionally, our findings emphasize the importance of using theoretical models that account for non-cognate factor binding when interpreting experimental measurements of gene expression. Altogether, this work provides a rigorous foundation for interrogating the role of energy dissipation in eukaryotic gene circuit regulation.

Results

A. A simple model for probing the interplay between energy and information in transcription

We sought to establish gene circuit models that capture two essential characteristics of eukaryotic transcription. First, gene regulation hinges upon interactions between specific and general transcription factors. Although salient regulatory information tends to reside exclusively in a few specific transcription factors targeted to binding sites within enhancers (Vincent et al., 2016), these proteins are not sufficient to give rise to transcription. Instead, transcription and transcriptional control depend on interactions between specific regulatory factors and other key molecular players at the gene locus, such as mediators (Grah et al., 2020; Malik and Roeder, 2016; Rybakova et al., 2015; Kagey et al., 2010), RNA polymerase (Tantale et al., 2016), nucleosomes (Shelansky and Boeger, 2020; Mirny, 2010), and various sub-units of the pre-initiation complex (Nogales et al., 2017). While these factors do not themselves carry biological information, they constitute key molecular steps within the transcriptional cycle. This multiplicity of molecular players implies that gene loci may exist in multiple distinct molecular states corresponding

to different binding configurations of specific and general molecules (e.g., (Biddle et al., 2019)). Moreover, some of these processes—e.g., nucleosome displacement (Zhou et al., 2016), pre-initiation complex assembly (Taatjes, 2017), and RNA polymerase initiation (Yan and Gralla, 1997)—entail the dissipation of biochemical energy, opening the door to non-equilibrium behaviors.

Second, it has recently become apparent that eukaryotic transcription is characterized by stochastic, episodic bursts of activity interspersed with periods of transcriptional silence (Bothma et al., 2014; Fukaya et al., 2016; Little et al., 2013; Zoller et al., 2018; Tantale et al., 2016; Lammers et al., 2020). Since the concentration of specific transcription factors can regulate burst dynamics (Lammers et al., 2020; Zoller et al., 2018; Xu et al., 2015), a simple model would suggest that transcriptional bursts originate from the binding and unbinding of specific transcription factors. Although this may be the case in some yeast genes (Donovan et al., 2019), recent *in vivo* measurements in higher eukaryotes have revealed that activators and repressors typically bind DNA for seconds, rather than minutes or hours (Lammers et al., 2020; Lionnet and Wu, 2021). This temporal disconnect between bursting and transcription factor binding suggests a model in which transcriptional burst cycles—corresponding to OFF \rightarrow ON \rightarrow OFF fluctuations in the locus conformation (Figure 1D)—are not determined by transcription factor binding alone, but entail additional molecular reactions that are decoupled from the timescale of activator binding.

Together, these observations support a Monod–Wyman–Changeux (MWC)-like framework (Phillips and Orme, 2020; Grah et al., 2020; Shelansky and Boeger, 2020; Mirny, 2010) for modeling transcription wherein specific transcription factors act as effector molecules, conditioning the frequency with which the gene locus fluctuates between active and inactive transcriptional conformations. The simplest model that meets this description is one where a transcriptional activator binds to a single binding site at the gene locus, and where a second molecular reaction dictates fluctuations between two conformations: an inactive (OFF) state where no mRNA is produced and a transcriptionally active (ON) state where mRNA is produced at rate r_0 .

If we neglect the binding of non-cognate transcription factors, this leads to the model shown in Figure 1C. This model contains four basal reaction rates: the transcription factor binding and unbinding rates (k_b and k_{u1}) and the locus activation and deactivation rates (k_a and k_{i1}). We leave the molecular identity of this locus activation step unspecified, but in principle, it may be any of the elements of the general transcriptional machinery mentioned above. In addition to these basal rates, the η terms in Figure 1C capture interactions between the transcription factor and activation step. Here, the first subscript indicates which molecular reaction the η term modifies (binding or unbinding; activation or inactivation), and the second subscript indicates the molecule

performing the modification (bound activator “b” or activated molecular step “a”). For instance, η_{ab} encodes the degree to which the rate of locus activation is modified by having a transcription factor bound at the locus ($\eta_{ab} > 1$ corresponds to an activating transcription factor). Lastly, the average rate of mRNA production in this model is simply equal to $\bar{r} = r_0(\pi_2 + \pi_3)$, where π_i is the steady-state probability of finding the system in state i .

B. Calculating energy dissipation rates and decision times

At equilibrium, all state transitions in our model must obey the law of microscopic reversibility. Energy dissipation along one or more of the microscopic transitions shown in Figure 1C lifts this strict equilibrium constraint and opens the door to novel forms of non-equilibrium gene regulatory logic. For the model shown in Figure 1C, the energy dissipated per unit time (Φ) can be expressed as

$$\Phi = J \ln \frac{\eta_{ab}\eta_{ua}}{\eta_{ib}\eta_{ba}}, \quad (1)$$

where the η terms are defined in Figure 1C and the net cycle flux, J , encodes the degree to which microscopic transitions in the system are biased in the clockwise ($J > 0$) or counterclockwise ($J < 0$) direction (Hill, 1989). See Appendix A.5 for further details. Φ is a strictly positive quantity with units of $k_B T$ per unit time that indicates how “near” or “far” a system is from thermodynamic equilibrium (Hill, 1989; Lang et al., 2014). For ease of comparison across different realizations of our model gene circuit, we express Φ in units of $k_B T$ per burst cycle (“energy per burst”).

Our central aim is to understand how energy dissipation impacts the rate at which gene loci transmit information and drive cellular decisions. For simplicity, we assume that c_0 and c_1 are constant over time. We also stipulate that the difference between these concentrations (δc) is relatively small, such that $\delta c = c_1 - c_0 = 0.1c^*$, where c^* is the midpoint concentration $c^* = (c_1 + c_0)/2$. This value of δc is equivalent, for example, to concentration differences for the activator Bicoid between adjacent nuclei in early fruit fly development (Gregor et al., 2007). Figure 1E shows trends indicating the predicted integrated transcriptional output of a gene locus when it is exposed to high or low activator concentrations. Intuitively, it should be easier to distinguish between these two scenarios when (i) the difference between average transcript production rates (slope of the lines in Figure 1E) is large or (ii) the noise (shaded regions) in the accumulated output is small.

IR codifies this intuition, providing a quantitative measure of a gene’s ability to read out and respond to different input activator concentrations. Formally, IR is defined as the rate of change in the Kullback–Leibler divergence (Cover and Thomas, 2006) between our two hypotheses ($C = c_0$ and $C = c_1$) given the expected transcriptional output of our model gene circuit. If we take the noise in the transcriptional output to be ap-

proximately Gaussian (see Appendix B), IR can be expressed as

$$\text{IR} = \frac{1}{2} \underbrace{\left(\frac{\delta c}{c^*}\right)^2}_{\text{input}} \times \underbrace{s^2 p^2}_{\text{output}}, \quad (2)$$

where IR is strictly positive and has units of information per unit time and s and p are the sharpness and precision of the transcriptional response, respectively, as defined in Figure 1B. See Appendix C for a full derivation of this expression. We note that the native units of Equation 2 are natural log units (“nats”). For simplicity, we give all informational quantities in the more familiar “bits,” such that IR has units of bits per burst cycle (“bits per burst”). Additionally, the precision term, p , pertains solely to noise from intrinsic fluctuations between microscopic states at the gene locus and does not account for Poisson noise resulting from mRNA synthesis. In general, this noise is expected to be small relative to the noise from locus fluctuations for the parameter regimes considered (see Appendix D for details).

Equation 2 contains two terms: an input component that encodes the size of the activator concentration gradient and an output component that depends on the sharpness and precision of the transcriptional input-output function (Figure 1A and B). This expression provides quantitative support for the intuitions outlined above. IR can be increased both by increasing the difference between the transcription rates driven by c_1 and c_0 (i.e., increasing the sharpness) and by decreasing the noise level (i.e., increasing precision). Moreover, since both s and p can be calculated analytically from the microscopic reaction rates in our gene circuit (see Appendix A), Equation 2 allows us to calculate and compare information rates for gene circuits with different microscopic reaction rates.

The IR, in turn, dictates how rapidly cells can distinguish between the two activator concentrations, c_0 and c_1 , based on the accumulated transcriptional output of a gene circuit. Previous works (Siggia and Vergassola, 2013; Desponds et al., 2020) have established that the theoretical lower limit for the time required to distinguish between c_0 and c_1 is given by

$$\bar{T} = \ln \left(\frac{1 - \varepsilon}{\varepsilon} \right) \frac{1 - 2\varepsilon}{\text{IR}}, \quad (3)$$

where ε is the probability of being wrong, i.e., choosing c_1 when the true value is c_0 (or vice versa) (see Appendix E and (Desponds et al., 2020) for details). We note the error-tolerance ε in Equation 3 is extrinsic to the gene circuit model and depends on the nature of the downstream cellular processes. Unless otherwise noted, we follow (Desponds et al., 2020) and set $\varepsilon = 0.32$, equivalent to an error level of “1 sigma.”

C. Energy dissipation increases the rate of information transmission

Utilizing our framework, we investigated whether increasing the energy dissipated by our model gene circuit, Φ , increases the rate at which this circuit drives

cellular decisions between c_0 and c_1 . We expanded methods employed in (Estrada et al., 2016; Eck et al., 2020) to develop an algorithm capable of systematically exploring how different transition rates dictate gene circuit features. This algorithm can determine the maximum IR achievable by different realizations of our gene circuit as a function of energy dissipation. See Appendices F and G for details regarding its implementation and validation.

Figure 2A shows the relation between IR and Φ resulting from our numerical analysis. Here, each circle represents IR and Φ values for a single realization of our gene circuit (Figure 1C), as defined by its complement of transition rate values. Near equilibrium, our analysis reveals that gene circuits can transmit information no faster than 0.035 bits per burst (far left-hand side of Figure 2A). According to Equation 3, this means that the best equilibrium gene circuits require at least 110 burst cycles to drive a decision between concentrations c_1 and c_0 with an error probability of 32% when these concentrations differ by 10% (Figure 2B). In the developing fruit fly embryo (*D. melanogaster*), where the burst timescale (τ_b) is approximately 2 minutes (Lammers et al., 2020), this translates to a decision time of 3.7 hours, far too long to meet the time constraints imposed by early nuclear cleavage cycles (8–60 minutes (et al. Alberts B, Johnson A, Lewis J, 2002)). Our equilibrium gene circuit would require even longer times in adult nematode (*C. elegans*) and mouse (*M. musculus*) cells, where τ_b is much higher, with measurements ranging from 61 to 105 minutes ($\bar{T} \geq 112$ hours, (Lee et al., 2019)) and 30 minutes to multiple hours ($\bar{T} \geq 55$ hours, (Lammers et al., 2020)), respectively. In each case, these timescales likely exceed decision time limits imposed by mRNA decay or cellular division times, which set upper limits on the time over which gene output can be averaged (horizontal lines in Figure 2B; see Appendix H for further details).

Our analysis indicates that energy dissipation opens the door to improved information transmission, leading to a fourfold increase in the upper IR limit from 0.0035 to 0.014 bits per burst cycle (Figure 2A). Moreover, this performance gain is realized at biologically plausible levels of energy consumption: IR reaches its maximum non-equilibrium value at $\Phi \approx 20$ k_BT per cycle, which is approximately equivalent to the hydrolysis of one to two ATP molecules (Milo and Phillips, 2015). This corresponds to an energy-dependent decrease in decision time from 110 to 29 burst cycles (red shaded region in Figure 2B). This reduction meets the upper decision limit for mouse cells (Figure 2B). Yet there remains an absolute speed limit that no amount of energy dissipation can overcome, as shown by the empty space below the red non-equilibrium boundary in Figure 2B.

How can gene circuits do better? Real transcriptional systems are typically far more complex than the simple four-state model in Figure 1C; gene enhancers typically feature multiple transcription factor binding sites (Vin-

cent et al., 2016), and transcriptional activation depends on the combined action of multiple molecular components at the gene locus (Lammers et al., 2020). Thus, to overcome this speed limit, we must examine the impact of tuning two molecular “knobs”: the number of specific activator binding sites in our model (N_B) and the number of molecular steps required to achieve productive transcription (N_A). For simplicity, we focus on systems in which all binding sites are identical and assume identical kinetics for all molecular transitions between locus conformations. While restrictive, this simple approach gives rise to rich, biologically salient behaviors. While we explore the effects of varying N_B and N_A separately, these mechanisms are mutually compatible and may act jointly in real biological systems. See Appendix I for details regarding the implementation of these higher-order models.

Adding binding sites improves information-energy trade-

offs. We first examined the performance of gene circuit models with multiple binding sites. In these models (as with the four-state model described above), activator binding does not directly dictate transitions into and out of transcriptionally active molecular states, but instead increases or decreases the likelihood of these transitions. Models with multiple binding sites also permit cooperative interactions between activator molecules, encoded by $\eta_{u,b}$ terms (see Appendix I and Figure A9A). With these assumptions, we employed our parameter sweep algorithm to explore tradeoffs between the rate of energy dissipation (Φ) and the IR for systems with 1–5 activator binding sites. In all cases, we held the number of activation steps constant at $N_A = 1$ (as in Figure 1C).

As illustrated in Figure 2C, adding activator binding sites shifts the IR vs. Φ tradeoff boundary from Figure 2A upwards, allowing for higher information transmission rates for a given energy dissipation rate. This leads to significant IR gains, even in gene circuits operating near the equilibrium limit (vertical dashed line in Figure 2C), with the upper equilibrium limit increasing by approximately a factor of 25 from 0.0035 bits per burst cycle for $N_B = 1$ to 0.090 bits per cycle for $N_B = 5$. As a result, equilibrium gene circuits with 5 binding sites need as little as 5 burst cycles to distinguish between c_1 and c_0 , easily satisfying the decision time constraints of the biological systems shown in Figure 2B (Figure S1A). More generally, the lower decision time limit scales as the inverse of the number of binding sites squared ($\bar{T} \sim N_B^{-2}$, see Figure S1A).

Adding molecular activation steps allows gene circuits to harness higher rates of energy dissipation.

Next, we expanded the four-state model by changing the number of activation steps ($1 \leq N_A \leq 4$) while holding the number of binding sites fixed at $N_B = 1$ (top panel of Figure 2D). To illustrate this model, let us first consider the baseline case, where $N_A = 1$. Here, locus activation depends on the state of a single molecular component (e.g., mediator), which can be disengaged (i.e., the locus is OFF)

499 or engaged (i.e., the locus is ON). Now, consider model
500 in which locus activation also depends on the state of
501 a second molecular component (e.g. PIC assembly)
502 that can, likewise, be either engaged or disengaged. If
503 we stipulate that both components must be engaged to
504 achieve RNA polymerase initiation, then two molecular
505 activation steps are required to reach the ON state and
506 $N_A = 2$. We use the same logic to extend the model
507 to the $N_A = 3$ and $N_A = 4$ cases to capture the impact
508 of the additional molecular components necessary for
509 transcription. See Appendix I and Figure A9B for details.
510

511 We conducted parameter sweeps to examine the
512 interplay between energy dissipation and information
513 transmission for these systems. As with adding binding
514 sites, the addition of activation steps leads to in-
515 creased rates of information transmission. Unlike in-
516 creasing N_B , however, these IR gains do not come for
517 free. Instead, the addition of activation steps extends
518 the Φ -IR boundary into higher-energy regimes, allowing
519 non-equilibrium gene circuits to achieve larger gains in
520 IR at the expense of increased energy dissipation rates
521 (Figure 2D).

522 This increased IR gain means that systems with mul-
523 tiple activation steps can drive decisions between c_1
524 and c_0 more rapidly than the simple four-state gene cir-
525 cuit. For example, non-equilibrium gene circuits with
526 four activation steps can drive decisions nearly four
527 times as rapidly as systems with a single step (8 vs. 29
528 burst cycles; see Figure S1B). This 8-burst-cycle limit
529 approaches what can be achieved by an equilibrium
530 gene circuit with 5 activator binding sites (5 burst cycles;
531 compare Figure S1A and B), suggesting a similarity be-
532 tween adding activator binding sites at equilibrium and
533 adding activation steps out of equilibrium. However, this
534 parity has an energetic cost: to approach the perfor-
535 mance of the five-binding-site model, the one-binding-
536 site system with five conformations must dissipate at
537 least $180 k_B T$ per burst.

538 D. Increases in non-equilibrium sharpness improve 539 information transmission

540 According to Equation 2, the energy-dependent in-
541 creases in IR uncovered in Figure 2 must result from in-
542 creased sharpness, increased precision, or some com-
543 bination thereof. Thus, to uncover how energy reshapes
544 the transcriptional input-output function to increase IR,
545 we used our numerical sweep algorithm to examine the
546 space of achievable sharpness and precision values for
547 our baseline four-state model (Figure 1C) both at and
548 away from thermodynamic equilibrium. One challenge
549 in comparing sharpness and precision levels across dif-
550 ferent gene circuits is that the upper bounds on both
551 s and p depend on the fraction of time, π_a , the sys-
552 tem spends in the transcriptionally active conformation,
553 which changes as the transition rates vary between dif-
554 ferent realizations of our gene circuit. Thus, for ease
555 of comparison across different model realizations, we

556 give all results in terms of normalized sharpness and
557 precision measures: $S = s/b$ and $P = pb$, where $b =$
558 $\pi_a(1 - \pi_a)$ is the binomial variance in the occupancy of
559 the transcriptionally active conformation. These metrics
560 have intuitive interpretations: the S value of a particular
561 gene circuit's input-output function gives the Hill coeffi-
562 cient of an equivalently sharp Hill function, and P is in-
563 versely proportional to the level of intrinsic noise in the
564 transcriptional output. See Appendix J for details.

565 Figure 3A shows the results of our analysis, with each
566 circle representing the S and P values for a single gene
567 circuit realization. For systems operating at equilibrium
568 (blue dots in Figure 3A), we find that both S and P are
569 bounded by “Hopfield barriers” (dashed lines) (Hopfield,
570 1974; Estrada et al., 2016) with values of 1 and $1/\sqrt{2}$,
571 respectively. These bounds place strict limits on infor-
572 mation transmission at equilibrium and have a straight-
573 forward interpretation: they are precisely equal to the
574 sharpness and precision of a simple two-state gene cir-
575 cuit with a single activator binding site and no molecu-
576 lar activation step, where the ON rate is concentration-
577 dependent ($k_{on} \propto [c]$, see Appendix K for details).

578 Energy dissipation permits gene circuits to overcome
579 these equilibrium performance bounds, increasing S
580 by up to a factor of 2 and P by up to a factor of
581 $\sqrt{2}$ with respect to their equilibrium limits (Figure 3A).
582 Yet, while energy can improve sharpness and preci-
583 sion individually, the absence of realizable gene circuits
584 in the upper-right-hand corner of Figure 3A indicates
585 that genes cannot maximize both simultaneously. This
586 tradeoff places inexorable limits on the degree to which
587 energy can boost IR and—as illustrated in Figure 3B—
588 arises because maximally sharp and maximally precise
589 gene circuits require distinct and incompatible underly-
590 ing molecular architectures (see Appendix L for details).

591 Because sharpness and precision cannot be maxi-
592 mized simultaneously, gene circuits that dissipate en-
593 ergy must “choose” which aspect to maximize. From
594 the perspective of IR maximization, the choice is clear:
595 Figure 3A shows the location of 100 gene circuits within
596 1% of the maximum of 0.014 bits per cycle (Figure 2A)
597 in S – P phase space (gray circles). Thus, the most in-
598 formative gene circuits maximize transcriptional sharp-
599 ness ($S = 2$) at the cost of retaining equilibrium preci-
600 sion levels ($P = 1/\sqrt{2}$), which makes sense given that
601 non-equilibrium systems can boost S by up to a factor of
602 2 while P is limited to a maximum gain of $\sqrt{2}$. As with the
603 equilibrium case, these S and P values have an intuitive
604 interpretation: they are simply equal to the expected
605 sharpness and precision of a two-state system, one in
606 which both the ON and OFF rates are concentration-
607 dependent (see Appendix M). Thus, although spend-
608 ing energy to overcome the constraints of detailed bal-
609 ance opens up a vast new space of possible regulatory
610 schemes, maximally informative non-equilibrium gene
611 circuits exhibit an emergent simplicity, converging upon
612 architectures in which their many molecular degrees of
613 freedom collapse into a few effective parameters that

E Energy dissipation is required for rapid cellular decisions at high non-cognate factor concentrations

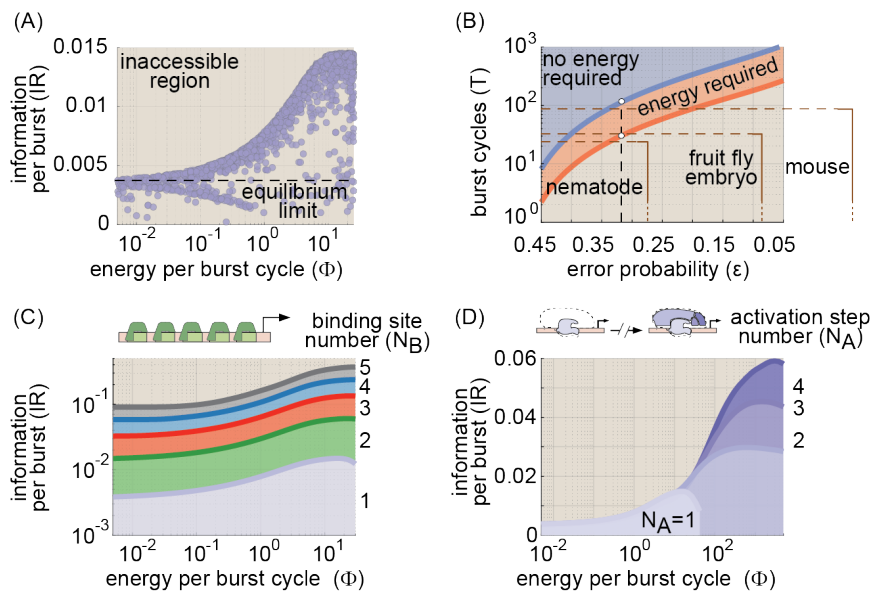


Figure 2. Energy dissipation increases the information transmission rate in gene circuits. (A) Information rate (IR from Equation 2) as a function of energy dissipation rate (Φ from Equation 1) for a parameter sweep exploring all possible model realizations. Modest energy dissipation rates can lead to a significant increase in the maximum amount of information that can be transmitted per burst cycle. (B) The amount of time needed to distinguish between c_0 and c_1 as a function of the probability of deciding incorrectly for equilibrium and non-equilibrium gene circuits. The decision time is given in terms of the number of transcriptional burst cycles required for a decision to be made. Note that the x-axis is arranged in order of *decreasing* error probability (i.e., increasing accuracy) from left to right. Horizontal lines indicate approximate upper bounds on decision times (in burst cycles) for different biological systems. (C) Parameter sweep results for achievable IR and Φ values for gene circuits with 1–5 activator binding sites. Achievable regimes for each molecular architecture are indicated as color-coded shaded regions. (D) Sweep results illustrating achievable IR vs. Φ regimes for gene circuits featuring 2–5 locus conformations. (For all parameter sweep results in A-D, transition rate and interaction term magnitudes, k and η , were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

614 define system behavior.

615 **Non-equilibrium gains in sharpness drive IR increases**
 616 **in more complex regulatory architectures.** To assess the
 617 generality of our results, we used our parameter sweep
 618 algorithm to examine equilibrium and non-equilibrium
 619 tradeoffs between sharpness and precision for more
 620 complex gene circuits with 2–5 activator binding sites
 621 and 2–4 molecular activation steps. In all cases, energy
 622 dissipation increases the upper limits of S and P, and as
 623 with our simple four-state model, these non-equilibrium
 624 performance gains cannot be realized simultaneously
 625 (Figure S2A and B). For all models considered, the
 626 gains in IR uncovered in Figure 2 are maximized by
 627 spending energy to increase sharpness, rather than
 628 precision (see Appendix N for further details). For the
 629 case of multiple activator binding sites ($N_B > 1$), the N_B -
 630 dependent increases in IR shown in Figure 2C arise be-
 631 cause increasing the number of binding sites increases
 632 the upper sharpness limit both at and away from equilib-
 633 rium (Figure S2A-C and Appendix N; (Grah et al., 2020;
 634 Estrada et al., 2016)).

635 More surprisingly, we find that increasing the number
 636 of molecular conformations (N_A) while holding the
 637 number of activator binding sites can increase transcrip-
 638 tional sharpness in systems operating out of equilib-
 639 rium. Figure 3C shows the range of achievable S val-
 640 ues for non-equilibrium systems as a function of N_A .
 641 The upper S limit scales linearly with N_A , such that
 642 $S_{\text{neq}} \leq N_A + 1$. This linear scaling is identical to the
 643 effect of adding activator binding sites at equilibrium,

644 where $S_{\text{eq}} \leq N_B$ (Figure S2C), providing intuition for why
 645 systems with multiple molecular steps can drive faster
 646 decisions: with respect to transcriptional sharpness, the
 647 regulation of multiple activation steps by a single bind-
 648 ing site in a non-equilibrium gene circuit is functionally
 649 equivalent to the effect of having multiple binding sites
 650 at equilibrium (Figure 3D).

651 **E. Energy dissipation is required for rapid cellular**
 652 **decisions at high non-cognate factor concentra-**
 653 **tions**

654 In real biological settings, cells do not contain only a
 655 single species of transcription factor. Therefore, to drive
 656 timely biological decisions, a gene circuit must not only
 657 sense and respond to its cognate transcription factor,
 658 but also efficiently filter out “irrelevant” signals from non-
 659 cognate factors. This process is inherently challeng-
 660 ing in eukaryotes, where short DNA-binding footprints
 661 lead to modest energetic differences between specific
 662 (correct) and non-specific (incorrect) transcription fac-
 663 tor binding events on the order of 4.6 k_BT (Maerkl and
 664 Quake, 2007), meaning that non-cognate transcription
 665 factors unbind from gene loci approximately 100-fold
 666 faster than cognate factors ($\alpha = k_{\text{off}}^w/k_{\text{off}} \approx 100$).

667 To understand whether this 100-fold difference in
 668 binding kinetics is sufficient to drive decisions in real bi-
 669 ological systems, we examined a stripped-down scenario
 670 in which cognate and non-cognate activators must com-
 671 pete to bind a single binding site (Figure 4A). We can
 672 quantify the severity of non-cognate factor interference
 673 by dividing the fraction of time the site is bound by a cog-

E Energy dissipation is required for rapid cellular decisions at high non-cognate factor concentrations

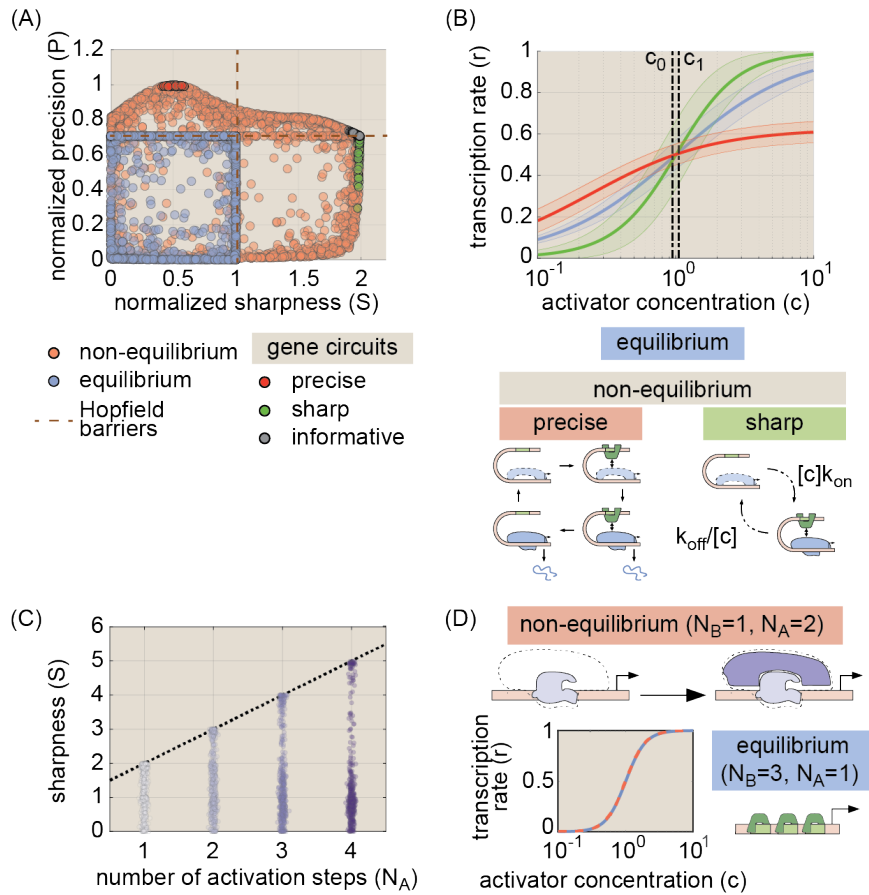


Figure 3. Increased transcriptional sharpness drives increased information transmission away from equilibrium. (A) Scatter plot of parameter sweep results showing the normalized sharpness and precision of 3,000 simulated gene circuits with and without equilibrium constraints. Energy expenditure overcomes Hopfield-like barriers, doubling the upper sharpness limit and increasing the precision limit by a factor of $\sqrt{2}$. The absence of gene circuits in the upper right quadrant indicates that no circuits can simultaneously maximize sharpness and precision. Calculations indicate that IR-maximizing systems (gray circles) spend energy to maximize non-equilibrium sharpness while maintaining precision at the maximum equilibrium level. (B) Illustrative input-output functions for a maximally informative equilibrium gene circuit (blue) from the parameter sweeps shown in (A) and maximally sharp and precise non-equilibrium gene circuits (green and red, respectively). The shaded region indicates predicted noise levels in gene expression patterns after 25 bursting cycles. Cartoons below illustrate molecular motifs for maximally precise and sharp non-equilibrium gene circuits. (C) Plot of achievable non-equilibrium sharpness levels for models with 2–5 locus conformations and one activator binding site. Each circle represents a single gene circuit model. Normalized sharpness is bounded by the number of locus conformations. (D) Cartoon illustrating functional equivalence between three binding sites at equilibrium and two activation steps out of equilibrium. The plot shows input-output functions for maximally sharp realizations of each case, demonstrating the equivalent sharpness levels driven by the two strategies. (For parameter sweep results in A and C, transition rate and interaction term magnitudes, k and η , were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

674 nate factor (π_c) by the total fraction of time it is bound by
 675 *either* the cognate or non-cognate species ($\pi_c + \pi_w$). If
 676 we assume equal basal binding rates (k_b) for cognate
 677 and non-cognate species, then the fraction of time the
 678 locus spends bound by a cognate transcription factor is
 679 given by

$$680 \quad p_c = \frac{\pi_c}{\pi_c + \pi_w} = \frac{f}{f + \frac{w}{c}}, \quad (4)$$

681 where we introduce a new quantity, the transcriptional
 682 specificity (f), defined as the (average) ratio of the
 683 probability of having cognate and non-cognate factors
 684 bound, normalized by the concentration, namely

$$685 \quad f = \frac{w}{c} \frac{\pi_c}{\pi_w}. \quad (5)$$

686 We note that Equation 5, which considers competition
 687 between two activator species to bind and activate a
 688 single gene, is distinct from and complements speci-
 689 ficity definitions employed in previous works, which ex-

690 amine the problem for a single activator species that
 691 regulates a cognate and a non-cognate locus (Shelansky
 692 and Boeger, 2020; Grah et al., 2020) (see Appendix
 693 O.1 for details).

694 From Equation 4, we see that f sets the scale for the
 695 severity of non-cognate factor interference. At equilib-
 696 rium, f is equal to the affinity factor α (see Appendix
 697 O.2), such that cognate factor binding dominates when
 698 $w/c < \alpha$ and non-cognate factors dominate when w/c
 699 exceeds α . For concreteness, we set $\alpha = 100$ through-
 700 out the remainder of this work. Where do actual bio-
 701 logical systems fall? A recent study pursuing synthetic
 702 enhancer design in the early fly embryo cited 47 per-
 703 tinent regulatory factors that were controlled to avoid
 704 off-target binding (Vincent et al., 2016), leading to an
 705 estimate of $w/c = 47$ (see also (Estrada et al., 2016)).
 706 Inserting this value into Equation 4, we predict that the
 707 cognate factor will be bound approximately 2/3 of the
 708 time in the fly embryo. At the other end of the spectrum,

709 we can use the genomic abundance of transcription factor
710 proteins to estimate upper bounds on w/c values
711 for adult nematode and mouse cells, yielding estimates
712 of $w/c \leq 698$ and $w/c \leq 1,426$, respectively (Charoen-
713 sawan et al., 2010). In this case, Equation 4 predicts
714 that cognate binding accounts for only a small fraction
715 of total binding interactions—as little as 1/8 in worms
716 and 1/15 in mice—suggesting that equilibrium affinity
717 differences alone may be insufficient in these cases.
718 To examine how these high interference levels impact
719 the timescale of biological decisions and to determine
720 whether energy dissipation can improve upon this equi-
721 librium baseline, we must extend our gene circuit model
722 to incorporate interference from non-cognate activator
723 binding.

724 To do this, we draw inspiration from (Cepeda-
725 Humerez et al., 2015), adding a second “wrong” acti-
726 vation cycle to our original four-state model (Figure 1C),
727 wherein the binding of a non-cognate factor to the gene
728 locus can also induce transitions to the active confor-
729 mation. This leads to the six-state model shown in Fig-
730 ure 4B, where, for simplicity, we have grouped all non-
731 cognate activators into a single concentration term: W .
732 Here, states 5 and 4 are identical to states 1 and 2,
733 except that a non-cognate activator species (blue circle)
734 is bound rather than the cognate activator (green square).
735 For notational convenience, we write the un-
736 binding rates of the non-cognate activator k_u^w as the un-
737 binding rate of the cognate factor k_u multiplied by an
738 affinity factor $\alpha = k_u^w/k_u$, with $\alpha = 100$.

739 We employed parameter sweeps to examine the up-
740 per limits on information transmission as a function
741 of the ratio of wrong-to-right activator concentrations
742 (w/c). We held the cognate factor concentration at
743 $C = c^*$, such that W was the only variable concentra-
744 tion parameter. Figure 4C presents the range of achiev-
745 able information rates as a function of the relative wrong
746 factor concentration. Our results reveal that the rate
747 of information transmission at equilibrium drops precipi-
748 tously once w/c exceeds α (blue circles in Figure 4C).
749 Away from equilibrium, the upper information limit like-
750 wise decreases with w/c ; however, we find that non-
751 equilibrium gene circuits are significantly more robust to
752 high non-cognate factor concentrations than equilibrium
753 systems. The relative IR gain from energy dissipation
754 increases from a factor of 4 when $w/c \approx 1$ to a factor
755 of 1,000 when $w/c = 10^5$ (Figure 4C, inset). This shift
756 in information gain suggests that a qualitative change
757 occurs in how energy is used once $w/c > \alpha$ (vertical
758 dashed line) (see Section F).

759 We next used Equation 3 to calculate the amount
760 of time required for a cell to decide between concen-
761 trations c_0 and c_1 of the cognate activator species for
762 different values of w/c , starting with gene circuits con-
763 strained to operate at equilibrium. As in Figure 2B, we
764 compared our model’s performance to the decision time
765 limits for different biological systems, this time with each
766 organism placed appropriately along the w/c axis. In

767 all organisms considered, gene circuits generally have
768 a few tens of burst cycles over which to transmit infor-
769 mation, with no organism exceeding 100 bursts (black
770 error bars in Figure 4D). This decision time limit is sig-
771 nificantly shorter than can be achieved by our simple
772 six-state model with one binding site and one activation
773 step at equilibrium, even in the presence of negligible
774 amounts of non-cognate transcription factor ($w/c = 1$,
775 purple shaded region corresponding to $N_B = 1$ in Fig-
776 ure 4D).

777 Next, we investigated the effect of having equilibrium
778 gene circuits with multiple sites. Figure 4D indicates
779 that equilibrium gene circuits with three or more acti-
780 vator binding sites (red, blue, and gray regions) are
781 sufficient to drive timely decisions in “low-interference”
782 systems such as the early fruit fly embryo. However,
783 we again observe a precipitous decline in performance
784 once $w/c > \alpha$. Indeed, the best equilibrium model
785 ($N_B = 5$) can drive decisions in no fewer than 1,100 burst
786 cycles—the equivalent of at least 550 hours (3 weeks)
787 for mouse cells—when $w/c \approx 1,400$ (the upper limit for
788 mice). This finding is over an order of magnitude too
789 slow for the mouse system’s decision time limit of 86
790 burst cycles (Figure 4D). Moreover, our analysis sug-
791 gests that at least 17 activator binding sites are needed
792 at equilibrium (see Figure S3A). Such a number is con-
793 ceivable for eukaryotic enhancers, but this analysis em-
794 phasizes that equilibrium systems—even those with bi-
795 ologically salient numbers of binding sites—struggle to
796 achieve realistic decision times in the presence of sig-
797 nificant non-cognate factor interference.

798 How do non-equilibrium gene circuits fare? The
799 dashed gray line in Figure 4D indicates the lower de-
800 cision time limit for *non-equilibrium* gene circuits with
801 five binding sites and one activation step. We ob-
802 serve a substantial improvement relative to the equilib-
803 rium case; however performance nonetheless suffers at
804 large values of w/c , falling short of the decision time
805 limit for the mouse system (209 vs. 86 burst cycles).
806 We used our parameter sweep algorithm to examine
807 the impact of increasing the number of molecular acti-
808 vation steps ($N_A > 1$) in non-equilibrium gene circuits
809 with a single activator binding site. This revealed sub-
810 stantial improvements, particularly at large w/c values.
811 Whereas the $N_A = 1$ system required at least 1,500
812 burst cycles when $w/c = 1,400$, gene circuits with two
813 activation steps can drive decisions between c_0 and c_1
814 in as little as 104 bursts (Figure 4E), a full order of
815 magnitude over equilibrium genes with five binding sites
816 and twice that of non-equilibrium gene circuits with five
817 binding sites and a single activation step (Figure 4D).
818 Adding a third step further improves this bound to 83
819 burst cycles, below the 86-burst limit for the mouse sys-
820 tem. Moreover, this $N_A = 3$ system exhibits remarkable
821 robustness to non-cognate factor interference, sustain-
822 ing the same level of performance up to $w/c \approx 10^4$ (Fig-
823 ure 4E).

824 These results suggest that, in biological contexts

where the ratio of wrong-to-right activator concentrations exceeds the intrinsic binding affinity difference (α), energy dissipation increasingly becomes a necessary precondition for driving cellular decisions within biologically salient timescales. Moreover, the presence of multiple molecular activation steps greatly amplifies non-equilibrium performance gains in these high-interference regimes. Yet Figure 4E also reveals that one-binding-site systems have a performance limit. To further improve, non-equilibrium gene circuits likely require multiple molecular steps ($N_A \geq 2$) and multiple activator binding sites ($N_B \geq 2$).

F. Non-cognate factor concentration defines performance tradeoffs between sharpness and specificity

Next, we investigated how much sharpness and precision each contribute to the IR gain depicted in the panel inset of Figure 4C. Figure 5A shows the relative non-equilibrium gains in S and P (S/S^{eq} and P/P^{eq}) as a function of w/c for information-maximizing realizations of the six-state gene circuit model shown in Figure 4B. The plot reveals that IR-maximizing gene circuits consistently utilize energy to drive sharpness above its equilibrium limit ($S/S^{eq} > 1$), while precision is maintained at or below its equilibrium limit ($P/P^{eq} \lesssim 1$). Moreover, the degree to which non-equilibrium gene circuits amplify S increases dramatically as w/c increases, from a factor of 2 when $w/c \approx 1$ to a factor of 100 when $w/c \approx 10^4$ (Figure 5A). Thus, the key to understanding how energy increases IR at large w/c values lies in understanding transcriptional sharpness.

The upper non-equilibrium limit on S can be expressed as a function of the specificity (f), such that

$$S \leq \underbrace{\frac{f}{\frac{w}{c} + f}}_{\text{specificity factor } (p_c)} \times \underbrace{S_0}_{\text{intrinsic sharpness}}, \quad (6)$$

where the observed sharpness (S) bound breaks naturally into two pieces: the specific bound fraction, p_c (defined in Equation 4), and the intrinsic sharpness (S_0), defined as a gene circuit's normalized sharpness absent non-cognate factor binding (i.e., $w = 0$).

To probe the interplay between intrinsic sharpness and specificity, we employed parameter sweeps for the six-state system in Figure 4B. At equilibrium, this analysis indicated that intrinsic sharpness is constrained such that $S_0 \leq 1$ (consistent with Figure 3A) and confirmed that specificity is fixed at α . Indeed, we find that $f^{eq} = \alpha$ applies for all gene circuits operating at equilibrium irrespective of the number of binding sites or activation steps, placing strict limits on information transmission at equilibrium when w/c is large (see Appendix O.3).

Away from equilibrium, systems can overcome these constraints, achieving up to a two-fold increase in S_0 and increasing specificity by up to an additional factor

of α to reach an upper limit of α^2 (Figure 5B). The observed 100-fold increase in f is comparable to the gain in the observed sharpness (S) in Figure 5A, suggesting that the sharpness gain at high w/c arises from non-equilibrium increases in specificity. Why not spend energy to simultaneously increase intrinsic sharpness by two-fold and specificity by 100-fold to achieve $S/S^{eq} = 2 \times \alpha = 200$? The simple answer is that non-equilibrium gains in intrinsic sharpness and specificity cannot be realized simultaneously. Instead, our analysis reveals a steep tradeoff between specificity and intrinsic sharpness away from equilibrium, with the maximum value of $S_0 = 2$ only realizable when specificity is at its equilibrium level ($f = \alpha$) and vice versa (Figure 5B). We find that the bound describing this tradeoff (black dashed line in Figure 5B) follows a simple analytic form, allowing us to express S as a function of the specificity, f , such that

$$S \leq \underbrace{\frac{f}{\frac{w}{c} + f}}_{\text{specificity factor } (p_c)} \times \underbrace{\left(\frac{\alpha^2 + \alpha f - 2f}{\alpha f - f} \right)}_{\text{intrinsic sharpness } (S_0)}, \quad (7)$$

where we assume that $\alpha \leq f \leq \alpha^2$. See Appendix P for a derivation of Equation 7. As with the non-equilibrium tradeoffs between sharpness and precision, this incompatibility stems from the fact that sharpness and specificity require distinct and incompatible underlying molecular architectures. Although we focused on the simple model shown in Figure 4B, we find similar non-equilibrium tradeoffs between f and S_0 for more complex molecular architectures (Figure S4B). Thus, we conclude that these specificity gains come at the cost of diminished intrinsic sharpness.

The inexorable tradeoff between the intrinsic sharpness S_0 and specificity f illustrated in Figure 5B means that gene loci must “choose” between allocating energy to maximize intrinsic sharpness and allocating energy to maximize specificity. To examine how the concentration of non-cognate factors shapes this tradeoff, we took IR-maximizing non-equilibrium gene circuits spanning the relevant range of w/c values for systems with 1–4 activation steps and calculated S_0 and f . Figure 5C illustrates the relative non-equilibrium gains in intrinsic sharpness and specificity, respectively, for these circuits as a function of w/c .

Figure 5C reveals that the relative non-cognate factor concentration, w/c , defines a shifting optimality landscape. At low non-cognate factor concentrations, maximally informative gene circuits spend energy exclusively to maximize intrinsic sharpness ($S_0/N_B > 1$ for all systems on the left-hand side of Figure 5C) at the cost of equilibrium specificity levels ($f/\alpha = 1$). Thus, our model predicts that at low levels of non-cognate factor interference—as would be experienced, for instance, in developing fruit fly embryos—non-equilibrium mechanisms are not required to buffer against non-cognate factor interference, and allocating energy to maximize

G Predicting experimental signatures of non-equilibrium processes in transcriptional regulation

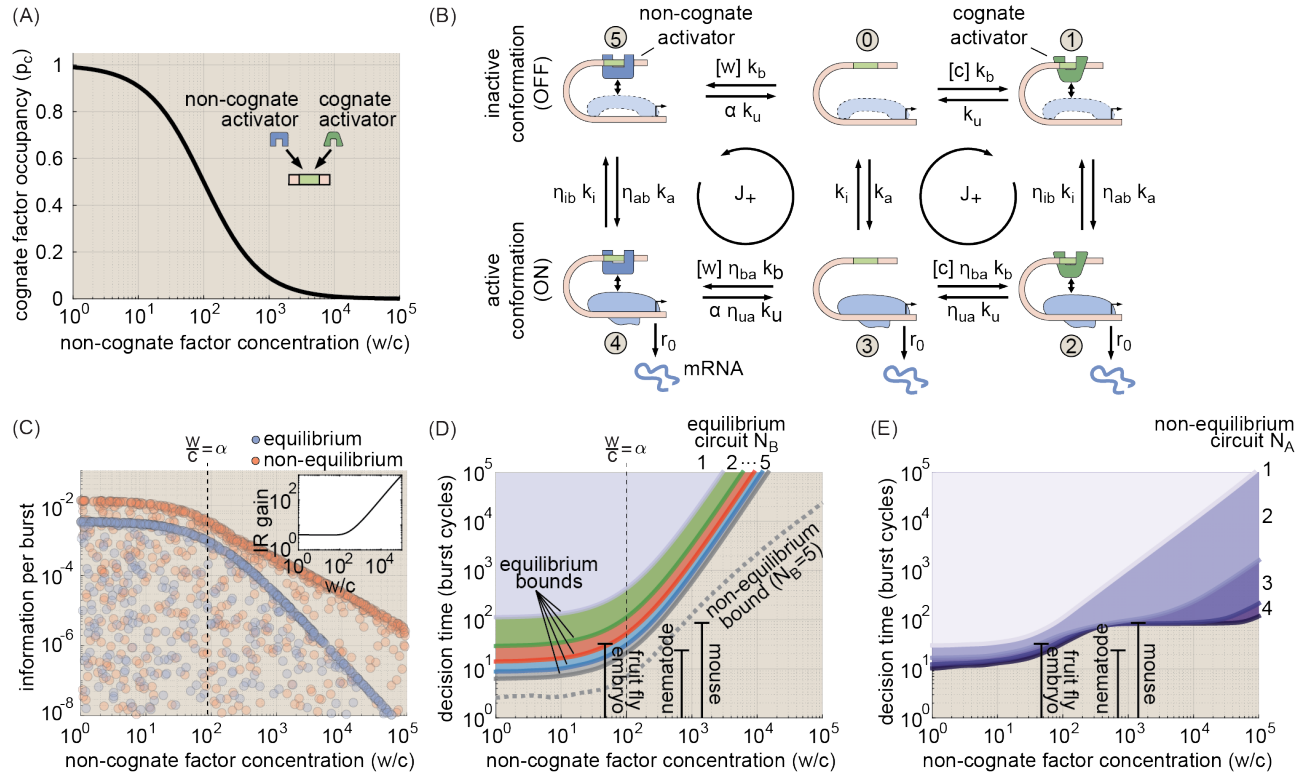


Figure 4. Energy dissipation is key to driving cellular decisions in the presence of non-cognate factor interference. (A) Cognate factor occupancy at a single binding site as a function of relative non-cognate factor concentration. (B) Incorporating non-cognate activator binding leads to a six-state model that features both a right and a wrong activation pathway. (C) Numerical results for the maximum achievable information rate for equilibrium (blue circles) and non-equilibrium (red circles) gene circuits with one activator binding site and one activation step (illustrated in (B)) as a function of the relative concentration of non-cognate activators w/c . The blue dashed line indicates the upper IR bound at equilibrium. The red line indicates the predicted non-equilibrium IR bound assuming quadratic scaling with w/c (see main text). The vertical dashed line indicates where the non-cognate factor concentration (w) equals the cognate factor concentration multiplied by the affinity factor (αc). Note how the optimal non-equilibrium systems begin to exceed the predicted bound beyond this point. The inset panel shows the non-equilibrium performance gain as a function of w/c . (D) Shaded regions indicate parameter sweep results for the range of achievable decision times for equilibrium gene circuits with 1–5 activator binding sites as a function of w/c . The dashed gray line indicates the lower bound for decision times driven by non-equilibrium gene circuits with five binding sites and one activation step. See Figure S3B for corresponding information rate ranges. (E) Decision times for non-equilibrium gene circuits with 1–4 activation steps. See Figure S3C for corresponding information rate ranges. (All results assume $\alpha = k_{uw}/k_{uw} = 100$. All decision time quantities assume $\varepsilon = 0.32$. For parameter sweep results in C–E, transition rate and interaction term magnitudes, k and η , were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

933 intrinsic sharpness constitutes the optimal regulatory
 934 strategy. However, once w/c surpasses the affinity
 935 factor α , IR maximization starts to disfavor sharpness
 936 (see decreasing S_0 near $w/c = 10^2$ in Figure 5C) and
 937 increasingly depends on enhancing specificity to non-
 938 equilibrium levels. Moreover, the presence of multiple
 939 activation steps dramatically increases the upper limit
 940 for non-equilibrium specificity, such that $f^{neq} \leq \alpha^{N_A+1}$
 941 (Figure S4B). Together, these results indicate that the
 942 optimal molecular strategy for transmitting information
 943 is not fixed, but changes according to a scale set by
 944 the relative amount of non-cognate factor interference,
 945 w/c , and the kinetic binding differences between cog-
 946 nate and non-cognate factors, α .

947 **G. Predicting experimental signatures of non-**
 948 **equilibrium processes in transcriptional regula-**
 949 **tion**

950 So far, we have demonstrated that energy dissipation
 951 can, in principle, increase the rate of information trans-
 952 mission in gene circuits. However, determining whether
 953 gene circuits mediating cellular decision making actu-

954 ally leverage energy dissipation to do so remains, to a
 955 large degree, an open challenge. Thus, we examined
 956 how simple experiments can identify signatures of non-
 957 equilibrium performance in real biological systems. For
 958 simplicity, we focused on the simple gene circuit in Fig-
 959 ure 4B with one binding site and one molecular activa-
 960 tion step, illustrating a broadly applicable set of exper-
 961 imental and analytical approaches that can be used to
 962 assess whether energy is harnessed to enhance trans-
 963 criptional performance in real biological systems.

964 Recent works have shown that strict equilibrium lim-
 965 its on transcriptional sharpness can be calculated if the
 966 number of activator binding sites is known, suggest-
 967 ing that sharpness might serve as an accessible signa-
 968 ture of non-equilibrium regulatory mechanisms (Estrada
 969 et al., 2016; Park et al., 2019). However, these studies
 970 did not consider off-target activation from non-cognate
 971 activator species. What happens when we account for
 972 the impact of such non-cognate factor binding? Equa-
 973 tion 7 predicts that the upper S limit should decrease as
 974 w/c increases (blue and red dashed lines in Figure 6A),
 975 as confirmed by numerical parameter sweeps of S vs.

G Predicting experimental signatures of non-equilibrium processes in transcriptional regulation

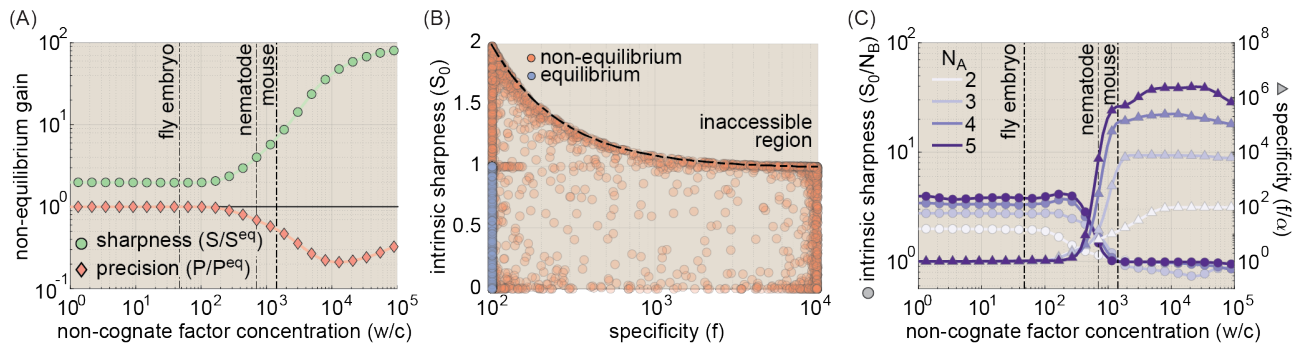


Figure 5. A shifting optimality landscape for information transmission. (A) Non-equilibrium gains in sharpness and precision as a function of w/c for six-state ($N_B = 1, N_A = 1$; Figure 4B) gene circuits found to drive maximum information rates. IR-maximizing gene circuits are drawn from optimal systems uncovered in the parameter sweeps from Figure 4E. Values above 1 indicate that the system is dissipating energy to enhance performance. The black line indicates a “break-even” point where the non-equilibrium value is equal to the equilibrium maximum. See Figure S4A for results for systems with $N_A > 2$. (B) Tradeoffs between intrinsic sharpness (S_0) and specificity (f) for equilibrium and non-equilibrium networks (blue and red circles, respectively). Note that equilibrium gene circuits have no horizontal dispersion because all are constrained to have $f = \alpha$. The black dashed line indicates the bound predicted by Equation 7. (C) Non-equilibrium gains in intrinsic sharpness and specificity for IR-maximizing gene circuits as a function of w/c . Values above 1 indicate that the system is dissipating energy to enhance sharpness or specificity. Note that the left and right axes have different scales. (α was set to 100 for all plots shown. For all parameter sweep results in A-C, transition rate and interaction term magnitudes, k and η , were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

976 w/c (blue and red circles). Thus, the upper sharpness
 977 limit is not absolute, but instead depends on the
 978 concentration of non-cognate factors in the cellular environment.
 979 This w dependence must be considered to accurately
 980 interpret experimental measurements.

981 For instance, consider the case where $w/c = 10^3$
 982 (black dashed vertical line in Figure 6A), a plausible
 983 value for mammalian systems (Friedlander et al., 2016;
 984 Cepeda-Humerez et al., 2015; Charoensawan et al.,
 985 2010). Our model predicts that the maximum achievable
 986 S for non-equilibrium gene circuits is 0.91, far exceed-
 987 ing the *true* equilibrium sharpness limit of 0.09 when
 988 accounting for the effects of non-cognate factor inter-
 989 ference (blue dashed line in Figure 6A). However, $S = 0.91$
 990 falls *below* the “naive” equilibrium bound of $S = 1$ that
 991 one would predict if w were not accounted for (see blue
 992 bound on far-left-hand side of Figure 6A, see also Fig-
 993 ure S5A). Thus, failing to account for non-cognate factor
 994 interference could mask strong non-equilibrium signa-
 995 tures, highlighting the importance of incorporating reg-
 996 ulatory cross-talk into transcription models. However,
 997 accurately measuring w/c may be challenging in many
 998 experimental settings, since w comprises the aggregate
 999 activity of all non-cognate activator species.

1000 In light of this challenge, we propose a complemen-
 1001 tary experimental approach to search for signatures of
 1002 non-equilibrium gene regulation that is more robust to
 1003 uncertainty regarding the precise value of w/c . As il-
 1004 lustrated in Figure 6B, this method involves measur-
 1005 ing changes in gene expression at $C = c^*$ that result
 1006 from point mutations to the activator binding site, which
 1007 thereby lead to a higher unbinding rate, k_u^{mut} , for cog-
 1008 nate activators ($k_u^{mut}/k_u > 1$). Whereas w/c may be
 1009 difficult to estimate in many biological contexts, robust
 1010 algorithms can predict changes in binding energies from
 1011 the DNA sequence of transcription factor binding sites
 1012 (Le et al., 2018), allowing for accurate predictions of
 1013 how much a particular mutation will perturb the relative

1014 binding kinetics of a specific activator species. We em-
 1015 ploy two metrics to quantify the resulting change in gene
 1016 expression: fold changes in the mRNA production rate
 1017 (\bar{r}^{mut}/\bar{r}) and in the normalized sharpness (S^{mut}/S),
 1018 each defined as the quantity corresponding to the mu-
 1019 tated binding site divided by its corresponding wild-type
 1020 value (Figure 6B).

1021 To illustrate the method, we used our model to pre-
 1022 dict outcomes for the case where the wild-type gene cir-
 1023 cuit is expressing at half its maximum rate ($\bar{r} = 0.5r_0$).
 1024 Overall, we find that IR-optimized non-equilibrium gene
 1025 circuits are highly sensitive to changes in cognate acti-
 1026 vator specificity and that this sensitivity can be used
 1027 to probe for non-equilibrium behavior. At low w/c lev-
 1028 els ($w/c \lesssim 10^3$), mutated non-equilibrium circuits ex-
 1029 hibit larger shifts in their transcription rate than can
 1030 be achieved at equilibrium (Figure S5B). Meanwhile,
 1031 when $w/c > 10^3$, IR-optimized non-equilibrium systems
 1032 experience a substantially larger sharpness decrease
 1033 than even maximally sensitive equilibrium circuits (Fig-
 1034 ure S5C). Consequently, when combined, S^{mut}/S and
 1035 \bar{r}^{mut}/\bar{r} define a perturbation response space in which
 1036 non-equilibrium gene circuits that transmit information
 1037 at optimal (or near-optimal) levels are completely dis-
 1038 joint from equilibrium systems. This is illustrated in
 1039 Figure 6C, which plots our model’s predictions for the
 1040 sharpness fold change (S^{mut}/S) vs. \bar{r}^{mut}/\bar{r} for three
 1041 binding site perturbation strengths for equilibrium and
 1042 non-equilibrium gene circuits (squares and circles, re-
 1043 spectively). Despite the wide range of perturbation
 1044 strengths and non-cognate factor concentrations exam-
 1045 ined, optimal non-equilibrium systems never cross the
 1046 equilibrium boundary (dashed line). Thus, by measur-
 1047 ing S^{mut}/S and \bar{r}^{mut}/\bar{r} , we can obtain clear-cut signa-
 1048 tures on non-equilibrium regulation, even when w/c is
 1049 unknown.

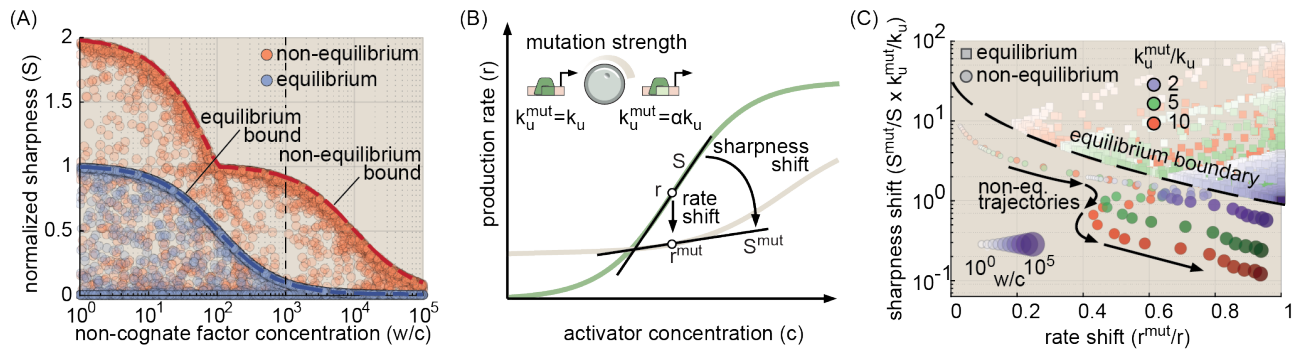


Figure 6. Experimental signatures of non-equilibrium processes in transcriptional regulation. (A) Observed sharpness as a function of w/c for equilibrium (blue circles) and non-equilibrium (red) gene circuits. The black dashed line indicates the point where $w/c = 10^3$. (B) Illustration of proposed binding site perturbation experiments. Reducing site specificity is predicted to reduce both the observed sharpness, S , and the mRNA production rate, \bar{r} . The strongest possible perturbation would entail a conversion from cognate specificity (k_{u}) to non-cognate specificity (αk_{u}). (C) Phase-space plot of predicted sharpness shift (normalized by $k_{\text{u}}/k_{\text{u}}^{\text{mut}}$) versus rate shift for equilibrium (squares) and non-equilibrium (circles) gene circuits at three binding site perturbation strengths. Note that we normalize the sharpness fold change by $k_{\text{u}}/k_{\text{u}}^{\text{mut}}$, which allows us to plot results for different mutation strengths on the same y-axis. Shading indicates the w/c value (darker shades correspond to higher values). Additionally, the circle size indicates the w/c magnitude for non-equilibrium circuits. We see that, regardless of non-cognate concentration and perturbation strength, non-equilibrium systems do not cross the equilibrium boundary (dashed line). Results assume the initial transcription rate of the wild-type gene is at half-maximum ($\bar{r} = 0.5r_0$). (For all parameter sweep results in A and C, transition rate and interaction term magnitudes, k and η , were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

Discussion

Gene regulation is central to cellular function. Yet, despite decades of biochemical and genetic studies that have established a reasonably complete “parts list” of the molecular components driving eukaryotic transcription (Kornberg, 2007), and despite recent advances in our ability to track how these pieces assemble in space (Nogales et al., 2017) and time (Lammers et al., 2020; Coulon et al., 2013; Lenstra et al., 2016), we nonetheless lack quantitative models that can predict how transcriptional control emerges from molecular interactions at the gene locus. Thermodynamic models of transcription, which assume that gene circuits operate at equilibrium, have been employed with considerable success to predict transcriptional control in the context of bacteria (Phillips et al., 2019). However, the presence of ATP-dependent processes—such as chromatin remodeling (Zhou et al., 2016), pre-initiation complex assembly (Taatjes, 2017), and Pol II initiation (Yan and Gralla, 1997)—within the eukaryotic transcriptional cycle suggests that equilibrium models may be insufficient to capture how eukaryotic gene circuits sense and respond to input transcription factor concentrations. Thus, there is an urgent need for theoretical frameworks that can probe how non-equilibrium mechanisms reshape the transcriptional input-output function and, ultimately, redefine the limits of transcriptional control.

Here, we employed simple kinetic models of transcription to investigate how energy dissipation within the transcriptional cycle impacts the rate at which a gene circuit drives cellular decisions. We found that biologically plausible rates of energy dissipation can drive significant gains in the information transmission rate and discovered that the regulatory mechanisms underlying these non-equilibrium gains change from increased sharpness to increased specificity depending on the level of interference in the cellular environment

from non-cognate factor binding.

Performance tradeoffs dictate limits of information transmission away from equilibrium. This work has established that, although energy dissipation can increase transcriptional sharpness, precision, and specificity *individually*, these gains cannot be realized simultaneously. For negligible non-cognate factor binding, we showed that IR is dictated by a tradeoff between sharpness (S) and precision (P). Although previous works have established that energy expenditure can boost sharpness (Estrada et al., 2016; Park et al., 2019) and, to a lesser extent, suppress transcriptional noise (Rieckh and Tkačik, 2014). As a result of this tradeoff, gene circuits must “choose” whether to spend energy to enhance sharpness or precision. For all models considered, we discovered that the information rate was maximized by systems that boosted transcriptional sharpness (not precision) above its equilibrium limit (Figure 3A, Figure S2A and B).

Similarly, our analysis revealed that non-equilibrium gains in specificity and sharpness cannot occur simultaneously (Figure 5B and Figure S4B). This incompatibility arises from the fact that intrinsically sharp systems are tuned to amplify concentration-dependent activator binding rates, whereas specific systems amplify differences in *unbinding* rates between cognate and non-cognate activator species. Our model predicts that w/c defines a shifting optimality landscape, wherein non-equilibrium gene circuits that maximize intrinsic sharpness drive the fastest decisions when $w/c \leq \alpha$, but the optimal strategy begins to shift from increasing sharpness to activator proofreading when $w/c > \alpha$ (Figure 5C). A recent study reported the potential for this kind of context-dependent shift from sharp to specific gene circuits (Grah et al., 2020), although sharpness was only investigated at its equilibrium limit. Here, we provide quantitative predictions for how IR-maximizing

1124 gene circuits navigate this sharpness-specificity trade-
1125 off far from equilibrium.

1126 **Activation steps amplify non-equilibrium performance**
1127 **gains.** Another key finding of this work is that the
1128 presence of multiple activation steps, wherein multiple
1129 molecular components must engage to achieve transcrip-
1130 tion, can amplify non-equilibrium gains in transcrip-
1131 tional sharpness (Figure 3C). Our result is evocative
1132 of a recent study (Biddle et al., 2020) demonstrating
1133 that systems with multiple conformational degrees of
1134 freedom can achieve sharper, more flexible transcrip-
1135 tional input-output functions, although these systems
1136 still adhere to the fundamental equilibrium limitation that
1137 sharpness cannot exceed the number of activator bind-
1138 ing sites ($S \leq N_B$). Thus, our findings further empha-
1139 size potential benefits of the conformational complexity
1140 of the eukaryotic gene cycle.

1141 Consistent with previous results in the kinetic proof-
1142 reading literature (Murugan et al., 2012), we also found
1143 that gene circuits with multiple activation steps can real-
1144 ize dramatic increases in transcriptional specificity when
1145 driven out of equilibrium (f), such that $f \leq \alpha^{N_A+1}$ (Fig-
1146 ure S4B). This result extends the findings of a recent
1147 work examining transcriptional specificity in systems
1148 with up to two activation steps (Shelansky and Boeger,
1149 2020). Yet there exists an important asymmetry be-
1150 tween sharpness and specificity: whereas the addition
1151 of activator binding sites can increase the sharpness S
1152 at equilibrium, energy dissipation constitutes the *only*
1153 route (short of altering activator binding sequences) for
1154 increasing specificity f above the intrinsic affinity factor
1155 α . Thus, for large w/c , energy dissipation overcomes a
1156 fundamental limitation of eukaryotic gene circuits—the
1157 lack of binding specificity—that no equilibrium mecha-
1158 nism can address.

1159 **Equilibrium regulatory schemes may be sufficient in**
1160 **many real biological systems.** While activator proofread-
1161 ing may be critical when w/c is large, our analysis sug-
1162 gests that it is unlikely to constitute a universal con-
1163 straint on gene regulatory architectures. Indeed, even
1164 relatively simple equilibrium architectures with 3–5 bind-
1165 ing sites should suffice to drive timely cellular deci-
1166 sions in “low-interference” systems such as the fruit fly
1167 embryo (Figure 4D). Moreover, while simple estimates
1168 based on genomic transcription factor abundances sug-
1169 gest that many eukaryotic systems can exceed the
1170 $w/c = \alpha$ interference limit, these estimates likely rep-
1171 resent upper bounds on w/c , since different cell types
1172 selectively express distinct subsets of transcription fac-
1173 tors (Choudhury and Ramsey, 2016; Lee et al., 2012;
1174 Henry et al., 2012). In addition, we note that the relative
1175 size of the concentration difference between c_1 and c_0
1176 ($\delta c/c$) plays a key role in dictating the information trans-
1177 mission rate (Equation 2) and varies across different bi-
1178 ological contexts. Thus, it would be interesting to use
1179 the quantitative tools presented in this work to enumer-
1180 ate the space of viable equilibrium and non-equilibrium

1181 gene circuit architectures for specific biological systems
1182 in which the relative magnitudes of w/c and $\delta c/c$ are
1183 well established.

1184 **Different frameworks for examining the impact of**
1185 **non-cognate factor binding.** In considering the impact of
1186 non-cognate factor binding, we drew inspiration from
1187 a previous study examining competition between cog-
1188 nate and non-cognate transcription factors to bind and
1189 activate a single gene locus (Cepeda-Humerez et al.,
1190 2015). This formulation of the problem is distinct from
1191 the approach taken in two recent works, which ad-
1192 dressed the problem of specificity from the perspec-
1193 tive of a single activator species that interacts with two
1194 different gene loci: a cognate (with specific binding
1195 sites) and a non-cognate locus (without specific bind-
1196 ing sites) (Shelansky and Boeger, 2020; Grah et al.,
1197 2020). While both approaches have proven fruitful, we
1198 favor the “single-locus” approach, since it captures the
1199 effects of competitive binding between different species,
1200 which are an unavoidable reality of crowded cellular en-
1201 vironments.

1202 Moreover, this shift in perspectives has meaningful
1203 consequences for our understanding of how off-target
1204 binding impacts gene regulation. A previous study
1205 found that the equilibrium limit of $f = \alpha$ could only be
1206 achieved at the cost of high levels of transcriptional
1207 noise (Shelansky and Boeger, 2020). Yet, we find that
1208 this tradeoff evaporates once competitive binding be-
1209 tween cognate and non-cognate factors is considered,
1210 since f is fixed at α in this case (Figure 5B). The upper
1211 limits of transcriptional sharpness also decrease as w/c
1212 increases (Equation 7 and Figure 6A). Previous studies
1213 have reported transcriptional sharpness as a key po-
1214 tential indicator of non-equilibrium optimization (Estrada
1215 et al., 2016; Park et al., 2019). Our analysis reaffirms
1216 this idea but, crucially, reveals that one must consider
1217 the relative concentration of non-cognate factors (w/c)
1218 to accurately assess whether a particular system is per-
1219 forming above the equilibrium limit (Figure 6A and B).
1220 For instance, a sharpness of 0.9 falls below the equi-
1221 librium limit for the six-state gene circuit shown in Fig-
1222 ure 4B when $w/c \approx 1$, but is an order of magnitude
1223 above the limit when $w/c \approx 10^3$ (Figure 6A).

1224 **Future directions.** While we have considered gene loci
1225 with varying numbers of *specific* activator binding sites,
1226 real enhancers also contain significant stretches of
1227 “neutral” DNA with no binding sites, as well as weak
1228 activator sites that fall below typical thresholds used
1229 to identify specific sites (Vincent et al., 2016; Shahein
1230 et al., 2021). This focus on specific sites is widespread
1231 in theoretical studies of transcription (Estrada et al.,
1232 2016; Park et al., 2019; Cepeda-Humerez et al., 2015;
1233 Lammers et al., 2020), despite the well-established im-
1234 portance of weak binding sites in the context of certain
1235 genes (Shahein et al., 2021; Crocker et al., 2015; Far-
1236 ley et al., 2015). Moreover, recent efforts on synthetic
1237 enhancer reconstitution have pointed to the importance

1238 of supposedly neutral stretches of regulatory DNA (Vincent et al., 2016), and it seems theoretically plausible
1239 that these stretches, where cognate and non-cognate
1240 activator species bind with equal affinity, could have important effects on the input-output function in systems
1241 when $w/c > \alpha$. We propose that the kinetic models utilized herein could readily be extended to feature some
1242 combination of specific and neutral sites. More ambitiously, the field would benefit from the introduction of
1243 continuous, rather than discrete, theoretical models that admit non-equilibrium phenomena while accounting for
1244 the reality that transcription factors interact with a continuum of sites along enhancer DNA.

1251 Ultimately, the key to unraveling the molecular mechanisms by which genes sense and respond to transcrip-
1252 tion factor concentrations lies in the coupling of theoretical models with careful experimental measurements.
1253 To this end, we advocate for the expanded use of theoretically tractable synthetic enhancer systems in which
1254 the number and identity of binding sites are well established and intervening DNA sequences are carefully
1255 engineered to minimize binding specificity (e.g., using SiteOut (Estrada et al., 2016)). Several recent studies
1256 constitute promising initial steps in this direction (Reimer et al., 2021; Park et al., 2019; Vincent et al.,
1257 2016; Kim et al., 2021). Additionally, synthetic transcription factor systems, which can act orthogonally to en-
1258 dogenous regulatory networks, represent an intriguing experimental platform for investigating questions relating
1259 to transcriptional specificity (Kabadi and Gersbach, 2014; Crocker and Stern, 2013). Lastly, statistical meth-
1260 ods that infer how transcription factor concentrations impact the kinetics of the transcriptional cycle (Zoller
1261 et al., 2018; Lammers et al., 2020; Corrigan et al., 2016; Bowles et al., 2022) hold promise for connecting macro-
1262 scopic experimental measurements to theoretical models of the microscopic processes driving transcription.
1263 Looking ahead, holistic research efforts that integrate cutting-edge experiments, statistical methods, and the-
1264 ory will be key to bridging the as yet yawning gap between enhancer sequence and gene regulatory func-
1265 tion.

1280 ACKNOWLEDGEMENTS

1281 We are grateful to Jane Kondev, Sara Mahdavi, and Vahe Galstyan for substan-
1282 tial comments and discussion on the manuscript. Thanks also to Rob Phillips,
1283 Muir Morrison, and Ben Kuznets-Speck for their helpful discussion and insights
1284 at various stages of this project's development. NCL was supported by NIH
1285 Genomics and Computational Biology training grant 5T32HG000047-18, the
1286 Howard Hughes Medical Institute, and by DARPA under award number N66001-
1287 20-2-4033. AIF was supported in part by an NSF Graduate Research Fellow-
1288 ship, NSF Grant No. PHY-1748958, the Gordon and Betty Moore Foundation
1289 Grant No. 2919.02, the Kavli Foundation, and by a Postdoctoral Fellowship
1290 from the Jane Coffin Childs Memorial Fund for Medical Research. HGG was
1291 supported by the Burroughs Wellcome Fund Career Award at the Scientific In-
1292 terface, the Sloan Research Foundation, the Human Frontiers Science Program,
1293 the Searle Scholars Program, the Shurl and Kay Curci Foundation, the Hellman
1294 Foundation, the NIH Director's New Innovator Award (DP2 OD024541-01) and
1295 NSF CAREER Award (1652236), an NIH R01 Award (R01GM139913) and the
1296 Koret-UC Berkeley-Tel Aviv University Initiative in Computational Biology and
1297 Bioinformatics. HGG is also a Chan Zuckerberg Biohub Investigator.

Supplementary Figures

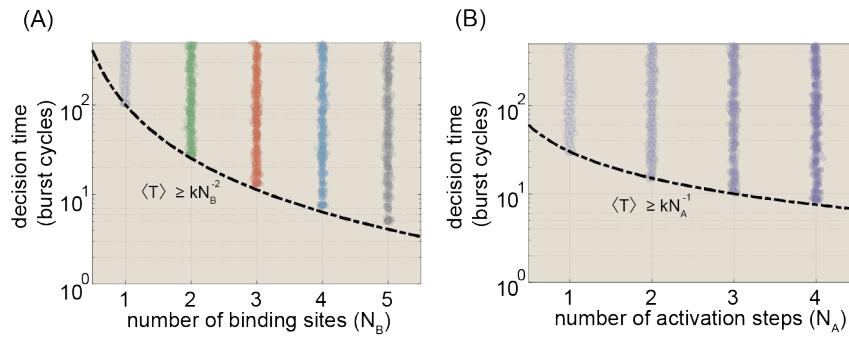


Fig. S1. Decision times for different gene circuit architectures. (A) Parameter sweep results for equilibrium gene circuits with different numbers of activator binding sites. Black dashed line indicates lower limit of the decision time and is a function of the form $\langle T \rangle = kN_B^{-2}$, where k is a proportionality constant. (B) Plot of range of achievable decision times for non-equilibrium gene circuits with a single activator binding site ($N_B = 1$) as a function of the number of activation steps, N_A . The dashed line indicates the lower decision time bound, and is a function of the form $\langle T \rangle = kN_A^{-1}$. (All results shown assume an error probability of 32%. For parameter sweep results in A and B, transition rate and interaction term magnitudes, k and η , were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

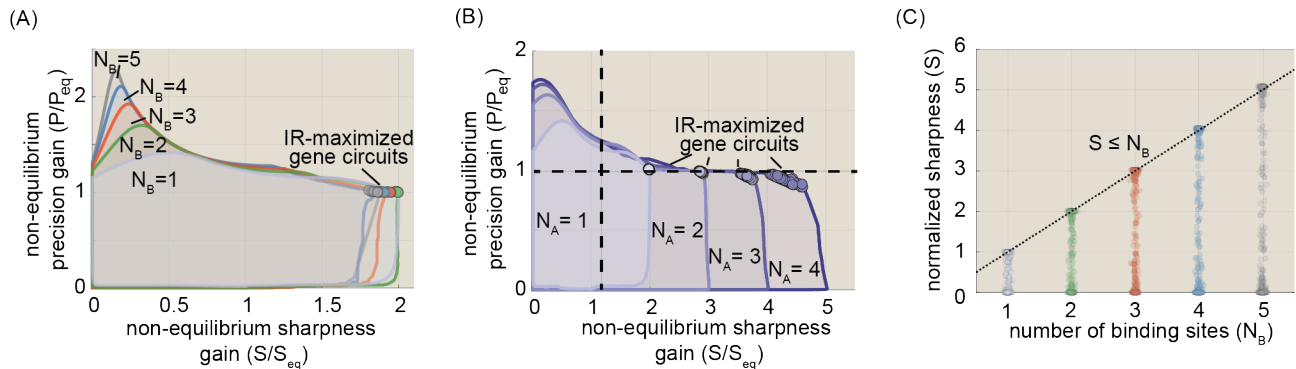


Fig. S2. Tradeoffs between sharpness and precision persist for more complex gene regulatory architectures. (A) Non-equilibrium gains in sharpness and precision for gene circuits with different numbers of activator binding sites (N_B) and one activation step. Shaded regions indicate achievable regimes for each system, as determined by no fewer than 10,000 unique simulated gene circuits. (B) Non-equilibrium gains in sharpness and precision for gene circuits with different numbers of activation steps (N_A) and one activator binding site. (C) Scatter plots indicate sharpness levels for equilibrium gene circuits as a function of the number of binding sites. Bounding line is for a function of the form $S = N_B$. (For parameter sweep results in A-C, transition rate and interaction term magnitudes, k and η , were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

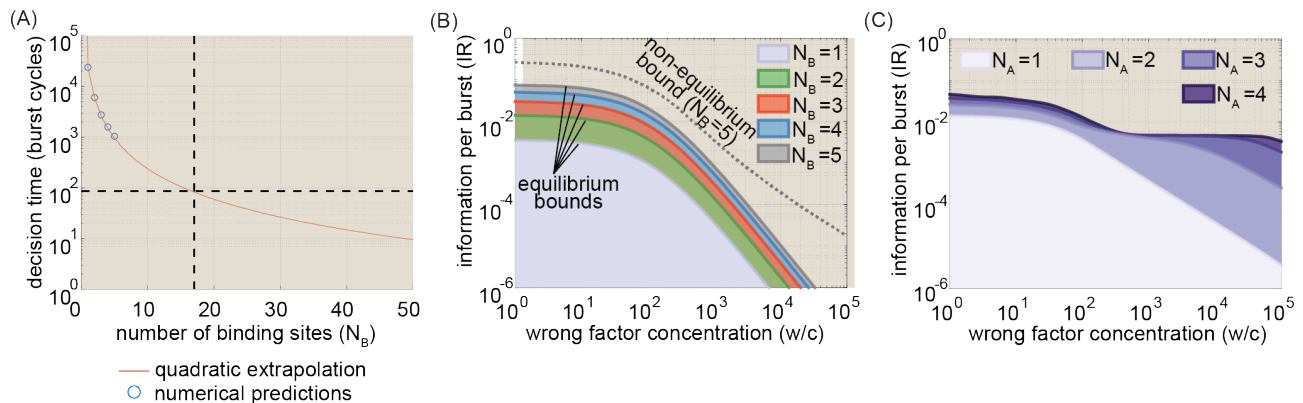


Fig. S3. Supplemental analyses for the dependence of IR with non-cognate transcription factor interference. (A) Parameter sweep results showing the range of achievable information rates as a function of w/c for equilibrium gene circuits with 1-5 activator binding sites and one molecular activation step. (B) Sweep results for non-equilibrium gene circuits with 1-4 activation steps and a single activator binding site. (C) Extrapolation of minimum decision times for equilibrium gene circuits as a function of number of activator binding sites based on numerical results for circuits with 1-5 binding sites. Analysis indicates that at least 17 sites would be required to achieve plausible decision times in the context of the mouse system. (For parameter sweep results in B and C, transition rate and interaction term magnitudes, k and η , were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

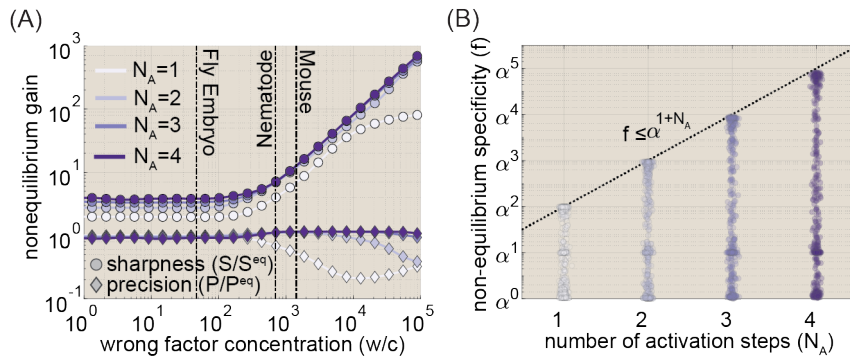


Fig. S4. Supplemental results for main text Figure 5. (A) Non-equilibrium sharpness and precision gains for IR-maximizing gene circuits with 1-4 activation steps. (B) Range of achievable specificity values for non-equilibrium gene circuits with 1-4 activation steps. (Transition rate and interaction term magnitudes, k and η , were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

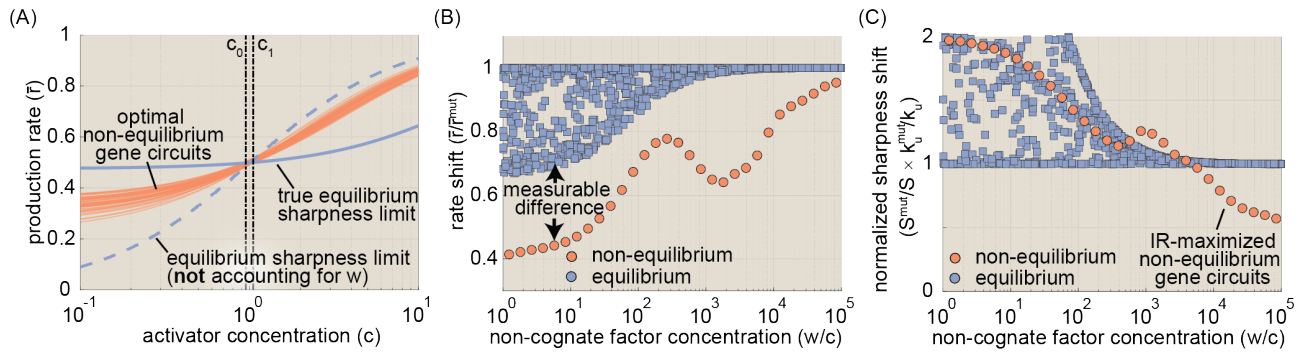


Fig. S5. Experimental signature of energy expenditure. (A) Predicted induction curves for 50 near-optimal non-equilibrium gene circuits when $w = 10^3 c^*$, as well as the *actual* induction curves for the sharpest achievable equilibrium curve (solid blue line) and the (incorrect) limit when that would be predicted if w was not accounted for (dashed line). Note that red curves fall above the true equilibrium limit but below the naive limit. (B) Predicted shift in the production rate resulting from a binding site perturbation that doubles the unbinding rate ($k_{i1}^{mut}/k_{i1} = 2$)—equivalent to an energetic difference of $0.7 k_B T$ —for equilibrium gene circuits (blue squares) and IR-maximizing non-equilibrium circuits (red circles). Note that non-equilibrium circuits are far more sensitive than equilibrium circuits when $w/c < 10^4$. The shift becomes negligible at higher values, thus providing a clear signature of energy dissipation. (C) Predicted sharpness shift upon perturbing the activator binding site. The non-equilibrium shift becomes markedly larger than equilibrium limit when $w/c > 10^3$. (For parameter sweep results in B and C, transition rate and interaction term magnitudes, k and η , were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

1299 Appendices

1300 A. Analytic expressions for key gene circuit characteristics

1301 This section lays out analytic expressions for key quantities that play a central role in the investigations undertaken
1302 over the course of the main text. We do not repeat derivations for expressions that are treated separately elsewhere
1303 in these Appendices, and avoid re-deriving expressions from scratch, unless they are novel to this work.

1304 **A.1. The transition rate matrix and activity vector.** Consider a gene circuit g that has K different microscopic states.
1305 We assume that microscopic transitions between the molecular states that make up g are Markovian, such that our
1306 system can be modeled as a continuous time Markov chain (CTMC). It follows that the steady-state behavior of g is
1307 fully determined by two quantities: the transition rate matrix, Q and the state activity vector, α .

1308 Q is a $K \times K$ matrix with off-diagonal elements that encode the rates with which the system switches between
1309 microscopic rates. For instance, q_{mn} —the element in the m th row and n th column of Q —gives the transition rate
1310 going from state n to state m . The diagonal elements of Q are negative, and are scaled such that each column of Q
1311 sums to 0. The activity vector α is a binary vector of length K that contains a “1” for each state that is transcriptionally
1312 active, and a “0” for inactive states. We assume that both Q and α are fixed in time.

1313 **A.2. State probabilities, transcription rate, and transcriptional noise.** A first step to calculating virtually all gene circuit
1314 characteristics of interest is to obtain the steady-state vector, π , which is a vector of length K that gives the steady
1315 state probability of finding the gene circuit of any one of the K microscopic states. We can obtain π by finding the
1316 right eigenvector (v_R) of Q with an eigenvalue of 0,

$$1317 \quad Qv_R = 0, \quad (8)$$

1318 and imposing the additional constraint that the elements of π sum to 1, such that

$$1319 \quad \pi = \frac{v_R}{\sum_{i=1}^K v_i}. \quad (9)$$

1320 With this the steady state probability vector in hand, we can calculate the average transcription rate by taking the
1321 dot product of α and π :

$$1322 \quad \bar{r} = r_0 \underbrace{\sum_{i=0}^K a_i \pi_i}_{\text{fraction of time active } (\pi_a)}, \quad (10)$$

1323 where we define the quantity indicated by the underbrace as the average fraction of time, π_a , that the system spends
1324 in the active state. Throughout the course of this work, we assume that r_0 is held fixed, such that the transcriptional
1325 activator may only impact transcription by modulating microscopic transition rates in Q to alter π . Further, since
1326 we take Poisson noise from mRNA synthesis to be negligible (see Appendix D), the absolute magnitude of r_0 is
1327 unimportant, and we set it to 1 for simplicity.

1328 Next, we turn to obtaining an expression for the variance (noise) in gene expression. From Whitt 1992 (Whitt,
1329 1992), we have that

$$1330 \quad \sigma^2 = 2 \sum_{i=1}^K \sum_{j=1}^K a_i \pi_i z_{ij} a_j, \quad (11)$$

1331 where z_{ij} is the element from i th row and j th column of what is known as the fundamental matrix, Z of our transition
1332 rate matrix, Q . Z is a $K \times K$ matrix that plays an integral role in the calculation of many key behaviors of a Markov
1333 chain. Once again drawing from Whitt, we can calculate Z using the formula

$$1334 \quad Z = (\mathbf{\Pi} - Q)^{-1} - \mathbf{\Pi}, \quad (12)$$

1335 where $\mathbf{\Pi}$ is a $K \times K$ matrix with each row equal to π .

1336 **A.3. Using the fundamental matrix to calculate first passage times.** First passage times provide a useful conceptual
1337 tool for connecting microscopic fluctuations, which often are unobservable, with emergent dynamical behaviors,
1338 such as transcriptional bursting. The fundamental matrix provides an invaluable tool for doing this in the context of
1339 arbitrarily complex transcriptional systems. Once again, we start with an expression from Whitt 1992 (Whitt, 1992)
1340 that relates off-diagonal elements of Z to first passage times between microscopic states:

$$1341 \quad z_{ji} = \pi_i [ET_{ei} - ET_{ij}], i \neq j. \quad (13)$$

Here, ET_{ij} is the mean expected first passage time from state j to state i and ET_{ei} the first passage time to state i at equilibrium, defined as

$$ET_{ei} = \pi_i \sum_{j=1}^K \pi_j ET_{ij}. \quad (14)$$

Now, from (Whitt, 1992) we also have that the diagonal elements of Z can be expressed as

$$z_{ii} = \pi_i ET_{ei}. \quad (15)$$

We can now combine Equations 13 and 15 to solve for the first passage time from state i to state j :

$$ET_{ij} = \frac{z_{ii} - z_{ji}}{\pi_i}. \quad (16)$$

A.4. Calculating the burst cycle time. First passage times are intimately related to a quantity of central importance throughout the text: the burst cycle time, τ_b , defined as the average time required for a system to complete one ON→OFF→ON cycle (Figure 1D). This is trivial in the case of a simple two state system with a single OFF and ON state and rates k_{on} and k_{off} (Figure A11). In this case, the burst cycle time is simply

$$\tau_b = \frac{k_{\text{on}} + k_{\text{off}}}{k_{\text{on}}k_{\text{off}}}. \quad (17)$$

The calculation becomes less trivial for systems with larger numbers of states, however. Fortunately, the concepts outlined above provide us with the tools necessary to derive a generic expression for τ_b that applies to systems of arbitrary complexity.

The essence of the procedure lies in calculating effective off and on rates (k_{off}^* and k_{on}^*) from Q using first passage times. We go through this procedure in detail for k_{on}^* and note that the same approach applies for k_{off}^* . The activity vector a partitions our system into M OFF states and N ON states. To calculate k_{on}^* , the first step is to estimate the expected amount of time it will take for the system to reach an ON state (any ON state) from each OFF state. We can do this by defining a new transition rate matrix, Q^{OFF} , that has dimensions $M+1 \times M+1$. The off-diagonal elements of the first M rows and M columns of Q^{OFF} are simply equal to the microscopic rates from Q that lead from one of the M OFF states to another OFF state. Together, these molecular states constitute a single coarse-grained OFF state.

The final row and column, however, are different and contain total fluxes into and out of all ON states from each OFF state. An element in the final row of Q^{OFF} is given by

$$q_{m+1,i}^{\text{OFF}} = \sum_{j=1}^K a_j q_{ji}, \quad (18)$$

where a_j is the j th element of the activity vector, q_{ij} is a microscopic rate from the original transition rate matrix, and we assume the state i is in the set of OFF states. Thus, we see that each element of the last row of Q^{OFF} gives the total flux from all OFF state into the ON conformation. The elements of the final column have a complementary definition:

$$q_{i,m+1}^{\text{OFF}} = \sum_{j=1}^K a_j q_{ij}. \quad (19)$$

With our condensed transition rate matrix thus defined, we can use Equations 8 and 9 to calculate π^{OFF} and Equation 12 to calculate Z^{OFF} . Then, we can use Equation 16 to obtain a vector et^{ON} of length M , where each element i is defined as the expected first passage time from OFF state i back into any of the ON states. Specifically, we have that each element, i , is given by

$$et_i^{\text{ON}} = \frac{z_{m+1,m+1}^{\text{OFF}} - z_{i,m+1}^{\text{OFF}}}{\pi_{m+1}}. \quad (20)$$

Thus, we have obtained a vector, et^{ON} , of expected mean first passage times out of each OFF state into the set of N active transcriptional states. But how do we weight the different passage times in this vector to arrive at an overall average expectation for the amount of time required for the system to turn back ON following a transition into an OFF state? It's tempting here to use the steady-state probabilities of each OFF state given by π , but this is actually not correct.

1383 Instead, the key is to recognize that each OFF state should be weighted by the rate at which ON states switch into
 1384 it. In other words, we weight OFF states by the probability that they are the initial state the system reaches upon
 1385 switching out of the ON conformation; the gateway into the OFF states. Mathematically, we encode these weights
 1386 using the flux vector f^{OFF} , which has M elements, each defined as

$$1387 \quad f_i^{\text{OFF}} = \sum_{j=1}^K a_j q_{ij} \pi_j, \quad (21)$$

1388 where a_j is the j th element of the activity vector \mathbf{a} (1 for ON states and 0 otherwise), q_{ij} is the transition rate from
 1389 state j to state i , and π_j is the steady-state probability of state j .

1390 Finally, we combine this expression with Equation 20 to obtain an expression for the average reactivation time as
 1391 a flux-weighted average of the first passage times out of each OFF state:

$$1392 \quad ET_{\text{OFF} \rightarrow \text{ON}} = \frac{1}{k_{\text{on}}^*} = \frac{\sum_{i=1}^M f_i e t_i^{\text{ON}}}{\sum_{i=1}^M f_i}. \quad (22)$$

1393 As noted above, the calculations for k_{off}^* follow precisely the same logic, with the roles of the OFF and ON states
 1394 switched. After this is done, the total burst cycle time, τ_b , is simply

$$1395 \quad \tau_b = ET_{\text{OFF} \rightarrow \text{ON}} + ET_{\text{ON} \rightarrow \text{OFF}}. \quad (23)$$

1396 Equation 22 is useful because it allows us to relate the (potentially quite complex) microscopic dynamics of a
 1397 transcriptional system to emergent bursting timescales observed in live imaging experiments (Lammers et al., 2020).
 1398 To our knowledge, this is the first time that take this flux-weighted first passage time approach is applied to the
 1399 modeling of burst dynamics. We hope that the expressions provided here will prove useful to others seeking to
 1400 pursue similar projects in the future.

1401 Finally, a useful feature implied by Equation 22 and Equation 23 is that the absolute size of τ_b scales inversely
 1402 with the microscopic rates in Q , such that we can decrease τ_b by some scaling factor λ by simply multiplying Q by
 1403 λ . We use this trick to renormalize all time-dependent metrics calculated over the course of our parameter sweeps
 1404 to have units of burst cycle time. This is done by calculating τ_b for each new model realization we generate, and then
 1405 multiplying its transition rate matrix by this quantity to generate a normalized rate matrix, namely

$$1406 \quad Q^* = \tau_b Q. \quad (24)$$

1407 The adjusted matrix, Q^* , is then used to calculate all relevant gene circuit characteristics.

1408 **A.5. A generic expression for the rate of energy dissipation.** Equation 1 gives an expression for the rate of energy
 1409 dissipation (also termed entropy production), Φ , in the context of the four-state model shown in Figure 1C. This is
 1410 a special case of a more general formula for Φ that applies to arbitrary molecular architectures. From (Lang et al.,
 1411 2014; Lebowitz and Spohn, 1999), we have

$$1412 \quad \Phi = \sum_{i=i}^K \sum_{j \neq i}^K \pi_i q_{ji} \ln \left(\frac{q_{ji}}{q_{ij}} \right). \quad (25)$$

1413 We use Equation 25 to calculate all energy dissipation rates given throughout the main text. In the case of the simple
 1414 four-state system shown in Figure 1B, we have from (Lang et al., 2014) that Equation 25 simplifies to

$$1415 \quad \Phi = J \ln \frac{k_b \eta_{ab} k_a \eta_{ua} k_u k_i}{k_b k_a \eta_{ba} k_u \eta_{ib} k_i}, \quad (26)$$

1416 which further simplifies to

$$1417 \quad \Phi = J \ln \frac{\eta_{ab} \eta_{ua}}{\eta_{ba} \eta_{ib}}. \quad (27)$$

1418 Here J is the net cycle flux, a quantity with units of inverse time which encodes the rate at which the system
 1419 completes extra cycles in the clockwise ($J > 0$) or counterclockwise ($J < 0$) directions. Mathematically, J is given
 1420 by

$$1421 \quad J = J_+ - J_-, \quad (28)$$

1422 where J_+ gives the average rate at which the system completes one full cycle in the clockwise direction (i.e., setting
 1423 out from state 0 to state 1, reaches state 0 from state 4), and J_- is defined analogously. In terms of microscopic
 1424 quantities, for any system with a single loop we can define J as

$$1425 \quad J = p_i k_{ji} - p_j k_{ij}, \quad (29)$$

1426 where k_{ji} denotes the transition rate from state i to state j and J_{ij} corresponds to the net transition flux between
 1427 the two states.

1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460

B. Gaussian noise approximation

Throughout this work, we make the simplifying assumption that the intrinsic noise in accumulated mRNA levels due to transcriptional bursting is approximately Gaussian. In this section, we use stochastic simulations to put this assumption to the quantitative test. The Markov chain central limit theorem states that the distribution of a quantity that is a function of a Markov chain (such as the transcription rate, \bar{r}), will become approximately Gaussian as the number of iterations becomes large (Geyer, 2011).

The question, then, is whether can expect the accumulated transcriptional output to approach this limiting Gaussian distribution within timescales that are relevant to the decision times discussed in this work. To determine this, we used stochastic simulations (Gillespie, 1977) to track the distribution of the accumulated output of 500 random realizations of the four-state system shown in Figure 1C for 5,000 burst cycles. Each realization had a unique set of transition rates and, correspondingly, a unique average rate of transcription, $\bar{r} = \pi_a r_0$, where π_a indicates the fraction of time the system spends in a transcriptionally active molecular state and r_0 is the rate of transcript initiation when active. For each model realization, we ran 100 stochastic simulations. We used these simulations to track the distribution of the apparent average transcription rate for each model realization as function of accumulation time. Figure A1A shows the apparent mean rate across 100 simulations for a single illustrative gene circuit realization. Inset histograms indicate distribution of apparent transcription rates at different time points. As expected, we see that the apparent rates are initially highly dispersed; however, even after 25 burst cycles, we see that $p(\bar{r})$ has become a much narrower, roughly symmetrical distribution that appears approximately Gaussian.

To systematically assess the rate of convergence to normality, we utilized the simple One-sample Kolmogorov-Smirnov test (“kstest”, (Massey, 1951)), which tests the null hypothesis that a vector of transcription outputs from realization i at time t , $r_i(t)$, is drawn from a normal distribution. The test returns a p value corresponding to the probability of observing $r_i(t)$ if the transcriptional output were truly Gaussian. In standard implementations $p \lesssim 0.05$ is taken to constitute strong evidence that the output is *not* Gaussian. Thus, to assess convergence to normality, we tracked this p value over time for each of the 500 gene circuit realizations.

Figure A1B shows the average kstest p -values across 10 different sets of gene circuits, grouped by their average rate of transcription. In all cases, we see that noise profiles rapidly converge towards normality, such that all systems cross the (relatively conservative) threshold of $p = 0.1$ within 5 burst cycles (dashed line in Figure A1B). Gene circuits near the tail ends of the induction curve ($\bar{r} \leq 0.1$ and $\bar{r} \geq 0.9$) take the longest to converge, which is likely because it takes longer for distributions near the boundaries to become symmetric about their mean; yet even these converge rapidly.

The fastest decisions discussed in the main text (Figure 4D and E), and most decision times considered are significantly longer than the time for Gaussian convergence revealed by Figure A1B). Thus, we conclude that the Gaussian noise approximation invoked throughout this work is justified.

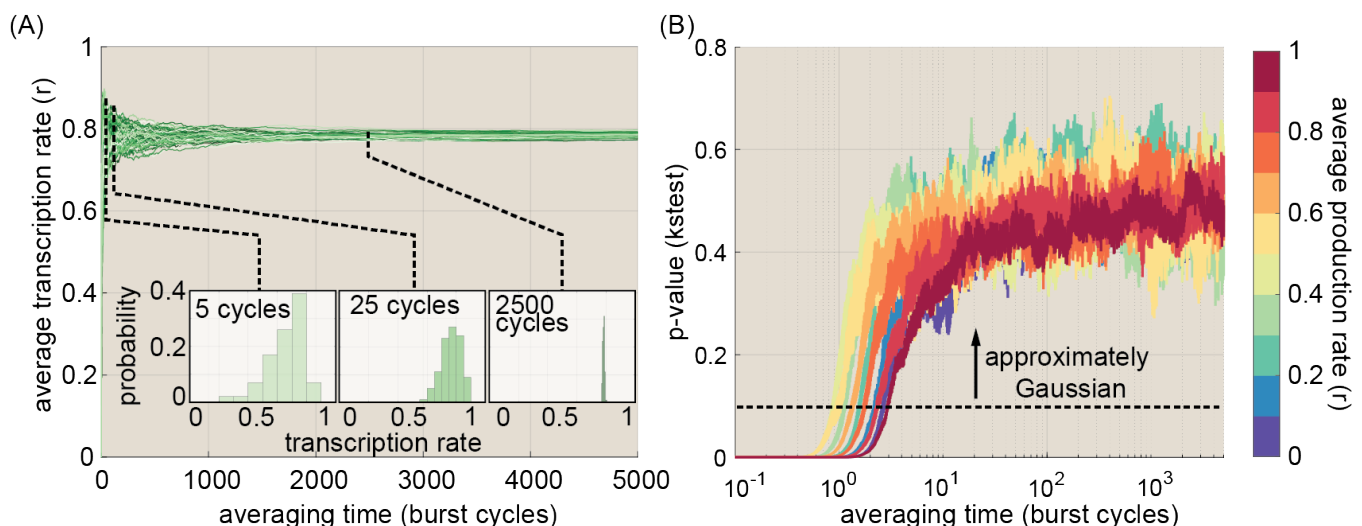


Fig. A1. Testing the validity of the Gaussian noise approximation. (A) Illustrative plot showing average transcription rate as a function of the averaging time across 100 stochastic simulations of one illustrative realization of the four-state model gene circuit. Inset histograms show distribution of apparent rates at three different time points. We see that, as the accumulation time increases, the distributions get tighter and appear more Gaussian in shape. (B) Plot showing p -values of one-sample Kolmogorov-Smirnov test. Different colors indicate average trends for systems with different average transcription rates. We see that systems near the low and high ends of the induction curve converge to Gaussian form most slowly, but even these cross the $p = 0.1$ line within a handful of burst cycles. Error bars indicate bootstrap estimates of standard error calculated for each group. (For stochastic simulations shown in A and B, transition rate and interaction term magnitudes, k and η , were constrained such that $10^{-2} \leq k\tau_b \leq 10^2$ and $10^{-2} \leq \eta \leq 10^2$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

1461 **C. Deriving the rate of information transmission for a gene locus**

1462 Motivated by (Siggia and Vergassola, 2013), we define the rate of information transmission as the time derivative
 1463 of the expected Kullback-Leibler (KL) divergence between the two hypotheses ($C = c_0$ and $C = c_1$), given some
 1464 accumulated mRNA level m , such that

$$1465 \quad IR = \frac{d}{dt} \left\langle D_{KL} [p(c_1|m) || p(c_0|m)] \right\rangle, \quad (30)$$

1466 where $P(c_0|m)$ and $P(c_1|m)$ indicate (respectively) the conditional likelihood that the true value of C is c_0 and c_1
 1467 given the observed output m , and where the angled brackets indicate that we are dealing with the expected value
 1468 of D_{KL} across many replicates. We refer readers to information theory reference materials for a formal definition
 1469 of D_{KL} (see, e.g., (Cover and Thomas, 2006)); however, at an intuitive level it can be regarded as measuring how
 1470 different two probability distributions are from one another. Thus, with Equation 30, we define the rate of information
 1471 production as the rate at which the two possibilities (c_1 or c_0 ?) become distinguishable from one another given the
 1472 observed “evidence” (m).

1473 We can write out the expected KL divergence from Equation 30 more explicitly as the weighted sum of log proba-
 1474 bility ratios:

$$1475 \quad IR = \frac{d}{dt} \left(p_0 \left\langle \ln \frac{p(c_0|m)}{p(c_1|m)} \right\rangle_0 + p_1 \left\langle \ln \frac{p(c_1|m)}{p(c_0|m)} \right\rangle_1 \right), \quad (31)$$

1476 where $\langle \dots \rangle_i$ indicates the expectation taken assuming the true value of C to be c_i and where p_0 and p_1 indicate the
 1477 priors on the true value of C , taken to be equal moving forward ($p_1 = p_0 = 1/2$). This formulation provides intuition
 1478 for the sense in which IR is the information rate: as the conditional probabilities of the observed output given the
 1479 true (numerators) and false (denominators) hypotheses about C diverge in favor of the true hypothesis, the log ratio
 1480 terms will become large and positive. Thus a positive derivative corresponds to positive information production.

1481 However, here we must recall that our focus here is to understand how the molecular architecture of gene loci
 1482 impacts the transcriptional response and, ultimately, IR. Thus we wish to work in terms of $p(m|c)$ —the conditional
 1483 distribution of observed mRNA outputs given some input—rather than $p(c|m)$. To do this, we make use of Bayes’
 1484 Theorem. We have:

$$1485 \quad \frac{p(c_0|m) p(m)}{p(c_1|m) p(m)} = \frac{p(m|c_0) p(c_0)}{p(m|c_1) p(c_1)}. \quad (32)$$

1486 This expression becomes an equality if we assumed equal prior probabilities for our two hypotheses ($p(c_0) = p(c_1)$):

$$1487 \quad \frac{p(c_0|m)}{p(c_1|m)} = \frac{p(m|c_0)}{p(m|c_1)}. \quad (33)$$

1488 Thus, we can use Equation 33 to rewrite Equation 31 as:

$$1489 \quad IR = \frac{d}{dt} \frac{1}{2} \left(\left\langle \ln \frac{p(m|c_0)}{p(m|c_1)} \right\rangle_0 + \left\langle \ln \frac{p(m|c_1)}{p(m|c_0)} \right\rangle_1 \right). \quad (34)$$

1491 We can think of the conditional probabilities, $p(m|c_i)$, in Equation 34 as representing the full *stochastic* transcriptional
 1492 response to some input activator concentration c_i . When these are approximately Gaussian (a condition discussed
 1493 above in Appendix B), it becomes a straightforward exercise to solve for the expected log ratios in Equation 34. We
 1494 will solve for the case when $C = c_1$ in full. The c_0 case proceeds in precisely the same fashion. To start, we have

$$1495 \quad \left\langle \ln \frac{p(m|c_1)}{p(m|c_0)} \right\rangle_1 = \int_0^\infty p(m|c_1) \ln p(m|c_1) dg - \int_0^\infty p(m|c_1) \ln p(m|c_0) dg. \quad (35)$$

1496 Recall that $m = rt$ is Gaussian with probability density function:

$$1497 \quad p(m|c_i) = \frac{e^{-\left(\frac{m - \bar{m}(c_i)}{2\sigma^2(c_i)}\right)^2}}{\sqrt{2\pi\sigma^2(c_i)}}. \quad (36)$$

1498 Plugging Equation 36 in for $\ln p(m|c_1)$ yields

$$1499 \quad \left\langle \ln \frac{p(m|c_1)}{p(m|c_0)} \right\rangle_1 = -\frac{1}{2} \ln(2\pi\sigma_m^2(c_1)) - \frac{1}{2} - \int_0^\infty p(m|c_1) \left[-\frac{1}{2} \ln(2\pi\sigma_m^2(c_0)) - \frac{1}{2} \left(\frac{\bar{m}(c_0) - g}{\sigma_m(c_0)} \right)^2 \right] dm, \quad (37)$$

1500 Where we’ve recognized that the first integral will simply yield the standard expression for the entropy of a Gaussian
 1501 random variable. Pulling constant factors out of the second integral leads to

$$1502 \quad \left\langle \ln \frac{p(m|c_1)}{p(m|c_0)} \right\rangle_1 = -\frac{1}{2} \ln(2\pi\sigma_m^2(c_1)) - \frac{1}{2} + \frac{1}{2} \ln(2\pi\sigma_m^2(c_0)) + \frac{1}{2\sigma_m^2(c_0)} \int_0^\infty p(m|c_1) [\bar{m}^2(c_0) - 2m\bar{m}(c_0) + m^2] dm. \quad (38)$$

Simplifying and recognizing that $\langle m^2 \rangle_1 = \bar{m}^2(c_1) + \sigma_m^2(c_1)$ leads to:

$$\left\langle \ln \frac{p(m|c_1)}{p(m|c_0)} \right\rangle_1 = \frac{1}{2} \ln \frac{\sigma_m^2(c_1)}{\sigma_m^2(c_0)} - \frac{1}{2} + \frac{1}{2\sigma_m^2(c_0)} \left[\bar{m}^2(c_0) - 2\bar{m}(c_0)\bar{m}(c_1) + \sigma_m^2(c_1)^2 + \bar{m}^2(c_1)^2 \right]. \quad (39)$$

Finally, we recall that $m = rt$ and $\sigma_m^2 = \sigma^2 t$, obtaining

$$\left\langle \ln \frac{p(m|c_1)}{p(m|c_0)} \right\rangle_1 = \frac{1}{2} \left[\ln \frac{\sigma_r^2(c_0)}{\sigma_r^2(c_1)} + t \frac{(\bar{r}(c_1) - \bar{r}(c_0))^2}{\sigma_r^2(c_0)} + \frac{\sigma_r^2(c_1)}{\sigma_r^2(c_0)} - 1 \right]. \quad (40)$$

Performing the same procedure for the case where $c = c_0$ yields:

$$\left\langle \ln \frac{p(m|c_0)}{p(m|c_1)} \right\rangle_0 = \frac{1}{2} \left[\ln \frac{\sigma_r^2(c_1)}{\sigma_r^2(c_0)} + t \frac{(\bar{r}(c_0) - \bar{r}(c_1))^2}{\sigma_r^2(c_1)} + \frac{\sigma_r^2(c_0)}{\sigma_r^2(c_1)} - 1 \right]. \quad (41)$$

Plugging [Equation 40](#) and [Equation 41](#) into [Equation 34](#) and taking the derivative with respect to time yields

$$\text{IR} = \frac{1}{4} \frac{(\bar{r}(c_1) - \bar{r}(c_0))^2 (\sigma(c_1)^2 + \sigma(c_0)^2)}{\sigma(c_0)^2 \sigma(c_1)^2}. \quad (42)$$

Next, if we assume that the difference between c_0 and c_1 is small (as stipulated in the main text), then $\sigma(c_0) \approx \sigma(c_1) \approx \sigma^2(c^*)$ and $\bar{r}(c_1) - \bar{r}(c_0) \approx \delta c dr/dc$, leading to

$$\text{IR} = \frac{1}{2} \left(\delta c \frac{dr}{dc} \right)^2 \frac{1}{\sigma(c^*)^2}. \quad (43)$$

Finally, we invoke the definitions of sharpness and precision given in [Figure 1B](#), which leads to [Equation 2](#) from the main text:

$$\text{IR} = \frac{1}{2} \left(\frac{\delta c}{c^*} \right)^2 s^2 p^2. \quad (44)$$

D. Poisson noise from mRNA synthesis is negligible relative to noise from bursting

In this section, we provide support for the claim, made in Main Text Section [B](#), that Poisson noise due to mRNA synthesis is negligible relative to noise from transcriptional bursting. We take as our starting point [Equation 68](#) from [Appendix J](#),

$$P = \frac{\pi_a(1 - \pi_a)}{\sigma}, \quad (45)$$

which relates the normalized precision, P , to the bursting noise, σ , and the fraction of time a gene circuit spends in transcriptionally active states, π_a . From [Figure 3A](#), we see that $P \leq 1$ for the four-state gene circuit shown in [Figure 1C](#) when the system is out of equilibrium, which, from [Equation 45](#), implies that

$$\sigma^2 \geq \pi_a^2 (1 - \pi_a)^2 \quad (46)$$

for the 4 state system.

Thus, [Equation 47](#) gives a lower bound for the intrinsic variance in gene expression that arises due to transcriptional burst fluctuations at the gene locus. To see how to relate this to noise from mRNA synthesis, we need to take two more steps. First, we must recall that we are working in units of the burst cycle time, τ_b . Second, we must further recall that we set the actual rate of mRNA synthesis, r_0 , equal to 1 throughout the main text. We must do away with these simplifications in order to relate σ^2 to synthesis noise. Accounting for these simplifications, the full expression for the noise floor, in “real” time units and accounting for the true rate of mRNA synthesis is

$$\sigma_{\text{burst}}^2 \geq \tau_b r_0^2 \pi_a^2 (1 - \pi_a)^2. \quad (47)$$

Now, if we assume mRNA synthesis to be a Poisson process (following, e.g., ([Shelansky and Boeger, 2020](#))), we have that this component of the variance is simply equal to

$$\sigma_{\text{Poisson}}^2 = r_0 \pi_a. \quad (48)$$

The key thing to notice about [Equation 48](#) is that mRNA synthesis noise is *independent* of the bursting timescale τ_b . Thus, as τ_b increases, σ_{burst}^2 will increase in magnitude relative to $\sigma_{\text{Poisson}}^2$. [Figure A2A](#) and [B](#) illustrate this fact, showing predicted bursting and mRNA synthesis variance components, respectively, as a function of the bursting time scale τ_b and the activity level (π_a). All calculations assume an mRNA synthesis rate of 20 mRNA per minute,

1541 a rate based off of estimates from the fruit fly (Lammers et al., 2020) and that is consistent with measurements
 1542 from other systems (Tantale et al., 2016). From Figure A2A, we see that σ_{burst}^2 peaks at $\pi_a = 0.5$ and increases
 1543 dramatically as we move rightward along the x-axis and the burst cycle time increases. We emphasize that this
 1544 represents a *lower* bound for maximally precise non-equilibrium gene circuits; most systems (including IR-optimized
 1545 systems) will lie above this bound. In contrast Figure A2B shows that noise from mRNA synthesis scales linearly
 1546 with π_a , and is constant in τ_b .

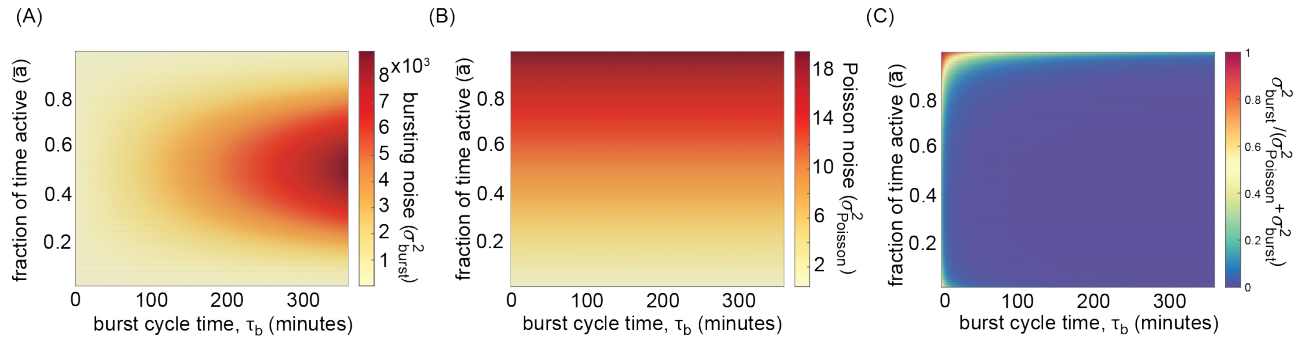


Fig. A2. Determining the contribution from mRNA synthesis noise. (A) Heatmap showing lower bound of bursting component of variance for the non-equilibrium four-state model shown in Figure 1C as a function of the fraction of time spent in the active state (π_a) and the burst cycle time (τ_b). (B) Heatmap showing predicted variance component arising from mRNA synthesis. (C) Predicted relative contribution of mRNA synthesis noise to total intrinsic noise levels in gene expression. Note that contribution is only significant for rapidly bursting systems near the saturation point. (All calculations assume an mRNA synthesis rate of 20 per minute, in keeping with estimates from (Lammers et al., 2020).)

1547 The total gene expression noise level is given by

$$1548 \quad \sigma_{\text{tot}}^2 = \sigma_{\text{Poisson}}^2 + \sigma_{\text{burst}}^2. \quad (49)$$

1549 We can use this expression to calculate a lower bound on the relative contribution of mRNA synthesis noise to the
 1550 overall intrinsic variance in gene expression. Figure A2C shows the results of this calculation. We see that, with the
 1551 exception of rapidly bursting systems near the saturation ($\pi_a \approx 1$), the contribution from Poisson noise due to mRNA
 1552 synthesis is negligible. Thus, we conclude that noise from transcriptional bursting constitutes the dominant source
 1553 of gene expression noise for the vast majority of the parameter regimes relevant for the investigations in this paper,
 1554 and that our decision to neglect Poisson noise from mRNA synthesis is reasonable.

1555 E. The Sequential Probability Ratio Test

1556 Over half a century ago, Wald conceived of the Sequential Probability Ratio Test (SPRT) as a solution to the problem
 1557 of making accurate decisions between two hypotheses, H_1 and H_0 in “real time” as relevant data is accruing (Wald,
 1558 1945). Shortly thereafter, it was established that SPRT represents the optimal approach to sequential decision
 1559 problems involving binary decisions (Wald and Wolfowitz, 1948), meaning that it requires the fewest observations to
 1560 achieve a desired level of accuracy. In this framework, a downstream receiver (in our case, downstream genes or
 1561 other cellular processes) tracks the accrual of some signal (mRNA, and eventually protein) over time and compares
 1562 how likely this accrued signal is under the two hypotheses to be distinguished (e.g., high or low activator concentra-
 1563 tion). In this work, we use the optimal nature of SPRT to set lower bounds on decision times that could be achieved
 1564 given the transcriptional output of model gene loci. The essence of the test lies in tracking the relative likelihoods of
 1565 our two hypotheses ($C = c_1$ and $C = c_0$) over time as more and more transcriptional output, m , accrues:

$$1566 \quad \frac{P_0}{P_1} = \frac{P(c_0|m)}{P(c_1|m)} \quad (50)$$

1567 Figure A3A shows a stochastic simulation of how this ratio evolves over time for the output of a single model gene
 1568 circuit. Although the true concentration in this case is c_0 , we see that the two hypotheses are essentially indistin-
 1569 guishable early on. This is because the range of possible outputs given high and low activator concentrations overlap
 1570 significantly early on (leftmost panel of Figure A3B). However, as more and more time passes, the expected outputs
 1571 (m) given the two possible inputs (c_1 and c_0) start to separate. We see that the ratio in their likelihoods diverges
 1572 more and more in favor of c_0 ($P_0/P_1 \gg 1$), corresponding to a higher and higher degree of certainty that c_0 is the
 1573 correct choice.

1574 This divergence, however, is non-monotonic and noisy, which reflects the stochastic nature of protein production
 1575 at a single gene locus. It has been shown that the noisy divergence of the log of the probability ratio (which we will
 1576 call \mathcal{L}) can modeled as a 1-D diffusive process with average drift IR (Siggia and Vergassola, 2013) given by

$$1577 \quad \text{IR} = \frac{d}{dt} \langle \mathcal{L} \rangle. \quad (51)$$

In this framework, a “decision” is made when \mathcal{L} crosses a so-called “decision boundary” (horizontal dashed lines in Figure A3A). Siggia et al showed that the Gaussian diffusion approximation could be used to obtain an analytic expression for the expected time needed to make a decision. From Equation 15 in the supplement of (Siggia and Vergassola, 2013), we have that:

$$\langle T \rangle = \frac{K}{2V \sinh \frac{VK}{D}} \left[e^{\frac{VK}{D}} + e^{-\frac{VK}{D}} - 2 \right], \quad (52)$$

where V is the same as IR from above (and in the main text), D encodes the diffusivity of decision process (essentially, how large the fluctuations are about its mean drift trajectory), and K is related to the log of the error tolerance parameter ε , such that

$$K = \log \left(\frac{1 - \varepsilon}{\varepsilon} \right). \quad (53)$$

We note that Equation 52 assumes equal priors regarding the likelihood of c_1 and c_0 , and also assumes equal error tolerances for choosing incorrectly in either case (Desponds et al., 2020).

If we take the accumulated transcriptional output of our gene circuit, $m = rt$, to be approximately Gaussian (see Appendix B), then it can be shown that D has the form:

$$D = \frac{(m_0 - m_1)^2 (\sigma_0^6 + \sigma_1^6)}{4\sigma_0^4 \sigma_1^4}, \quad (54)$$

where m_i and σ_i give the mean and variance in the accumulated transcriptional output, given that $C = c_i$. From Equation 42 in Appendix C, we also have that

$$V = \text{IR} = \frac{(m_0 - m_1)^2 (\sigma_0^2 + \sigma_1^2)}{4\sigma_0^2 \sigma_1^2}. \quad (55)$$

In a different context (exponential distributions, rather than Gaussian), Desponds and colleagues (Desponds et al., 2020) demonstrated that $D \approx V$ when the difference between hypothesis— $\delta c/c^*$ in our case—is small. From Equations 54 and 55, we see that this also holds for the Gaussian case: when c_1 and c_0 are sufficiently close, σ_1 and σ_0 will be approximately equal, such that:

$$D \approx V \approx \frac{(m_0 - m_1)^2}{2\sigma^2}. \quad (56)$$

As demonstrated by (Desponds et al., 2020), when $D \approx V$, Equation 52 simplifies dramatically, yielding

$$\langle T \rangle = \log \left(\frac{1 - \varepsilon}{\varepsilon} \right) \frac{1 - 2\varepsilon}{\text{IR}}, \quad (57)$$

which is Equation 3 from the main text. For correctness, we use the full expression (Equation 52) to calculate all decision time quantities shown in the main text. However, since Equation 57 holds quite well for the 10% concentration difference considered here, we give the simpler expression in the main text to aid the reader’s intuition.

F. Implementation of parameter sweep algorithm

In this section, we describe the parameter sweep algorithm employed throughout this work to enumerate the performance bounds of gene circuit models. We note that this approach is based off of an algorithm previously employed by Eck & Liu et al. (Eck et al., 2020) to explore the behavior of non-equilibrium models of transcription (see also, (Estrada et al., 2016)). Figure A4A illustrates the key steps in this numerical procedure. First, an initial set of gene circuit realizations (typically comprised of 1,000 variants) is generated by sampling random values for each transition rate in the system. We then calculate the performance metrics of interest (S and P for the example in Figure A4A) for each gene circuit realization. This defines an initial set of points (Figure A4A, Panel i) that collectively span some region in 2D parameter space with area a_1 .

Next (Panel ii), we subdivide parameter space into N different bins along the X and Y axes, with N dictated by the total number of points ($10 \leq N \leq 50$). We subsequently calculate the maximum and minimum point in each X and Y slice (Panel iii). Finally, we randomly select candidate gene circuit models from these boundary points and apply small perturbations to each transition rate to generate a new set of random variants (iv). In general, these variants will lie close to the original model in 2D parameter space and, thus, close to the current outer boundary of parameter space. The key to the algorithm’s success is that some of these variants will lie *beyond* the current boundary (blue points in Figure A4A, Panel iv). This has the effect of extending the boundary outward, leading to an increase in the surface area spanned by our sample points (panel iv). As a result, cycling through steps ii-iv amounts to a stochastic edge-finding algorithm that will iteratively expand the boundary spanned by sample points outward in 2D parameter space until some analytic boundary is reached.

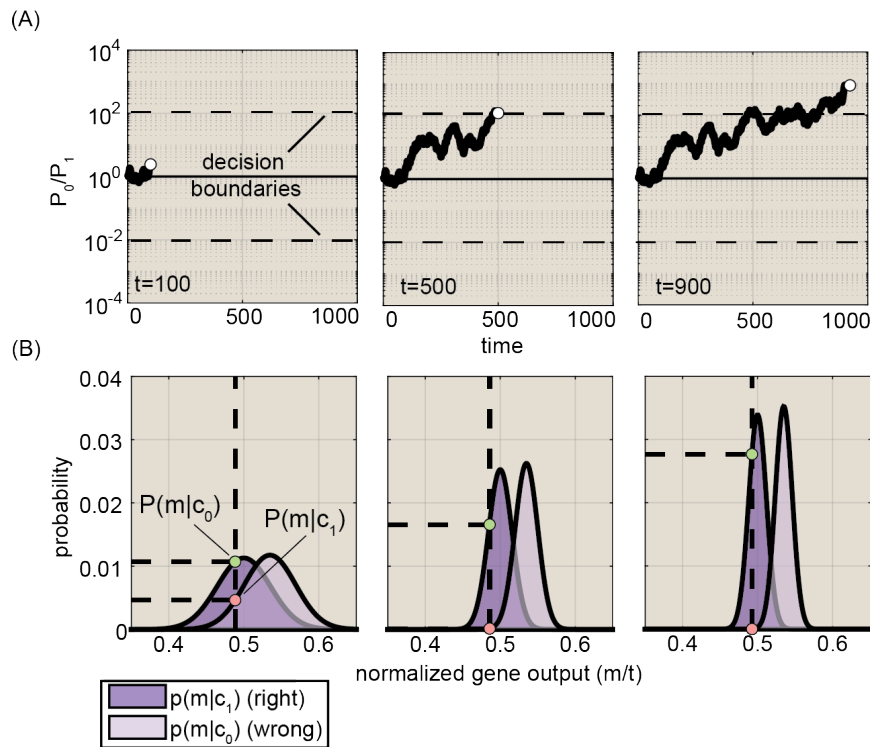


Fig. A3. The Sequential Probability Ratio Test. (A) Panels show stochastic trajectory of the relative probabilities of c_0 and c_1 over time, given the observed output of some gene circuit (illustrated in Figure 1E). (B) Panels illustrating expected distributions of transcriptional outputs, m , for each concentration at different time points normalized by the total time over which the gene circuit has been active, t . Note how the distributions narrow and separate as time progresses.

The panels in Figure A4B show snapshots of the sweep algorithm's progress exploring sharpness vs. precision parameter space for non-equilibrium realizations of the four-state gene circuit (Figure 1C). Figure A4C shows the total area spanned by the sample points for this run as a function of sweep iteration. By eye it appears that most of salient parameter space has been explored by step 10 of the algorithm, but we are quite strict with our convergence criteria. We will only terminate a sweep at step t if $(a_t - a_{t-2})/a_{t-2} \leq 0.001$ and $(a_{t-1} - a_{t-3})/a_{t-3} \leq 0.001$. In this case, this convergence criterion is met following step 25, leading to the final set of sample points in Figure A4D. In general, we run all sweeps until the above criterion is met or some pre-specified maximum number of iterations (usually 50) is reached.

F.1. Numeric vs. symbolic metric calculations. The algorithm outlined in Figure A4A is predicated upon the ability to rapidly calculate performance metric quantities (e.g., S and IR) given a set of transition rate magnitudes. Wherever possible, we use symbolic expressions to perform these calculations; however, this is only feasible for the simple four and six state systems depicted in Figure 1C and Figure 4B. For more complex models, it is infeasible to perform the symbolic operations required to obtain closed-form symbolic expressions. As a result, we use numerical calculations to arrive at performance metrics for all higher-order models.

F.2. Enforcing equilibrium constraints. In this work, we make frequent use of comparisons between equilibrium and non-equilibrium gene circuits in order to elucidate how energy expenditure alters gene-regulatory performance. A key step in performing parameter sweeps for equilibrium gene circuits is ensuring that transition rates adhere to the constraints imposed by detailed balance. For the simple four state model shown in Figure 1C, this process boils down to ensuring that the product of the four transition rates moving in a clockwise direction about the square is equal to the product of the four counterclockwise rates. As shown in Appendix A.5, this amounts to enforcing the constraint that

$$\lambda = \frac{\eta_{ab}\eta_{ua}}{\eta_{ba}\eta_{ib}} = 1, \quad (58)$$

where the η factors on the top and bottom of the left-hand-side expression correspond to regulatory interaction terms that modify transition rates in the clockwise and counterclockwise directions, respectively, and where λ is the flux factor that captures the relative magnitudes of clockwise and counterclockwise transitions.

To enforce this constraint during the course of a parameter sweep, we add a step to the process outlined above. New gene circuit realizations are generated as before, but now, following its generation, we calculate the initial flux factor, λ_0 for each new realization using Equation 58. In general this quantity will not equal one for the new realizations ($\lambda^* \neq 1$). To fix this, we then multiply η_{ba} and η_{ib} each by a factor of $\lambda^{\frac{1}{2}}$, which leads to a modified

G Testing the convergence characteristics of the parameter sweep algorithm

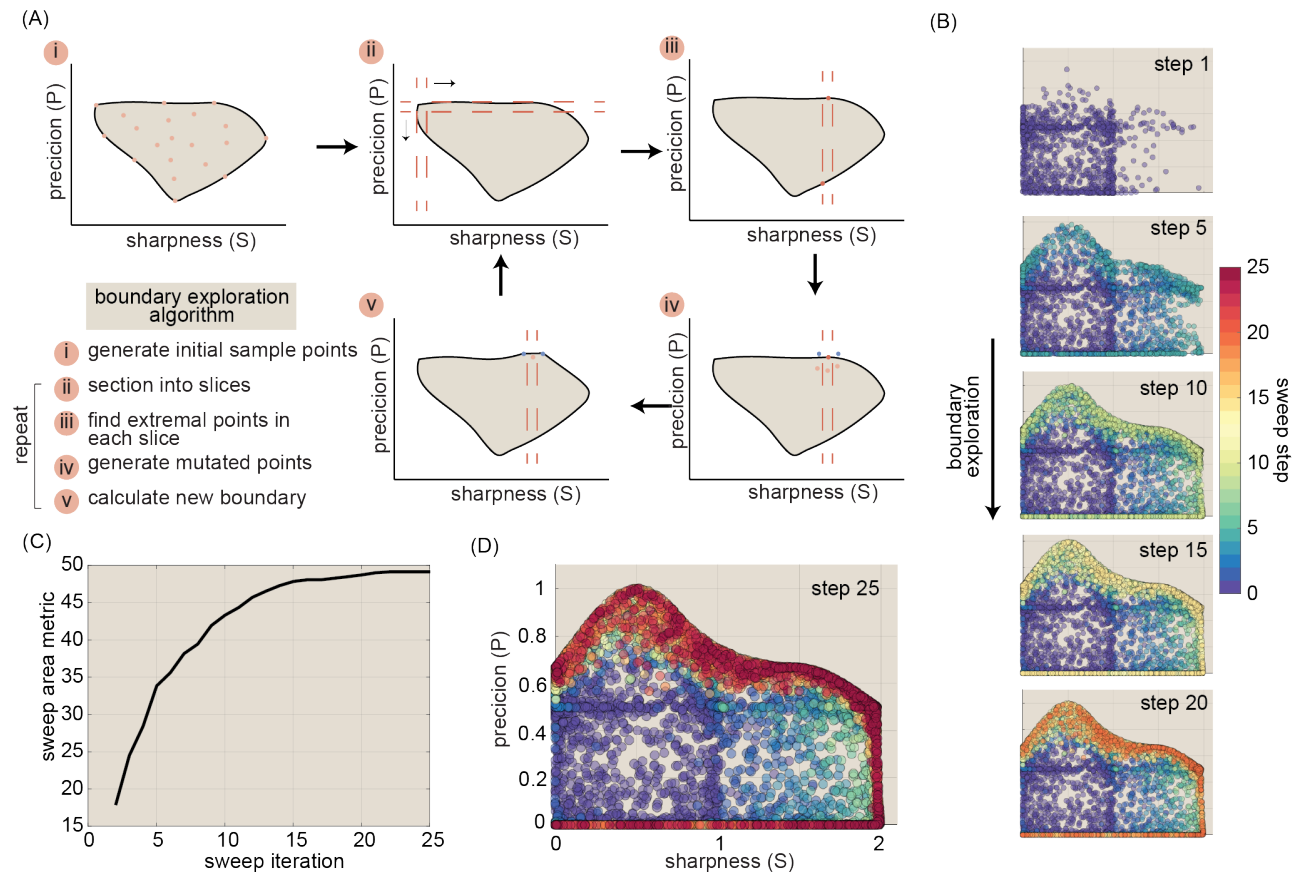


Fig. A4. A simple stochastic edge-finding algorithm for numerical parameter sweeps. (A) Schematic illustrating key steps in our parameter sweep approach (see text for details). This panel has been adapted with permission from (Eck et al., 2020). (B) Sequence of snapshots showing progress of sweep algorithm across a single run for the case of normalized sharpness (S) versus normalized precision (P). Circle color indicates the sweep step on which it was generated. (C) Plot showing 2D surface area spanned by sample points over time. (E) Plot showing final set of sample points obtained by the sweep algorithm.

system that adheres to the constraint laid out in Equation 58. Next, we check the modified terms to ensure that they adhere to magnitude constraints (typically $10^{-5}/\tau_b \geq \eta_i 10^{-5}/\tau_b$) and pass all qualifying rates along to the next step in the sweep iteration (step ii in Figure A4A). Finally, we note that, although we have focused on the simple four state system, our assumption that all binding and activation reactions are identical (see Appendix I) ensures that the exact same approach holds for all higher-order models ($N_A > 1$ or $N_B > 1$) considered in this work.

G. Testing the convergence characteristics of the parameter sweep algorithm

Here, we discuss results from a series of tests designed to assess the convergence of our sweep algorithm for key scenarios examined in the main text. This task is the most straight-forward when the algorithm is employed for “two-boundary” sweeps, such as S vs. P (Figure 3A) and S_0 vs. f (Figure 5B), where both parameters examined adhere to finite performance bounds and, thus, where the 2D region of accessible parameter space has a finite area. In this case, our general approach will be to assess whether independent runs of the algorithm (i) converge prior to the 50 run limit and (ii) reach a consistent final estimate for the area of 2D space that is attainable for different model architectures. The task becomes more complicated for “one-boundary” sweeps, such as IR vs. Φ (Figure 2A, C, and D) and IR vs. w/c , where only a single parameter (IR in each case) has a finite upper bound and the other (Φ and w/c) is limited only by bounds imposed externally as a part of sweep specification. We will begin by assessing convergence for the simpler two-sided case, and will turn thereafter to examining one-sided cases.

G.1. Sharpness vs. Precision sweeps. Figure 3A and Figure S2A and B show results for parameter sweeps examining tradeoffs between normalized sharpness (S) and normalized precision (P) for systems with 1-5 binding sites and 1-4 activation steps. We note that Figure 3C and Figure S2C also derive from these parameter sweep results. Across the board, we find that nearly all independent runs of the sweep algorithm converge according to the definition laid out above (Figure A5A and B). Moreover, for simpler architectures, we find that all independent sweep runs converge to essentially the same total area. For instance, Figure A5B shows normalized area as a function of sweep step for 500 non-equilibrium realizations of the baseline four-state model, indicating that all runs terminate near the global maximum found across all runs (dashed line). We take this as strong evidence that the algorithm is consistently exploring the full extend of 2D parameter space.

G Testing the convergence characteristics of the parameter sweep algorithm

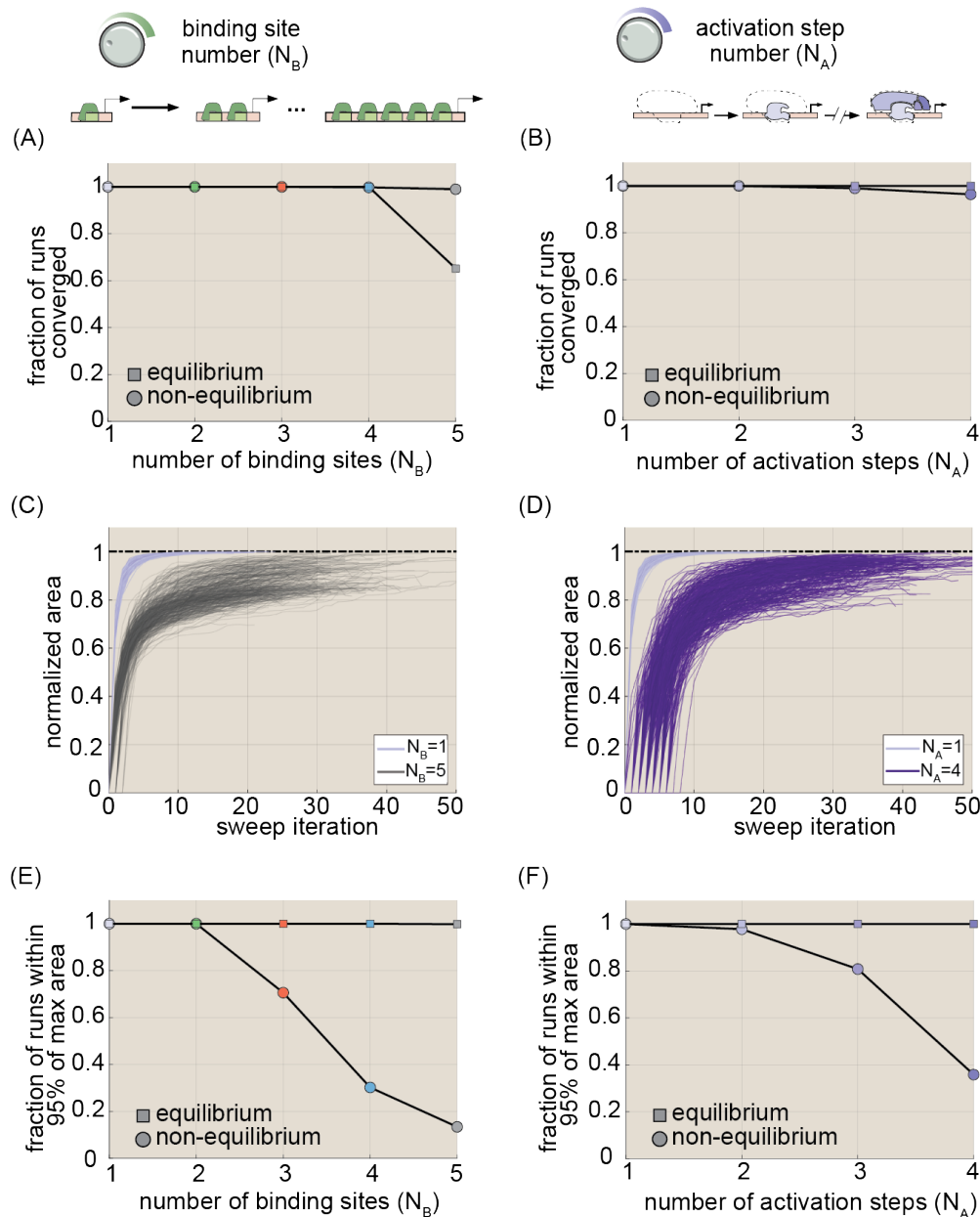


Fig. A5. Convergence results for S vs. P parameter sweeps. (A-B) Plots showing fraction of parameter sweeps that met convergence criteria for multi-binding site and multi-activation step models, respectively. Squares indicate results for equilibrium models and circles indicate non-equilibrium models. (C-D) Plots of area vs. sweep step for different model architectures. Note that the area corresponding to the first (iteration=1) is not recorded by the algorithm, and so has been estimated in each case via linear interpolation. Staggered starts apparent for $N_B = 5$ and $N_A = 5$ models indicate cases where model initialization were aborted one or more times due to an insufficient number of gene circuits meeting quality control criteria. (E-F) Fraction of parameter sweeps having a final area within 95% of the global maximum for multi-binding site and multi-step models, respectively. (All results were calculated using 500 independent runs of the sweep algorithm for each model architecture. Transition rate and interaction term magnitudes (k and η) were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

1678 As might be expected, the task of exhaustively exploring parameter space becomes more difficult as models
 1679 become more complex. Note the larger spread in outcomes for the non-equilibrium five binding site ($N_B = 5$) and 4
 1680 activation steps ($N_A = 4$) models in Figure A5C and D, respectively. Nonetheless, we find that a significant number of
 1681 sweeps converge to a consistent maximum area, even for the most complex models considered. Figure A5E and F
 1682 give the total fraction of sweeps having a final area within 95% of the global maximum as a function of binding
 1683 site number and activation step number, respectively. First, we see that 100% of sweeps for equilibrium models
 1684 uniformly meet this standard for all model architectures considered (squares in Figure A5E and F). Second, our
 1685 analysis indicates that, even for the extrema ($N_B = 5$ and $N_A = 5$), 13% and 36% of total runs, respectively (67 and
 1686 179 sweeps), still achieve final areas comparable to the global maximum, suggesting that the algorithm still does an
 1687 adequate job of exploring parameter space in these cases.

G Testing the convergence characteristics of the parameter sweep algorithm

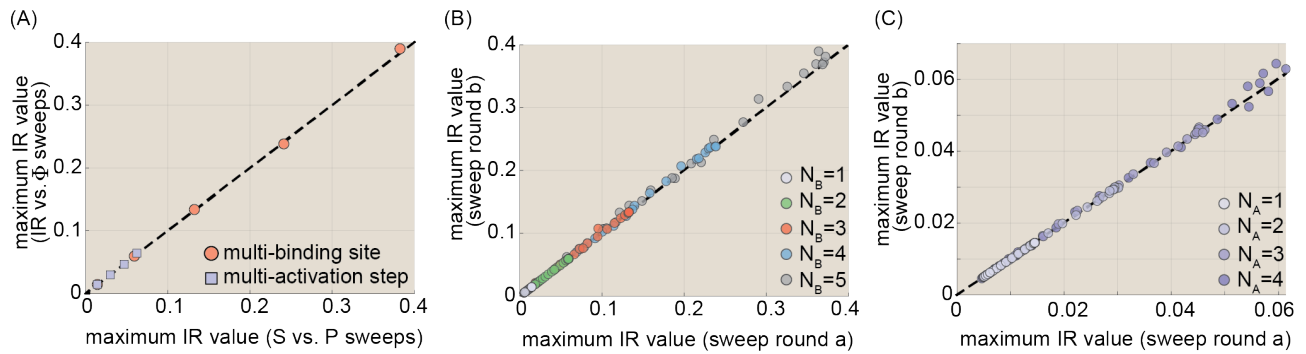


Fig. A6. Convergence results for IR vs. Φ parameter sweeps. (A) Scatter plot comparing maximum information rate estimated from S vs. P and from IR vs. Φ sweeps. (B-C) Scatter plots comparing results for the upper IR bound at different points along the curves shown in Figure 2C and D from two independent rounds of parameter sweeps comprising 200 and 500 separate runs, respectively. Points reflect IR maxima for Φ values ranging from 0.1k_BT to 5000k_BT. (Transition rate and interaction term magnitudes (k and η) were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

G.2. Information vs. energy sweeps. Next, we turn to the one-sided sweeps. First, let's consider the IR vs. Φ sweep results shown in Figure 2A, C, and D. Because Φ has no natural barrier in parameter space, the convergence metrics considered above do not provide reliable indicators of model convergence. Instead, we make use of the fact that the IR vs. Φ and the S vs. P parameter sweeps should function (either directly or indirectly) to uncover the maximum achievable non-equilibrium information rate for each model architecture. Thus, as a basic test of sweep performance, we checked for the consistency between IR estimates derived from these different sweep modalities. As illustrated in Figure A6A, we find excellent agreement between the maximum IR values derived from the S vs. P (x-axis) and IR vs. Φ (y-axis) parameter sweeps for all model architectures considered. This provides one indication the IR vs. Φ sweeps are fully exploring the relevant parameter space.

As a second check, we compared the IR vs. Φ bounds derived for two separate rounds of parameter sweeps ("round a" and "round b") comprised of 200 and 500 independent parameter sweeps, respectively. We reasoned that, if our algorithm is accurately recovering the true IR vs. Φ bound for each model architecture, this bound (i) should be replicable across different parameter sweep rounds and (ii) should be insensitive to the precise number of sweep runs per round. For each model architecture, we calculated the maximum IR value returned by sweep rounds a and b for 30 different rates of energy dissipation ranging from 0.1k_BT (close to equilibrium) to 5000k_BT (upper limit of x axis in Figure 2D). Figure A6B and C show the results of this exercise for multi-binding site and multi-activation step models, respectively, indicating excellent agreement between different sweep round for all model architectures. This demonstrates that our information vs. energy bounds are highly replicable across different rounds of sweeps. The consistency across round comprised of significantly different numbers of runs provides further evidence that we are conducting a sufficient number of independent sweeps (≥ 200) per run. Taken together, these results and the results from the preceding paragraph provide strong evidence that our algorithm is robustly recovering accurate IR vs. Φ bounds for all models considered.

G.3. Information vs. w/c sweeps. Finally, we turn to the parameter sweep results for information (and, correspondingly, decision time) as a function of wrong-to-right activator concentration (w/c) shown in Figure 4C-E. We note that the results shown in Figure 5A and C are also derived from these sweeps. Like Φ , w/c has no intrinsic boundary in parameter space and, thus, swept area provides a poor indication of convergence. Fortunately, in addition to treating w/c as a sweep parameter, we can also conduct 2D parameter sweeps where w/c is set at a constant value (e.g., $w/c = 1000$ in Figure S4B). Thus we cross-validate the IR vs. w/c bounds returned by the sweeps from Figure 4 by conducting separate sweeps of IR vs. \bar{r} (the mean transcription rate) at different w/c values (illustrated in Figure A7A). These sweeps do converge, with an average of 80% of runs reaching 93% of the global maximum.

Figure A7B and C show the results of this comparison for three different values of w/c : 10, 10^2 , and 10^3 . We focus on the architectures depicted in Figure 4, namely equilibrium systems with 1-5 binding sites (and one activation step) and non-equilibrium systems with 1-4 activation steps (and one binding site). We also test convergence for the non-equilibrium gene circuit with 5 binding sites and 1 activation step shown as a dashed line in Figure 4D. In most cases, we find good agreement between the two methods, suggesting that the IR vs. w/c sweeps are generally returning accurate estimates for the IR vs. w/c bound. We do note a couple of exceptions, however. First, we see that that IR vs. w/c sweeps appear to underestimate the upper IR bound to a significant degree for the non-equilibrium model with 4 activation steps when $w/c = 10$ (circle in upper right-hand corner of Figure A7C). This indicates that the IR vs. w/c sweep is performing sub-optimally in this case. However, since this deviation occurs in the extreme low interference regime and our focus in Section E lies on model performance at higher w/c levels ($w/c \gtrsim 100$), where our sweep algorithm performs reliably, it does not impact any conclusions drawn throughout the course of the main

H Estimating decision time ranges for different biological systems

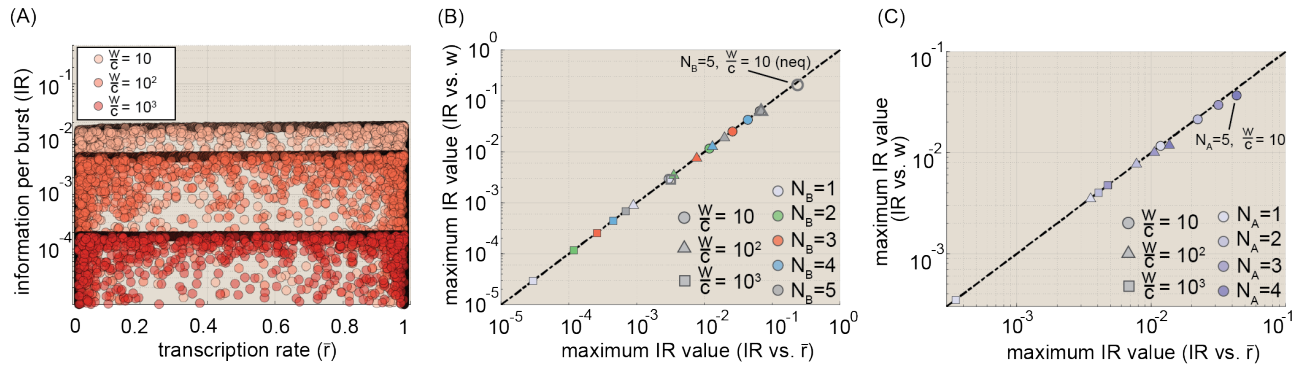


Fig. A7. Convergence results for IR vs. w/c parameter sweeps. (A) Illustrative scatter plot showing IR vs. \bar{r} sweep results for the three binding site model at equilibrium for three different values of w/c . (B-C) Scatter plots comparing parameter sweep results for the upper IR bound at three different w/c levels (10 , 10^2 , and 10^3) derived from IR vs. \bar{r} sweeps (x-axis) and IR vs. w/c sweeps (y-axis) for equilibrium multi-binding site and non-equilibrium multi-activation step models, respectively. Circles, triangles, and squares indicate w/c values of 10 , 10^2 , and 10^3 , respectively. Hollow markers in (B) indicate non-equilibrium systems. All other results in (B) are for equilibrium gene circuits (in keeping with Figure 4D). All results in (C) correspond to non-equilibrium gene circuits (in keeping with Figure 4E). (Transition rate and interaction term magnitudes (k and η) were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

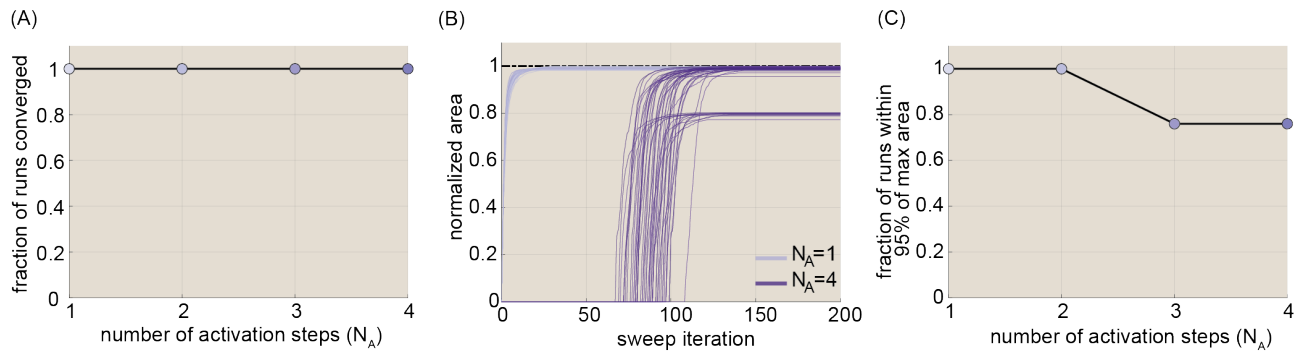


Fig. A8. Convergence results for f vs. \bar{r} parameter sweeps. (A) Plot showing fraction of runs converged as a function of the number of activation steps. All 50 runs converged for each of the four gene circuit models considered. (B) Plot showing area spanned in parameter space as a function of iteration number for all 50 runs for the $N_A = 1$ and $N_A = 4$ models. The delayed rise for $N_A = 4$ models reflects the fact that repeated initializations were required to find a sufficient number of gene circuit realizations that adhered magnitude and quality control constraints. (C) Fraction of total runs for each model time that reached a final area greater than or equal to the 95% of the global maximum across all runs. (w/c was set to 10^3 for all runs. Transition rate and interaction term magnitudes (k and η) were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{ab} \geq 1$ and $\eta_{ib} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

1729 text. We note that the IR vs w/c sweeps similarly underestimate the IR bound non-equilibrium realizations of the
 1730 5 binding site model when $w/c = 10$ (hollow gray circle in upper right-hand corner of Figure A7B). In this case,
 1731 however, even the IR vs. \bar{r} parameter sweeps do not converge reliably, with only 3-4% of sweeps reaching 95%
 1732 of the global maximum. Thus, we are unable to assess the full extent to which the IR vs w/c is sub-optimal in this case.
 1733 Once again, though, this claims in the main text rely only on the IR bound when w/c is large ($w/c \gtrsim 10^3$); a regime
 1734 in which we find that the sweeps perform reliably (hollow gray square in Figure A7B). Thus, we conclude that the IR
 1735 vs. w/c sweeps provide a viable basis for the investigations undertaken in this study.

1736 **G.4. Specificity vs. N_A results.** We claim in the main text that, out of equilibrium, the specificity is bounded by the
 1737 number of activation steps, such that $f \leq \alpha^{N_A+1}$. Here α is the affinity factor (set to 100) that reflects intrinsic differ-
 1738 ences in the binding kinetics between cognate and non-cognate factors ($\alpha = k_u^w/k_u$). Figure S4B shows parameter
 1739 sweep results in support of this claim. These results are derived from 2D f vs. \bar{r} sweeps. Figure A8 shows conver-
 1740 gence statistics for these runs for non-equilibrium systems with 1 to 4 activation steps and 1 binding site. We find
 1741 that all 50 sweep runs met their convergence criteria for each run (Figure A8A) and, further, that no fewer than 76%
 1742 of runs converged to a 2D area that was within 95% of the global maximum. This indicates that these parameter
 1743 sweep results converge reliably to consistent overall values for specificity and a function of transcription rate and,
 1744 thus, that they provide a sound basis for assessing the maximum achievable non-equilibrium specificity as a function
 1745 of N_A .

1746 **H. Estimating decision time ranges for different biological systems**

1747 ***Caenorhabditis elegans* decision time estimation.** A recent study by Lee and colleagues (Lee et al., 2019) used live
 1748 imaging to examine Notch-dependent burst dynamics in the *sygl-1* gene in the germ line of young adult nematodes.
 1749 Their results indicate that the gene exhibits burst cycle times ranging from 60.5 minutes up to 105.3 minutes (see

Figure 2 E and F in (Lee et al., 2019)). Meanwhile, a review article indicated potential values for the cell cycle time for adult germ-line cells in *C. elegans* as ranging from 16 to 24 hours (Hubbard, 2007). A separate study examining nonsense-mediated mRNA decay in *C. elegans* reported a half life of approximately 6 hours for the *rpl-7A* gene (Figure 4k in (Son et al., 2017)). If we take the cell cycle time as the upper time limit for cellular decision-making, this leads to an estimate of $1440/60.5 = 23.8$ burst cycles.

Mus musculus decision time estimation. Burst cycle time estimates were taken from Table A.1 in Appendix A of (Lammers et al., 2020), which indicates times ranging from 30 minutes to a “few hours”. mRNA half life estimates were taken from Table 1 of (Pérez-Ortín et al., 2013), which indicates a range of 30 minutes to 30 hours for mouse cells. To estimate the effective decision time corresponding to an mRNA half-life of 30 hours (1,800 minutes), we recognize that, once mRNA levels have reached a steady state, they will reflect (in effect) a weighted average of the preceding transcriptional activity, where weights moving backward in time contribute

$$w(t) = e^{-\frac{t}{\tau_{\text{mRNA}}}}, \quad (59)$$

where t indicates temporal distance from the present and τ_{mRNA} is the exponential time constant, given by $\tau_{\text{mRNA}} = t_{1/2} / \ln(2)$. Integrating Equation 59, we find that τ_{mRNA} time steps are effectively present in steady-state mRNA levels. Taking 30 minutes as the lower bound for bursting timescales, this yields an upper bound of $(1800 / \log 2) / 30 = 86.6$ burst cycles. We note that this estimate is not materially different from the 60 cycle estimate that would be obtained by simply dividing 1,800 by 30.

Drosophila melanogaster decision time estimation. We take the duration of nuclear cycle 14, which follows the thirteenth (and final) round of synchronous cellular divisions in early *Drosophila melanogaster* development, as the relevant timescale for cellular decisions in early fruit fly development. Studies have found that the duration of this developmental period varies along the embryo, with a minimum duration of 65 minutes (Foe and Alberts, 1983). To estimate bursting timescales, we use burst inference results from our previous work (Lammers et al., 2020), which indicate a burst cycle time of approximately 2 minutes for the *even-skipped* gene. Thus, we arrive at an upper limit of $65/2 = 32.5$ cycles.

I. Higher-order molecular models

Here we provide an overview of key modeling assumptions underlying our approach to modeling gene circuits with multiple activator binding sites or multiple activation steps.

1.1. Gene circuits with multiple activator binding sites. A key feature of eukaryotic enhancers is the presence of multiple distinct binding sites for regulatory factors (Vincent et al., 2016; Erokhin et al., 2015). To better understand the impact of variable numbers of binding sites on information transmission, this work examines gene circuit models with between 1 and 5 activator binding sites. In so doing, we maintain the same basic MWC architecture outlined in the context of the simple 4 state model with one activator binding site shown in Figure 1B. No number of bound activators is alone sufficient for mRNA production, but each contributes an extra factor of η_{ab} and η_{ib} to impact locus activation dynamics. In all cases, we assume a single molecular activation step for multi-binding site models ($N_A = 1$). Finally, we also allow for cooperative interactions between bound activator molecules, which are captured by the interaction term η_{ub} .

Figure A9A illustrates what this looks like for a model gene circuit with two activator binding sites. The model has eight total states, with four inactive states (top) and four active states (bottom). There are several features to note. First, the transitions between states 2 and 6, which feature two bound activator molecules, are weighted by squared interaction terms, η_{ab}^2 and η_{ib}^2 , to reflect the regulatory influence of two activators on locus activation dynamics. More generally, if n activators are bound, these weights are raised to the n th power (i.e. η_{ab}^n and η_{ib}^n). Second, note that unbinding reactions leading out of states 2 and 6 are multiplied by the additional η_{ub} mentioned above. This reflects interactions between bound molecules. For simplicity, we assume that η_{ub} is the same for both cognate and non-cognate activator species, as well as for interactions between cognate and non-cognate activators. In general, unbinding reactions out of states with n activators bound will be weighted by η_{ub}^{n-1} to reflect interactions from the remaining bound factors.

Lastly, a key simplifying assumption that we make in this work is that each activator binding site is identical with respect to its regulatory influence on the gene locus. As a result, it does not matter *which* binding sites are bound, only *how many* are bound. In the context of Figure A9A, this means that states 1 and 5 are functionally identical to states 3 and 7, respectively. Thus, these states can be combined into single coarse-grained states, which leads to an effective model with 6 states, rather than 8. This ability to coarse grain is invaluable for more complex architectures, since it means that the total number of unique molecular states scales as $N_A N_B$, rather than $N_A^{N_B}$.

1802 **1.2. Gene circuits with multiple molecular activation steps.** Setting $N_A > 1$ is intended to reflect the reality that multi-
1803 ple distinct molecular reactions—e.g., mediator engagement, PIC assembly, nucleosome displacement, etc.—are
1804 necessary preconditions for achieving productive transcription. In the main text we investigate the performance of
1805 gene circuits whose transcriptional activity is dictated by 1-4 molecular components, each of which can be either
1806 engaged (compatible with transcription) or disengaged (incompatible with transcription). In their simplest interpreta-
1807 tion, “engaged” and “disengaged” states might correspond to the presence or absence of some critical component
1808 of the transcriptional machinery at the gene locus; however, we remain intentionally non-committal about their phys-
1809 ical interpretation, since these generic states are meant to capture a broad swath of potential molecular reactions.
1810 For instance, in the case of a nucleosome, the “engaged” state would correspond to the *absence* of the nucleosome
1811 (Mirny, 2010). The terms could also capture conformational shifts in key macromolecules such as mediator (Nogales
1812 et al., 2017), or in the topology of the gene locus itself.

1813 We assume that each component is required for transcription, such that, in a model with n molecular components
1814 only molecular states with all n components engaged are transcriptionally active, and $N_A = n$ activation steps are
1815 required to achieve locus activation. Furthermore, while in reality each molecular component is likely characterized
1816 by heterogeneous dynamics (see, e.g., (Lammers et al., 2020)) we again make the simplifying assumption that each
1817 molecular step is identical. As a result, it does not matter which molecular components are engaged, only how many.
1818 Figure A9B shows how this logic plays out for the case where $N_A = 2$. As with the $N_B = 2$ case, the model gene circuit
1819 has 8 states; however, in this case, only two states (5 and 6)—the ones in which both components are engaged—
1820 are transcriptionally active. Note that the binding and unbinding reactions connecting these states are weighted by
1821 factors of η_{ba}^2 and η_{ua}^2 , respectively, to reflect the influence of each molecular factor. In general, if n components
1822 are engaged, these factors are raised to the n th power. In addition, we allow for cooperative interactions between
1823 molecular components (curved arrow in states 1,2 and 4-7), captured by the η_{aa} and η_{ia} terms in Figure A9B. In
1824 general these terms are raised to the power of $n - 1$, where n is the number of engaged components at the initial
1825 molecular states.

1826 **1.3. Future directions.** Throughout this work, we have treated activator binding sites and activation steps as orthogonal
1827 axes of gene circuit complexity. In reality, of course, both elements are likely at play in gene regulatory architectures.
1828 We choose to investigate the impact of each independently for two chief reasons: first it greatly simplifies exposition
1829 and allows us to more easily isolate how each aspect of gene locus architecture interacts with energy dissipation
1830 to dictate rates of information transmission. Second, since model complexity scales as $N_A N_B$, we are limited in our
1831 ability to accurately explore the performance of models where both N_A and N_B are large. Improving computational
1832 and numerical techniques to permit such explorations represents an interesting future direction. We note also that
1833 such models should be tractable without need for additional development if limited to operate at equilibrium.

1834 In addition, we wish to emphasize the potential importance of allowing for heterogeneity, both in the properties of
1835 different binding sites along the enhancer and between different molecular components within the activation pathway.
1836 This question seems especially interesting in the context of the molecular activation steps. Our simple model with
1837 identical steps likely represents the floor of system performance. How much is to be gained when each reaction can
1838 adhere to its own kinetics, and exert a distinct kind of regulatory influence over the gene locus?

1839 J. Deriving normalized sharpness and precision metrics

1840 In Figure 1B, the transcriptional sharpness, s , is defined as the first derivative of the transcriptional input-output
1841 function multiplied by the activator concentration, c^* , such that

$$1842 \quad s = \frac{dr}{dc} c^*. \quad (60)$$

1843 The transcriptional precision, p , is defined as the inverse of the intrinsic noise in the transcriptional input-output
1844 function:

$$1845 \quad p = \frac{1}{\sigma}, \quad (61)$$

1846 where σ is as defined in Equation 11 in the main text. Under these definitions, a key challenge in comparing sharp-
1847 ness and precision levels across different gene circuits is that the upper bounds on both s and p depend on the frac-
1848 tion of time the system spends in the transcriptionally active state, π_a (defined in Equation 10). Figure A10A and B
1849 illustrate this π_a -dependence for equilibrium and non-equilibrium realizations of the four-state system defined in Fig-
1850 ure 1C. As an example: the equilibrium bound on s is 0.25 when $\pi_a = 0.5$, but only 0.09 when $\pi_a = 0.1$ (Figure A10A).
1851 Since we allow gene circuits to take on different transcription rates ($r = \pi_a r_0$) at $C = c^*$, this π_a -dependence thus
1852 confounds our efforts to understand how the molecular architecture of gene circuits—the number of binding sites,
1853 number of molecular steps, and presence or absence of energy dissipation—dictates transcriptional performance.

1854 To overcome this issue, we need to normalize s and p such that they are independent of π_a . Focusing first on
1855 sharpness, we were inspired by previous works (Estrada et al., 2016; Grah et al., 2020) to leverage Hill Function as

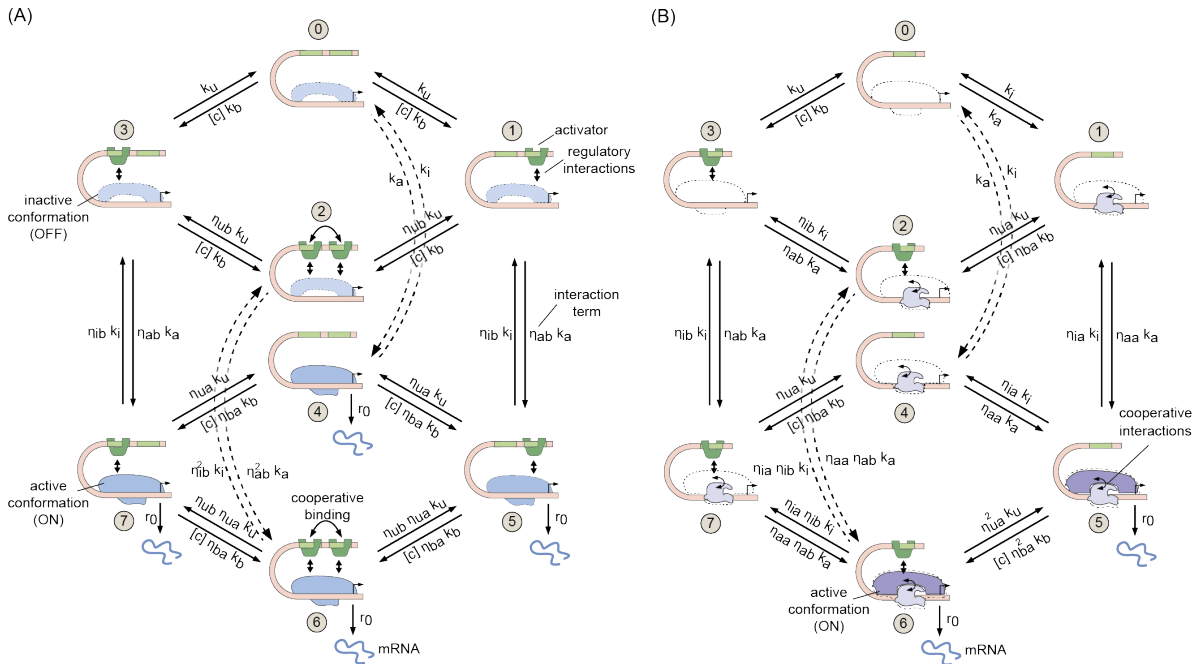


Fig. A9. Higher order gene circuit models. (A) Cartoon indicating the molecular architecture of a model gene circuit with one activation step and two activator binding sites ($N_A = 1$ and $N_B = 2$). (B) Molecular architecture of a model gene circuit with two activation steps and one activator binding site ($N_A = 1$ and $N_B = 2$).

a flexible conceptual tool for extracting generic sharpness measures. The Hill function is defined as:

$$\pi_a = \frac{c^S}{c^S + K_d^S}, \quad (62)$$

where c is the activator concentration, S is the Hill coefficient, and K_d is a constant that dictates the location of the function's half-max point. In general, the input-output functions generated by our model gene circuits will have more complex functional forms, but nonetheless, Equation 62 indicates that we can relate these more complex functions to the Hill function via the shared parameters π_a and c .

The sharpness of the Hill function has the form:

$$s_H = S \frac{c^S K_d^S}{(c^S + K_d^S)^2}. \quad (63)$$

To better relate this to our input-output function, we need to re-express K_d in terms of C and π_a . Solving Equation 62 for K_d yields

$$K_d = c \left(\frac{1 - \pi_a}{\pi_a} \right)^{\frac{1}{S}}. \quad (64)$$

Plugging this in to Equation 63 we obtain, after simplification:

$$s = \pi_a (1 - \pi_a) S. \quad (65)$$

This expression tells us that the sharpness (s) of a Hill function with activity level π_a at $C = c^*$ is equal to the Hill coefficient, S , multiplied by the term $\pi_a(1 - \pi_a)$. By rearranging, we can obtain the Hill coefficient as a function of s and π_a

$$S = \frac{s}{\pi_a(1 - \pi_a)}. \quad (66)$$

Thus, for a generic gene circuit input-output function with sharpness s and expression level π_a at $C = c^*$ we can invoke Equation 66 to calculate the Hill coefficient for the equivalently sharp Hill function (Figure A10C). This provides us with a generic measure of transcriptional sharpness that is independent of π_a and thus can facilitate comparisons across gene circuits that drive differing activity levels at $C = c^*$ (Figure A10D). We refer to this independent sharpness metric as the "normalized sharpness" in the main text, and denote it with the variable S .

This leads us to the question of transcriptional precision. The two key considerations in defining the normalized precision metric, P , are that (i) we want it to yield a quantity proportional to the information rate when multiplied with S (Equation 66), where

$$IR = \left(\frac{\delta c}{c^*} \right)^2 S^2 P^2, \quad (67)$$

K Optimal equilibrium four-state gene circuits behave like effective two state systems

and (ii) we want it to adhere to a single upper bound, regardless of π_a . There is only one definition that satisfies the first constraint:

$$P = p(\pi_a(1 - \pi_a)) = \frac{\pi_a(1 - \pi_a)}{\sigma}. \quad (68)$$

Happily, Equation 68 exhibits consistent upper bounds for all π_a values, and thus satisfies our second constraint (Figure A10E).

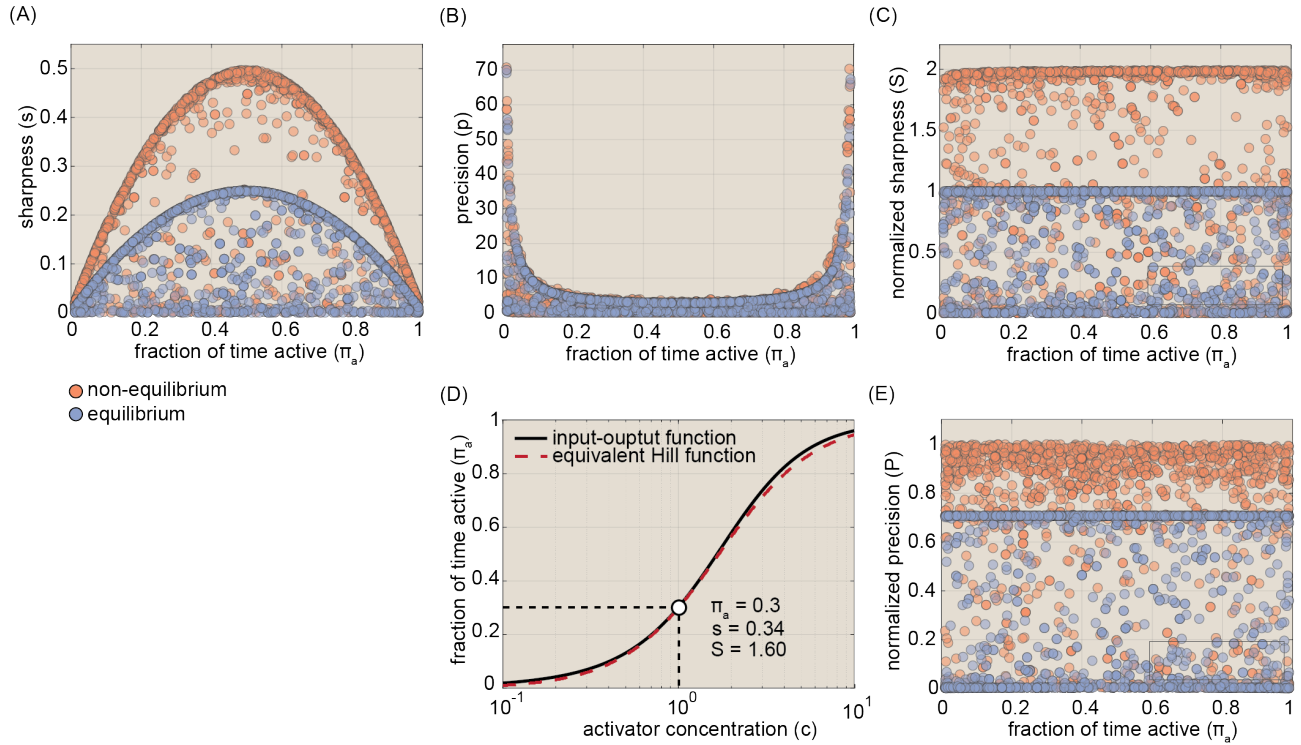


Fig. A10. Defining normalized sharpness and precision. (A) Plot depicting the upper sharpness limit for equilibrium (blue) and non-equilibrium (red) realizations of the four-state system depicted in Figure 1C. The upper limit depends on the fraction of time spent in the active state, π_a . (B) Plot of precision as a function of the transcription rate. Here again, the upper bounds depend on π_a . (C) Plot of normalized sharpness as a function of the transcription rate. In the case, the upper limits are invariant. (D) Illustration of normalized sharpness concept. For a given input-output curve, we identify normalized sharpness, S, as the Hill coefficient of an equivalently sharp Hill function with the same expression level at $C = c^*$. (E) On the other hand, the normalized precision, P, exhibits invariant performance bounds.

K. Optimal equilibrium four-state gene circuits behave like effective two state systems

In this section, we calculate the normalized sharpness (S) and precision (P) for a simple 2 state gene circuit (Figure A11A) with one ON state and one OFF state and two transition rates, k_{off} and k_{on} . We assume that activator binding dictates fluctuations into and out of the ON state, such that k_{on} is proportional to c ($k_{\text{on}} = ck_{\text{on}}^0$). For this simple system, the rate of transcription is given by

$$\bar{r} = r_0 \frac{ck_{\text{on}}^0}{ck_{\text{on}}^0 + k_{\text{off}}} = r_0 \pi_a. \quad (69)$$

Differentiating this expression with respect to c and setting $r_0 = 1$ (as in main text), we find that

$$s = \frac{ck_{\text{on}}^0 k_{\text{off}}}{(ck_{\text{on}}^0 + k_{\text{off}})^2}. \quad (70)$$

Finally, dividing through by $b = \pi_a(1 - \pi_a)$ yields the normalized sharpness, which is simply given by

$$S = 1. \quad (71)$$

Thus, we see that the two state model is constrained to a normalized sharpness level that represents the upper performance limit for the four-state model operating at equilibrium (blue circles in Figure 3A).

Next, we turn to precision. From Equation 11, we find that

$$\sigma^2 = \frac{2ck_{\text{on}}^0 k_{\text{off}}}{(ck_{\text{on}}^0 + k_{\text{off}})^3}. \quad (72)$$

L Sharp and precise non-equilibrium networks exhibit distinct and incompatible microscopic topologies

Inverting and multiplying by b^2 gives

$$P^2 = \frac{ck_{\text{on}}^0 + k_{\text{off}}}{2(ck_{\text{on}}^0 + k_{\text{off}})}. \quad (73)$$

Finally, multiplying through by τ_b (Equation 17) and taking the square root gives

$$P = \frac{1}{\sqrt{2}}, \quad (74)$$

which, again, is equivalent to the upper limit of the four-state gene circuit at equilibrium (Figure 3A).

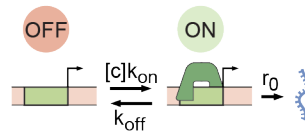


Fig. A11. A simple 2 state model of transcription. Cartoon of a simple 2 state gene circuit model in which activator binding and unbinding dictate transitions into and out of a transcriptionally active state.

L. Sharp and precise non-equilibrium networks exhibit distinct and incompatible microscopic topologies

One simple way to probe the microscopic architectures of different gene circuits is to measure the degree of heterogeneity (or dispersion) in (a) transition rates and (b) state probabilities. We developed entropy-based dispersion metrics ranging from 0 to 1 to quantify how uniform (0) or heterogeneous (1) transition rates and state probabilities were for different realizations of the four-state network shown in Figure 1C. While crude, these measures can provide useful microscopic insights. For instance, in gene circuits with a state probability score of 0 each microscopic state must be equiprobable ($\pi_1 = \pi_2 = \pi_3 = \pi_4 = 1/4$), while those with a 1 are maximally heterogeneous. In general, maximal heterogeneity corresponds to the case when one and only one state has a nonzero probability; however, since, for simplicity, we have elected here to focus on gene circuits where $\bar{r} = 0.5$, the maximum instead corresponds to a case when two molecular states (one OFF and one ON) have probability $\pi_i = 0.5$. Similar considerations hold for the transition rate axis. We conducted parameter sweeps to explore the space of achievable dispersion values for 10,000 non-equilibrium gene circuits (gray circles in Figure A12A).

From Figure A12A, we can see immediately that precise and sharp gene circuits occupy opposite extremes of dispersion space. Specifically, precise systems exhibit highly uniform state probability and transition rate values, while sharp networks are highly heterogeneous, both with respect to the fraction of time spent in each state and the relative magnitudes of their transition rates. These stark differences, as well as the tight clustering of each motif, suggest that sharpness and precision arise from distinct and non-overlapping microscopic topologies.

Detailed examination of maximally precise gene circuits from our parameter sweeps indicated that these systems exhibit highly uniform molecular architectures wherein each microscopic state is equiprobable, all clockwise transition rates are uniform, and all counterclockwise rates are negligible. This results in a “clock-like” system that maximizes the regularity of molecular transitions. Maximally sharp gene circuits, on the other hand, exhibit an all-or-none character, behaving as effective two state systems that spend most of their time either activator-bound and active (0), or unbound and inactive (1), and which have effective ON and OFF rates that are concentration-dependent (see Appendix M for further details).

M. A hierarchy of microscopic transition rates underpins non-equilibrium sharpness gain

Figure 3A shows that energy dissipation opens up a broad spectrum of S and P values that are not attainable at equilibrium. It is difficult to formulate general statements that apply to all gene circuit models inhabiting these spaces beyond the upper equilibrium limit; however, we can learn much by examining the architecture of gene circuits lying at the outer limits of non-equilibrium performance, since these systems tend to distill the logic underpinning non-equilibrium performance gains into relatively simple regulatory motifs.

Such is the case for the IR-optimized non-equilibrium four-state systems depicted as gray circles in Figure 3A. In Main Text Section D, we found that the driver of this IR is a twofold increase in sharpness relative to the upper equilibrium limit. To realize this twofold sharpness gain, we find that non-equilibrium driving is harnessed to facilitate effective one-way transitions between the active and inactive conformations—specifically, from states 1 to 2 and 3 to 0 in Figure 1C—ensuring that the system will have a strong tendency to complete transcriptional cycles in the clockwise direction ($J > 0$).

In addition to this non-equilibrium driving, sharpness maximization places strict constraints on the relative magnitudes of microscopic transition rates within the network. To understand these constraints, it is instructive to consider

1944 a coarse-grained representation of our network with a single ON state (2) and a single OFF state (0). We can obtain
 1945 expressions for the two effective transition rates in the network by recognizing that they are equal to the inverse of
 1946 the mean first passage times between states 2 and 0, which we can calculate using Equation 16 from Appendix A.

1947 If we neglect the energetically disfavored transitions from 2 to 1, the effective ON rate (k_{on}^* in Figure 3D) takes on
 1948 a relatively simple form

$$1949 \quad k_{\text{on}}^* = \frac{[c]k_b k_a \eta_{\text{ab}}}{[c]k_b + \eta_{\text{ab}}k_a + k_u}. \quad (75)$$

1950 From Equation 75, we see that the effective ON rate becomes proportional to the concentration, c , when the factor
 1951 of $[c]k_b$ becomes negligible in the denominator. The limit where $k_u \gg [c]k_b, \eta_{\text{ab}}k_a$ represents a scenario in which
 1952 the activator K_d is larger when the network is in the inactive conformation than when in the active conformation
 1953 such that the activator must bind multiple times (on average) before it succeeds in driving the system into the active
 1954 conformation. The other limit, when $\eta_{\text{ab}}k_a \gg [c]k_b, k_u$, corresponds to a system where locus activation happens
 1955 rapidly upon activator binding.

1956 In similar fashion, the effective OFF rate can be expressed as the inverse of the first passage time from 3 to 1

$$1957 \quad k_{\text{off}}^* = \frac{\eta_{\text{ua}}k_u k_i}{\eta_{\text{ba}}[c]k_b + k_i + \eta_{\text{ua}}k_u}. \quad (76)$$

1958 Interestingly, we see that the effective k_{off}^* becomes *inversely* proportional to c when activator binding rate exceeds
 1959 both the unbinding rate and the rate of locus deactivation ($\eta_{\text{ba}}[c]k_b \gg k_u, k_{\text{off}}^*$). This imbalance causes the system
 1960 to become kinetically trapped in the active conformation for multiple cycles of activator unbinding and rebinding, with
 1961 an average duration inversely proportional to $\eta_{\text{ba}}[c]k_b$.

1962 Thus, when the proper hierarchy of microscopic rates is realized, our four-state network behaves *as though* it were
 1963 a two state system in which both the on and off rates are concentration dependent, such that

$$1964 \quad \bar{r} \approx \frac{k_{\text{on}}^*}{k_{\text{on}}^* + k_{\text{off}}^*} \approx \frac{[c]\lambda}{[c]\lambda + \frac{\gamma}{[c]}}, \quad (77)$$

1965 where λ and γ are coarse-grained transition rates with units of $s^{-1}[c]^{-1}$ and $s^{-1}[c]$, respectively. Repeating the
 1966 calculations from Appendix K for the above effective two state system will yield an S value of 2 and a P value of 1, in
 1967 agreement with our numerical results from Figure 3A. We propose that this doubled concentration dependence can
 1968 be conceptualized as a kind of “on rate-mediated” proofreading. In contrast to classical kinetic proofreading, which
 1969 works by amplifying intrinsic differences in ligand off rates (Hopfield, 1974; Ninio, 1975), sharp networks amplify the
 1970 concentration-dependence carried by binding rates, effectively “checking” C twice per cycle since both k_{on}^* and k_{off}^*
 1971 are functions of the activator concentration c .

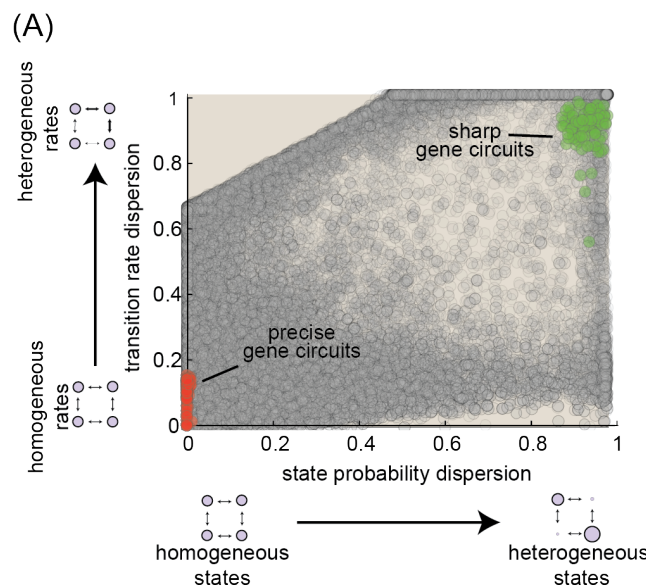


Fig. A12. Sharp and precise non-equilibrium networks exhibit distinct and incompatible microscopic topologies. Plot showing dispersion scores for state probabilities and transition rates for 50,000 non-equilibrium networks. Here, a score of 0 indicates maximal uniformity (all rates or probabilities are equal) and a 1 indicates maximal heterogeneity. Green and red circles indicate the scores for the 100 gene circuits within 2% of the maximum achievable non-equilibrium sharpness and precision levels, respectively. (Transition rate and interaction term magnitudes (k and η) were constrained such that $10^{-5} \leq k\tau_b \leq 10^5$ and $10^{-5} \leq \eta \leq 10^5$, where τ_b is the burst cycle time. η_{ab} and η_{ib} were further constrained such that $\eta_{\text{ab}} \geq 1$ and $\eta_{\text{ib}} \leq 1$, consistent with our assumption that the transcription factor activates the gene locus.)

N. Non-equilibrium gains in sharpness drive IR increases in more complex regulatory architectures

This appendix section contains additional discussion relating to sharpness-precision tradeoffs for higher-order model architectures with multiple binding sites or multiple activation steps.

N.1. Sharpness maximization remains optimal for systems with multiple binding sites. To assess whether sharpness-maximization remains the optimal strategy for more complex architectures featuring multiple activator binding sites, we employed parameter sweeps to examine the space of achievable S and P values for gene circuits with 1-5 activator binding sites (and N_B fixed at 1). Figure S2A shows the results of this analysis. For ease of comparison across different models, we plot the relative gains in S and P for each model with respect to their maximum equilibrium values. For instance, the maximum equilibrium S value for the $N_B = 2$ model is 2, so a non-equilibrium gene circuit model with two binding sites that exhibits an S value of 2.5 will be calculated to have a sharpness gain of $2.5/2 = 1.25$.

Figure S2A reveals that the sharpness-precision tradeoff observed for the one-binding site model persists and, indeed, becomes more severe for systems with additional activator binding sites. We see that the non-equilibrium gain in S is fixed at approximately 2. And while the non-equilibrium gain in P increases from $\sqrt{2}$ for $N_B = 1$ to approximately 2.25 for $N_B = 5$, these P maxima (peaks in the upper left quadrant of Figure S2A) occur at lower and lower values of S, which renders them more and more disadvantageous from an IR perspective. As a result, when we plot IR-optimal gene circuits for each value of N_B (colored circles in Figure S2A), we find that they are invariably located in regions where $S/S_{eq} \approx 2$ and $P/P_{eq} \approx 1$. These results demonstrate that spending energy to maximize sharpness remains the key to maximizing transcriptional information transmission, irrespective of the number of activator binding sites.

N.2. Multiple activation steps increases upper sharpness bound away from equilibrium. Figure S2B shows the range of achievable non-equilibrium gains in S and P for systems with 1-4 activation steps (and $N_B = 1$). Once again we observe a strong tradeoff between sharpness and precision, which suggests that this incompatibility is a general feature of transcriptional systems. And, once again, we find that IR-maximizing gene circuits (colored circles) lie at or near the right-most edge of achievable parameter space, indicating that dissipating energy to enhance transcriptional sharpness (rather than precision) remains the best strategy for maximizing the IR.

Yet unlike the systems examined in Figure S2A, Figure S2B reveals that the non-equilibrium gain in transcriptional sharpness (S) is not fixed but, rather, increases with the number of molecular steps from a factor of two when $N_A = 1$ to a factor of five when $N_A = 4$. This indicates that increasing the number of dissipative molecular steps in the activation pathway raises the upper limit on the sharpness of the transcriptional input-output function, even when the number of binding sites is held constant.

O. Specificity definitions and details

This Appendix Section uses a simple two state gene circuit model to compare and contrast the specificity definition employed in two recent works (Shelansky and Boeger, 2020; Grah et al., 2020), which compares how a single transcription factor (“TF”) activates at two different gene loci (the “TF-centric” approach)—a target locus with specific binding sites, and a non-cognate locus that lacks binding site—with the definition employed in this work, which focuses on cognate and non-cognate factors competing to activate a single locus (the “gene-centric” approach).

O.1. A detailed comparison of specificity definitions for a simple 2-state model of transcription. Figure A13A illustrates the second “Tf-centric” scenario for the case of a simple two state network with a single binding site and no possibility of a conformation change at the locus; however the same idea applies equally well for the 4 state network we considered above, as well as more complicated architectures. Here transcriptional specificity is defined as the ratio of the average steady state transcription rates at on- and off-target gene loci:

$$f_{TF} = \frac{\bar{r}_r}{\bar{r}_w}, \quad (78)$$

where f_{TF} is the specificity under the TF-centric framing of the problem, and r_r and r_w indicate the transcription rates at the cognate (right) and non-cognate (wrong) loci, respectively. In (Shelansky and Boeger, 2020), the authors show that specificity for the two state system shown in Figure A13A is given by:

$$f_{TF} = \frac{\alpha k_u + [c]k_b}{k_u + [c]k_b}. \quad (79)$$

From Equation 79, we see that the activator specificity is bounded from above by α . Moreover, this upper performance limit is achieved only in an off rate-dominated regime where $k_u \gg [c]k_b$, which the authors in (Shelansky

2021 and Boeger, 2020) note leads to a runaway increase in transcriptional noise with increasing specificity under the con-
 2022 straint that the mean transcription rate must remain constant. As a result, the authors conclude that non-equilibrium
 2023 network architectures are necessary in order to improve specificity and minimize transcriptional noise (Shelansky
 2024 and Boeger, 2020).

2025 In analogy to the parallel case outlined above, we employ a “gene-centric” definition (Figure A13B), which takes
 2026 specificity as the ratio of the average number of cognate and non-cognate factors bound while the locus is in a
 2027 transcriptionally productive state, normalized by concentration:

$$2028 \quad f = \frac{w}{c} \frac{\pi_c}{\pi_w}. \quad (80)$$

2029 In the case of the two state model shown in Figure A13B, this is simply given by the ratio of fractional occupancies
 2030 of states 1 and 1*:

$$2031 \quad f = \frac{w}{c} \frac{\pi_2}{\pi_{2^*}}. \quad (81)$$

2032 Since in steady state this is necessarily at equilibrium (note the absence of cycles), we can express this ratio as a
 2033 function of the difference between the energies of cognate and non-cognate factor binding, ε_c and ε_w , which leads
 2034 to

$$2035 \quad f = \frac{w}{c} e^{-\frac{(\varepsilon_c - \varepsilon_w)}{k_B T}}. \quad (82)$$

2036 Next, we note that the energies can be expressed as ratios of binding and unbinding rates, such that

$$2037 \quad \varepsilon_c = -k_B T \ln \frac{ck_b}{k_u} \quad (83)$$

2038 and

$$2039 \quad \varepsilon_w = -k_B T \ln \frac{wk_b}{\alpha k_u}. \quad (84)$$

2040 Plugging these two expressions into Equation 82, we have

$$2041 \quad f = \frac{w}{c} e^{\ln \frac{ck_b k_u}{\alpha k_u w k_b}}, \quad (85)$$

2042 which simplifies to a simple equality

$$2043 \quad f = \alpha. \quad (86)$$

2044 From Equation 86, we see that f is simply equal to the binding specificity factor α for our three state network,
 2045 *irrespective of binding kinetics*. Thus, in contrast to (Shelansky and Boeger, 2020), we find that equilibrium gene
 2046 circuits need not shift towards a noisy, off rate-dominated regime to achieve maximum fidelity; indeed *all* systems
 2047 necessarily achieve precisely $f = \alpha$. Intuitively, this difference stems from the fact that our model captures the effects
 2048 of kinetic competition between cognate and non-cognate activators: whenever the cognate activator (green square in
 2049 Figure A13B) is bound, non-cognate factors cannot bind.

2050 A key limitation of this approach is that it neglects the presence of non-specific stretches of regulatory DNA, even
 2051 at cognate gene enhancers. Thus, to more accurately reflect the specificity challenges faced by real gene loci, a
 2052 synthesis of the two approaches summarized above will be necessary, which considers competition between cognate
 2053 and non-cognate factors to bind and activate a gene locus that features both specific binding sites (which favor the
 2054 cognate activator) and neutral sites (to which all activator species bind non-specifically). One expectation for such a
 2055 scenario is that the simple equality stated in Equation 86 will no longer hold, and tradeoffs similar to those observed
 2056 in (Shelansky and Boeger, 2020) will again emerge; although, this time, the severity of these tradeoffs will depend
 2057 on w/c .

2058 **O.2. Calculating equilibrium specificity for a gene circuit with one binding site and one activation step.** Here we extend
 2059 the arguments from the previous section to show that, at equilibrium, the transcriptional specificity of the six state
 2060 model gene circuit shown in Figure 4B is fixed at $f^{eq} = \alpha$, irrespective of molecular details. For this system, the
 2061 specificity is simply equal to the concentration-normalized ratio of the occupancies of states 2 and 4:

$$2062 \quad f = \frac{w}{c} \frac{\pi_2}{\pi_4}. \quad (87)$$

2063 Since we’re assuming equilibrium conditions, we can re-express this as a difference between state energies, such
 2064 that

$$2065 \quad f^{eq} = \frac{w}{c} e^{-\frac{(\varepsilon_2 - \varepsilon_4)}{k_B T}}. \quad (88)$$

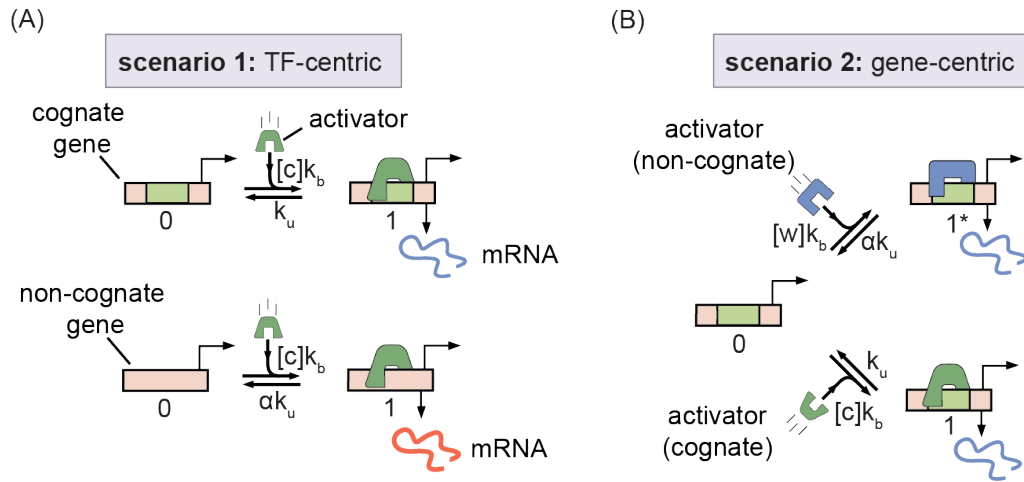


Fig. A13. Accounting for the influence of off-target activation. (A) An illustration of the parallel definition of activation fidelity. This approach considers the relative amounts of transcription driven by a transcriptional activator at its target locus and at an off target locus. (B) Cartoon illustrating “gene-centric” specificity definition, which considers competition between cognate and non-cognate factors to bind and activate a single gene locus.

In each case, we can express the state energies as the sum of the energy due to cognate or non-cognate factor binding with the energetic contributions from being in the active conformation, ε_a , and from interactions between the activator and the locus conformation, ε_{ab} . This leads to

$$\varepsilon_2 = -k_B T \ln \frac{ck_b}{k_u} + \varepsilon_a + \varepsilon_{ab} \quad (89)$$

and

$$\varepsilon_4 = -k_B T \ln \frac{wk_b}{\alpha k_u} + \varepsilon_a + \varepsilon_{ab}. \quad (90)$$

The key is to note the the first terms on the right-hand side of the above expressions are identical to Equations 83 and 84. Since the remaining energy terms are identical, they will cancel out, such that we once again have

$$f^{eq} = \frac{w}{c} e^{\ln \frac{ck_b k_u}{\alpha k_u w k_b}}, \quad (91)$$

which simplifies to

$$f^{eq} = \alpha. \quad (92)$$

0.3. Calculating equilibrium specificity for gene circuits with multiple binding sites. The above arguments can be extended to apply to more complex model architectures with multiple activator binding sites. To do this, we first need to generalize the definition of specificity put forward in the main text (Equation 5) for the case when there is more than one binding site. We define multi-site specificity as the ratio of the number of cognate and non-cognate activators bound to the gene locus (on average) while the gene is in the transcriptionally active (ON) conformation, such that:

$$f = \frac{w \langle n_c \rangle}{c \langle n_w \rangle} = \frac{w \sum_{i \in ON} n_i^c \pi_i}{c \sum_{i \in ON} n_i^w \pi_i}, \quad (93)$$

where $i \in ON$ stipulates that state i is part of the ON conformation, π_i is the probability of finding the gene locus in state i , and where n_i^c and n_i^w indicate the number of cognate and non-cognate factors bound to the gene locus in state i . Note also that we retain the normalizing prefactor of w/c .

Now, let’s calculate f for an equilibrium gene circuit with two binding sites. Once again, we work with energies since the system is at equilibrium. From Equation 93 we see that only states with at least one cognate or non-cognate factor bound contribute to the numerator and denominator, respectively. As a result, in each case, there are just three distinct molecular states to consider. For the cognate case (numerator), these are 1 cognate bound and 0 non-cognate, 1 cognate and 1 non-cognate, and 2 cognate. The non-cognate case (denominator) follows the same pattern. Drawing from the expression in the previous section, this leads to

$$f^{eq} = \frac{w \frac{ck_b}{k_u} \eta_a \eta_{ab} + 2 \frac{ck_b}{k_u} \frac{wk_b}{\alpha k_u} \eta_a \eta_{ab}^2 \eta_{ub} + 2 \left(\frac{ck_b}{k_u} \right)^2 \eta_a \eta_{ab}^2 \eta_{ub}}{c \frac{ck_b}{\alpha k_u} \eta_a \eta_{ab} + 2 \frac{ck_b}{k_u} \frac{wk_b}{\alpha k_u} \eta_a \eta_{ab}^2 \eta_{ub} + 2 \left(\frac{ck_b}{\alpha k_u} \right)^2 \eta_a \eta_{ab}^2 \eta_{ub}}, \quad (94)$$

2094 where η_a is a weight factor corresponding to the active conformation ($\eta_a = e^{-\frac{\epsilon_a}{k_B T}}$), η_{ab} is a weight factor capturing
 2095 cooperative interactions between the bound activator and the active conformation, and η_{ub} captures cooperative
 2096 interactions between bound activator molecules. Note that the three terms in the numerator and denominator of
 2097 Equation 94 match the ordering of the scenarios given above the equation. Factoring out common multipliers leads
 2098 to

$$2099 \quad f^{eq} = \frac{w \frac{ck_b}{k_u} 1 + 2 \frac{wk_b}{\alpha k_u} \eta_{ab} \eta_{ub} + 2 \frac{ck_b}{k_u} \eta_{ab} \eta_{ub}}{c \frac{wk_b}{\alpha k_u} 1 + 2 \frac{ck_b}{k_u} \eta_{ab} \eta_{ub} + 2 \frac{ck_b}{\alpha k_u} \eta_{ab} \eta_{ub}}, \quad (95)$$

2100 where we now see that the numerator and denominator are identical in the right-most ratio. Thus, we find that

$$2101 \quad f^{eq} = \alpha. \quad (96)$$

2102 Similar patterns repeat for systems with more binding. See the Mathematica notebook entitled “[specificity_](#)
 2103 [multi_site.nb](#)” in this paper’s git repository (<https://github.com/nlammers371/noneq-gene-regulation.git>) for a full
 2104 treatment of the 3 and 5 binding site cases.

2105 P. Deriving non-equilibrium tradeoff bound between intrinsic sharpness and specificity

2106 In this section, we lay out the key steps in deriving the non-equilibrium tradeoff bound between sharpness and
 2107 specificity given in Equation 7 in the main text. To do so, we make use of insights gained in Appendix M, where
 2108 we used first passage times to examine the key microscopic conditions for the twofold gain in sharpness away
 2109 from equilibrium observed in Figure 3A. Even for the simple six state system illustrated in Figure 4B, our system
 2110 has eight degrees of freedom when operating away from equilibrium. As such, a key part of our approach will be
 2111 to first reduce this complexity as much as possible while preserving the salient behaviors, namely the possibility
 2112 for non-equilibrium gains in sharpness and specificity. After this, we identify a tuning parameter, β , that can be
 2113 used to interpolate between maximally sharp to maximally specific non-equilibrium gene circuit architectures. Since
 2114 the expressions for non-equilibrium gene circuits are, in general, quite complex, we sketch the key steps here and
 2115 direct the reader to the Mathematica notebook entitled “[sharpness_specificity_bound_derivation.nb](#)” on the project
 2116 git repository for additional details: <https://github.com/nlammers371/noneq-gene-regulation.git>. Note that we work in
 2117 units of c throughout, such that $c = 1$.

2118 To begin, we strip unnecessary dimensions from our system. We set $\eta_{ib}k_i$ and k_a to the same generic rate, k_1 .
 2119 Next, we set $\eta_{ab}k_a$, $\eta_{ua}k_u$, and k_b to a second rate parameter, k_2 . Finally, we set k_i equal to $\beta\eta_{bs}k_b$, where β is our
 2120 interpolation parameter. This leaves us with a system with five free parameters, rather than eight.

2121 In Appendix M, we saw that maximally sharp non-equilibrium gene circuits (i) only switch into the active tran-
 2122 scriptional conformation when the activator is bound and (ii) only switch *out* of the ON states when the activator
 2123 is unbound. This amounts to effective one-way transitions from states $1 \rightarrow 2$ (equivalently, $5 \rightarrow 6$) and $3 \rightarrow 1$. We
 2124 impose this condition by taking the limit where $k_1 \rightarrow 0$. Next, we impose the condition uncovered by examination of
 2125 Equation 75,

$$2126 \quad k_u \gg [c]k_b, \eta_{ab}k_a, \quad (97)$$

2127 by taking the limit where k_u approaches infinity.

2128 These limits lead to a further simplified system that can be used to investigate fundamental tradeoffs between
 2129 intrinsic sharpness and specificity. For this stripped-down system, we find that the expression for specificity, f , is
 2130 quite simple:

$$2131 \quad f = \frac{\alpha(\alpha + \alpha\beta + w)}{\alpha + \beta + w}, \quad (98)$$

2132 where we see that all dependence on microscopic transition rates has dropped out, with the exception of our inter-
 2133 polation parameter, β . Furthermore, tuning β causes Equation 98 to shift from equilibrium levels ($f = \alpha$ when $\beta = 0$)
 2134 to the non-equilibrium limits revealed by Figure 5B ($f = \alpha^2$ when $\beta \gg \alpha, w$).

2135 The normalized sharpness, S , has a slightly more complicated functional form, given by

$$2136 \quad S = \frac{\alpha[\alpha(k_2 + (2 + \beta)\eta_{ba}k_b) + 2\eta_{ba}k_b w]}{\alpha^2(k_2 + \eta_{ba}k_b + \beta\eta_{ba}k_b) + \alpha w(k_2 + 2\eta_{ba}k_b) + w\eta_{ba}k_b(\beta + w)}. \quad (99)$$

2137 To obtain an expression for the intrinsic sharpness, S_0 , we divide through by the specificity prefactor (p_c) from
 2138 Equation 4:

$$2139 \quad S_0 = \frac{f + \frac{w}{c}}{f} S. \quad (100)$$

2140 Simplifying and applying the condition that $k_2 \approx 0$ leads to

$$2141 \quad S_0 = 2 - \frac{\alpha\beta}{\alpha + \alpha\beta + \frac{w}{c}}. \quad (101)$$

Here again, as with Equation 98, we see that all dependence on the on rate parameters drops away. Further, it is easy to see that this expression goes to 2 when $\beta = 0$ and 1 when $\beta \gg \alpha, w$. Thus, when β is small, our system exhibits equilibrium levels of specificity and non-equilibrium levels of intrinsic sharpness and, when β is large, it exhibits non-equilibrium specificity and equilibrium sharpness levels. Thus, we have succeeded in our initial aim to establish a simplified model that can capture the tradeoffs between sharpness and specificity revealed by our numerical parameter sweeps (Figure 5B).

As a final step, we can solve Equation 98 to obtain an expression for β in terms of f :

$$\beta = \frac{(f - \alpha)(\alpha + w)}{\alpha^2 - f}. \quad (102)$$

Plugging this expression into Equation 100 and simplifying yields an expression for S_0 as a function of f :

$$S_0 = \frac{\alpha^2 + \alpha f - 2f}{\alpha f - f}, \quad (103)$$

where we assume that $\alpha \leq f \leq \alpha^2$. Thus, we have obtained the final S_0 expression depicted in Equation 7. Observe that $S_0 \approx 2$ when $f = \alpha$ and $S_0 \approx 1$ when $f = \alpha^2$. Equation 103 gives the dashed black curve bounding f vs. S_0 sweep results shown in Figure 5B, confirming that it represents the limiting behavior of intrinsic sharpness and specificity for non-equilibrium realizations of the six state model from Figure 4B.

References

- Biddle, J. W., Nguyen, M., and Gunawardena, J. Negative reciprocity, not ordered assembly, underlies the interaction of Sox2 and Oct4 on DNA. pages 1–30, 2019.
- Biddle, J. W., Martinez-Corral, R., Wong, F., and Gunawardena, J. Allosteric conformational ensembles have unlimited capacity for integrating information, dec 2020. ISSN 26928205. URL <https://doi.org/10.1101/2020.12.10.420117>.
- Bothma, J. P., Garcia, H. G., Esposito, E., Schlissel, G., Gregor, T., and Levine, M. Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living Drosophila embryos. *Proceedings of the National Academy of Sciences*, 111(29):10598–10603, 2014. doi: 10.1073/pnas.1411002111.
- Bowles, J. R., Hoppe, C., Ashe, H. L., and Rattray, M. Scalable inference of transcriptional kinetic parameters from MS2 time series data. *Bioinformatics*, 38(4):1030–1036, jan 2022. doi: 10.1093/BIOINFORMATICS/BTAB765.
- Cepeda-Humerez, S. A., Rieckh, G., and Tkačik, G. Stochastic Proofreading Mechanism Alleviates Crosstalk in Transcriptional Regulation. *Physical Review Letters*, 115(24):1–5, 2015. doi: 10.1103/PhysRevLett.115.248101.
- Charoensawan, V., Wilson, D., and Teichmann, S. A. Genomic repertoires of DNA-binding transcription factors across the tree of life. *Nucleic Acids Research*, 38(21):7364, nov 2010. doi: 10.1093/NAR/GKQ617.
- Choudhury, M. and Ramsey, S. A. Identifying Cell Type-Specific Transcription Factors by Integrating ChIP-seq and eQTL Data—Application to Monocyte Gene Regulation. *Gene Regulation and Systems Biology*, 10:105, dec 2016. doi: 10.4137/GRSB.S40768.
- Corrigan, A. M., Tunnaciffe, E., Cannon, D., and Chubb, J. R. A continuum model of transcriptional bursting. *eLife*, 5:e13051, 2016. doi: 10.7554/eLife.13051.
- Coulon, A., Chow, C. C., Singer, R. H., and Larson, D. R. Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nature reviews. Genetics*, 14(8):572–84, aug 2013. doi: 10.1038/nrg3484.
- Cover, T. M. and Thomas, J. A. *Elements of Information Theory 2nd Edition (Wiley Series in Telecommunications and Signal Processing)*. 2006. ISBN 0471241954. URL <https://www.bibsonomy.org/bibtex/22e9bfa879286689a14feb55b69d326c1/ywhuanghttp://www.amazon.com/Elements-Information-Edition-Telecommunications-Processing/dp/0471241954>.
- Crocker, J. and Stern, D. L. TALE-mediated modulation of transcriptional enhancers in vivo. *Nature methods*, 10(8):762–767, aug 2013. doi: 10.1038/NMETH.2543.
- Crocker, J., Abe, N., Rinaldi, L., McGregor, A. P., Frankel, N., Wang, S., Alswadi, A., Valenti, P., Plaza, S., Payne, F., Mann, R. S., and Stern, D. L. Low Affinity Binding Site Clusters Confer Hox Specificity and Regulatory Robustness. *Cell*, 160(0):191, jan 2015. doi: 10.1016/J.CELL.2014.11.041.
- Desponds, J., Vergassola, M., and Walczak, A. M. A mechanism for hunchback promoters to readout morphogenetic positional information in less than a minute. *eLife*, 9, jul 2020. doi: 10.7554/elife.49758.
- Desponds, J., Vergassola, M., and Walczak, A. M. A mechanism for hunchback promoters to readout morphogenetic positional information in less than a minute. *eLife*, 9, jul 2020. doi: 10.7554/elife.49758.
- Donovan, B. T., Huynh, A., Ball, D. A., Patel, H. P., Poirier, M. G., Larson, D. R., Ferguson, M. L., and Lenstra, T. L. Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting. *The EMBO journal*, 38(12), jun 2019. doi: 10.15252/EMBJ.2018100809.
- Eck, E., Liu, J., Kazemzadeh-Atoufi, M., Ghoreishi, S., Blythe, S., and Garcia, H. G. Quantitative dissection of transcription in development yields evidence for transcription factor-driven chromatin accessibility. *bioRxiv*, page 2020.01.27.922054, mar 2020. doi: 10.1101/2020.01.27.922054.
- Erokhin, M., Vassetzky, Y., Georgiev, P., and Chetverina, D. Eukaryotic enhancers: common features, regulation, and participation in diseases. *Cellular and molecular life sciences : CMLS*, 72(12):2361–2375, jun 2015. doi: 10.1007/S00018-015-1871-9.
- Estrada, J., Wong, F., DePace, A., and Gunawardena, J. Information Integration and Energy Expenditure in Gene Regulation. *Cell*, 166(1):234–44, jun 2016. doi: 10.1016/j.cell.2016.06.012.
- Estrada, J., Ruiz-Herrero, T., Scholes, C., Wunderlich, Z., and DePace, A. H. SiteOut: An Online Tool to Design Binding Site-Free DNA Sequences. *PLOS ONE*, 11(3):e0151740, mar 2016. doi: 10.1371/JOURNAL.PONE.0151740.
- et al. Alberts B, Johnson A, Lewis J. *Drosophila and the Molecular Genetics of Pattern Formation: Genesis of the Body Plan*. In *Molecular Biology of the Cell*. 2002. URL <https://www.ncbi.nlm.nih.gov/books/NBK26906/>.
- Farley, E. K., Olson, K. M., Zhang, W., Brandt, A. J., Rokhsar, D. S., and Levine, M. S. Suboptimization of developmental enhancers. *Science (New York, N.Y.)*, 350(6258):325, oct 2015. doi: 10.1126/SCIENCE.AAC6948.
- Foe, V. E. and Alberts, B. M. Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in Drosophila embryogenesis. *Journal of cell science*, 61:31–70, 1983. doi: 10.1242/JCS.61.1.31.
- Friedlander, T., Prizak, R., Guet, C. C., Barton, N. H., and Tkačik, G. Intrinsic limits to gene regulation by global crosstalk. *Nature Communications* 2016 7:1, 7(1):1–12, aug 2016. doi: 10.1038/ncomms12307.
- Fukaya, T., Lim, B., and Levine, M. Enhancer Control of Transcriptional Bursting. *Cell*, 166(2):358–368, jul 2016. doi: 10.1016/j.cell.2016.05.025.
- Geyer, C. J. Introduction to Markov Chain Monte Carlo. Technical report, 2011.
- Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. In *Journal of Physical Chemistry*, volume 81, pages 2340–2361. American Chemical Society, 1977. doi: 10.1021/j100540a008. URL <https://pubs.acs.org/sharingguidelines>.
- Grah, R., Zoller, B., and Tkačik, G. Nonequilibrium models of optimal enhancer function. *Proceedings of the National Academy of Sciences of the United States of America*, 117(50):31614–31622, dec 2020. doi: 10.1073/PNAS.2006731117/-/DCSUPPLEMENTAL.
- Gregor, T., Tank, D. W., Wieschaus, E. F., and Bialek, W. Probing the Limits to Positional Information. *Cell*, 130(1):153–164, jul 2007. doi: 10.1016/J.CELL.2007.05.025.
- Hammar, P., Walldén, M., Fange, D., Persson, F., Baltekin, Ö., Ullman, G., Leroy, P., and Elf, J. Direct measurement of transcription factor dissociation excludes a simple operator occupancy model for gene regulation. *Nature genetics*, 46(4):405, 2014. doi: 10.1038/NG.2905.
- Henry, G. L., Davis, F. P., Picard, S., and Eddy, S. R. Cell type-specific genomics of Drosophila neurons. *Nucleic Acids Research*, 40(19):9691, oct 2012. doi: 10.1093/NAR/GKS671.
- Hill, T. L. *Free Energy Transduction and Biochemical Cycle Kinetics*. Springer New York, 1989. ISBN 9781461235583.

P Deriving non-equilibrium tradeoff bound between intrinsic sharpness and specificity

- 2211 Hopfield, J. J. Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity (protein synthesis/DNA replication/amino-acid recognition). 71
2212 (10):4135–4139, 1974.
- 2213 Hubbard, E. J. A. The C. elegans germ line: a model for stem cell biology. *Developmental Dynamics: an official publication of the American Association of Anatomists*, 236(12):3343, dec 2007.
2214 doi: 10.1002/dvdy.21335.
- 2215 Kabadi, A. M. and Gersbach, C. A. Engineering Synthetic TALE and CRISPR/Cas9 Transcription Factors for Regulating Gene Expression. *Methods (San Diego, Calif.)*, 69(2):188, 2014. doi:
2216 10.1016/j.ymeth.2014.06.014.
- 2217 Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., van Berkum, N. L., Ebmeier, C. C., Goossens, J., Rahl, P. B., Levine, S. S., Taatjes, D. J., Dekker, J., and Young, R. A.
2218 Mediator and cohesin connect gene expression and chromatin architecture. *Nature*, 467(7314):430–435, sep 2010. doi: 10.1038/nature09380.
- 2219 Kim, Y. J., Rhee, K., Liu, J., Jeammet, P., Turner, M., Small, S., and Garcia, H. G. Predictive modeling reveals that higher-order cooperativity drives transcriptional repression in a synthetic
2220 developmental enhancer. *bioRxiv*, pages 1–55, 2021.
- 2221 Kornberg, R. D. The molecular basis of eukaryotic transcription. *Proceedings of the National Academy of Sciences of the United States of America*, 104(32):12955–12961, aug 2007. doi:
2222 10.1073/PNAS.0704138104/ASSET/573DB0B8-41E5-4EFE-A261-EFFE5771389F/ASSETS/GRAPHIC/ZPQ0260768180018.JPEG.
- 2223 Lammers, N. C., Kim, Y. J., Zhao, J., and Garcia, H. G. A matter of time: Using dynamics and theory to uncover mechanisms of transcriptional bursting, dec 2020. ISSN 18790410.
- 2224 Lammers, N. C., Galstyan, V., Reimer, A., Medin, S. A., Wiggins, C. H., and Garcia, H. G. Multimodal transcriptional control of pattern formation in embryonic development. *Proceedings of the
2225 National Academy of Sciences*, 117(2):836–847, jan 2020. doi: 10.1073/PNAS.1912500117.
- 2226 Lang, A. H., Fisher, C. K., Mora, T., and Mehta, P. Thermodynamics of statistical inference by cells. *Physical Review Letters*, 113(14), 2014. doi: 10.1103/PhysRevLett.113.148103.
- 2227 Le, D. D., Shimko, T. C., Aditham, A. K., Keys, A. M., Longwell, S. A., Orenstein, Y., and Fordyce, P. M. Comprehensive, high-resolution binding energy landscapes reveal context dependencies
2228 of transcription factor binding. *Proceedings of the National Academy of Sciences*, 115(16):E3702–E3711, apr 2018. doi: 10.1073/PNAS.1715888115.
- 2229 Lebowitz, J. L. and Spohn, H. A gallavotti-cohen-type symmetry in the large deviation functional for stochastic dynamics. *Journal of Statistical Physics*, 95(1-2):333–365, 1999. doi: 10.1023/a:
2230 1004589714161.
- 2231 Lee, B. K., Bhingre, A. A., Battenhouse, A., McDaniel, R. M., Liu, Z., Song, L., Ni, Y., Birney, E., Lieb, J. D., Furey, T. S., Crawford, G. E., and Iyer, V. R. Cell-type specific and combinatorial usage
2232 of diverse transcription factors revealed by genome-wide binding studies in multiple human cells. *Genome Research*, 22(1):9, jan 2012. doi: 10.1101/GR.127597.111.
- 2233 Lee, C. H., Shin, H., and Kimble, J. Dynamics of Notch-Dependent Transcriptional Bursting in Its Native Context. *Developmental Cell*, 50(4):426–435.e4, aug 2019. doi: 10.1016/J.DEVCEL.
2234 2019.07.001.
- 2235 Lenstra, T. L., Rodriguez, J., Chen, H., and Larson, D. R. Transcription Dynamics in Living Cells. *Annual Review of Biophysics*, 45:25–47, 2016. doi: 10.1146/
2236 annurev-biophys-062215-010838.
- 2237 Lionnet, T. and Wu, C. Single-molecule tracking of transcription protein dynamics in living cells: seeing is believing, but what are we seeing? *Current Opinion in Genetics & Development*, 67:
2238 94–102, apr 2021. doi: 10.1016/J.GDE.2020.12.001.
- 2239 Little, S. C., Tikhonov, M., and Gregor, T. Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell*, 154(4):789–800, aug 2013. doi: 10.1016/j.cell.
2240 2013.07.025.
- 2241 Maerkl, S. J. and Quake, S. R. A systems approach to measuring the binding energy landscapes of transcription factors. *Science*, 315(5809):233–237, jan 2007. doi: 10.1126/science.
2242 1131007.
- 2243 Malik, S. and Roeder, R. Mediator: A Drawbridge across the Enhancer-Promoter Divide. *Molecular Cell*, 64(3):433–434, nov 2016. doi: 10.1016/J.MOLCEL.2016.10.024.
- 2244 Massey, F. J. The Kolmogorov-Smirnov Test for Goodness of Fit. *Source: Journal of the American Statistical Association*, 46(253):68–78, 1951.
- 2245 Milo, R. and Phillips, R. *Cell Biology by the Numbers*. Garland Science, dec 2015. doi: 10.1201/9780429258770. URL <https://www.taylorfrancis.com/https://www.taylorfrancis.com/books/mono/10.1201/9780429258770/cell-biology-numbers-ron-milo-rob-phillips>.
- 2246 Mirny, L. A. Nucleosome-mediated cooperativity between transcription factors. *Proceedings of the National Academy of Sciences of the United States of America*, 107(52):22534–22539, dec
2247 2010. doi: 10.1073/pnas.0913805107.
- 2248 Murugan, A., Huse, D. A., and Leibler, S. Speed, dissipation, and error in kinetic proofreading. *Proceedings of the National Academy of Sciences*, 109(30):12034–12039, jul 2012. doi:
2249 10.1073/PNAS.1119911109.
- 2250 Nicolas, D., Phillips, N. E., Naef, F., Levy, M., Landen, S., and Elinav, E. What shapes eukaryotic transcriptional bursting? *Mol. BioSyst.*, 13(7):1280–1290, jun 2017. doi: 10.1039/
2251 C7MB00154A.
- 2252 Ninio, J. Kinetic amplification of enzyme discrimination. *Biochimie*, 57(5):587–95, 1975.
- 2253 Nogales, E., Louder, R. K., and He, Y. Structural Insights into the Eukaryotic Transcription Initiation Machinery. *Annual Review of Biophysics*, 46(1):59–83, 2017. doi: 10.1146/
2254 annurev-biophys-070816-033751.
- 2255 Obst, R. The Timing of T Cell Priming and Cycling. *Frontiers in Immunology*, 6(NOV):1, 2015. doi: 10.3389/FIMMU.2015.00563.
- 2256 Park, J., Estrada, J., Johnson, G., Vincent, B. J., Ricci-tam, C., Bragdon, M. D. J., Shulgina, Y., Cha, A., Wunderlich, Z., Gunawardena, J., and Depace, A. H. Dissecting the sharp response of a
2257 canonical developmental enhancer reveals multiple sources of cooperativity. pages 1–25, 2019.
- 2258 Pérez-Ortín, J. E., Alepuz, P., Chávez, S., and Choder, M. Eukaryotic mRNA Decay: Methodologies, Pathways, and Links to Other Stages of Gene Expression. 2013. doi: 10.1016/j.jmb.
2259 2013.02.029.
- 2260 Phillips, R. and Orme, N. *The Molecular Switch: Signaling and Allostery*. Princeton University Press, 2020. ISBN 9780691200248. doi: 10.2307/j.ctvx5w8pf. URL [http://www.jstor.
2261 org/stable/j.ctvx5w8pf](http://www.jstor.org/stable/j.ctvx5w8pf).
- 2262 Phillips, R., Belliveau, N. M., Chure, G., Garcia, H. G., Razo-Mejia, M., and Scholes, C. Figure 1 Theory Meets Figure 2 Experiments in the Study of Gene Expression. *Annu. Rev. Biophys.*, 48:
2263 121–163, 2019. doi: 10.1146/annurev-biophys-052118.
- 2264 Reimer, A., Alamos, S., Westrum, C., Turner, M. A., Talleo, P., Zhao, J., and Garcia, H. G. Minimal Synthetic Enhancers Reveal Control of the Probability of Transcriptional Engagement and its
2265 Timing by a Morphogen Gradient. *SSRN Electronic Journal*, 2021. doi: 10.2139/ssrn.3887798.
- 2266 Rieckh, G. and Tkačik, G. Noise and information transmission in promoters with multiple internal states. *Biophysical Journal*, 106(5):1194–1204, mar 2014. doi: 10.1016/j.bpj.2014.01.014.
- 2267 Rybakova, K. N., Bruggeman, F. J., Tomaszewska, A., Moné, M. J., Carlberg, C., and Westerhoff, H. V. Multiplex Eukaryotic Transcription (In)activation: Timing, Bursting and Cycling of a Ratchet
2268 Clock Mechanism. *PLoS Computational Biology*, 11(4):e1004236, apr 2015. doi: 10.1371/journal.pcbi.1004236.
- 2269 Shahein, A., Maria, L., Istomin, I., Olson, E. J., Cheng, S., and Maerkl, S. J. Systematic analysis of low-affinity transcription factor binding site clusters in vitro and in vivo establishes their
2270 functional relevance. *bioRxiv*, 2021.
- 2271 Shelansky, R. and Boeger, H. Nucleosomal proofreading of activator–promoter interactions. *Proceedings of the National Academy of Sciences of the United States of America*, 117(5):2456–2461,
2272 2020. doi: 10.1073/pnas.1911188117.
- 2273 Siggia, E. D. and Vergassola, M. Decisions on the fly in cellular sensory systems. *Proceedings of the National Academy of Sciences*, 110(39):E3704–E3712, sep 2013. doi: 10.1073/pnas.
2274 1314081110.
- 2275 Son, H. G., Seo, M., Ham, S., Hwang, W., Lee, D., An, S. W. A., Artan, M., Seo, K., Kaletsky, R., Arey, R. N., Ryu, Y., Ha, C. M., Kim, Y. K., Murphy, C. T., Roh, T.-Y., Nam, H. G., and Lee,
2276 S.-J. V. RNA surveillance via nonsense-mediated mRNA decay is crucial for longevity in daf-2/insulin/IGF-1 mutant C. elegans. *Nature Communications* 2017 8:1, 8(1):1–11, mar 2017. doi:
2277 10.1038/ncomms14749.
- 2278 Taatjes, D. J. The Continuing SAGA of TFIID and RNA Polymerase II Transcription. *Molecular Cell*, 68(1):1–2, oct 2017. doi: 10.1016/J.MOLCEL.2017.09.028.
- 2279 Tantale, K., Mueller, F., Kozulic-Pirher, A., Lesne, A., Victor, J.-M., Robert, M.-C., Capozzi, S., Chouaib, R., Bäcker, V., Mateos-Langerak, J., Darzacq, X., Zimmer, C., Basyuk, E., and Bertrand, E.
2280 A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nature Communications*, 7(1):12248, nov 2016. doi: 10.1038/ncomms12248.
- 2281 Vincent, B. J., Estrada, J., and DePace, A. H. The appeasement of Doug: a synthetic approach to enhancer biology. *Integrative Biology*, 8(4):475–484, apr 2016. doi: 10.1039/C5IB00321K.
- 2282 Wald, A. and Wolfowitz, J. Optimum Character of the Sequential Probability Ratio Test. *The Annals of Mathematical Statistics*, 19(3):326–339, sep 1948. doi: 10.1214/aoms/1177730197.
- 2283 Wald, A. Sequential Tests of Statistical Hypotheses. *The Annals of Mathematical Statistics*, 16(2):117–186, may 1945.
- 2284 Whitt, W. Asymptotic Formulas for Markov Processes with Applications to Simulation. *Operations Research*, 40(2):279–291, apr 1992. doi: 10.1287/opre.40.2.279.
- 2285 Wong, F. and Gunawardena, J. Gene Regulation in and out of Equilibrium, may 2020. ISSN 19361238. URL <https://pubmed.ncbi.nlm.nih.gov/32375018/https://pubmed.ncbi.nlm.nih.gov/32375018/?dopt=Abstract>.
- 2286 Xu, H., Sepúlveda, L. A., Figard, L., Sokac, A. M., and Golding, I. Combining protein and mRNA quantification to decipher transcriptional regulation. *Nature Methods*, 12(8):739–742, aug 2015.
2287 doi: 10.1038/nmeth.3446.
- 2288 Yan, M. and Gralla, J. D. Multiple ATP-dependent steps in RNA polymerase II promoter melting and initiation. *The EMBO Journal*, 16(24):7457, dec 1997. doi: 10.1093/EMBOJ/16.24.7457.
- 2289 Zhou, C. Y., Johnson, S. L., Gamarrá, N. I., and Narlikar, G. J. Mechanisms of ATP-Dependent Chromatin Remodeling Motors. *Annual Review of Biophysics*, 45:153–181, jul 2016. doi:
2290 10.1146/annurev-biophys-051013-022819.
- 2291 Zoller, B., Little, S. C., and Correspondence, T. G. Diverse Spatial Expression Patterns Emerge from Unified Kinetics of Transcriptional Bursting In Brief Multiple gene expression boundaries in a
2292 developing embryo arise from the same strategy of modulated transcriptional bursting rates. *Cell*, 175:835–847, 2018. doi: 10.1016/j.cell.2018.09.056.
- 2293 Zoller, B., Gregor, T., and Tkačik, G. Eukaryotic gene regulation at equilibrium, or non? oct 2021.