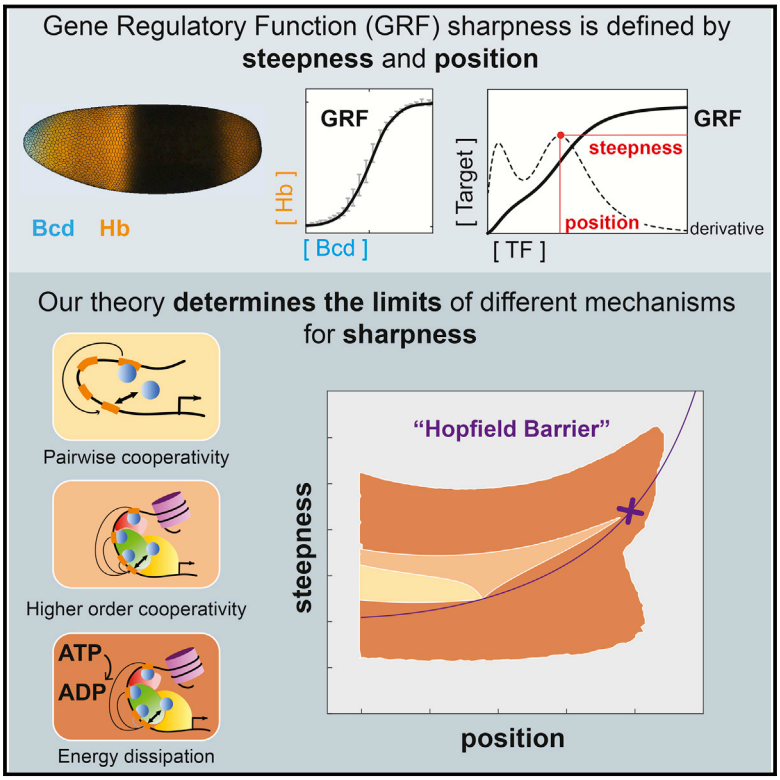
Cell

Information Integration and Energy Expenditure in Gene Regulation

Graphical Abstract



Highlights

- Gene regulation is understood quantitatively in terms of a bacterial paradigm
- This paradigm cannot account for the sharpness of gene expression in development
- Information integration or energy expenditure can explain the sharpness
- Hill functions form a "Hopfield barrier" for sharpness at thermodynamic equilibrium

Authors

Javier Estrada, Felix Wong, Angela DePace, Jeremy Gunawardena

Correspondence

jeremy@hms.harvard.edu

In Brief

The quantitative principles governing gene regulation in bacteria cannot explain the sharpness of gene expression in eukaryotes; energy use and information integration have to be taken into account as well.

Estrada et al., 2016, Cell *166*, 234–244
June 30, 2016
http://dx.doi.org/10.1016/j.cell.2016.06.012

CellPress

Cell

Information Integration and Energy Expenditure in Gene Regulation

Javier Estrada,[1] Felix Wong,[1,2] Angela DePace,[1] and Jeremy Gunawardena[1,*]
[1]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
[2]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
*Correspondence: jeremy@hms.harvard.edu
http://dx.doi.org/10.1016/j.cell.2016.06.012

SUMMARY

The quantitative concepts used to reason about gene regulation largely derive from bacterial studies. We show that this bacterial paradigm cannot explain the sharp expression of a canonical developmental gene in response to a regulating transcription factor (TF). In the absence of energy expenditure, with regulatory DNA at thermodynamic equilibrium, information integration across multiple TF binding sites can generate the required sharpness, but with strong constraints on the resultant "higher-order cooperativities." Even with such integration, there is a "Hopfield barrier" to sharpness; for *n* TF binding sites, this barrier is represented by the Hill function with the Hill coefficient *n*. If, however, energy is expended to maintain regulatory DNA away from thermodynamic equilibrium, as in kinetic proofreading, this barrier can be breached and greater sharpness achieved. Our approach is grounded in fundamental physics, leads to testable experimental predictions, and suggests how a quantitative paradigm for eukaryotic gene regulation can be formulated.

INTRODUCTION

The molecular machinery which transcribes DNA into RNA is general purpose. Deciding which gene to transcribe requires regulatory DNA sequence information, which is interpreted by sequence-specific, DNA-binding transcription factors (TFs). Quantitative measurements of TF-DNA and TF-TF interactions in bacteria (Ptashne, 2004), together with analysis of the underlying physics (Ackers et al., 1982), have introduced fundamental quantitative concepts like "affinity" and "cooperativity" to explain the regulated recruitment of RNA polymerase to a gene. This bacterial paradigm has been widely used to interpret experimental results even outside the bacterial domain. However, eukaryotic transcription differs considerably from bacterial transcription; as a result, this raises the question of whether the bacterial paradigm is sufficient to explain how eukaryotic genes are regulated.

Bacterial TF sequence motifs have an average length of 16 base pairs, and those in eukaryotes are only half as long (Wunderlich and Mirny, 2009), suggesting that eukaryotes depend on combinatorial integration of many small packets of information. Such information integration might be implemented through nucleosomes or by multi-protein co-regulators, such as Mediator or CBP/p300, that make multiple contacts between TFs and the transcriptional machinery (Spitz and Furlong, 2012). Also, while bacterial gene regulation appears not to require energy from donors like ATP, making it reasonable to assume that it takes place at thermodynamic equilibrium, eukaryotic gene regulation depends on energy expenditure to reorganize chromatin, displace nucleosomes, post-translationally modify regulatory proteins, and methylate DNA. This qualitative appreciation of eukaryotic complexity has been difficult to translate into rigorous, well-defined concepts and new kinds of experiments that can explain the role of these molecular mechanisms in gene regulation.

Quantitative models grounded in physics could fill this critical gap. The physics-based "thermodynamic formalism" developed for bacteria assumes that regulation takes place at thermodynamic equilibrium. This formalism has been codified (Bintu et al., 2005) and applied to gene regulation in *Drosophila*, yeast, and human cells (Segal and Widom, 2009; Sherman and Cohen, 2012). However, the molecular complexity found in eukaryotes, especially the complexity that implements the information integration and energy expenditure described above, has not been incorporated into these models.

Questions about the sufficiency of the bacterial paradigm have been accumulating (Coulon et al., 2013), but the absence of a compelling example and the lack of appropriate concepts make it easy to fall back on what is familiar. Here, we present a compelling example of insufficiency and introduce appropriate quantitative concepts, rigorously based on the underlying physics, with which to reason about eukaryotic gene regulation.

We bring together three ingredients that exemplify a general approach to the problem. First, we focus on a property of gene regulation that can be described quantitatively; second, we identify a biological system in which that property has been measured; and, third, we exploit a mathematical framework that allows us to analyze both equilibrium and non-equilibrium systems.

The quantitative property on which we focus is the sharpness of gene expression in response to a TF, or the extent to which a small change in TF concentration can lead to a larger change in gene expression. Sharpness has been investigated in several biological systems, but is particularly evident in developmental patterning. The zygotic gap gene *hunchback* (*hb*) is expressed in an anterior region of the early *Drosophila* embryo under

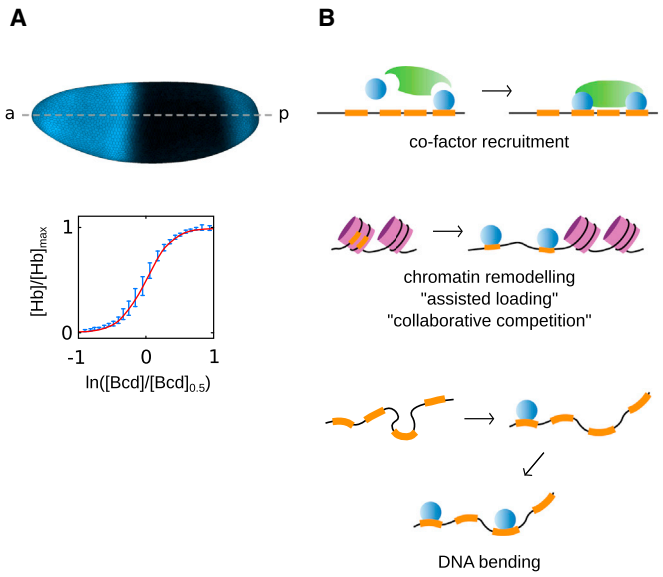


Figure 1. Sharpness in Development and Cooperativity Mechanisms

(A) Top: *Drosophila* embryo stained for Hb expression. Bottom: plot adapted from Figure 4A of Gregor et al. (2007) showing mean ± SE of Hb and Bcd from several embryos (blue) and a fit to the Hill function $\mathcal{H}_5$ (red).

(B) Examples of indirect, long-distance cooperativity, adapted from Figure 1 of Spitz and Furlong (2012).

(C) Top: pairwise cooperativity between two sites. Bottom: higher-order cooperativity of order three.

regulation by the maternal morphogen Bicoid (Bcd) (Figure 1A). The expression levels of Hb and Bcd proteins are related to each other in a way closely approximated by a simple algebraic expression (Gregor et al., 2007):

$$\frac{[\mathrm{Hb}]}{[\mathrm{Hb}]_{max}} \approx \frac{x^5}{1+x^5}. \quad (1)$$

Here, the concentration of Hb, denoted [Hb], is normalized to its maximal level, and x denotes Bcd concentration, normalized to the value at which half-maximal Hb expression is reached, so that $x = [\mathrm{Bcd}]/[\mathrm{Bcd}]_{0.5}$. Equation 1 describes the *hb* gene regulation function, which quantitatively expresses how the output of *hb* depends on [Bcd].

The expression in Equation 1 is a Hill function, $\mathcal{H}_a(x) = x^a / (1 + x^a)$, for which the Hill coefficient, a, has the value $a = 5$. Increasing Hill coefficients imply increasing sharpness. In Equation 1, the sharpness represented by $a = 5$ reflects the precision with which individual nuclei use Bcd to determine their position along the anterior-posterior axis and create the tight boundary between Hb "on" and Hb "off."

Gregor et al. (2007) explain how the sharpness in Equation 1 arises by saying that it is "consistent with the idea that Hb transcription is activated by cooperative binding of effectively five Bcd molecules." This reflects the conventional bacterial paradigm, in which sharpness is accounted for at thermodynamic equilibrium by pairwise cooperativity between TFs, whereby TF binding at one site influences the affinity of TF binding at another site (Ptashne, 2004). With n binding sites and pairwise cooperativity, it is widely believed, as Gregor et al. (2007) suggest, that sharpness corresponding to a Hill coefficient of n can be achieved, without requiring any expenditure of energy.

To examine this idea, we use a recently introduced mathematical framework that generalizes the thermodynamic formalism to accommodate mechanisms that expend energy (Ahsendorf et al., 2014). We show that for regulatory DNA at thermodynamic equilibrium with only pairwise cooperativity, the experimentally measured sharpness described in Equation 1 cannot be biochemically realized, no matter how many TF binding sites are present. The widely held belief that the bacterial paradigm can be extrapolated in this way is not rigorously justified. We believe this is a compelling example of its insufficiency.

Information integration through nucleosomes or co-regulators could yield indirect, long-distance forms of cooperativity (Figure 1B), which could link multiple TF binding sites. To account for this, we introduce the concept of "higher-order cooperativity" at thermodynamic equilibrium (Figure 1C). If such cooperativities are present, greater levels of sharpness become possible. With n binding sites and higher-order cooperativities, a Hill coefficient of n or more still remains out of reach, but a Hill coefficient less than n can be achieved. Furthermore, not just the Hill coefficient but also the overall shape of the gene regulation function (GRF) can match what is found experimentally: with enough binding sites, GRFs can be found that are statistically indistinguishable in shape from the Hill functions in Equation 1. However, these GRFs lie on the edge of what can be biochemically achieved and impose stringent quantitative constraints on the mechanisms responsible for higher-order cooperativity.

Higher-order cooperativities improve sharpness but reveal fundamental barriers to what can be achieved without the expenditure of energy. The existence of such barriers was first suggested in Hopfield's work on kinetic proofreading (Hopfield, 1974; Ninio, 1975). He showed, in effect, that if a biochemical system operates at thermodynamic equilibrium, then physics imposes a barrier to how well a given information processing task can be accomplished. (In his case, the task was achieving fidelity in transcription and translation.) The only way to bypass this barrier is to expend energy and maintain the system away from equilibrium. Kinetic proofreading is one way to do this.

Here, we identify a "Hopfield barrier" for sharpness in gene regulation. With n binding sites, a Hill coefficient of n sets the Hopfield barrier; at thermodynamic equilibrium, no GRF can reach it, even with higher-order cooperativities. If, however, energy is expended to maintain regulatory DNA away from equilibrium, then much greater sharpness can be achieved.

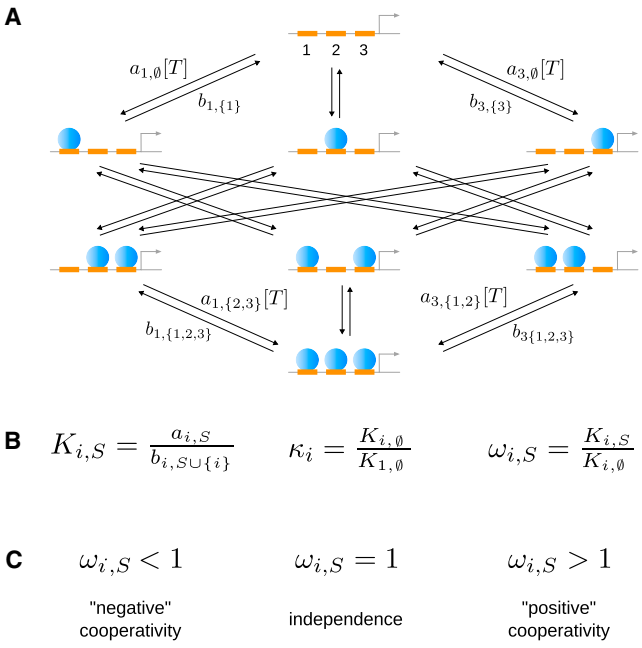


Figure 2. Linear Framework Model

(A) The graph G_3 showing the 8 microstates and the associated labeled, directed edges, with middle-layer labels omitted for clarity.

(B) The essential parameters at thermodynamic equilibrium. The association constants $K_{i,S}$ have units of $(\text{concentration})^{-1}$, and κ_i and $\omega_{i,S}$ are non-dimensional. $S \cup \{i\}$ is the set in which site i has been added to the sites in S.

(C) The higher-order cooperativity $\omega_{i,S}$ measures whether binding of T to site i, when T is already bound to the sites in S, shows reduced affinity $(\omega_{i,S} < 1)$, unchanged affinity $(\omega_{i,S} = 1)$, or enhanced affinity $(\omega_{i,S} > 1)$, as compared to binding to site i, when no other sites are bound.

RESULTS

The Rationale for the Model

We introduce a mathematical model for analyzing gene regulation. As with all models, the conclusions depend on the assumptions (Gunawardena, 2014). Our assumptions are guided by the example of *hb*, but the model is general and not restricted to this example. The anterior expression pattern of *hb* is believed to be regulated by, at least, three enhancers (Perry et al., 2011). Both the classical P2 enhancer, which is promoter proximal, and a shadow enhancer, located ~3 kb upstream, drive broad anterior patterns early in embryo development. Later, the central stripe enhancer drives expression near the middle of the embryo. Bcd is a transcriptional activator for the P2 and shadow enhancers; the stripe enhancer is also targetted by transcriptional repressors. The stripe enhancer has no effect on sharpness early in nuclear cycle 14 (Perry et al., 2012), when the data on which Equation 1 is based (Gregor et al., 2007) were acquired.

Accordingly, we focus on a single TF, binding to a specified but arbitrary number of sites and functioning solely as a transcriptional activator. TF binding sites can be anywhere on the genome and are not assumed to be confined to a single enhancer; thus, our analysis is not limited to the 5–7 Bcd binding sites thought to be present in the *hb* P2 enhancer.

Molecular mechanisms other than TF binding and unbinding, such as nucleosomes or co-regulators, are not directly represented in our model, but their influence is captured through their effects on rate constants and the dependence of these constants on the state of DNA ("microstate", see below). This permits general conclusions to be drawn without knowing the specific mechanisms at work in a particular gene but does not allow us to assign mechanisms to the effects we find. Other molecular features, such as post-transcriptional mechanisms or network effects like feedback, could influence sharpness, but these are not thought to be relevant for Bcd regulation of *hb*. Addressing such features in future work may yield further insights into sharpness.

A Graph-Based Model of Gene Regulation

We recently developed a graph-based "linear framework" for modeling gene regulation (Ahsendorf et al., 2014). We use this to formulate a general model of a gene responding to a TF, called T, binding as a monomer to a number, n, of sites. Oligomerization of a TF in solution can contribute to gene-expression sharpness, but it is not thought to be significant in Bcd regulation of *hb* (Lebrecht et al., 2005; Gregor et al., 2007), and we do not consider it here. In this section and the next, we discuss the quantitative details of how T binds and unbinds, how cooperativity is defined, and how T influences transcription.

The model consists of the labeled, directed graph, G_n (Figure 2A). The vertices of G_n represent the microstates, or patterns of T bound to DNA, with the binding sites labeled by the numbers $1, \cdots, n$. The edges represent binding or unbinding of T from the microstates. Each edge has a label describing the rate of the corresponding reaction. The label on a binding edge is the product of the concentration of T, $[T]$, and an on-rate for binding, $a_{i,S}$, where i is the binding site and S is the subset of sites at which T is already bound. Subsets are denoted $\{i_1, \cdots, i_k\}$, where the site indices, $i_1, \cdots, i_k$, are drawn from the numbers $1, \cdots, n$. The label on an unbinding edge is an off-rate, $b_{i,S'}$, where S' is the subset of sites to which T is bound and i is one of the sites in S'.

Importantly, the on-rates, $a_{i,S}$, and the off-rates, $b_{i,S'}$, can depend on the site of binding or unbinding, i, as well as on the pattern of existing binding to a subset of sites, S or S'. This reflects the potential influence of background mechanisms, such as nucleosomes or co-regulators, and allows higher-order cooperativities to be introduced below.

The linear framework describes how such a graph gives rise to a stochastic master equation for the probabilities of the microstates. As in the thermodynamic formalism, we make the basic assumption that regulatory DNA is at steady state. However, unlike the thermodynamic formalism, the linear framework allows steady-state probabilities to be calculated regardless of whether or not the system is at thermodynamic equilibrium (Ahsendorf et al., 2014); see the Experimental Procedures and the Supplemental Information.

Higher-Order Cooperativities and the Exchange Formula

Higher-order cooperativity between multiple TF binding sites may be important in gene regulation, but thermodynamic formalism models have usually been limited to pairwise cooperativity. This

is not a fundamental limitation but arises from technical difficulties with the principle of detailed balance, which imposes algebraic constraints on higher-order cooperativities (Supplemental Information) that have not been worked out within the thermodynamic formalism. Detailed balance, or "microscopic reversibility," is a fundamental requirement arising from the time-reversal symmetry of the laws of physics (Mahan, 1975). The constraints are a serious obstacle because they mean that the numerical values of higher-order cooperativities cannot be chosen independently. Thus, it is important to determine these constraints (Equation 2) and to thereby identify a subset of cooperativities for which the numerical values are independent (Equation 3).

If the regulatory system described by G_n can reach thermodynamic equilibrium, the relevant parameters are the association constants $K_{i,S}$ (Figure 2B), of which there are $n2^{n-1}$ (Supplemental Information). To define higher-order cooperativities at equilibrium, we compare the binding of T to site i when T is already bound at the sites in S ($K_{i,S}$) to the binding of T to site i when T not bound elsewhere ($K_{i,\varnothing}$, where $\varnothing$ denotes the empty set). This yields a non-dimensional higher-order cooperativity, $\omega_{i,S} = K_{i,S}/K_{i,\varnothing}$, for which the value indicates whether or not there is positive or negative cooperativity or independence (Figure 2C). To non-dimensionalize the remaining association constants, we define $\kappa_i = K_{i,\varnothing}/K_{1,\varnothing}$,

The number of sites in S is called the order of $\omega_{i,S}$ and denoted $\#S$; it specifies how many sites collaborate to influence binding. Pairwise cooperativity corresponds to order 1. Thermodynamic formalism models set $\omega_{i,S} = 1$ for $\#S > 1$. In this case, detailed balance reduces to a symmetry requirement on pairwise cooperativities, $\omega_{i,\{j\}} = \omega_{j,\{i\}}$ (see Equation 2). With only pairwise cooperativity, there are only $n(n-1)/2$ parameters, instead of $n2^{n-1}$; this greatly simplifies thermodynamic formalism calculations.

We prove that, because of detailed balance, higher-order cooperativities must satisfy the "exchange formula" (Supplemental Information),

$$\omega_{i,S\cup\{j\}}\omega_{j,S} = \omega_{j,S\cup\{i\}}\omega_{i,S}, \tag{2}$$

which summarizes the algebraic constraints among the cooperativities. Here, i, j are sites not in S, while the notation $S\cup\{v\}$, for $v = i$ or $v = j$, denotes the addition of v to the sites in S. We further prove that, if we retain only those $\omega_{i,S}$ for which i is less than all the sites in S (abbreviated $i < S$), then the parameters

$$K_{1,\varnothing}, \quad \kappa_i\ (i > 1), \quad \omega_{i,S}\ (i < S), \tag{3}$$

of which there are $2^n - 1$, are algebraically independent, and all the $K_{i,S}$ can be calculated from them using Equation 2 (Supplemental Information). We can thus vary the parameters in Equation 3 independently and be confident that detailed balance holds. These fundamental results provide the basis for the equilibrium calculations that follow.

Equilibrium Gene Regulation Functions

To calculate a gene regulation function (GRF), we make the same basic assumption as in the thermodynamic formalism and consider the overall rate of transcription to be an average over the steady-state probabilities of the microstates. For this, we must specify the rate of transcription in each microstate, about which surprisingly little is known for eukaryotic genes. As explained above, we assume that T acts as a transcriptional activator (Supplemental Information), so that the binding of T does not reduce the expression level. We consider three expression strategies which work for any number of sites (Figure 3A): all-or-nothing, in which transcription only occurs when all sites are bound; one-or-more, in which transcription occurs when at least one site is bound; and average binding, in which transcription is proportional to the number of bound sites.

It is computationally infeasible to explore all expression strategies, but these three strategies broadly sample the spectrum of possibilities (Supplemental Information). All-or-nothing and one-or-more are extreme opposites, while average binding is an intermediate strategy. All-or-nothing is widely used in thermodynamic formalism models, and average binding corresponds to the "fractional saturation" used in models of protein allostery (Monod et al., 1965; Mirny, 2010).

The level of protein expression after normalization to its asymptotic maximum is a rational function of $x = [T]$, denoted $f_n(x)$, which has the form, for the all-or-nothing strategy (Supplemental Information),

$$f_n(x) = \frac{c_n x^n}{1 + c_1 x + \cdots + c_n x^n}. \tag{4}$$

The coefficients c_k are given (Supplemental Information) by a sum of products

$$c_k = \left(\sum_{1 \le i_1 < \cdots < i_k \le n} \left(\prod_{j=1}^{k} \kappa_{i_j} \omega_{i_j,\{i_{j+1},\cdots,i_k\}} \right) \right) (K_{1,\varnothing})^k, \tag{5}$$

which involves only the independent parameters in Equation 3 and allows higher-order cooperativity of any order up to the maximum of $n - 1$. For the other strategies (Figure 3A), only the numerator of Equation 4 changes (Supplemental Information). The GRFs discussed in this paper are strictly increasing functions (Supplemental Information). For the all-or-nothing strategy at equilibrium, they also appear to be sigmoidal ("S shaped"), so that the derivative of the GRF has only a single maximum, but this is not so in general (Figure 3B).

The algebraic form of the Hill function, $\mathcal{H}_a(x) = x^a / (1 + x^a)$, is closest to that of the GRF for the all-or-nothing strategy in Equation 4 (Supplemental Information). If $a < n$, it is clear that the former cannot algebraically resemble the latter because the degrees of their respective denominator polynomials are different. If $a = n$, algebraic resemblance is only possible, and then only approximately, if the parameters in the GRF are given implausible numerical values (Supplemental Information). We see that Hill functions are not GRFs. However, the question posed by Equation 1 is whether a GRF can match the shape of a Hill function. This is a more delicate problem.

Position and Steepness as Quantitative Measures of Shape

To determine the match between a GRF and a Hill function, we introduce two quantitative measures of shape. We first normalize the concentration scale of $x = [T]$, in a similar way to Equation 1,

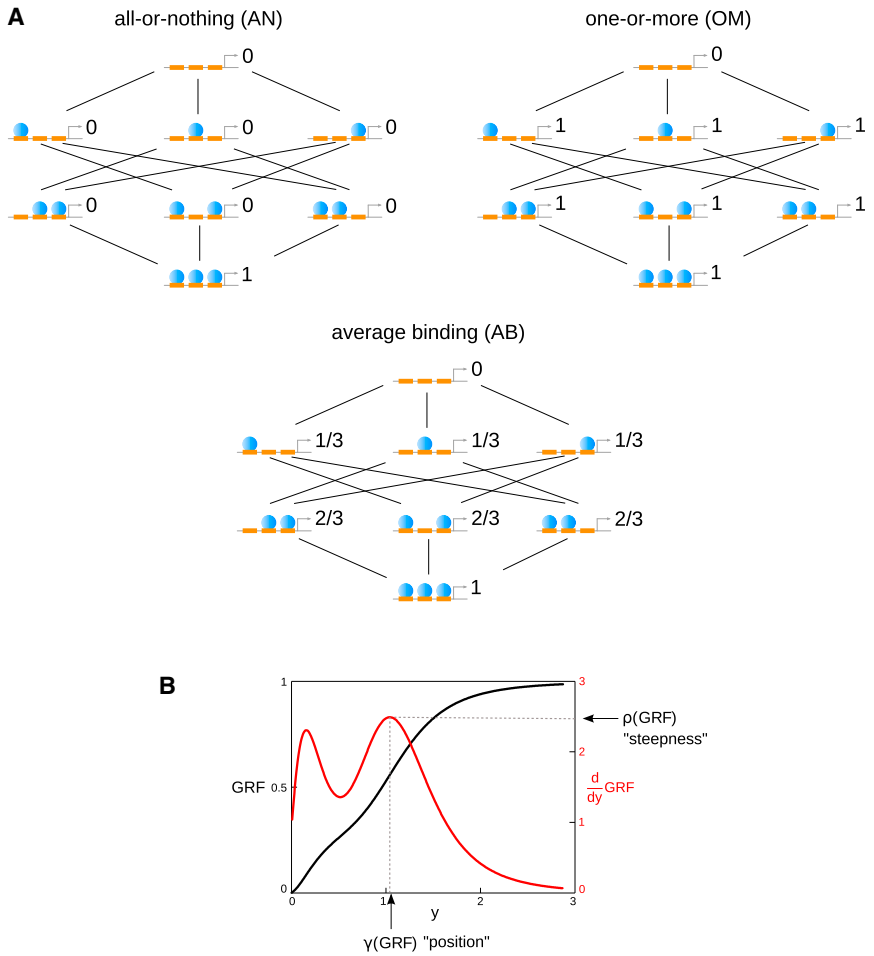


Figure 3. Gene Expression Strategies and Shape Measures

(A) The graph G_3, with pairs of reversible edges as unlabelled single lines for clarity, illustrating three expression strategies for a transcriptional activator. Each microstate is annotated with a number describing the corresponding rate of gene expression, with the maximal rate normalized to 1. (B) Plot of a hypothetical GRF (black), together with its derivative (red) showing steepness (ρ) and position (γ), as defined in Equation 6. The derivative can have multiple local maxima, and ρ and γ are defined at the global maximum.

by taking $x_{0.5}$ to be the concentration at which f_n is half-maximal, so that $f_n(x_{0.5}) = 0.5$, and setting $y = x/x_{0.5}$. The normalized GRF is $g_n(y) = f_n(yx_{0.5})$, where y and $g_n(y)$ are now both non-dimensional quantities. Note that for $\mathcal{H}_a(x)$, $x_{0.5} = 1$, so that Hill functions are already normalized.

To quantify shape, we take the maximum derivative, $\rho(g_n)$ ("steepness"), and the position of the maximum derivative $\gamma(g_n)$ ("position"),

$$\rho(g_n) = \max_{y \geq 0} \frac{dg_n}{dy}, \quad \gamma(g_n) = z \text{ such that } \left.\frac{dg_n}{dy}\right|_{y=z} = \rho(g_n), \qquad (6)$$

which are also non-dimensional quantities (Figure 3B). The advantage of γ and ρ is that they can be calculated from g_n, in contrast to a numerical fit to a Hill function, which is subject to statistical noise. Two GRFs with the same γ and ρ ("matched") are not identical but have similar sharpness. Considering only one measure of sharpness, such as ρ, can be misleading (below).

Impact of Higher-Order Cooperativity on Sharpness

We first determined the position and steepness of a GRF in the all-or-nothing strategy with $n = 5$ sites, allowing higher-order cooperativity of any order. Although γ and ρ depend only on the coefficients c_k in Equation 5, these aggregated parameters have no biochemical meaning and we seek instead to understand how γ and ρ depend on the affinities and cooperativities in Equation 3, which are defined in terms of molecular interactions. Because of normalization, γ and ρ do not depend on $K_{1,\varnothing}$ (Supplemental Information), so we set $K_{1,\varnothing} = 1$ in units of $(\text{concentration})^{-1}$ and chose κ_i and $\omega_{i,S}$ to be in the range $[10^{-3}, 10^{3}]$ by random logarithmic sampling. We believe this range is generous, but most of our results do not depend on it (below; Supplemental Information).

A sample of 10^5 GRFs chosen in this way reveals that position and steepness are not independent but are constrained within a crescent-shaped region in which the highest steepness is found at the extremes of position (Figure 4A). On the left of the region, high steepness occurs only for very low position and the resultant GRFs are highly degenerate: when these GRFs are fitted to Hill functions, they yield a Hill coefficient of $a = 1$ (inset at top and caption). The marginal distribution (top) shows that these degenerate GRFs account for nearly half of all GRFs in this parameterisation. Degeneracy underscores the importance of considering γ together with ρ.

The upper edge of the crescent-shaped region has low probability. We therefore used, for Figure 4 and those that follow, a biased sampling algorithm to identify the boundary of the region (Supplemental Information), with the same parameter range but with the GRFs filtered so that $\gamma(g_n)$ lies in the interval $[0.5\gamma(\mathcal{H}_n), 1.5\gamma(\mathcal{H}_n)]$. This focuses on the GRFs of interest and avoids the degeneracy near $\gamma(g_n) = 0$. We found the gray boundary in Figure 4A.

The right-hand edge of the gray boundary coincides with that of the randomly sampled region in a series of line segments. Strikingly, the "Hill line" on which Hill functions are located (magenta curve) lies just to the right of this boundary. We see that a GRF cannot have greater position than a Hill function of the same steepness, so that the Hill functions define a barrier. The line segments on the boundary of the region touch the Hill line at their corners and these occur, surprisingly, at

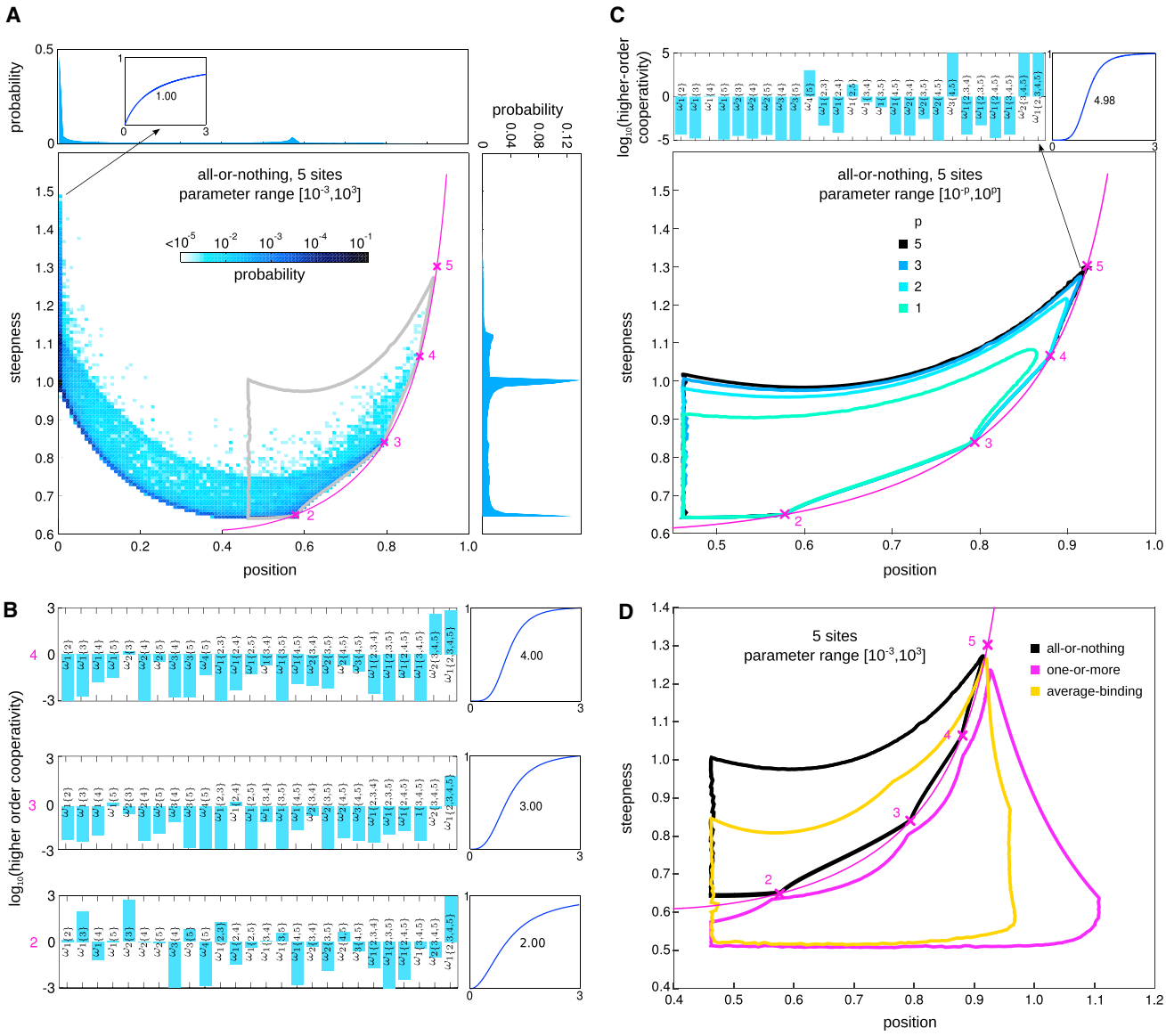


Figure 4. Position and Steepness for $n = 5$ Sites

(A) Probability density function (blue points) of $(\gamma(g_5), \rho(g_5))$, obtained by random sampling, with the respective marginal distributions (top and right). The inset (top) shows the GRF of the marked point, annotated with the value *a* obtained by fitting to $\mathcal{H}_a$. The gray line marks the boundary of the position-steepness region, obtained by a biased sampling algorithm (Supplemental Information). The magenta line is the locus of $(\gamma(\mathcal{H}_a), \rho(\mathcal{H}_a))$ for varying *a* (the "Hill line"), with the integer values of *a* marked by magenta crosses and numbers.

(B) Higher-order cooperativities (left), plotted on a logarithmic scale, for the GRFs closest in (γ, ρ) distance to the integer Hill coefficients (magenta crosses) in (A), with the Hill coefficient annotated on the left (magenta). The corresponding curve (right) is annotated with the value *a*, obtained by fitting to $\mathcal{H}_a$.

(C) Position-steepness boundaries with the parameter range $[10^{-p}, 10^{p}]$ for varying *p*. At the top are the higher-order cooperativities (left) and curve (right) for the marked GRF closest to $\mathcal{H}_5$ within the $p = 5$ region, plotted as in (B).

(D) Position-steepness regions for all three expression strategies; see also Figures 6 and S1.

See also Figure S1.

exactly the integer values, 2, 3, 4, of the Hill coefficient. When the GRFs which are closest to these points are fitted to Hill functions, the estimated Hill coefficients correspond very closely to the integer values (Figure 4B, right). Such a close correspondence in fitted shape is unexpected in view of the lack of algebraic resemblance between GRFs and Hill functions, as discussed above. The emergence of bona fide GRFs which closely match the shape of Hill functions with integer Hill coefficients is intriguing in view of the coefficient 5 found in Equation 1. However, this shape matching requires high levels of positive and negative higher-order cooperativity of all orders (Figure 4B, left).

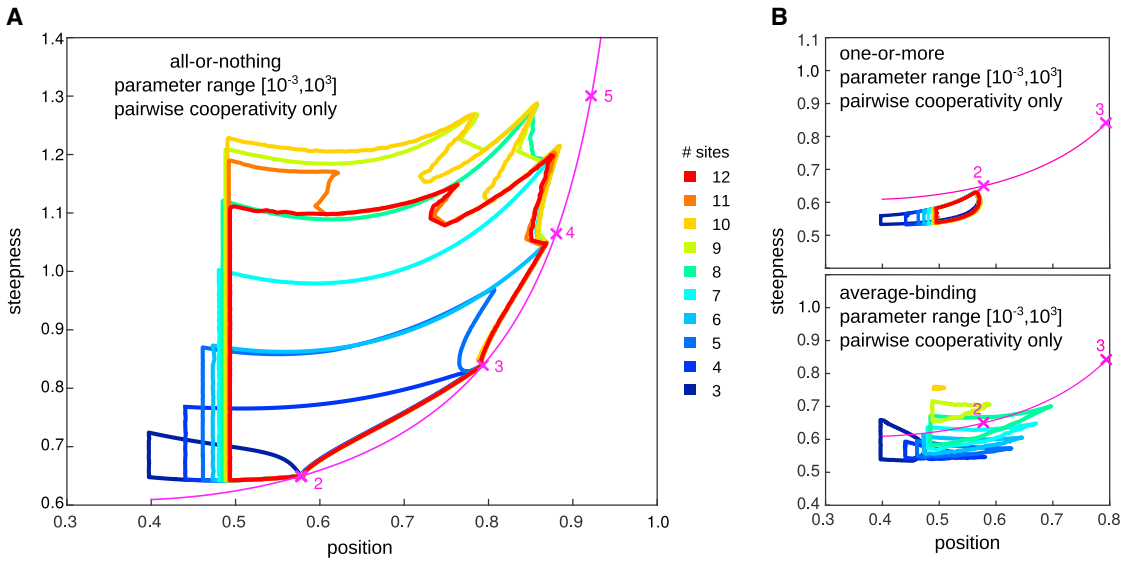


Figure 5. Pairwise Cooperativity Only

(A and B) Each panel uses the color-code in the center and shows the Hill line in magenta. (A) Position-steepness regions for the all-or-nothing strategy. (B) Position-steepness regions for the one-or-more (top) and average-binding (bottom) strategies. The biased sampling algorithm had to be modified to find the average-binding regions (Supplemental Information).

The boundary of the position-steepness region lies below $\mathcal{H}_5$ and approaches it at the tip of a cusp. Changing the parameter range does not alter the line segments in the boundary, but the tip of the cusp approaches closer to $\mathcal{H}_5$ as the range is increased (Figure 4C). The GRF closest to $\mathcal{H}_5$ has a fitted Hill coefficient close to 5 (top, right), although not as close as for integer values less than 5. This still requires high levels of positive and negative higher-order cooperativity of all orders (top, left).

Each expression strategy reveals a different trade-off between position and steepness (Figure 4D). The Hill line also presents a barrier to the one-or-more strategy but from the opposite side, while the average-binding strategy straddles the Hill line. For the one-or-more strategy, the position-steepness region approaches at cusps the Hill functions whose coefficients are integers less than 5 (Figure 4D) but, in contrast to the all-or-nothing strategy, the region does not touch the Hill line and the integer-valued Hill functions are not closely matched to the nearest GRFs unless the parameter range is increased (data not shown). The barrier presented by the Hill line seems, therefore, to act differently in the all-or-nothing and one-or-more strategies. Regardless of the expression strategy, $\mathcal{H}_5$ offers a barrier to all strategies with $n = 5$ sites: each region lies below it and only approaches it at the tip of a cusp as the parameter range is increased.

The features found above are reproduced for different numbers of sites (Figure S1 for $n = 7$ sites).

Pairwise Cooperativity Alone Permits Limited Sharpness

Thermodynamic formalism models have typically been limited to pairwise cooperativity. We restricted ourselves to pairwise cooperativity by setting $\omega_{i,S} = 1$ for $\#S > 1$. We found that the position-steepness region for the all-or-nothing strategy increases initially with increasing n but then shrinks in extent and no GRF approaches close to $\mathcal{H}_5$ (Figure 5A). The one-or-more and average-binding strategies do not even get close to $\mathcal{H}_3$ (Figure 5B).

Thus, in contrast to common assumptions, pairwise cooperativity alone is insufficient for sharp responses in eukaryotic genes. None of the expression strategies considered here can account for $\mathcal{H}_5$ in Equation 1 with only pairwise cooperativity, no matter how many sites are available.

Non-equilibrium GRFs Exceed the Equilibrium Sharpness Barriers

If the system is maintained away from thermodynamic equilibrium by energy expenditure, detailed balance no longer holds. The non-equilibrium GRF for the all-or-nothing strategy then takes the form (Supplemental Information)

$$f_n^{ne}(x) = \frac{d_n x^n + \cdots + d_{2^n-1} x^{2^n-1}}{e_0 + e_1 x + \cdots + e_{2^n-1} x^{2^n-1}}, \qquad (7)$$

where the coefficients of the highest order term, x^{2^n-1}, in the numerator and the denominator are equal, so that $d_{2^n-1} = e_{2^n-1}$. GRFs for the other strategies differ only in the numerator (Supplemental Information). The denominator of Equation 7 shows a striking increase in degree, from n to $2^n - 1$, in comparison to that of the equilibrium f_n in Equation 4, despite the number of sites being the same.

The parameters in Figure 2B are no longer meaningful away from equilibrium and the coefficients d_i and e_i in Equation 7 are expressions in the rate constants $a_{i,S}$ and $b_{i,S'}$. For reasons discussed below, the largest number of sites that we can feasibly analyze is $n = 3$ (Supplemental Information).

We took non-dimensional parameters $a_{i,S}/a_{1,\varnothing}$ and $b_{i,S'}/b_{1,\{1\}}$ in the range $[10^{-2}, 10^2]$, deliberately restricting the range so that,

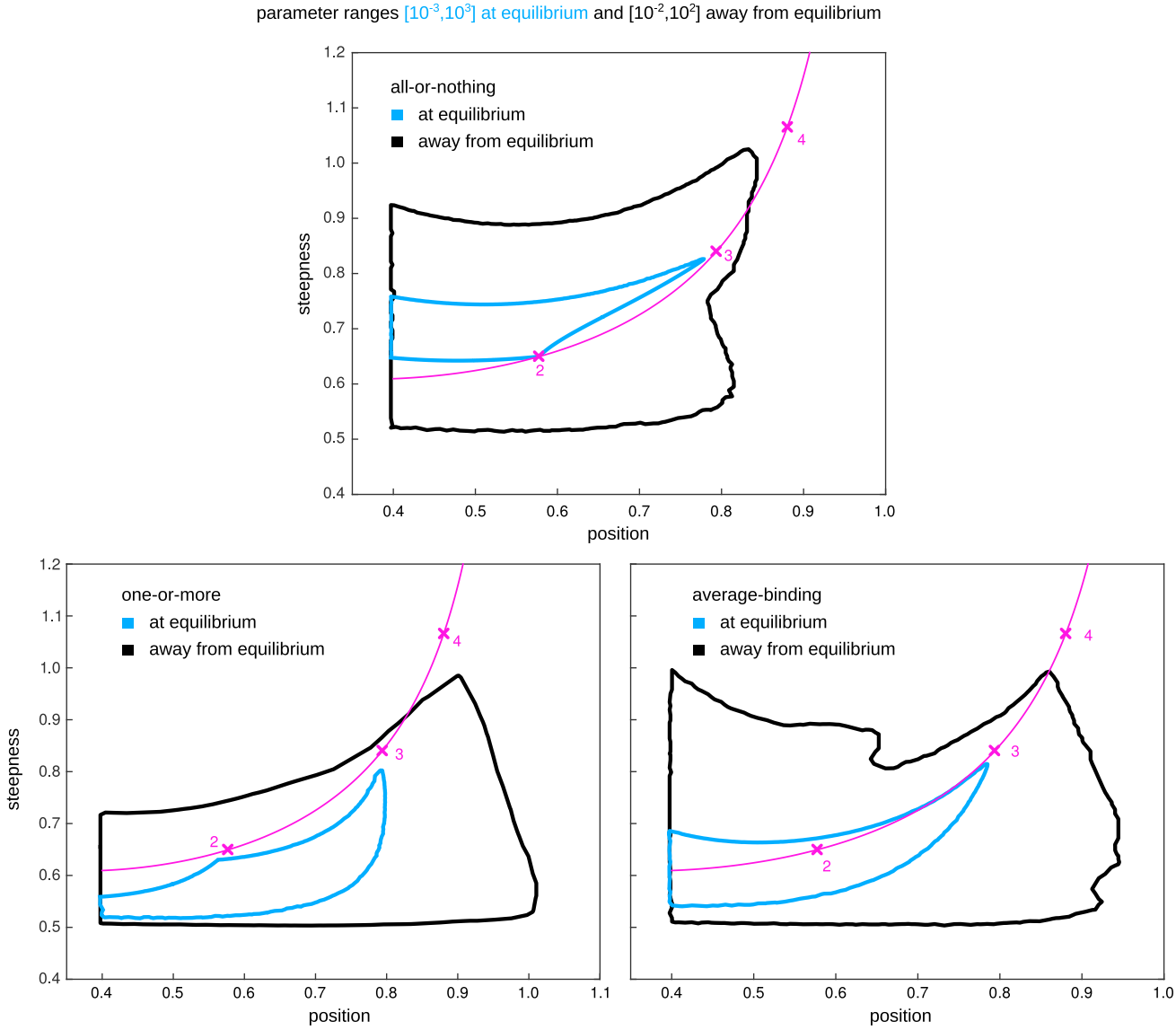


Figure 6. The Non-equilibrium Case for $n = 3$ Sites

Position-steepness regions for all expression strategies, showing the equilibrium (blue) and non-equilibrium (black) boundaries. The horizontal scale for the one-or-more strategy is extended.

See also Figure S2.

if the system were at equilibrium, it would be comparable with the previous equilibrium analysis (Supplemental Information). Because of normalization, the steepness and position of g_n^{ne} are independent of the values of $a_{1,\varnothing}$ and $b_{1,\{1\}}$ (Supplemental Information), so we set $a_{1,\varnothing} = 1$ and $b_{1,\{1\}} = 1$ in their respective units. For each expression strategy, we found (Figure 6) that the non-equilibrium position-steepness region is much enlarged (black boundary) compared to the corresponding equilibrium region for the same number of sites (blue boundary). The non-equilibrium regions now include the Hill line up to $\mathcal{H}_3$ and the all-or-nothing region can reach as far as $\mathcal{H}_5$ if the parameter range is increased (Figure S2A).

The limitation to $n = 3$ sites arises from loss of detailed balance, which leads to a dramatic increase in the complexity of the coefficients in Equation 7 (Supplemental Information; see the Discussion). This complexity is algebraic, not numerical. To compute position-steepness regions, cooperativities are treated as symbols whose numerical values are assigned by sampling. Symbolic calculation of the GRF is extremely expensive away from equilibrium but numerical calculation of individual GRFs presents no particular difficulty.

Symbolic treatment of parameters is informative because it reveals the structure of the non-equilibrium GRF (Equation 7). The denominator of this GRF increases in degree exponentially

with n but the denominator of the equilibrium GRF (Equation 4) increases only linearly. This discrepancy arises from loss of detailed balance (Supplemental Information). With $n = 3$ sites, the non-equilibrium position-steepness region comfortably exceeds the equilibrium region (Figure 6). Because of the exponential increase in the degree of the GRF denominator, when there are $n = 5$ sites, or however many sites are relevant for Bcd regulation of *hb*, the discrepancy in the position-steepness regions will be even greater and the non-equilibrium region will extend well beyond $\mathcal{H}_5$.

In confirmation of this, we used numerical parameter values to find a non-equilibrium GRF on $n = 5$ sites, with parameters in the same range, $[10^{-2}, 10^{2}]$, as in Figure 6, whose position and steepness match that of $\mathcal{H}_6$ (Figure S2B). If the parameter range is increased to $[10^{-3}, 10^{3}]$, then there is a GRF on four sites, for which position and steepness match that of $\mathcal{H}_{5.7}$ (Figure S2B). Being away from equilibrium makes it much easier to achieve the sharpness required for Equation 1.

DISCUSSION

Eukaryotic gene regulation lies at the nexus of many of the central issues in modern biology, including multi-cellular development (Davidson, 2006), the evolution of complexity (Carroll, 2008), cellular reprogramming (Takahashi and Yamanaka, 2016) and synthetic biology (Keung et al., 2015). The extraordinary molecular complexity implicated in such regulation continues to present a formidable challenge. It has made it difficult to see the wood for the trees, to discern general principles and to unravel how different molecular mechanisms contribute to specific forms of information processing.

In this paper, we have presented compelling evidence that the bacterial paradigm, upon which it has been so convenient to default, is not sufficient for reasoning about eukaryotic genes and we have introduced appropriate quantitative concepts for doing so. We have done this by taking seriously the lessons of the bacterial paradigm itself. The paradigm relied on analyzing the physics of interaction between TFs and DNA. What we have done here is to update this foundation for two of the key processes that influence gene-expression sharpness in eukaryotes: information integration and energy expenditure.

While information integration is often acknowledged, it has not been defined sufficiently clearly to know how to find it, making experimental analysis problematic. We have introduced the concept of "higher-order cooperativity," $\omega_{i,S}$ (Figure 2B), for a system at thermodynamic equilibrium, as a measure of how the affinity of TF binding at site i is influenced by the presence of TF bound at the sites in S. This is a precisely defined quantity that allows experimentally-testable hypotheses to be framed.

Different TFs often work together and the definition of higher-order cooperativities that we have given here for homotypic interactions of a single TF can be readily extended to heterotypic interactions between different TFs. Such cooperativities could arise from nucleosomes (Mirny, 2010; Voss et al., 2011) or from co-regulators like Mediator and CBP/p300 (Borggrefe and Yue, 2011; Wang et al., 2013). Mediator is especially provocative as a potential mechanism of higher-order cooperativity. Mediator has around 30 subunits and the Med1 subunit alone interacts with up to 20 different TFs (Borggrefe and Yue, 2011). Some TFs interact with multiple subunits and the overall Mediator complex exhibits different conformations, which suggests how local information may be globally integrated (Nussinov et al., 2013). Our analysis shows that if gene expression follows an all-or-nothing strategy, high levels of positive and negative higher-order cooperativity of all orders are needed to yield high levels of sharpness (Figures 4B and 4C). Experimental measurements will show whether co-regulators like Mediator or CBP/p300 can meet these requirements.

When energy is expended during gene regulation, much higher sharpness can be achieved for the same number of TF binding sites (Figure 6). The concept of a "Hopfield barrier" offers a way to articulate this rigorously. With n binding sites, the Hopfield barrier to sharpness is set by the Hill function $\mathcal{H}_n$, whose steepness cannot be exceeded by any equilibrium GRF (Figure 4D). For GRFs in the all-or-nothing strategy, the Hill line itself forms a Hopfield barrier (Figure 4C). However, both these barriers are readily breached away from equilibrium (Figures 6 and S2).

There are many routes through which energy can be expended, including chromatin reorganization, nucleosome displacement, protein post-translational modification and DNA methylation. Experiments which perturb these routes and assay the impact on sharpness can bring to light which energy expending mechanisms are particularly relevant. For the specific example of *hb* regulation by Bcd, we note that Bcd binds to *Drosophila* Mediator in a way that affects early embryonic patterning (Park et al., 2001; Bosveld et al., 2008) and that Bcd also binds to the Sin3/Rp3 histone-deacetylase complex (Singh et al., 2005). The impact of these interactions on sharpness appears not to have been previously studied. The early *Drosophila* embryo provides an unrivalled experimental context for testing the hypotheses made here and this is now work in progress.

Real-time studies have already confirmed the importance of non-equilibrium kinetics in gene regulation (Voss et al., 2011; Hammar et al., 2014). Coulon et al. (2013) have also argued for the importance of a non-equilibrium perspective. Our results strongly endorse this but an important challenge lies ahead. When a system is at thermodynamic equilibrium, detailed balance implies that any path to a microstate can be used to calculate the steady-state probability of the microstate. The history of the system is irrelevant. Away from equilibrium, detailed balance no longer holds and all possible paths to a microstate must be examined to calculate its steady-state probability. Non-equilibrium systems are history dependent. The resultant combinatorial explosion results in a profound increase in algebraic complexity, which manifests itself in the striking difference between the equilibrium GRF in Equation 4 and the non-equilibrium GRF in Equation 7. It is this which leads to the breaching of the Hopfield barrier. We suspect that further insights into non-equilibrium gene regulation are concealed within this algebraic complexity. We are only just learning how to uncover them (Ahsendorf et al., 2014).

The concepts introduced here encourage us to examine other forms of genetic information processing. If energy expenditure is important, what is it buying? What could not be achieved if regulatory DNA is at thermodynamic equilibrium? Such questions

can be answered by quantifying each information processing task, as we have done here for sharpness (Equation 6), and developing experimental systems in which it can be measured. We may look forward in this way to a quantitative classification of the kinds of information processing that genes undertake and an understanding in molecular terms of how energy expenditure breaks the corresponding Hopfield barriers.

We note two further implications of the present paper. First, Hill functions emerge in an unexpected light. When Archibald Vivian (A.V.) Hill first introduced them in 1913, he recognized that they had no biochemical justification and were only a convenient fit to the data on oxygen binding to haemoglobin (Hill, 1913). The empirical nature of Hill functions has been repeatedly pointed out (Engel, 2013; Weiss, 1997). We were all the more surprised, therefore, to find that, when the Hill coefficient is an integer, there are bona fide GRFs which are statistically indistinguishable from Hill functions (Figure 4B). In this sense, the Hill functions appear to be closer to biochemistry than Hill, or anyone else, could have imagined. We hope to clarify the mathematical reasons for this in subsequent work.

Second, we have exploited mathematics differently here to what is sometimes expected of it. We have relied on data to frame the question but we have not fitted any mathematical models to data. In recent years, experimental biologists have become more comfortable with the idea that theory can follow experiment, as a way to analyze and understand data. Our colleague Rob Phillips calls this "Figure 7 theory" (Phillips, 2015). Here, we have used mathematics to introduce concepts and to determine the limits of what can be expected, thereby providing a foundation for designing new kinds of experiments. Experiment will follow theory. This is "Figure 1 theory" (Phillips, 2015). We believe it has much to recommend it as we face the daunting molecular complexity of eukaryotic gene regulation. We need to think about how such regulation works using concepts which are not just based on intuition and induction but are also grounded in the underlying physics, which is the bedrock on which all biology rests. This is an old lesson (Gunawardena, 2013; Bialek, 2015). If we have lost sight of it in the press of mastering the molecular details, now is the time to revisit it and construct a new paradigm for eukaryotic gene regulation.

EXPERIMENTAL PROCEDURES

The mathematical model and the results presented here are based on the "linear framework" for gene regulation (Ahsendorf et al., 2014) (see the Supplemental Experimental Procedures for full details). The framework starts from a labeled, directed graph, $\mathcal{G}$, which gives rise to a stochastic master equation,

$$\frac{du}{dt} = L(\mathcal{G}).u,$$

for the vector of microstate probabilities, $u = (u_1, \cdots u_n)^t$. Here, $L(\mathcal{G})$ is the Laplacian matrix of $\mathcal{G}$ and N is the number of microstates in the graph. For the graph G_n used here, $N = 2^n$, where n is the number of TF binding sites. Provided $\mathcal{G}$ is strongly connected, which is the case for G_n, the steady state, u^*, at which $(du/dt)\,|_{u=u^*} = 0$, is unique up to a scalar multiple. A basis vector may be calculated in terms of the edge labels in one of two ways, depending on whether or not the system is at equilibrium. If the system reaches thermodynamic equilibrium, u_i^* can be calculated by choosing any path of reversible edges from the reference vertex 1 to i and taking the product, over all reversible edges in the path, of the ratio of the label on the forward edge, in the direction from 1 to i, to the label on the reverse edge. The principle of detailed balance ensures that this result is independent of the chosen path because of the cycle condition: on any cycle of reversible edges, the product of the labels going clockwise around the cycle equals the product of the labels going counterclockwise. The cycle condition leads to the exchange formula in Equation 2, which allows the algebraically independent set of parameters in Equation 3 to be chosen. For G_n, a path of reversible edges can be chosen from 1, the vertex with no sites bound, to i, such that u_i^* is expressed in terms of the independent parameters. Away from equilibrium, u_i^* has to be calculated using the matrix-tree theorem as a sum, over all directed spanning trees rooted at i, of the product of the labels on the edges of each spanning tree. Once u^* is known, the state-state probability of microstate i is given by

$$\Pr(i) = \frac{u_i^*}{u_1^* + \cdots + u_N^*}.$$

For a system that reaches thermodynamic equilibrium, the denominator in this formula is the partition function of equilibrium statistical mechanics, but the formula equally holds for a system away from thermodynamic equilibrium with u^* calculated as above. The gene regulation function for mRNA production rate as output is defined as an average over the steady-state probabilities,

$$\frac{d}{dt}[\mathrm{mRNA}] = \sum_{1 \le i \le N} r(i)\Pr(i).$$

The expression rate in microstate i, given by $r(i)$, depends on the gene expression strategy being followed, as specified in Figure 3. To obtain protein level as output, assume that mRNA is linearly degraded and that the steady-state protein level is proportional to the steady-state mRNA level. The proportionality constants are absorbed in the normalization that underlies the definitions of position and steepness in Equation 6.

In the equilibrium GRFs, the quantities u_i^* depend on paths of reversible edges from 1 to i, which can incur up to n factors of $x = [T]$, so that the degree of the denominator polynomial in $\Pr(i)$ is n (Equation 4). In contrast, in the non-equilibrium GRFs, the quantities u_i^* depend on directed spanning trees rooted at i, which each have $N - 1$ edges and can incur up to $N - 1$ factors of x, so that the degree of the denominator polynomial becomes $2^n - 1$ (Equation 7).

The numerical results presented in Figures 4, 5, and 6 are obtained by a biased sampling algorithm in which the boundary of the position-steepness region is found by successive approximation. An initial region is found by independently selecting parameter values for GRFs by logarithmic random sampling within the specified range, calculating the (γ, ρ) coordinates of these GRFs, and determining the enclosing boundary. This initial boundary is then successively improved by randomly altering GRFs on the current boundary until the area of the region ceases to increase. The details, along with the tests that were used to confirm convergence and to check the numerical accuracy of the results, are given in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and two figures and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2016.06.012.

AUTHOR CONTRIBUTIONS

A.D. and J.G. formulated and supervised the project. J.E. developed the algorithms and carried out the numerical calculations, which J.E. and F.W. separately tested. F.W. and J.G. carried out the mathematical derivations. J.E., A.D., and J.G. made the figures. J.G. wrote the paper with the assistance of all authors.

ACKNOWLEDGMENTS

We thank the editor, Robert Kruger, and the anonymous reviewers for their help. We thank Rebecca Ward for editorial assistance. We gratefully acknowledge the Department of Systems Biology for supporting J.E. F.W. was

supported by a National Science Foundation (NSF) grant (GRF DGE1144152); A.D. was supported by an NIH grant (U01 GM103804) and NSF CAREER grant (1452557); and J.G. was supported by an NSF grant (1462629).

Received: July 3, 2015
Revised: March 8, 2016
Accepted: June 1, 2016
Published: June 30, 2016

REFERENCES

Ackers, G.K., Johnson, A.D., and Shea, M.A. (1982). Quantitative model for gene regulation by lambda phage repressor. Proc. Natl. Acad. Sci. USA *79*, 1129–1133.

Ahsendorf, T., Wong, F., Eils, R., and Gunawardena, J. (2014). A framework for modelling gene regulation which accommodates non-equilibrium mechanisms. BMC Biol. *12*, 102.

Bialek, W. (2015). Perspectives on theory at the interface of physics and biology. arXiv, arXiv:1512.08954.

Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., and Phillips, R. (2005). Transcriptional regulation by the numbers: models. Curr. Opin. Genet. Dev. *15*, 116–124.

Borggrefe, T., and Yue, X. (2011). Interactions between subunits of the Mediator complex with gene-specific transcription factors. Semin. Cell Dev. Biol. *22*, 759–768.

Bosveld, F., van Hoek, S., and Sibon, O.C.M. (2008). Establishment of cell fate during early Drosophila embryogenesis requires transcriptional Mediator subunit dMED31. Dev. Biol. *313*, 802–813.

Carroll, S.B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell *134*, 25–36.

Coulon, A., Chow, C.C., Singer, R.H., and Larson, D.R. (2013). Eukaryotic transcriptional dynamics: from single molecules to cell populations. Nat. Rev. Genet. *14*, 572–584.

Davidson, E.H. (2006). The Regulatory Genome: Gene Regulatory Networks in Development and Evolution (Academic Press).

Engel, P.C. (2013). A hundred years of the Hill equation. Biochem. J. *2013*, 1–4.

Gregor, T., Tank, D.W., Wieschaus, E.F., and Bialek, W. (2007). Probing the limits to positional information. Cell *130*, 153–164.

Gunawardena, J. (2013). Biology is more theoretical than physics. Mol. Biol. Cell *24*, 1827–1829.

Gunawardena, J. (2014). Models in biology: 'accurate descriptions of our pathetic thinking'. BMC Biol. *12*, 29.

Hammar, P., Walldén, M., Fange, D., Persson, F., Baltekin, O., Ullman, G., Leroy, P., and Elf, J. (2014). Direct measurement of transcription factor dissociation excludes a simple operator occupancy model for gene regulation. Nat. Genet. *46*, 405–408.

Hill, A.V. (1913). The combinations of haemoglobin with oxygen and with carbon monoxide. I. Biochem. J. *7*, 471–480.

Hopfield, J.J. (1974). Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. Proc. Natl. Acad. Sci. USA *71*, 4135–4139.

Keung, A.J., Joung, J.K., Khalil, A.S., and Collins, J.J. (2015). Chromatin regulation at the frontier of synthetic biology. Nat. Rev. Genet. *16*, 159–171.

Lebrecht, D., Foehr, M., Smith, E., Lopes, F.J., Vanario-Alonso, C.E., Reinitz, J., Burz, D.S., and Hanes, S.D. (2005). Bicoid cooperative DNA binding is critical for embryonic patterning in Drosophila. Proc. Natl. Acad. Sci. USA *102*, 13176–13181.

Mahan, B.H. (1975). Microscopic reversibility and detailed balance. J. Chem. Educ. *52*, 299–302.

Mirny, L.A. (2010). Nucleosome-mediated cooperativity between transcription factors. Proc. Natl. Acad. Sci. USA *107*, 22534–22539.

Monod, J., Wyman, J., and Changeux, J.P. (1965). On the nature of allosteric transitions: a plausible model. J. Mol. Biol. *12*, 88–118.

Ninio, J. (1975). Kinetic amplification of enzyme discrimination. Biochimie *57*, 587–595.

Nussinov, R., Tsai, C.-J., and Ma, B. (2013). The underappreciated role of allostery in the cellular network. Annu. Rev. Biophys. *42*, 169–189.

Park, J.M., Gim, B.S., Kim, J.M., Yoon, J.H., Kim, H.-S., Kang, J.-G., and Kim, Y.-J. (2001). Drosophila Mediator complex is broadly utilized by diverse gene-specific transcription factors at different types of core promoters. Mol. Cell. Biol. *21*, 2312–2323.

Perry, M.W., Boettiger, A.N., and Levine, M. (2011). Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. Proc. Natl. Acad. Sci. USA *108*, 13570–13575.

Perry, M.W., Bothma, J.P., Luu, R.D., and Levine, M. (2012). Precision of hunchback expression in the Drosophila embryo. Curr. Biol. *22*, 2247–2252.

Phillips, R. (2015). Theory in biology: Figure 1 or Figure 7? Trends Cell Biol. *25*, 723–729.

Ptashne, M. (2004). A Genetic Switch: Phage Lambda Revisited, Third Edition (Cold Spring Harbor Laboratory Press).

Segal, E., and Widom, J. (2009). From DNA sequence to transcriptional behaviour: a quantitative approach. Nat. Rev. Genet. *10*, 443–456.

Sherman, M.S., and Cohen, B.A. (2012). Thermodynamic state ensemble models of cis-regulation. PLoS Comp. Biol. *8*, e1002407.

Singh, N., Zhu, W., and Hanes, S.D. (2005). Sap18 is required for the maternal gene bicoid to direct anterior patterning in Drosophila melanogaster. Dev. Biol. *278*, 242–254.

Spitz, F., and Furlong, E.E.M. (2012). Transcription factors: from enhancer binding to developmental control. Nat. Rev. Genet. *13*, 613–626.

Takahashi, K., and Yamanaka, S. (2016). A decade of transcription factor-mediated reprogramming to pluripotency. Nat. Rev. Mol. Cell Biol. *17*, 183–193.

Voss, T.C., Schiltz, R.L., Sung, M.-H., Yen, P.M., Stamatoyannopoulos, J.A., Biddie, S.C., Johnson, T.A., Miranda, T.B., John, S., and Hager, G.L. (2011). Dynamic exchange at regulatory elements during chromatin remodeling underlies assisted loading mechanism. Cell *146*, 544–554.

Wang, F., Marshall, C.B., and Ikura, M. (2013). Transcriptional/epigenetic regulator CBP/p300 in tumorigenesis: structural and functional versatility in target recognition. Cell. Mol. Life Sci. *70*, 3989–4008.

Weiss, J.N. (1997). The Hill equation revisited: uses and misuses. FASEB J. *11*, 835–841.

Wunderlich, Z., and Mirny, L.A. (2009). Different gene regulation strategies revealed by analysis of binding motifs. Trends Genet. *25*, 434–440.

Cell

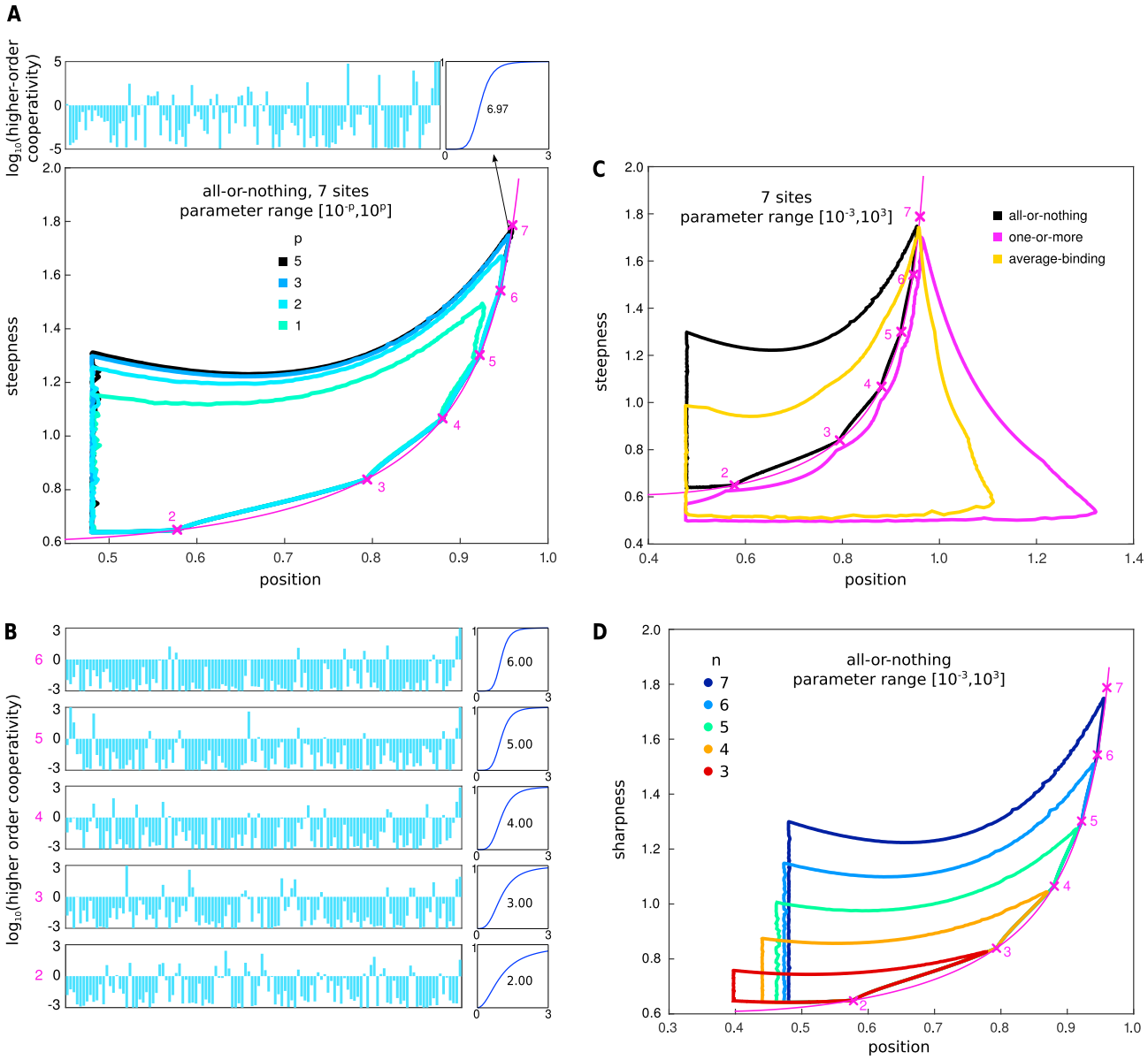


Figure S1. Equilibrium Position-Steepness Regions for up to $n = 7$ Sites, Related to Figure 4

(A) The all-or-nothing strategy for $n = 7$ sites and varying parameter ranges, as shown. At the top, higher-order cooperativities (left) and curve (right) for the GRF closest to $\mathcal{H}_7$ in the region for $p = 5$. The 120 higher-order cooperativities cannot be individually annotated but are listed in blocks of increasing cooperativity order from left to right and within each block of cooperativity order $k - 1$, the $\omega_{i_1,\{i_2,\cdots i_k\}}$ are listed in lexicographic order for the sequence $i_1, \cdots, i_k$.

(B) Higher-order cooperativities (left) for the GRF, chosen from the region in panel (A) corresponding to $p = 5$, that is closest to $\mathcal{H}_k$, where k is the integer value in magenta on the left. The cooperativities are in the same order as for panel (A). The corresponding curve (right), is annotated with the value a obtained by fitting to $\mathcal{H}_a$.

(C) Position-steepness regions for all three expression strategies.

(D) Position-steepness regions for the all-or-nothing strategy for n varying from 3 to 7.

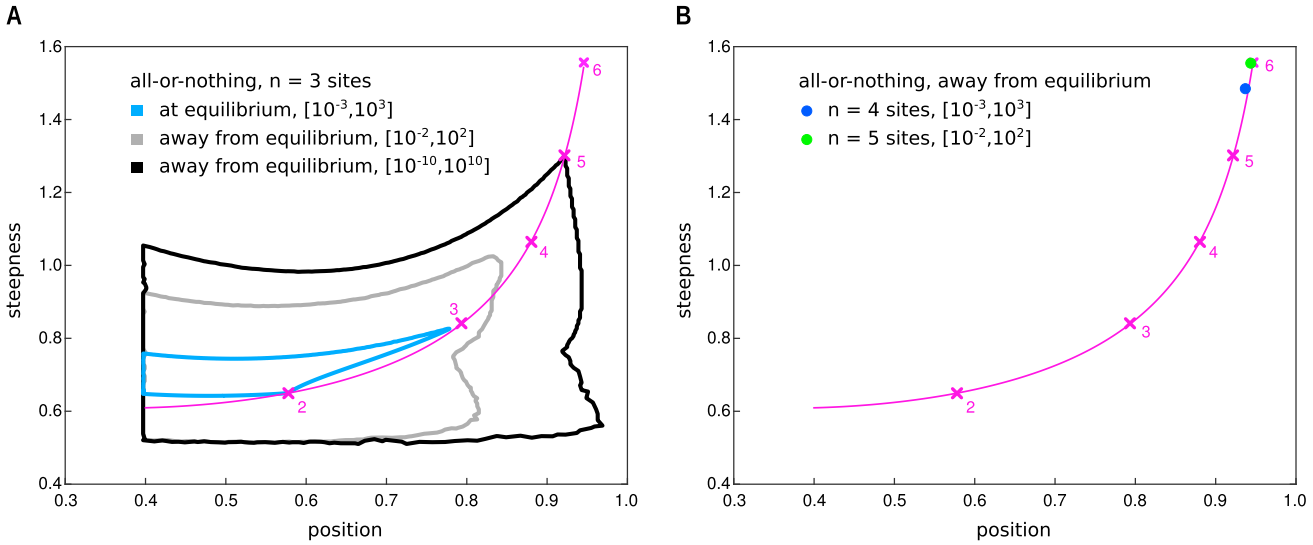


Figure S2. Non-equilibrium Position-Steepness Regions and GRFs for the All-or-Nothing Strategy, Related to Figure 6

(A) Position-steepness regions for $n = 3$ sites. The two inner boundaries, in blue and gray, are as shown in Paper Figure 6 for the all-or-nothing strategy. The outer black boundary gives the non-equilibrium region when the parameter range is pushed to $[10^{-10}, 10^{10}]$.

(B) Two non-equilibrium GRFs. The blue point shows a GRF on $n = 4$ sites with the parameter range $[10^{-3}, 10^{3}]$. This is close to $\mathcal{H}_{5.7}$. The green point shows a GRF on $n = 5$ sites with the parameter range $[10^{-2}, 10^{2}]$. This is close to $\mathcal{H}_{6}$.

Cell, Volume *166*

Supplemental Information

Information Integration and Energy Expenditure in Gene Regulation

Javier Estrada, Felix Wong, Angela DePace, and Jeremy Gunawardena

SUPPLEMENTAL INFORMATION

Contents

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Sections §1 to §7 which follow provide supplemental procedures which give full details of how the mathematical results in the Paper were derived. For the convenience of readers, the following is a list, in order of occurrence in the Paper, of each citation to the Supplemental Information (SI), with pointers to the relevant parts of this document.

- **"the linear framework allows steady-state probabilities to be calculated $\cdots$; see Experimental Procedures and Supplemental Information (SI)"**. See §1.
- **"This is not a fundamental limitation but arises from technical difficulties with the Principle of Detailed Balance, which imposes algebraic constraints on higher-order cooperativities (SI)"**. See §1.2.
- **"the relevant parameters are the association constants $K_{i,S}$ (Figure 2B), of which there are $n2^{n-1}$ (SI)"**. See §3.1 and Eq. 8.
- **Paper Equation 2** and **"We prove that, because of detailed balance, higher-order cooperativities must satisfy the "exchange formula" (SI)"**. See Eq. 13 in §3.4.
- **"all the $K_{i,S}$ can be calculated from them using Eq. 2 (SI)"**. See Eq. 9 and §3.2.
- **"we assume that T acts as a transcriptional activator (SI)"**. See §5.1.1.
- **"these three strategies broadly sample the spectrum of possibilities (SI)"**. See §5.1.1.
- **Paper Equation 4** and **"The level of protein expression after normalisation $\cdots$ has the form, for the all-or-nothing strategy (SI)"**. See Eq. 23 in §5.2.
- **"The coefficients c_k are given (SI) by a sum of products"**. See Eq. 22 in §5.2.
- **"For the other strategies (Figure 3A), only the numerator of Eq. 4 changes (SI)"**. See §5.1.1 and Eqs. 23 to 25 in the equilibrium case.
- **"The GRFs discussed in this paper are strictly increasing functions (SI)"**. See §5.5.
- **"The algebraic form of the Hill function, $\mathcal{H}_a(x) = x^a/(1+x^a)$, is closest to that of the GRF for the the all-or-nothing strategy in Eq. 4 (SI)"**. See §5.3.
- **"If $a = n$, algebraic resemblance is only possible, and then only approximately, if the parameters in the GRF are given implausible numerical values (SI)"**. See §5.3.
- **"Because of normalisation, γ and ρ do not depend on $K_{1,\emptyset}$ (SI)"**. See §6.2.
- **"We believe this range is generous but most of our results do not depend on it (below and SI)"**. See §7.1.
- **"We therefore used $\cdots$ a biased sampling algorithm to identify the boundary of the region (SI)"**. See §7.3.
- **"The non-equilibrium GRF for the all-or-nothing strategy then takes the form (SI)"**. See Eq. 26 in §5.4.
- **"GRFs for the other strategies differ only in the numerator (SI)"**. See §5.1.1 and Eqs. 26 to 28 for the non-equilibrium case.
- **"the largest number of sites that we can feasibly analyse is $n = 3$ (SI)"**. See §4.2.

- **"if the system were at equilibrium, it would be comparable with the previous equilibrium analysis (SI)"**. See §7.2.
- **"the steepness and position of g_n^{ne} are independent of the values of $a_{1,\emptyset}$ and $b_{1,\{1\}}$ (SI)"**. See §6.2.
- **"which leads to a dramatic increase in the complexity of the coefficients in Eq. 7 (SI; see Discussion)"**. See §4.2.
- **"This discrepancy arises from loss of detailed balance (SI)"**. See §4.3 and §5.4.
- **"The mathematical model and the the results presented here are based on the "linear framework" $\cdots$ see the SI for full details"**. See §1.
- **"The details are given in the SI, along with the tests"**. See §7.3 for the details of the algorithm and §7.4 for the tests.
- Paper Figure 4 caption, **"obtained by a biased sampling algorithm (SI)"**. See §7.3.
- Paper Figure 5 caption, **"The biased sampling algorithm had to be modified to find the average-binding regions (SI)"**. See Step 2 of the algorithm described in §7.3.

1 The linear framework

The linear framework is a graph-based method for time-scale separation in biochemical systems. It was introduced in previous work [2, 5] and reviewed in [3]. It was applied to gene regulation in [1], which provides the foundation on which the present Paper is based. As these ideas are relatively new, we provide here the background needed to understand the present Paper but [1, 5] should be consulted for more details and proofs.

1.1 Laplacian dynamics

The framework starts from a finite, connected, labelled, directed graph, G, as in Paper Figure 2A. The vertices are microstates, or "snapshots" of chromatin together with bound proteins, the edges are transitions between microstates and the edge labels are infinitesimal transition rates for a Markov process, with units of (time)$^{-1}$. Assume that the vertices are indexed $1, \cdots, N$. The stochastic master equation of the Markov process is given by the equation

$$\frac{du}{dt} = \mathcal{L}(G).u\,, \tag{1}$$

where, for $1 \leq i \leq N$, $u_i(t)$ is the probability of the system being in microstate i at time t and $\mathcal{L}(G)$ is the $N \times N$ Laplacian matrix of G. Eq. 1 is a linear differential equation, hence "linear framework". The nonlinearities in the biochemistry are absorbed into the edge labels. For any graph and any initial condition, the Laplacian dynamics in Eq. 1 always converges to a steady state $u^* \in \mathbb{R}^N$ where

$$\left.\frac{du}{dt}\right|_{u=u*} = 0\,,$$

so that $u^* \in \ker \mathcal{L}(G)$. An important basic result in the linear framework is that if G is strongly connected, then

$$\dim \ker \mathcal{L}(G) = 1\,.$$

It follows that, if ρ^G is any basis element of $\ker \mathcal{L}(G)$, then the steady-state probabilities, u^*, can be calculated by normalising ρ^G to its total,

$$u^* = \frac{\rho^G}{1.\rho^G}\,. \tag{2}$$

There are two basic methods for calculating a Laplacian steady state, ρ^G.

1.2 Steady states at thermodynamic equilibrium

If the system can reach thermodynamic equilibrium, then detailed balance must be satisfied. This requires, first, that each edge $i \overset{a}{\to} j$ is accompanied by a reverse edge, $j \overset{b}{\to} i$ and, second, in any cycle of reversible edges, the product of the labels going clockwise around the cycle equals the product going counterclockwise ("cycle condition"). These requirements guarantee that G is strongly connected and that, in any steady state, each pair of reversible edges is independently at steady state, irrespective of any other edges involving the vertices. If vertex 1 is chosen as a reference vertex and j is any other vertex, then there is always a path of reversible edges from 1 to j

$$1 = i_1 \underset{b_1}{\overset{a_1}{\rightleftharpoons}} i_2 \underset{b_2}{\overset{a_2}{\rightleftharpoons}} \cdots \underset{b_{p-1}}{\overset{a_{p-1}}{\rightleftharpoons}} i_p \underset{b_p}{\overset{a_p}{\rightleftharpoons}} i_{p+1} = j\,.$$

A basis element $\xi^G \in \ker \mathcal{L}(G)$ is then given, for each vertex j, by the corresponding product of label ratios

$$\xi^G_j = \left(\frac{a_p}{b_p}\right)\left(\frac{a_{p-1}}{b_{p-1}}\right)\cdots\left(\frac{a_2}{b_2}\right)\left(\frac{a_1}{b_1}\right)\,. \tag{3}$$

The cycle condition ensures that the quantity in Eq. 3 does not depend on the choice of path from 1 to j.

It can be seen from van't Hoff's formula that ξ_j^G is the Boltzmann factor of microstate j, as calculated in equilibrium statistical mechanics. The denominator of Eq. 2 is the partition function. The linear framework reduces to equilibrium statistical mechanics for systems which can reach thermodynamic equilibrium.

1.3 Independent generators at equilibrium

The cycle condition also implies that the edge labels are not independent quantities. Consider any cycle of reversible edges from vertex i to itself,

$$i = i_1 \underset{b_1}{\overset{a_1}{\rightleftharpoons}} i_2 \underset{b_2}{\overset{a_2}{\rightleftharpoons}} \cdots \underset{b_{p-1}}{\overset{a_{p-1}}{\rightleftharpoons}} i_p \underset{b_p}{\overset{a_p}{\rightleftharpoons}} i_{p+1} = i\,.$$

According to the cycle condition,

$$a_1 a_2 \cdots a_{p-1} a_p = b_p b_{p-1} \cdots b_2 b_1\,, \tag{4}$$

which implies an algebraic relationship between the labels. Eq. 3 shows that it is not the labels themselves which are important at equilibrium but the label ratios. If $i \overset{a}{\to} j$ and $j \overset{b}{\to} i$ are a pair of reversible edges, define $K_{i\to j} = a/b$, so that $K_{j\to i} = (K_{i\to j})^{-1}$. Eq. 4 can be rewritten in terms of these ratios as the algebraic relationship

$$\prod_{j=1}^{p-1} K_{i_j \to i_{j+1}} = 1\,, \tag{5}$$

The relations among the $K_{i\to j}$ arise in this way from cycles of reversible edges. For the applications below, it is important to know which sets of $K_{i\to j}$ constitute independent generators, so that, on the one hand, there are no algebraic relations between these generators ("independent"), while, on the other hand, all the other $K_{i\to j}$ can be "generated" as rational functions of the generators using Eq. 5. If T is any directed spanning tree of G, the label ratios over the edges of T, $\{K_{i\to j} \,|\, i \to j \in T\}$, constitute a set of independent generators. They are independent because there can be no cycle of reversible edges among edges which belong to a tree. Furthermore, if $i \to j \notin T$, then adding this edge to T cannot add new vertices to T because T is spanning. It must therefore create a cycle of reversible edges. If $u \to v$ is an edge on this cycle, which is not the same edge as $i \to j$, then either $u \to v$ is an edge in T or $v \to u$ is an edge in T. If the former, then $K_{u\to v}$ is a generator. If the latter, then $u \rightleftharpoons v \rightleftharpoons u$ is a cycle of reversible edges, so by Eq. 5, $K_{u\to v} = (K_{v\to u})^{-1}$ and $K_{u\to v}$ is the inverse of a generator. Note that this inverse property can be deduced from the cycle condition without having to know how the association constants are defined in terms of the labels. The $K_{i\to j}$ can now be written as a rational function of the generators using Eq. 5. Since G is connected, the number of edges in a spanning tree is one less than the number of vertices in G, so the number of independent generators is $N - 1$.

1.4 Steady states away from equilibrium

If the system does not reach thermodynamic equilibrium but the graph G is still strongly connected, then a basis element $\nu^G \in \ker \mathcal{L}(G)$ is given by the Matrix-Tree Theorem. Let $\Theta_j(G)$ be the set of spanning trees of G which are rooted at microstate j. Then, ν_j^G is given by taking $T \in \Theta_J(G)$, multiplying together the labels on the edges of T to form a monomial and adding up these monomials over all spanning trees in $\Theta_j(G)$,

$$\nu_j^G = \sum_{T \in \Theta_j(G)} \left(\prod_{k \overset{a}{\to} l \in T} a \right)\,. \tag{6}$$

The structure of ν^G is considerably more complicated than that of ξ^G. A spanning tree contains many paths that reach the root and there are usually many spanning trees rooted at any vertex. In particular, G is strongly

connected if, and only if, it has a rooted spanning tree at any vertex, so that $\nu_j^G > 0$ for $1 \leq j \leq N$. Moreover, each spanning tree has $N-1$ edges, so that the monomials which appear in ρ^G have degree $N-1$ in the edge labels. It can be shown that if G satisfies detailed balance then $\nu^G = \lambda \xi^G$, for an appropriate scalar factor λ, which is a polynomial in the edge labels (Wong, Gunawardena, unpublished results). Accordingly, in calculating steady-state probabilities using Eq. 2, identical results are obtained at equilibrium with either $\rho^G = \xi^G$ and Eq. 3 or $\rho^G = \nu^G$ and Eq. 6. The linear framework extends equilibrium statistical mechanics to systems that are far from equilibrium.

1.5 Gene regulation functions

The linear framework can treat transcription as a complex process with multiple irreversible steps [1] but the Paper follows the thermodynamic formalism and assumes that gene expression is a fast process compared to gene regulation, so that the overall rate of gene expression can be treated as an average over the steady-state probabilities of the microstates, as given by u^* in Eq. 2. Let $r(i)$ be the rate of gene expression in microstate i. The overall rate of mRNA synthesis is given by the average

$$\frac{d}{dt}[\text{mRNA}] = \sum_{1 \leq i \leq N} r(i) u_i^* = \left(\sum_{1 \leq i \leq N} r(i) \rho_i^G \right) \left(\frac{1}{1.\rho^G} \right) , \tag{7}$$

where we have used Eq. 2 in the second step. If the system is at equilibrium then Eq. 7 is calculated from $\rho^G = \xi^G$ using Eq. 3 and if it is away from equilibrium then Eq. 7 is calculated from $\rho^G = \nu^G$ using Eq. 6. In either case, the rate of mRNA synthesis can be expressed in terms of the edge labels, which may themselves contain the concentrations of transcription factors (TFs), as in the graph G_n in Paper Figure 2A.

Eq. 7 describes a gene regulation function for the rate of mRNA expression. The Paper uses a GRF for protein level. The additional assumptions required to obtain such a GRF are discussed in §5.1.

2 The graph G_n and background assumptions

We follow the notation introduced in the Paper. A single monomeric TF, T, is assumed to bind to n sites. The microstates are the $N = 2^n$ patterns of TF binding. It will be helpful to use the notation $\langle S \rangle$ for a microstate, where $S = \{i_1, \cdots, i_k\} \subseteq \{1, \cdots, n\}$, is the subset of sites bound by T. We will work with this notation rather than converting from $\langle S \rangle$ to an index $1, \cdots, N$, as in §1. The edges in G_n correspond to binding of T to a site $i \notin S$ with label $a_{i,S}[T]$ or, in reverse, the unbinding of T from $\langle S \cup \{i\} \rangle$ with label $b_{i,S \cup \{i\}}$. To make the text more readable, we will avoid repeating requirements like $i \notin S$ from now on but will use the notation $S \cup \{i\}$ where necessary, as a reminder.

As discussed in the Paper, molecular mechanisms other than binding and unbinding of T are not directly represented in G_n but can exert their influence through the rate constants and through the dependence of the rate constants on the microstates. Such indirect influences could be exerted, for instance, by other transcription factors, co-regulators, nucleosomes or multi-protein complexes like Mediator. The main assumption is that the total levels of these indirect factors do not change with changes in $[T]$ but the factors may participate in the dynamics which gives rise to the edges of G_n and, thereby, to the numerical values of the labels on the edges.

Since edge labels have units of (time)$^{-1}$, the rate constant $a_{i,S}$ has units of (concentration.time)$^{-1}$, while $b_{i,S}$ is a pure rate constant with units of (time)$^{-1}$. The factor $[T]$ brings the nonlinearity of binding into the edge label, thereby allowing the dynamics to be treated in a linear manner, as in Eq. 1. To avoid additional complexity, it is conventional, in both the thermodynamic formalism and the linear framework to make the "no depletion" assumption, so that the binding of T is presumed not to deplete the free concentration of T. Accordingly, $[T] \approx T_{tot}$ and the labels are time-independent constants in the Laplacian dynamics, provided the total concentration of the TF is constant.

In the non-equilibrium case, a gene regulatory system can be maintained away from thermodynamic equilibrium through continual dissipation of free energy by the background dissipative mechanisms discussed in the Paper. The transitions between microstates due to binding and unbinding of T can become effectively irreversible and the graph G_n can thereby lose some of its edges. For simplicity, we consider only the case where the structure of G_n remains the same, so that each edge, which may represent an irreversible transition, has a reverse edge, which also represents an irreversible transition. For instance, if background phosphorylation of a nucleosome is required for a binding reaction, then dephosphorylation is assumed to be available for the unbinding reaction; phosphorylation and dephosphorylation are each individually irreversible reactions. With that simplification, the non-equilibrium case can be based on the same labelled, directed graph G_n as the equilibrium case and differs only in that detailed balance is no longer satisfied.

The number of vertices in G_n is 2^n. The number of edges can be determined as follows. There are $\binom{n}{i}$ microstates of order $\#S = i$ and each of these gives rise to a binding edge to $n - i$ microstates of order $\#S + 1$. Hence, the total number of binding edges is

$$\sum_{i=0}^{n}(n-i)\binom{n}{i} = \sum_{i=0}^{n}(n-i)\frac{n!}{i!(n-i)!} = n\sum_{i=0}^{n-1}\binom{n-1}{i} = n2^{n-1}\,.$$

The unbinding edges are in one-to-one correspondence with the binding edges. Accordingly, the total number of edges is $n2^n$.

If labels are ignored and the reversible edges of G_n are coalesced into a single undirected edge, the resulting undirected graph is the well-studied hypercube graph [4].

3 Equilibrium parameters for G_n

3.1 Equilibrium parameters $K_{i,S}$ and $\omega_{i,S}$

At equilibrium, it follows from the discussion in §1.3 that it is not the individual labels which are relevant but only the label ratios. In the binding direction, these take the form

$$K_{\langle S\rangle\to\langle S\cup\{i\}\rangle} = \frac{a_{i,S}[T]}{b_{i,S\cup\{i\}}}\,.$$

The factor $[T]$ has a different role to that of the constants $a_{i,S}$ or $b_{i,S\cup\{i\}}$. It occurs in every label ratio and becomes the independent variable in the gene regulation functions defined in §5.2. It must also cancel out of the relationships that come from the cycle condition in Eq. 5 since, if it did not, this would imply that $[T]$ was determined by the rate constants at equilibrium, which is impossible. It therefore only plays the role of a scale factor and it is easier to work with ratios from which it has been removed, which leads to the association constants defined in Paper Figure 2B,

$$K_{i,S} = \frac{a_{i,S}}{b_{i,S\cup\{i\}}} = \frac{K_{\langle S\rangle\to\langle S\cup\{i\}\rangle}}{[T]}\,. \tag{8}$$

The association constants have units of (concentration)$^{-1}$ and, from the previous calculation, there are $n.2^{n-1}$ such parameters.

For numerical calculations, it is more convenient and meaningful to use non-dimensional parameters. It is also helpful to identify higher-order cooperativities, since these are biologically significant. Both goals are realised by defining the higher-order cooperativities, $\omega_{i,S}$, to be

$$\omega_{i,S} = \frac{K_{i,S}}{K_{i,\emptyset}}\,.$$

These non-dimensional parameters capture how the binding of T to site i depends on the higher-order influence of the sites in S, as compared to the "bare" affinity in the absence of anything bound. Note that $\omega_{i,\emptyset} = 1$. To non-dimensionalise the bare association constants $K_{i,\emptyset}$, we normalise them to $K_{1,\emptyset}$,

$$\kappa_i = \frac{K_{i,\emptyset}}{K_{1,\emptyset}} \,.$$

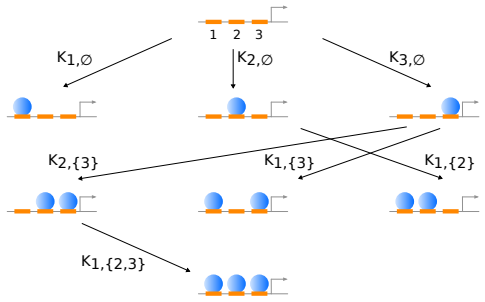
These are the non-dimensional parameters introduced in Paper Figure 2B, leaving $K_{1,\emptyset} = K_1$ as the only dimensioned parameter. The non-dimensional parameters are simply an alternative way of writing the $K_{i,S}$, which may be recovered from them and K_1 by noting that

$$K_{i,S} = \left(\frac{K_{i,S}}{K_{i,\emptyset}}\right)\left(\frac{K_{i,\emptyset}}{K_{1,\emptyset}}\right) K_{1,\emptyset} = \omega_{i,S}\kappa_i K_1 \,. \tag{9}$$

3.2 Independent generators for the $K_{i,S}$

For numerical calculations, it is necessary to identify a set of independent generators. We do this first for the $K_{i,S}$ and then deduce the results for their non-dimensional equivalents, κ_i and $\omega_{i,S}$, which are used in the Paper.

Following the method described in §1.3, consider the directed subgraph P_n of G_n obtained by retaining only those binding edges, $\langle S\rangle \to \langle S \cup \{i\}\rangle$, such that, if $j \in S$ then $i < j$. As in the Paper, this last condition is abbreviated to $i < S$. The diagram below shows the directed spanning tree P_3 for G_3, as described in §3.2. The edges are annotated with their corresponding association constants, which constitute a set of independent generators for the $K_{i,S}$. Note that this spanning tree is not rooted in the sense of Eq. 6.



Suppose that $\langle S\rangle$ is any microstate and that $S = \{i_1, \cdots, i_k\}$, where the indices in S are ordered so that $1 \le i_1 < \cdots < i_k \le n$. We can always find a path of reversible edges from the empty microstate to $\langle S\rangle$ of the form

$$\langle\emptyset\rangle \underset{b_{i_k,\{i_k\}}}{\overset{a_{i_k,\emptyset}[T]}{\rightleftharpoons}} \langle\{i_k\}\rangle \underset{b_{i_{k-1},\{i_{k-1},i_k\}}}{\overset{a_{i_{k-1},\{i_k\}}[T]}{\rightleftharpoons}} \langle\{i_{k-1},i_k\}\rangle \quad \cdots \quad \langle\{i_2,\cdots,i_k\}\rangle \underset{b_{i_1,S}}{\overset{a_{i_1,\{i_2,\cdots,i_k\}}[T]}{\rightleftharpoons}} \langle S\rangle \,. \tag{10}$$

The binding edges in this path all satisfy the ordering condition just introduced. Hence P_n contains all microstates and is, therefore, a spanning subgraph. Moreover, P_n has the property that each microstate is the target of exactly one edge: if the microstate is $\langle S\rangle$, where $S = \{i_1, \cdots, i_k\}$ with $1 \le i_1 < \cdots < i_k \le n$, then the unique edge corresponds to the association constant $K_{i_1,\{i_2,\cdots,i_k\}}$. Hence, P_n cannot have an undirected cycle. Accordingly, P_n is a directed spanning tree of G_n. It follows from §1.3, ignoring the factor $[T]$ for the reasons explained in §3.1, that the association constants $K_{i,S}$ with $i < S$, which we refer to as "ordered", constitute a set of independent generators for the $K_{i,S}$.

As a check on this, the number of pairs (i, S) with $i < S \subseteq \{1, \cdots, n\}$ can be counted as follows. For any given $1 \le i \le n$, there are 2^{n-i} subsets S satisfying the constraint $i < S$, so that the number of generators is

$$2^{n-1} + \cdots + 2 + 1 = 2^n - 1 \,,$$

as expected from the discussion in §1.3 since $N = 2^n$. The number of ordered association constants $(2^n - 1)$ becomes substantially smaller than the total number of association constants $(n2^{n-1})$ as soon as $n \geq 3$.

3.3 The exchange formula for $K_{i,S}$

The discussion in §1.3 provides a method for expressing any $K_{i,S}$ in terms of the ordered ones. This is not straightforward to apply in practice as it requires locating where the edge corresponding to $K_{i,S}$ occurs in relation to P_n. The following alternative method provides a formula that is independently useful. Suppose that $i, j \notin S \subseteq \{1, \cdots, n\}$. There is then a cycle of reversible edges from $\langle S \rangle$ to itself,

$$\langle S \rangle \underset{b_{i,S\cup\{i\}}}{\overset{a_{i,S}[T]}{\rightleftharpoons}} \langle S \cup \{i\} \rangle \underset{b_{j,S\cup\{i,j\}}}{\overset{a_{j,S\cup\{i\}}[T]}{\rightleftharpoons}} \langle S \cup \{i,j\} \rangle \underset{a_{i,S\cup\{j\}}[T]}{\overset{b_{i,S\cup\{i,j\}}}{\rightleftharpoons}} \langle S \cup \{j\} \rangle \underset{a_{j,S}[T]}{\overset{b_{j,S\cup\{j\}}}{\rightleftharpoons}} \langle S \rangle .$$

Applying Eq. 5 to this and cancelling $[T]$ (§3.1), leads to the exchange formula for association constants

$$K_{i,S}K_{j,S\cup\{i\}} = K_{j,S}K_{i,S\cup\{j\}} . \tag{11}$$

Eq. 11 is sufficient to calculate any $K_{i,S}$ in terms of the ordered ones. The proof of this is by induction on $\#S$. If $\#S = 0$ then $i < S$ for all $1 \leq i \leq n$, so the parameters $K_{i,\emptyset}$ are all ordered and there is nothing to prove. Now suppose that $\#S = k$, where $k > 0$, and that all higher-order cooperativities of the form $\omega_{i,U}$ with $\#U < k$ have been expressed in terms of the ordered parameters. Consider the k-th order cooperativity $\omega_{i,S}$ for any $1 \leq i \leq n$. Since $k > 0$, $S = \{i_1, \cdots, i_k\}$ with $1 \leq i_1 < \cdots < i_k \leq n$, then i could occur anywhere along this sequence. The exchange formula can be used to move i by a succession of swaps to before the left-hand end of the sequence, at which point the association constant becomes one of the ordered ones. This requires a second induction.

Let $\mu(i, S)$ denote the number of elements of S which precede i in the sequence:

$$\mu(i, S) = \#\{i_j \in S \mid i_j < i\} .$$

If $\mu(i, S) = 0$, then $\omega_{i,S}$ is one of the ordered parameters so there is nothing to prove. Suppose, as a second inductive hypothesis, that $\omega_{v,U}$ has been expressed in terms of the ordered parameters for all $\#U = k$ and for all $\mu(v, U) < p$. Suppose that $\mu(i, S) = p$, so that

$$i_1 < \cdots < i_p < i < i_{p+1} < \cdots < i_k .$$

Let $V = \{i_1, \cdots, i_{p-1}, i_{p+1}, \cdots, i_k\}$, so that $\#V = k - 1$ and $S = V \cup \{i_p\}$. It follows from the exchange formula in Eq. 11 that

$$K_{i,S} = \left(\frac{K_{i,V}}{K_{i_p,V}} \right) K_{i_p,V\cup\{i\}} .$$

Of the terms on the right, $K_{i,V}$ and $K_{i_p,V}$ have $\#V < k$ so that, by the first inductive hypothesis, they have been expressed in terms of the ordered parameters. Furthermore, $K_{i_p,V\cup\{i\}}$ has $\#(V\cup\{i\}) = k$ with $\mu(i_p, V\cup\{i\}) < p$ so that, by the second inductive hypothesis, it has also been expressed in terms of the ordered parameters. It follows that $K_{i,S}$ can be expressed in terms of the ordered parameters, thereby completing both inductions. The proof offers an algorithm for expressing $K_{i,S}$ in terms of the ordered association constants.

3.4 Independent non-dimensional κ_i and $\omega_{i,S}$

Paper Equation 3 introduces the following set of $2^n - 1$ parameters

$$K_1 , \quad \kappa_i \, (i > 1) , \quad \omega_{i,S} \, (i < S \neq \emptyset) , \tag{12}$$

in which only K_1 is dimensioned and where the higher-order cooperativities $\omega_{i,S}$ are restricted to those which are ordered, as above for the $K_{i,S}$. It follows from Eq. 9 that the parameters in Eq. 12 are merely a rewriting of the

ordered $K_{i,S}$ and are therefore also a set of independent generators from which any $K_{i,S}$ can be recovered. The equilibrium GRFs described in §5.2 will be expressed in terms of the parameters in Eq. 12.

For calculations with higher-order cooperativities, a non-dimensional version of Eq. 11 can be obtained by dividing through by $K_{i,\emptyset}K_{j,\emptyset}$ to give

$$\omega_{i,S}\omega_{j,S\cup\{i\}} = \omega_{j,S}\omega_{i,S\cup\{j\}}\,, \tag{13}$$

Eq. 13 is the exchange formula for higher-order cooperativities given in **Paper Equation 2**. Eq. 13 provides an effective method for calculating any higher-order cooperativity in terms of the ordered ones by following the same algorithm as described above for the $K_{i,S}$.

4 Laplacian steady states of G_n

4.1 Steady states of G_n at equilibrium

At equilibrium, $\rho^{G_n} = \xi^{G_n}$ and the prescription in §1.2 can be used to calculate the latter. Eq. 10 provides a path of reversible edges from $\langle\emptyset\rangle$ to $\langle S\rangle$, where $S = \{i_1, \cdots, i_k\}$ with the indices ordered so that $1 \leq i_1 < \cdots < i_k \leq n$. It follows from Eq. 3 that, at thermodynamic equilibrium, the Laplacian steady state of $\langle S\rangle$ is given by

$$\xi^{G_n}_{\langle S\rangle} = K_{i_k,\emptyset}K_{i_{k-1},\{i_k\}}\cdots K_{i_1,\{i_2,\cdots,i_k\}}[T]^k\,. \tag{14}$$

Eq. 14 can be non-dimensionalised using Eq. 9 to yield

$$\xi^{G_n}_{\langle S\rangle} = \kappa_{i_k}(\kappa_{i_{k-1}}\omega_{i_{k-1},\{i_k\}})\cdots(\kappa_{i_1}\omega_{i_1,\{i_2,\cdots,i_k\}})(K_1[T])^k\,.$$

Noting that $\omega_{i_k,\emptyset} = 1$, this may be rewritten more concisely as

$$\xi^{G_n}_{\langle S\rangle} = \left(\prod_{j=1}^{k}\kappa_{i_j}\omega_{i_j,\{i_{j+1},\cdots,i_k\}}\right)(K_1[T])^k\,, \tag{15}$$

which is expressed in the independent parameters of Paper Equation 3 (Eq. 12). As an illustration of Eq. 15, the Laplacian steady states of the 8 microstates of G_3 in Paper Figure 2A are listed in the table below. Note that K_1 in Eq. 15 plays the role of a scale factor (§6.2).

microstate	**Laplacian steady state**
$\langle\emptyset\rangle$	1
$\langle\{1\}\rangle$	$K_1[T]$
$\langle\{2\}\rangle$	$\kappa_2 K_1[T]$
$\langle\{3\}\rangle$	$\kappa_3 K_1[T]$
$\langle\{1,2\}\rangle$	$\kappa_2\kappa_1\omega_{1,\{2\}}(K_1[T])^2$
$\langle\{1,3\}\rangle$	$\kappa_3\kappa_1\omega_{1,\{3\}}(K_1[T])^2$
$\langle\{2,3\}\rangle$	$\kappa_3\kappa_2\omega_{2,\{3\}}(K_1[T])^2$
$\langle\{1,2,3\}\rangle$	$\kappa_3\kappa_2\kappa_1\omega_{2,\{3\}}\omega_{1,\{2,3\}}(K_1[T])^3$

4.2 History dependence away from equilibrium

At thermodynamic equilibrium, the Laplacian steady state can be calculated from a single path to each microstate, as in Eq. 14. All paths give the same answer because of detailed balance, which guarantees that the steady state does not depend on the history through which that state was reached. Equilibrium is fundamentally reversible. In marked contrast, away from thermodynamic equilibrium, Eq. 6 shows that all paths to a microstate must be

used to calculate the Laplacian steady state. The rooted spanning trees of the Matrix-Tree Theorem provide the bookkeeping for this calculation. Non-equilibrium systems are history dependent [1].

The number of spanning trees increases worse than exponentially in the size of a graph and the resulting combinatorial explosion is one of the fundamental challenges in analysing systems away from equilibrium [1]. In the case of G_n, the number of spanning trees rooted at any vertex (the number is the same for any vertex because G_n is reversible), is equal to the number of spanning trees in the undirected hypercube graph (§2), which is known to be [4],

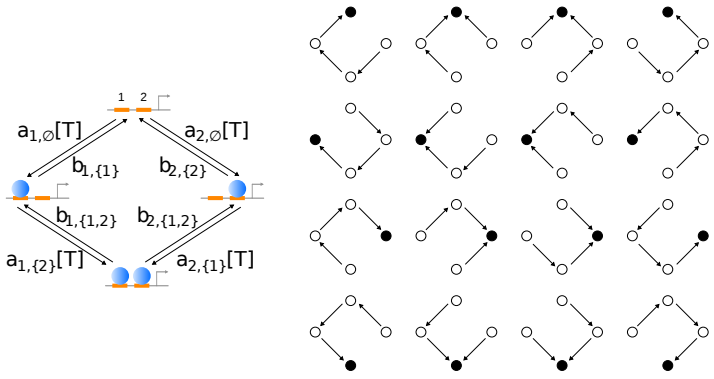
$$2^{2^n-n-1}\prod_{k=2}^{n} k^{\binom{n}{k}}\,.$$

For n = 2, this formula gives 4 spanning trees rooted at each vertex but for n = 3, it gives 384. The increase with n is extraordinarily rapid: G_4 has no less than 42,467,328 spanning trees rooted at each vertex. Accordingly, the calculation of position-steepness regions had to be restricted to G_3, the complexity of whose GRF is already substantial.

It is important to appreciate that the size of a graph is not the dominant factor in the complexity of the Laplacian steady state. If a graph is forced to satisfy detailed balance, by making every edge reversible (which only increases the complexity of the graph) and imposing the cycle condition on the edge labels, then its Laplacian steady state is given by Eq. 22, which can be readily calculated even for very large graphs. Complexity arises primarily from non-equilibrium history-dependence.

4.3 Steady states of G_n away from equilibrium

Away from equilibrium, $\rho^{G_n} = \nu^{G_n}$, with the attendant history-dependent complexity described in §4.2. The algebraic structure also becomes different to that at equilibrium. The diagram below shows the graph G_2 on the left, annotated with its edge labels, and shows on the right the rooted spanning trees, with each row giving the spanning trees rooted at the vertex in black. Each tree has 3 edges, as expected for a graph with $2^n - 1$ vertices for $n = 2$ (§1.4). Each vertex has the same number of rooted spanning trees, a property that holds for any labelled, directed graph in which each edge is accompanied by a reverse edge.



Several properties can be inferred from this diagram which hold for general n. According to Eq. 6, the spanning trees rooted at a vertex give rise to monomials of the form

$$a_{i_1,S_1}\cdots a_{i_p,S_p} b_{j_1,S_{p+1}\cup\{j_1\}}\cdots b_{j_q,T_{p+q}\cup\{j_q\}}[T]^p\,, \tag{16}$$

where p is the number of binding edges and q is the number of unbinding edges and $p+q = 2^n-1$. Notice that the power of $x = [T]$ can vary depending on the spanning tree, so that the Laplacian steady-state $\nu^{G_n}_{\langle S\rangle}$ is a polynomial in x, in contrast to a monomial at equilibrium.

It is not difficult to show that $\nu^{G_n}_{\langle S\rangle}$ has the following structure as a polynomial in x,

$$\nu^{G_n}_{\langle S\rangle} = a_1 x^{q_1} + \cdots + a_k x^{q_k}\,, \tag{17}$$

where $q_1 < \cdots < q_k$ and the minimum and maximum degrees are given by

$$q_1 = \#S \quad \text{and} \quad q_k = 2^n - (n+1) + \#S\,.$$

Note, in particular, that $\nu^{G_n}_{\emptyset}$ has a constant term which, unlike the equilibrium case, is different from zero, while $\nu^{G_n}_{\langle S\rangle}$ has the highest degree of $2^n - 1$ only when $S = \{1, \cdots, n\}$.

4.3.1 Parameterisation away from equilibrium

The monomial structure in Eq. 16 makes it clear that the ratios of rate constants are no longer the appropriate parameters for the steady states away from equilibrium. However, non-dimensionalisation is still helpful for numerical calculation. We chose the non-dimensional parameters

$$\epsilon_{i,S} = \frac{a_{i,S}}{a_{1,\emptyset}} \quad \text{and} \quad \delta_{i,S\cup\{i\}} = \frac{b_{i,S\cup\{i\}}}{b_{1,\{1\}}}\,, \tag{18}$$

so that $a_{1,\emptyset}$ and $b_{1,\{1\}}$ are the only two dimensioned parameters. It follows from Eq. 16 that the monomial arising from a rooted spanning tree can be rewritten

$$\epsilon_{i_1,S_1}\cdots\epsilon_{i_p,S_p}\delta_{j_1,T_1\cup\{j_1\}}\cdots\delta_{j_q,T_q\cup\{j_q\}}(a_{1,\emptyset})^p(b_{1,\{1\}})^q[T]^p\,,$$

where $p + q = 2^n - 1$. We can now divide by the same factor $(b_{1,\{1\}})^{2^n-1}$ in each monomial, to rewrite it as

$$\epsilon_{i_1,S_1}\cdots\epsilon_{i_p,S_p}\delta_{j_1,T_1\cup\{j_1\}}\cdots\delta_{j_q,T_q\cup\{j_q\}}\left(\frac{a_{1,\emptyset}}{b_{1,\{1\}}}[T]\right)^p\,. \tag{19}$$

This division does not change any calculated GRF because the factor $(b_{1,\{1\}})^{2^n-1}$ cancels between the numerator and the denominator. The quantities $a_{1,\emptyset}$ and $b_{1,\{1\}}$ act as scale factors (§6.2).

5 Gene regulation functions for G_n

5.1 The normalised GRF for protein level

The discussion the follows applies to both the equilibrium and non-equilibrium contexts. Eq. 7 shows how the rate of mRNA synthesis can be calculated in terms of the edge labels of G_n and, thereby, in terms of $[T]$. Call this function $s([T])$. Paper Equation 1 describes how Hb protein concentration level, [Hb], depends on $[T]$, after normalisation. To go from mRNA synthesis rate to protein level, we implicitly assume that post-transcriptional mechanisms do not contribute to sharpness. Explicitly, we assume, first, that mRNA is linearly degraded at rate δ, so that the overall rate of mRNA production is given by

$$\frac{d}{dt}[\text{mRNA}] = s([T]) - \delta[mRNA]\,.$$

Hence, the steady-state concentration of mRNA, $[\text{mRNA}]^*$, is

$$[\text{mRNA}]^* = \frac{1}{\delta}s([T])\,.$$

Second, we assume that the steady-state protein level, [protein]*, is proportional to the steady-state level of mRNA, so that

$$[\text{protein}]^* = \frac{\alpha}{\delta} s([T])\,.$$

It follows that the steady-state protein level is proportional to $s([T])$. The gene regulation function for protein can now be calculated using Eq. 7,

$$[\text{protein}]^* = \left(\frac{\alpha}{\delta}\right)\left(\sum_{S\subseteq\{1,\cdots,n\}} r(S)\rho^{G_n}_{\langle S\rangle}\right)\left(\frac{1}{1.\rho^{G_n}}\right)\,, \tag{20}$$

where, to avoid excessive notation, we use the abbreviation $r(S)$ to denote the rate of gene expression in microstate $\langle S\rangle$.

5.1.1 T acts only as a transcriptional activator

In the Paper, it is assumed that T acts only as a transcriptional activator. This means that the binding of T to a microstate cannot decrease its rate of expression: if $\langle U\rangle$ and $\langle V\rangle$ are microstates and V includes all the sites of S, so that $U \subseteq V$, then the rate of expression in $\langle V\rangle$ cannot be smaller than that in $\langle U\rangle$, so that $r(U) \leq r(V)$. Note that this the case for each of the expression strategies in Paper Figure 3A. It follows that the maximal rate of expression is $r(\{1,\cdots,n\})$. This maximal rate is taken to be 1 and $r(\emptyset) = 0$, so that there is no baseline expression. Hence, $0 \leq r(S) \leq 1$. The maximum steady-state level of protein expression is then given in Eq. 20 by α/δ. For the gene expression strategies described in Paper Figure 3A, this maximum value is attained asymptotically as $[T] \to \infty$ (§5.2). If we normalise the steady-state protein concentration to this asymptotic maximum, the normalised gene regulation function, $f_n(x)$, is given by

$$f_n(x) = \left(\sum_{S\subseteq\{1,\cdots,n\}} r(S)\rho^{G_n}_{\langle S\rangle}\right)\left(\frac{1}{1.\rho^{G_n}}\right)\,. \tag{21}$$

Here, $x = [T]$ and $[T]$ appears through the expressions for ρ^{G_n} either at equilibrium using Eq. 3, as discussed in §5.2, or away from equilibrium using Eq. 6, as discussed in §5.4.

Because the components of both the equilibrium ρ^{G_n} in Eq. 15 and the non-equilibrium ρ^{G_n} in Eq. 17 are polynomials in $x = [T]$, $f_n(x)$ is a rational function of x (ie: a fraction in which the numerator and denominator are each polynomials in x). Furthermore, the denominator in Eq. 21, $1.\rho^{G_n}$, does not depend on the expression rate $r(S)$. Hence, the denominator of $f_n(x)$ is the same for different expression strategies at equilibrium and away from equilibrium.

Very little is known about transcription rates in individual microstates, making it difficult to assign values to $r(S)$ in Eq. 21, even after assuming that T is a transcriptional activator. Numerical exploration is challenging because, for each set of choices, $\{0 \leq r(S) \leq 1 \,|\, S \subseteq \{1,\cdots,n\}\}$, a position-steepness region similar to Paper Figure 4 would have to be calculated by randomly varying the parameters. Instead, we chose three expression strategies (Paper Figure 3A)—two extreme (all-or-nothing, one-or-more) and one middle-of-the-road (average-binding)—which sample the range of possibilities. Our expectation is that the position-steepness regions for other sets of values of $r(S)$ would resemble that of average-binding (Paper Figure 4D) in lying on both sides of the Hill line but would approach $\mathcal{H}_n$ at the tip of a cusp, as is the case for each of the three strategies analysed in the Paper.

5.2 Equilibrium gene regulation functions

At equilibrium, $\rho^{G_n} = \xi^{G_n}$, and the denominator, $1.\xi^{G_n}$, in the expression for $f_n(x)$ in Eq. 21 is the partition function of equilibrium statistical mechanics. The partition function is the same for all equilibrium GRFs, which

differ only in the numerator of the expression. According to Eq. 15, a microstate S with $\#S = k$ gives rise to a monomial in x^k. Hence, the partition function is given by the polynomial

$$1.\xi^{G_n} = 1 + c_1 x + \cdots + c_n x^n \,,$$

where the coefficient c_k of x^k collects together, using Eq. 15, all microstates of order k,

$$c_k = \left(\sum_{1 \le i_1 < \cdots < i_k \le n} \left(\prod_{j=1}^{k} \kappa_{i_j} \omega_{i_j, \{i_{j+1}, \cdots, i_k\}} \right) \right) (K_1)^k \,. \tag{22}$$

This gives **Paper Equation 5**. The three gene expression strategies described in Paper Figure 3A then have the following GRFs, where c_k is defined in Eq. 22,

all-or-nothing $r(S) = 1$ if $S = \{1, \cdots, n\}$ and $r(S) = 0$ otherwise. This gives **Paper Equation 4**,

$$f_n(x) = \frac{c_n x^n}{1 + c_1 x + \cdots + c_n x^n} \,. \tag{23}$$

one-or-more $r(S) = 1$ if $S \neq \emptyset$ and $r(\emptyset) = 0$. Hence,

$$f_n(x) = \frac{\sum_{k=1}^{n} c_k x^k}{1 + c_1 x + \cdots + c_n x^n} \,. \tag{24}$$

average-binding $r(S) = (\#S)/n$. Hence,

$$f_n(x) = \frac{(\sum_{k=1}^{n} k c_k x^k)/n}{1 + c_1 x + \cdots + c_n x^n} \,. \tag{25}$$

5.3 Similarity of $f_n(x)$ in the all-or-nothing strategy to $\mathcal{H}_n(x)$

The Hill function $\mathcal{H}_a(x) = x^a/(1 + x^a)$ has a numerator, x^a, which is identical to the highest-degree monomial in the denominator. Its algebraic form therefore corresponds most closely to a GRF in the all-or-nothing strategy (Eq. 23), at least when the Hill coefficient a is an integer. In this case, we can ask if a GRF $f_n(x)$ can be found which coincides in algebraic form with $\mathcal{H}_a(x)$. For obvious reasons, it is necessary that $a = n$. It is also clear from Eq. 23 that $c_1, \cdots, c_{n-1}$ must be made 0, while c_n must remain at 1. This cannot be done exactly because it follows from Eq. 22 that if $c_1 = 0$, then $K_1 = 0$, so that $c_i = 0$ for $1 \le i \le n$. However, it can be done approximately. If we set $\kappa_i = 1$ for $i > 1$, so that $K_{i,\emptyset} = K_1$, and $\omega_{i,S} = 1$ for $S \neq \{2, \cdots, n\}$, then it follows from Eq. 22 that

$$c_k = \begin{cases} \binom{n}{k} (K_1)^k & \text{if } k < n \\ \omega_{1,\{2,\cdots,n\}} (K_1)^n & \text{if } k = n \,. \end{cases}$$

Hence, if we choose K_1 in units of (concentration)$^{-1}$ to be very small compared to $\binom{n}{\lceil n/2 \rceil}^{-1}$, then, for $1 \le k \le n$,

$$(K_1)^k < K_1 \ll \frac{1}{\binom{n}{\lceil n/2 \rceil}} \le \frac{1}{\binom{n}{k}} \le 1 \,,$$

so that c_k remains very small compared to 1 for $k < n$. However, we must then set

$$\omega_{1,\{2,\cdots,n\}} = \frac{1}{(K_1)^n} \,,$$

to achieve $c_n = 1$, which means that $\omega_{1,\{2,\cdots,n\}} \gg 1$. The combination of the bare association constants, $K_{i,\emptyset}$, being very small and the highest-order cooperativity, $\omega_{1,\{2,\cdots,n\}}$, being very large is biochemically implausible. A similar difficulty is encountered no matter what scheme is used to make c_k small for $k < n$. The Hill function is, therefore, not similar in algebraic form to a GRF for realistic parameter values.

5.4 Non-equilibrium gene regulation functions

Away from equilibrium, $\rho^{G_n} = \nu^{G_n}$. As in the equilibrium case, the denominator in Eq. 21 is the same in all strategies. Using Eq. 17, it can be seen to take the form,

$$1.\nu^{G_n} = e_0 + e_1 x + \cdots + e_{2^n-1} x^{2^n-1} \, .$$

Note the exponential increase in the degree from the partition function at equilibrium. Also, unlike the equilibrium case, $e_0 \neq 0$, as is seen from the table below, which gives the non-equilibrium steady states for the graph G_2.

microstate	**Laplacian steady state**
$\langle\emptyset\rangle$	$b_{1,\{1,2\}}b_{1,\{1\}}b_{2,\{2\}} + b_{1,\{1\}}b_{2,\{2\}}b_{2,\{1,2\}} + (a_{2,\{1\}}b_{1,\{1,2\}}b_{1,\{1\}} + b_{1,\{1\}}b_{2,\{2\}}a_{1,\{2\}})[T]$
$\langle\{1\}\rangle$	$(b_{1,\{1,2\}}b_{2,\{2\}}a_{1,\emptyset} + b_{2,\{1,2\}}b_{2,\{2\}}a_{1,\emptyset})[T] + (b_{1,\{1,2\}}a_{2,\{1\}}a_{2,\emptyset} + b_{1,\{1,2\}}b_{2,\{2\}}a_{1,\emptyset})[T]^2$
$\langle\{2\}\rangle$	$(b_{1,\{1,2\}}b_{1,\{1\}}a_{2,\emptyset} + b_{2,\{1,2\}}b_{1,\{1\}}a_{2,\emptyset})[T] + (b_{2,\{1,2\}}a_{1,\{2\}}a_{2,\emptyset} + b_{2,\{1,2\}}a_{1,\{2\}}a_{1,\emptyset})[T]^2$
$\langle\{1,2\}\rangle$	$(b_{2,\{2\}}a_{1,\{2\}}a_{1,\emptyset} + b_{1,\{1\}}a_{2,\{1\}}a_{1,\emptyset})[T]^2 + (a_{2,\{1\}}a_{1,\{2\}}a_{1,\emptyset} + a_{2,\{1\}}a_{1,\{2\}}a_{2,\emptyset})[T]^3$

It is difficult to give an expression for the coefficients e_i similar to the equilibrium one in Eq. 22 because each coefficient has contributions from microstates of different orders.

all-or-nothing $r(S) = 1$ if $S = \{1, \cdots, n\}$ and $r(S) = 0$ otherwise. Using the information about the structure of $\nu^{G_n}_{\langle\{1,\cdots,n\}\rangle}$ in Eq. 17, it can be seen that the the GRF has the form given in **Paper Equation 7**,

$$f_n^{ne}(x) = \frac{d_n x^n + \cdots + d_{2^n-1} x^{2^n-1}}{e_0 + e_1 x + \cdots + e_{2^n-1} x^{2^n-1}} \, , \tag{26}$$

where $d_{2^n-1} = e_{2^n-1}$.

one-or-more $r(S) = 1$ if $S \neq \emptyset$ and $r(\emptyset) = 0$. Hence, using Eq. 17 again,

$$f_n^{ne}(x) = \frac{g_1 x + \cdots + g_{2^n-1} x^{2^n-1}}{e_0 + e_1 x + \cdots + e_{2^n-1} x^{2^n-1}} \, , \tag{27}$$

where, again, $g_{2^n-1} = e_{2^n-1}$.

average-binding $r(S) = (\#S)/n$. Hence, using Eq. 17 again,

$$f_n^{ne}(x) = \frac{h_1 x + \cdots + h_{2^n-1} x^{2^n-1}}{e_0 + e_1 x + \cdots + e_{2^n-1} x^{2^n-1}} \, , \tag{28}$$

where, again, $h_{2^n-1} = e_{2^n-1}$. Note that $h_i \neq g_i$.

For all these cases, history-dependent complexity and the overlapping contributions of different microstates make it hard to calculate the coefficients, d_i, g_i and h_i, in the numerator polynomials. However, for all expression strategies, the GRF $f_3^{ne}(x)$, which is the case discussed in the Paper, can be calculated using the software described in [1].

5.5 GRFs are increasing but not always sigmoidal

Gene regulation functions for a transcriptional activator T are expected to be increasing functions of $x = [T]$. More precisely, under the assumptions that we have made, we expect the following.

$$f_n(0) = 0 \, , \quad f_n(x) \to 1 \ \text{ as } \ x \to \infty \, , \quad f_n(x) < f_n(y) \ \text{ if } \ 0 \leq x < y \, .$$

The first two parts of this are readily proved for all 6 GRFs (ie: for the all-or-nothing, one-or-more and average-binding strategies at equilibrium and away from equilibrium). The third part, that $f_n(x)$ is strictly increasing, is also readily proved for equilibrium GRFs by differentiating the expressions in Eqs. 23 to 25. However, strict increase is not so readily proved for the non-equilibrium GRFs in Eqs. 26 to 28. We therefore tested it numerically as follows. If f_n ceases to be increasing in $(0, \infty)$, then df_n/dx has a positive zero. This condition can be checked by considering only the numerator polynomial of df_n/dx. For each position-steepness plot, once a final converged boundary had been numerically found (§7.3), the corresponding parameters were passed to Mathematica, the coefficients of the relevant GRF were converted to numbers and the numerator polynomial of df_n/dx was calculated. Mathematica's native *CountRoots* function was used, with the polynomial variable specified to lie in the interval $(0, \infty)$, to determine the number of positive roots of the numerator polynomial. We found this to be zero in all cases, confirming that the non-equilibrium GRFs were all strictly increasing. It would be interesting from a mathematical point of view to find a rigorous proof of strict increase for any GRF arising from a transcriptional activator.

GRFs for a transcriptional activator are often assumed to be sigmoidal ("S-shaped"), by which is meant that df_n/dx increases to a maximum and then decreases steadily towards zero: $df_n/dx \to 0$ as $x \to \infty$ and df_n/dx has a single maximum for $x \in [0, \infty)$. The Hill functions exhibit this behaviour. Sigmoidality can be numerically checked by showing that df_n^2/dx^2 has a single zero for $x \in [0, \infty)$, which we did in a similar way to the previous paragraph. We found that GRFs for the equilibrium all-or-nothing strategy were always sigmoidal but that equilibrium GRFs for the one-or-more and average-binding strategies and non-equilibrium GRFs for any strategy were not always sigmoidal. Paper Figure 3B shows a typical GRF which is increasing but not sigmoidal.

6 Properties of the shape measures γ and ρ

6.1 γ and ρ for the Hill function $\mathcal{H}_a$

The position, γ, and the steepness, ρ, of a GRF are defined in Paper Equation 6 and these definitions apply equally well to the Hill functions $\mathcal{H}_a(x)$. Note that the Hill coefficient can be any real number $a > 0$. A straightforward calculation shows that

$$\frac{d\mathcal{H}_a}{dx} = \frac{ax^{a-1}}{(1+x^a)^2}\,, \qquad \frac{d^2\mathcal{H}_a}{dx^2} = \frac{ax^{a-2}(a-1-(a+1)x^a)}{(1+x^a)^3}\,.$$

It follows that,

$$\gamma(\mathcal{H}_a) = \left(\frac{a-1}{a+1}\right)^{1/a} \quad \text{and} \quad \rho(\mathcal{H}_a) = \frac{a^2-1}{4a\gamma(\mathcal{H}_a)}\,.$$

6.2 γ and ρ do not depend on the scale factor

Recall from the Paper that the normalised GRF, $g_n(y)$, is defined by $g_n(y) = f_n(y.x_{0.5})$ where $y = x/x_{0.5}$ and $f_n(x_{0.5}) = 0.5$. It follows that

$$\frac{dg_n}{dy} = \frac{df_n}{dx}\frac{dx}{dy} = x_{0.5}\frac{df_n}{dx}\,. \tag{29}$$

Suppose that we scale the x value so that we replace x by cx in f_n. Let us call the resulting scaled GRF $f_n^c(x) = f_n(cx)$ and use superscript c in a similar way to denote other entities obtained from f_n^c. If follows that $0.5 = f_n^c(x_{0.5}^c) = f_n(cx_{0.5}^c)$, so that $x_{0.5}^c = x_{0.5}/c$. Then, using Eq. 29 twice,

$$\frac{dg_n^c}{dy} = x_{0.5}^c\frac{df_n^c}{dx} = \left(\frac{x_{0.5}}{c}\right)\left(\frac{df_n(cx)}{dx}\right) = x_{0.5}\frac{df_n}{dx} = \frac{dg_n}{dy}\,.$$

Hence, both the maximum derivative of dg_n/dy, which is the steepness $\rho(g_n)$, and its position, which is $\gamma(g_n)$, do not depend on the scale factor c.

This argument applies equally well to the scale factor K_1 at equilibrium (Eq. 15) and to the scale factors $a_{1,\emptyset}$ and $b_{1,\{1\}}$ away from equilibrium (Eq. 19).

7 Parameter ranges and search algorithms

7.1 Equilibrium parameter ranges

For numerical calculations at equilibrium, we chose non-dimensional parameters $\kappa_i, \omega_{i,S} \in [10^{-a}, 10^a]$ by logarithmic sampling. The default choice of $a = 3$ which was used for many of the regions shown in the Paper allows the bare association constants $K_{i,\emptyset}$ to vary by a factor of 10^3 and the higher-order association constants $K_{i,S}$ to vary by a factor 10^6, both in comparison to K_1. We felt this was sufficiently broad to reveal most properties while giving good numerical stability and computation time. We explored the parameter range as far as $a = 5$ for the all-or-nothing strategy at equilibrium but found no new features (Paper Figure 4C and Figure S1A).

7.2 Non-equilibrium parameter ranges

For numerical calculations away from equilibrium, we chose the non-dimensional parameters $\epsilon_{i,S}$, $\delta_{i,S\cup\{i\}} \in [10^{-2}, 10^2]$, for $i > 1$ or $S \neq \emptyset$, by logarithmic sampling (Paper Figure 6). If the gene regulatory system were at equilibrium, so that the association constants $K_{i,S} = a_{i,S}/b_{i,S\cup\{i\}}$ became meaningful, the chosen parameter range would allow $K_{i,S}$ to vary by a factor of 10^4 with respect to $K_{1,\emptyset}$. This allows the bare association constants, $K_{i,\emptyset}$, more leeway than the default factor of 10^3 used at equilibrium (§7.1) but the higher-order association constants, $K_{i,S}$, are now more constrained than the factor of 10^6 used at equilibrium. We felt that this permits a reasonable comparison between the equilibrium and non-equilibrium regions.

When the parameter range is increased, the non-equilibrium region for the all-or-nothing strategy behaves similarly to the equilibrium region and approaches $\mathcal{H}_a$ at the tip of a cusp (Figure S2). However, in the non-equilibrium case, $a = 5$ when $n = 3$, unlike the equilibrium case, in which $a = 3$ when $n = 3$.

7.3 Biased sampling algorithm for position-steepness regions

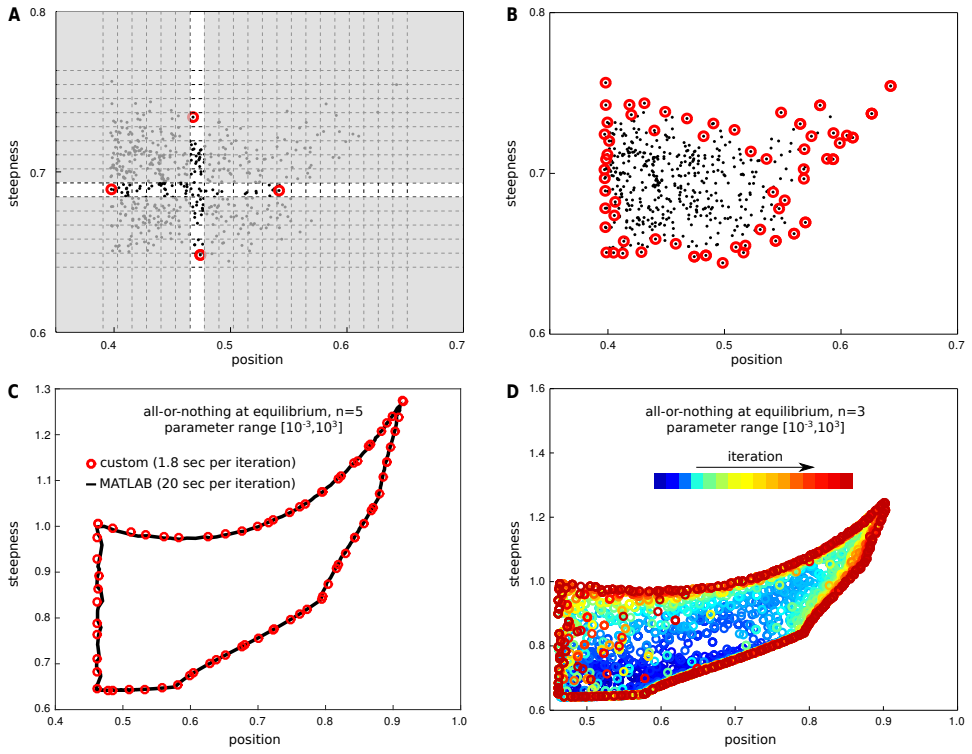
The random sampling which yielded the probability density function in Paper Figure 4A does not identify the shape of the sampled region in (γ, ρ) space. We therefore developed an algorithm which biases the sampling towards the boundary of the region and used this to find the position-steepness regions shown in Paper Figures 4, 5 and 6. The algorithm has the following steps. The non-dimensional parameters are assumed to be sampled from the range $[10^{-a}, 10^a]$ for some a, while the scale factors are kept at 1 in their respective units.

1. **Generate the symbolic GRF**. We used previously-developed software[1] described in [1] for calculating ξ^{G_n} at equilibrium or ν^{G_n} away from equilibrium and assembled the resulting GRF, $f_n(x)$, in symbolic form in Mathematica. This symbolic GRF was passed to Matlab.

2. **Create an initial set of GRFs, $\mathbf{\Phi_{ran}}$.** An initial set of numerically-specified GRFs, $\mathbf{\Phi_{ran}}$, was generated by randomly sampling the non-dimensional parameters from the uniform distribution on $[10^{-a}, 10^a]$ and instantiating the symbolic GRF at these parameter values. For each such numerically-specified GRF, $\gamma(g_n)$ and $\rho(g_n)$ were computed as described in §7.3.1 and those GRFs which failed the position filter $\gamma(g_n) \notin [0.5\mathcal{H}_n, 1.5\mathcal{H}_n]$ were rejected. Typically, 100 GRFs were chosen to create an initial set, $\mathbf{\Phi_{ran}}$, of (γ, ρ) values in position-steepness space. A different method was used for the average-binding strategy with pairwise cooperativity only (Paper Figure 5B, bottom) because we found no GRFs which passed the position filter. They appear to very rare. We therefore ran the present algorithm until we found 20 GRFs which passed the position filter and then used these to form the initial set, $\mathbf{\Phi_{ran}}$. The variable $\mathbf{\Phi_X}$ is set equal to $\mathbf{\Phi_{ran}}$.

[1] The software is freely and publicly available from `http://vcp.med.harvard.edu/software.html`.

3. **Identify the boundary of the set Φ_X, $\Phi_{ext} = \texttt{extreme}[\Phi_X]$.** We did this in two ways using a fast, custom subroutine, `extreme.m`, for regions without re-entrant cusps and a slower Matlab subroutine, `boundary.m`, for regions with re-entrant cusps (as in Paper Figure 5A). In the custom subroutine, the γ and ρ axes were divided into small slices and the GRFs with minimum and maximum values of γ and ρ were found within each slice, as shown in part **A** of the diagram below. Typically, 20 equally spaced slices across the point cloud in each axis were sufficient to find enough extreme points for boundary finding, as in part **B** of the diagram. The advantage of this method is that it is not sensitive to differences in the density of points over the boundary. The disadvantage is that any re-entrant cusp on the boundary may only be partially revealed by horizontal and vertical slicing. In contrast, Matlab's `boundary.m` subroutine finds the entire boundary, including those parts with re-entrant cusps. However, the non-uniform distribution of points in the steepness-position plots causes this subroutine to oversample the high-density (and low steepness) regions and to slow down significantly. We tested `extreme.m` against `boundary.m` and found no difference between the boundaries for regions without re-entrant cusps, as shown in part **C** of the diagram, but with a 10X difference in performance. We therefore used our custom subroutine `extreme.m` as the default and switched to Matlab's `boundary.m` only when we encountered re-entrant cusps.



4. **Generate "mutated" GRFs from the extremes, $\Phi_{new} = \texttt{mutate}[\Phi_{ext}]$.** Once the boundary is identified, for each GRF in the boundary, a new "mutated" GRF is constructed by sampling each non-dimensional parameter from the uniform distribution on $[0.5, 1.5]$ times its original value, while ensuring that the sampled value also lies in $[10^{-a}, 10^{a}]$. $\gamma(g_n)$ and $\rho(g_n)$ are then computed for the mutated GRF as described in §7.3.1 and GRFs which fail the position filter are rejected. Because there are many GRFs on the boundary, it was usually sufficient to generate just one mutated GRF for each GRF on the boundary. However, for pairwise-cooperativity only with the all-or-nothing strategy (Paper Figure 5A), we found that the regions grew very

slowly in discrete jumps, indicating that GRFs beyond the boundary were rare and were being found very slowly. We therefore generated 30 mutated GRFs for each GRF on the boundary and this substantially improved the convergence rate.

5. **Create an enlarged set of extreme GRFs, $\boldsymbol{\Phi_{all} = \Phi_{ext} \cup \Phi_{new}}$.** The mutated GRFs could be less extremal than their parents so they are adjoined to the previous ones to form an enlarged set of extremal GRFs. The variable $\boldsymbol{\Phi_X}$ is then set equal to $\boldsymbol{\Phi_{all}}$, control is returned to Step 3 and the the loop is repeated a fixed number of times.

6. **Generate a smooth boundary for plotting**. Once the algorithm had converged (see below), we used Matlab's `boundary.m` algorithm to re-compute the final boundary, which Matlab provides in a data structure that is more convenient for plotting. The boundaries shown in Paper Figures 4, 5 and 6 were generated in this way.

With each iteration of the algorithm, the extreme points found in Step 3 approach closer to the boundary and eventually condense at the boundary, as shown in part **D** of the diagram above. Because the algorithm chooses random GRFs by mutation in Step 4, we found that the smoothness of the boundary was a good test of whether the algorithm had converged. However, we also undertook further tests as described in §7.4

7.3.1 Calculating γ and ρ

The position, γ, and steepness, ρ, have to be calculated for each GRF that is generated by the algorithm above. This entails multiple evaluations of the derivative and we found it necessary for performance reasons to optimise this calculation. A numerically-specified GRF, $f_n(x)$, is given as a rational function of x with numerical coefficients. Its derivative, df_n/dx, was evaluated at $x = x_1$ as follows.

- At equilibrium, if $n \leq 7$, df_n/dx was calculated by symbolically differentiating f_n with respect to x and then evaluating this expression at $x = x_1$.
- At equilibrium, if $n > 7$, $f_n(x)$ was expressed as a fraction, $f_n(x) = p(x)/q(x)$, where $p(x)$ and $q(x)$ are polynomials in x, and the expression for the derivative of a fraction,
$$\frac{p'(x_1)q(x_1) - p(x_1)q'(x_1)}{q(x_1)^2},$$
was evaluated, where $p'(x)$ and $q'(x)$ are the symbolic derivatives of the polynomials.
- Away from equilibrium, the approximation for the derivative
$$\left.\frac{df_n}{dx}\right|_{x=x_1} \approx \frac{f_n(x_1+\Delta) - f_n(x_1-\Delta)}{2\Delta}$$
was calculated with $\Delta = 10^{-8}$.

To then determine γ and ρ, $x_{0.5}$ was first estimated by using the Matlab subroutine `fzero` to solve the equation $f_n(x_{0.5}) = 0.5$. The derivative df_n/dx, calculated as described above, was then evaluated at the points
$$x_i = i\left(\frac{x_{0.5}}{10^3}\right),$$
for i running from $i = 0$ to 10^4, thereby covering the interval $[0, 10x_{0.5}]$. The maximum derivative in this list was taken to be $\rho(f_n)$ and the corresponding x at which the maximum was found was taken to be $\gamma(f_n)$. Finally, the position and steepness of the normalised GRF was calculated by using Eq. 29, so that
$$\gamma(g_n) = \frac{\gamma(f_n)}{x_{0.5}} \quad \text{and} \quad \rho(g_n) = x_{0.5}\rho(f_n).$$
The choices made here were determined after several numerical experiments but the overall accuracy of the algorithm was assessed by the tests described next.

7.4 Testing convergence and numerical accuracy

We tested convergence of the biased sampling algorithm as follows. Starting from the converged boundary to be tested, we iterated the algorithm in §7.3 10 times but now generated 10^3 mutated GRFs for each boundary GRF in Step 4, thereby oversampling the boundary. For each region, $0 \leq i \leq 10$, where 0 is the starting region, we computed its area a_i using Matlab's `boundary.m` subroutine and then calculated the MAI, or "Mean Area Increase" over the iterations, as given by

$$\frac{1}{10}\left(\sum_{0\leq i<10}\frac{a_{i+1}-a_i}{a_i}\right). \tag{30}$$

Since mutated GRFs are added to the existing GRFs in Step 5, the area can only increase, so that $a_i \leq a_{i+1}$. The MAI scores for each of the boundaries shown in the Paper are show in the table below.

case	#sites	#GRFs	#iterations	MAI	MDI
Paper Figure 4C					
$p=1$	5	1099	1000	4.8×10^{-4}	1.6×10^{-4}
$p=2$	5	1995	1000	1.9×10^{-4}	1.8×10^{-4}
$p=3$	5	22418	1000	8.4×10^{-4}	$6.6\times10^{-3*}$
$p=5$	5	4836	1000	5.1×10^{-4}	3.7×10^{-3}
Paper Figure 4D					
AN	5	22418	1000	8.4×10^{-4}	$6.6\times10^{-3*}$
OM	5	15060	1000	2.7×10^{-4}	1.5×10^{-3}
AB	5	15044	1000	1.1×10^{-3}	9.9×10^{-4}
Paper Figure 5A					
AN	3	4704	1000	9.3×10^{-7}	2.5×10^{-5}
AN	4	4886	1000	3.7×10^{-5}	1.5×10^{-5}
AN	5	4773	1000	4.0×10^{-6}	3.1×10^{-5}
AN	6	2878	2000	1.6×10^{-4}	4.9×10^{-5}
AN	7	2157	2000	8.4×10^{-5}	5.4×10^{-5}
AN	8	1338	6000	7.8×10^{-5}	2.7×10^{-3}
AN	9	782	6000	9.3×10^{-4}	6.8×10^{-3}
AN	10	833	6000	1.3×10^{-4}	2.7×10^{-3}
AN	11	553	6000	2.8×10^{-5}	3.6×10^{-3}
AN	12	644	6000	3.1×10^{-4}	3.3×10^{-3}
Paper Figure 5B, top					
OM	3	3748	1000	7.3×10^{-3}	1.0×10^{-4}
OM	4	3692	1000	1.4×10^{-2}	4.7×10^{-5}
OM	5	7926	1000	1.1×10^{-3}	1.0×10^{-4}
OM	6	7502	1000	1.1×10^{-2}	6.3×10^{-4}
OM	7	23068	1000	9.0×10^{-3}	1.1×10^{-4}
OM	8	1818	1000	2.7×10^{-2}	1.0×10^{-4}
OM	9	11014	1000	1.1×10^{-2}	1.1×10^{-4}
OM	10	5528	1000	2.0×10^{-2}	1.0×10^{-4}
OM	11	6516	1000	8.1×10^{-2}	1.0×10^{-4}
OM	12	3041	1000	2.8×10^{-2}	1.1×10^{-4}
Paper Figure 5B, bottom					
AB	3	37376	$1000^{\#}$	6.3×10^{-6}	2.3×10^{-2}
AB	4	4065	$1000^{\#}$	4.0×10^{-6}	4.9×10^{-4}
AB	5	37405	$1000^{\#}$	$0^{\dagger}$	1.7×10^{-2}
AB	6	37401	$1000^{\#}$	$0^{\dagger}$	1.0×10^{-2}
AB	7	37362	$1000^{\#}$	$0^{\dagger}$	8.6×10^{-3}
AB	8	37384	$1000^{\#}$	$0^{\dagger}$	1.2×10^{-2}
AB	9	2893	$1000^{\#}$	$0^{\dagger}$	8.9×10^{-3}
AB	10	130	$1000^{\#}$	$0^{\dagger}$	1.2×10^{-3}
Paper Figure 6					
AN	3	292	3000	4.3×10^{-3}	1.0×10^{-3}
OM	3	293	6000	7.0×10^{-3}	5.0×10^{-3}
AB	3	355	6000	6.2×10^{-3}	4.8×10^{-3}

The abbreviations AN for all-or-nothing, OM for one-or-more and AB for average-binding have been used in this table. The column for #iterations gives the number of iterations of the algorithm after the initial set was found. For the average-binding strategy with pairwise cooperativity only, we used a different strategy to generate the initial set of GRFs (see Step 2 of the algorithm in §7.3) and this is indicated by the symbol #; also in this case, when the number of sites (#sites) was 5 or more, the convergence test created no new GRFs on the boundary, so that the MAI score was identically zero and this is indicated by the symbol †. Where possible, we checked all GRFs generated by the algorithm in §7.3 but when there were too many because of high numbers of iterations, we restricted attention to the GRFs on the final converged boundary. The symbol * in the table signifies a duplicated dataset.

The table shows that the MAI scores are never worse than 0.1—these worst cases occur for the one-or-more strategy with pairwise cooperativity only, as in Paper Figure 5B, top, for which the regions are very small—and are usually less then 0.001, which we considered to be good evidence that the boundary had converged. We note, however, that boundary finding depends on randomly generating GRFs in Step 4 of the algorithm and we cannot rule out the possibility that increased iterations would occasionally find extremely rare GRFs outside the boundaries that we have shown.

As a further test of the algorithm and its numerical accuracy, we undertook an independent calculation of the region using the same parameter values. We recorded all the parameter values generated by the algorithm in §7.3 and passed them to Mathematica. We instantiated the symbolic GRF at each set of parameter values to generate a GRF, f_n, with numerical coefficients. We used the native `Solve[]` capability in Mathematica to normalize x to the half-maximal concentration $x_{0.5}$ and defined $g_n(y) := f_n(x_{0.5}y)$. Like f_n, g_n has numerical, not symbolic, coefficients. We calculated d^2g_n/dy^2 as a function of y and used the `Solve[]` function again to obtain values of y where $d^2g_n/dy^2 = 0$. This process sometimes returned very large values of y, corresponding to where $g_n(y)$ is nearly at its asymptotic limit and $dg_n/dy \to 0$. To avoid this numerical artifact, we imposed the constraint that $y < 10^3$. We note that y is non-dimensional and the position of the maximum derivative was never found beyond $y = 2$, so the limit of 10^3 seems safe. Having found the values of y at which dg_n/dy has a turning point (ie: either a local maximum or a local minimum), we substituted these values into dg_n/dy and took their maximum, thereby obtaining $\rho(g_n)$, and took the corresponding value of y to be $\gamma(g_n)$. To compare this with the values obtained by the biased sampling algorithm, we took, for each GRF, the Euclidean distance between the position-steepness coordinates found by the algorithm in Matlab, (γ^A, ρ^A), and those found by the test just described in Mathematica, (γ^T, ρ^T). We defined the MDI, or Maximum DIstance, to be the maximum of this value over all the GRFs generated in calculating a particular boundary:

$$\max_{g \in GRFs} \sqrt{(\gamma^A(g) - \gamma^T(g))^2 + (\rho^A(g) - \rho^T(g))^2}\,. \tag{31}$$

The MDI scores for each of the boundaries shown in the Paper are given in the table above, with the column for #GRFs showing the number of GRFs that were checked to evaluate the MDI.

The MDI scores are never worse than 1.0×10^{-2}—these worst cases occur for the average-binding strategy with pairwise cooperativity only, as in Paper Figure 5B, bottom—and are usually an order of magnitude better. We considered this to be evidence of good numerical accuracy.

The MDI test not only assessed numerical accuracy, it was also an effective method for identifying mistakes in our overall work flow, including in file handling, software coding, numerical implementation and algebraic manipulation, which we then corrected. The algorithm calculations in Matlab were carried out by JE and the independent tests in Mathematica were carried out by FW and this cross-checking also helped with the detection of errors.

SUPPLEMENTAL REFERENCES

[1] T. Ahsendorf, F. Wong, R. Eils, and J. Gunawardena. A framework for modelling gene regulation which accommodates non-equilibrium mechanisms. *BMC Biol.*, 12:102, 2014.

[2] J. Gunawardena. A linear framework for time-scale separation in nonlinear biochemical systems. *PLoS ONE*, 7:e36321, 2012.

[3] J. Gunawardena. Time-scale separation: Michaelis and Menten's old idea, still bearing fruit. *FEBS J.*, 281:473–88, 2014.

[4] F. Harary, J. P. Hayes, and H.-J. Wu. A survey of the theory of hypercube graphs. *Comput. Math. Applic.*, 15:277–89, 1988.

[5] I. Mirzaev and J. Gunawardena. Laplacian dynamics on general graphs. *Bull. Math. Biol.*, 75:2118–49, 2013.