Precision in a rush: trade-offs between reproducibility and steepness of the hunchback expression pattern

Huy Tran[1,5], Jonathan Desponds[2], Carmina Angelica Perez Romero[3], Mathieu Coppey[4], Cecile Fradin[3], Nathalie Dostatni[5], Aleksandra M. Walczak[1*]

1 Ecole Normale Supérieure,
PSL Research University,
CNRS, Sorbonne Université,
Laboratoire de Physique Théorique, Paris, France
2 Dept. of Physics,
University of California, San Diego, USA
3 Dept. of Physics and Astronomy,
McMaster University, Canada
4 Institut Curie, PSL Research University,
CNRS, Sorbonne Université,
Physico Chimie, Paris, France
5 Institut Curie, PSL Research University,
CNRS, Sorbonne Université,
Nuclear Dynamics, Paris, France

Fly development amazes us by the precision and reproducibility of gene expression, especially since the initial expression patterns are established during very short nuclear cycles. Recent live imaging of *hunchback* promoter dynamics shows a stable steep binary expression pattern established within the three minute interphase of nuclear cycle 11. Considering expression models of different complexity, we explore the trade-off between the ability of a regulatory system to produce a steep boundary and minimize expression variability between different nuclei. We show how a limited readout time imposed by short developmental cycles affects the gene's ability to read positional information along the embryo's anterior posterior axis and express reliably. Comparing our theoretical results to real-time monitoring of the *hunchback* transcription dynamics in live flies, we discuss possible regulatory strategies, suggesting an important role for additional binding sites, gradients or non-equilibrium binding and modified transcription factor search strategies.

I. INTRODUCTION

During development reproducible cell identity is determined by expressing specific genes at the correct time and correct location in space in all individuals. How is this reproducible expression pattern encoded in the noisy expression of genes [1, 2], and read out in a short amount of time? We study this question in one of the simplest and the best understood developmental examples – the Bicoid-*hunchback* system in *Drosophila melanogaster*. In the fly embryo, the exponentially decaying Bicoid (Bcd) gradient [3–5] acts as a maternal source of positional information along the embryo's Anterior-Posterior (AP) axis [6]. The *hunchback* (*hb*) gene extracts this positional information from the local Bicoid concentration and forms a steep binary-like expression pattern, observed as early as in nuclear cycle (nc) 10 (see Fig. 1 A) [3, 7–9]. This Hb pattern later becomes a source of positional information

for the formation of other gap gene patterns [10, 11], forming the first step in the differentiation of cecullar phenotypes.

From real-time monitoring of the *hb* transcription dynamics [12, 13] using the MS2-MCP RNA-tagging system [14], we observed that from nc11 to nc13 the positional readout process of the *hb* gene is interrupted by mitosis, leaving a window of 5-10 minutes for gene expression in each cycle. Once the pattern stabilizes 2-3 minutes after mitosis, as we describe in detail in a companion experimental paper [13], the boundary between regions of high and low *hb* transcription is already steeper than even the Hb protein concentration profile in nc14 [13, 15, 16].

Several studies have proposed that the steep boundary between regions of high and low *hb* expression, given the smooth Bcd transcription factor (TF) gradient, is due to the cooperativity between the TF binding sites (Fig. 1B) [7, 15, 17–20]. This cooperativity diversifies gene expression levels given small changes in the input [21–23]. Conventionally,
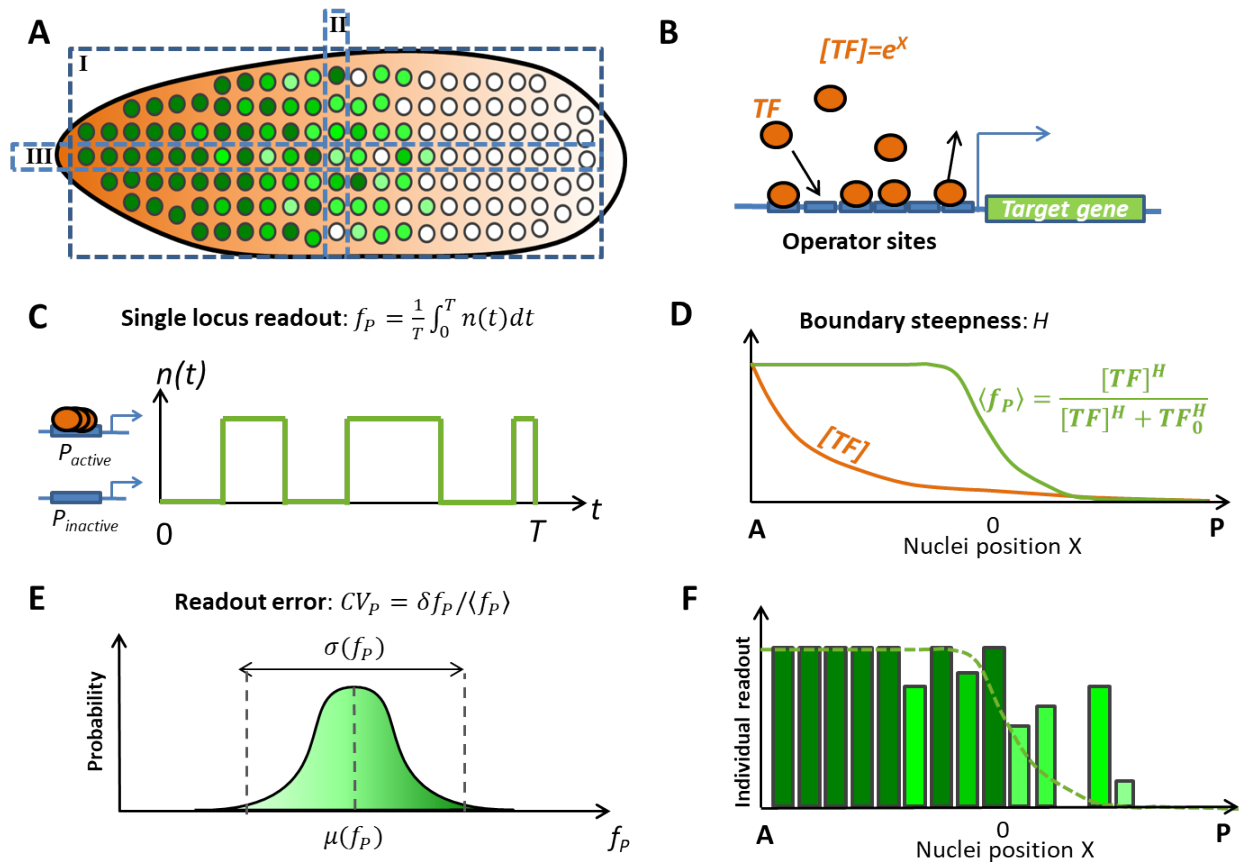
FIG. 1: **Setup of the problem: features of the hb transcription pattern in early fly development.**
(A) A cartoon of the side section of the fly embryo in nuclear cycles (nc) 11-13. Nuclei at different positions along the anterior posterior (AP) axis express different concentrations of hb mRNA, represented by different shades of green (dark green denotes larger concentrations). The hb gene expression pattern can be studied from different perspectives: (I) Side view of the whole embryo. (II) Single columns of nuclei at similar position along the AP axis. (III) Single rows of nuclei along the AP axis.
(B) The expression of hb mRNA from the hb gene is regulated by Bcd transcription factor (TF) binding. We consider a model of gene expression regulation via the binding and unbinding of Bcd proteins (orange) to multiple operator sites (blue) of the promoter. Bcd forms an exponentially decaying gradient along the AP axes and the concentration of Bcd TF in the nuclei depends on the position of the nucleus along the AP axis, X.
(C) During each nuclear cycle each hb loci switches from periods of activation to inactivation by binding and unbinding Bcd TF. The distribution of these periods depends on the binding and unbinding rates of Bcd TF. Within the model, each hb loci produces a readout f_P defined as the average promoter activity level $n(t)$ during the steady state expression interval T of the interphase of a given nuclear cycle.
(D) By observing the whole embryo (perspective I in Fig. 1 A), we are able to calculate the average of the expression pattern f_P of the nuclei, $\langle f_P \rangle$, as a function of the nuclei's position along the AP axes (green line), from which the boundary steepness (denoted by H) is quantified by fitting a Hill function of Bcd concentration $[TF]$ (orange line). TF_0 denotes the concentration of Bcd TF at half-maximal expression.
(E) By observing nuclei at similar position X along the AP axis (perspective II in Fig. 1 A), we can make a distribution of the readout, f_P and use it to calculate the readout errors CV_P in the single locus readout f_P. CV_P is defined as the standard variation of the readout f_P divided by its mean.
(F) The detailed expression pattern f_P obtained by observing a single row of nuclei (perspective III in Fig. 1 A) along the AP axis, depends both on the averaged pattern and the errors in the readout of this mean value.

the pattern steepness is represented by the Hill coefficient H. We define the hb gene readout as the hb gene transcription state of one locus in a single nucleus averaged over a given transcription window, f_P, (Fig. 1C). We can evaluate this quantity as a function of the TF concentration $[TF]$ (Fig. 1D), and thanks to the exponential nature of the Bcd gradient [15], uniquely associate a position along the AP axis to a Bcd concentration [TF]. The Hill coefficient is then estimated by fitting the mean readout value averaged over all nuclei at a specific position along the AP axis, $\langle f_P \rangle$, to a sigmoidal function:

$$\langle f_P \rangle \sim \frac{[TF]^H}{TF_0^H + [TF]^H}, \qquad (1)$$

where TF_0 is the Bcd concentration that results in half-maximal hb expression, $\langle f_P|_{[TF]=TF_0} \rangle = 0.5$ (Fig. 1D). TF_0 defines the middle of the boundary, which we will call the mid-boundary point, that separates the highly expressing "ON" nuclei in the anterior region and the minimally expressing "OFF" nuclei in the posterior region of the embryo. Within a simple model where hb expression depends only on the binding and unbinding of Bcd to the hb promoter, the maximal steepness of the hb expression pattern was shown to depend on the number of operator binding sites in the promoter region of the gene N. Depending on whether this process conserves detailed balance or not, the maximal Hill coefficient is $2N - 1$ or N, respectively [18].
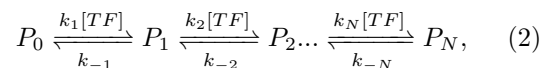
These studies did not address whether such a steep boundary is achievable within the limited time window of 3 to 15 minutes in nuclear cycles 11-13, in which the Bcd concentration is read. The effects of the time constrained readout are further aggravated by the fact that transcription is stopped before and during each mitosis [13, 24], suggesting that the hb expression pattern needs to be re-established in each nuclear cycle. In addition, the intrinsic noise in chemical processes leads to inherent errors in the Bcd concentration readout [16, 25]. This noise results in a lower bound for the Bcd concentration readout error, defined as the standard deviation of the concentration of the measured molecule divided by its mean (Fig. 1E), that depends on the readout integration time and the diffusion constant of ligand molecules [5, 26–29]. Extending the original work that considered a single or an array of non-interacting receptors [5, 26], other work pointed out that cooperativity from receptor arrays increases the readout noise [30, 31]. Given these effects, it is unclear how the readout precision of the Bcd concentration (or nuclei position) changes quantitatively given the highly cooperative readout process by the

promoter observed as the steep hb expression pattern [9, 15, 16] (Fig. 1F) and what are the consequences for the ability of neighboring nuclei to take on different cell fates. In this work, we investigate how the constraints coming from short cell cycles affect the steepness and errors in the hb expression pattern.

II. THE MODEL

In the early stage of development, the hb transcription pattern is steep, despite relying mostly on the exponential Bcd gradient as the source of positional information [8]. It was hypothesized that Bcd molecules can bind cooperatively to the many Bcd binding sites on the hb promoter, enabling the gene to have diverse expression levels in response to gradual changes in the Bcd concentration [7, 18]. We use a simple model of gene expression regulation by binding of Bcd transcription factors (TF) to the operator sites (OS) of the target promoter [18] (Fig. 1B, SI Fig. 1). The promoter activity depends on the occupancy state of the operator sites and we consider different activation schemes, which we specify below. The binding rates are functions of the position-dependent TF concentration and we further assume their value is bounded by the promoter search time of individual TFs (see SI section 2). The promoter readout, f_P, is defined as the mean of the promoter activity level $n(t)$, calculated from the temporal average of the promoter state $n(t)$ over the steady state expression interval T of a given nuclear cycle interphase (see Fig. 1C).

We first focus on a simplified version of the general model of gene regulation for binding of Bcd TF to the N OS [18], where all the binding sites of the target promoter are identical. This assumption gives a Markov model of TF binding/unbinding to the many identical OS of the target promoter:

$$P_0 \underset{k_{-1}}{\overset{k_1[TF]}{\rightleftharpoons}} P_1 \underset{k_{-2}}{\overset{k_2[TF]}{\rightleftharpoons}} P_2... \underset{k_{-N}}{\overset{k_N[TF]}{\rightleftharpoons}} P_N, \qquad (2)$$

where P_i denotes the promoter state with i bound OS and $N - i$ free OS. $[TF]$ is the relative Bcd TF concentration with respect to that at the mid-boundary position. Since Bcd concentration decays exponentially along the embryo AP axis, we estimate the relative nuclei position X measured in terms of the gradient decay length from the TF concentration ($X = \ln([TF])$), such that at mid-boundary $X = 0$ and $[TF] = 1$. The binding and unbinding of TF to the promoter occur with rate constants k_i and

k_{-i}. If all the rates are non-zero, all reactions are reversible and Eq. 2 defines an equilibrium model.

Throughout the paper, we randomize the binding and unbinding rates to explore the behavior of the model (see Methods section V B for details). When comparing models with different parameters we rescale unbinding rate values k_{-i} in order to keep the binding rate at the mid-boundary position constant. In order to best align to experimental observations, we estimate this fixed binding rate at -5% embryo length (EL) (-50% EL and 50% EL are the embryo's anterior and posterior poles), which is the typical boundary position in the analyzed wild type embryos [13] (see SI - section 3.) .

We first consider the "all-or-nothing" case, i.e. the promoter is active when the OS are fully bound by TF ($P_N \equiv P_{\text{active}}$), although the qualitative conclusions remain the same for the "K-or-more" scenario [18, 30], where the promoter is active if at least K sites are occupied (see section III F). At steady state, we find the probability that the promoter is in the active state given the nucleus position X (see SI - section 1):

$$P(P_{\text{active}}, X) = \frac{\tilde{K}_N e^{N \cdot X}}{\sum_{i=0}^{N} \tilde{K}_i e^{iX}}, \qquad (3)$$

where for convenience of notation we define the effective equilibrium constant $\tilde{K}_i = \prod_{j=1}^{i} k_j / \prod_{j=i}^{1} k_{-j}$ and $\tilde{K}_0 = 1$. We assume the target gene transcription rate at steady-state is proportional to the probability of the promoter to be in the active state (Eq. 3).

The steepness of the expression pattern is quantified by the Hill coefficient H [32], calculated as the slope of the expression pattern at the mid-boundary position (see SI - section 4) $H = N - (\sum_{i=1}^{N-1} i \cdot \tilde{K}_i)/\tilde{K}_N$. H is bounded from below by 1, and from above by N – the OS number, confirming previous results [18]. Maximum steepness ($H = N$) is achieved when the system spends most of the time in the fully free (P_0) or fully bound states (P_N) while $H = 1$ when the system spends most of the time in highly occupied states P_{N-1} and P_N (see SI - section 5).

Lastly, we also consider a full non-equilibrium binding model (defined in SI Fig. 1 and SI section 8), in which not all binding reactions are reversible, and reversible equilibrium models with two different types of TF factors (defined in SI – section 10). To explore the properties of all of these models, we solve the time dependent equations of motion for the stochastic binding models numerically and, when possible analytically in steady state, considering dif-

ferent expression schemes ("all-or-nothing" and "K-or-more"), different numbers of TF binding sites and randomizing binding and unbinding parameters (see Methods section V B).

III. RESULTS

A. The expression pattern formation time

The hb expression pattern in the early phase of development is always formed under rigorous time constraints: the total time of transcription during an interphase of duration T_{full} varies in nc 10-13 from ~ 100 seconds to ~ 520 seconds (Fig. 2A). During mitosis, Bcd molecules leave the nuclei and only re-enter at the beginning of the interphase [5]. The steep expression pattern takes time to reestablish. Assuming that at the beginning of the interphase all OS of the hb promoter in all nuclei are free, the mean probability $\mu_P(t, X)$ for the promoter to be active at position X at time t following the entering of the TF to the nuclei is initially large only in the anterior of the embryo (see SI Fig. 3). By propagating the time dependent equations of motion for the stochastic equilibrium binding models (SI Eq. 5) in time, we see that with time $\mu_P(t, X)$ increases also in other regions of the embryo, to reach its steady-state form ($P(P_{\text{active}}, X)$) with a border between low and high expressing nuclei that defines the mid-boundary position.

Whether the interphase duration T_{full} in a given nuclear cycle is long enough for the system to reach steady state depends on the parameters of binding and unbinding of the TF to the operator sites. However, the binding and unbinding rates also determine the expression pattern steepness, leading to constraints between expression pattern steepness and formation time. Considering the "all-or-nothing" equilibrium model, when the promoter is active ($P_N \equiv P_{\text{active}}$), any unbinding from the promoter inactivates the promoter. We find that a steep expression pattern requires the promoter to stay in the active state for a long time or to be regulated by binding of TF to many OS (see SI - section 5.3):

$$H \leq N - (N-1)\frac{\tau_{\text{bind}}}{\tau_{\text{active}}}, \qquad (4)$$

where $\tau_{\text{active}} = (k_{-N})^{-1}$ is the mean time for the promoter to switch from the active state to the inactive state and $\tau_{\text{bind}} = (k_N)^{-1}$ is the mean time for a TF to bind to the last unoccupied OS. The maximum value of H is reached when the promoter spends most of the time in the P_0, P_{N-1} and P_N
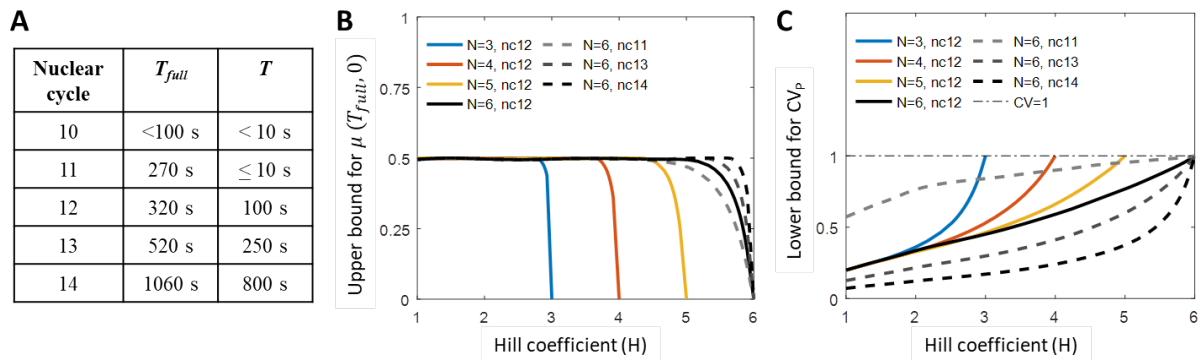
FIG. 2: **Equilibrium model predictions for the pattern formation time and readout error.**
(A) The early nuclear cycles have a short interphase T_{full} and even shorter steady state periods T when the average transcription rate is neither increasing after, nor decreasing before, transcription shut-off during mitosis. The total transcription window T_{full} and the time window when the transcription rate is at steady state T of hb transcription in nuclear cycle 10, 11, 12, 13 and early nuclear cycle 14 (before cellularization) at 25°C obtained from 8 MS2-MCP movies [13]. The short periods of transcription inactivity right before and after mitosis are excluded.
(B) Steep steady state expression profiles (large H) cannot be reached in short nuclear cycles. Since transcription is shut-off during mitosis, the sigmoidal expression pattern (as in Fig. 1D), characterized by the mean promoter activity in nuclei positioned at mid-boundary $\mu_P(T_{\text{full}}, 0) = 0.5$, needs to be re-established in each nuclear cycle. We randomize the binding and unbinding rates of the equilibrium model to calculate the upper bound for the mean promoter activity in nuclei positioned at mid-boundary $\mu_P(T_{\text{full}}, 0)$, and the corresponding Hill coefficient H, for varying OS number N and nuclear cycle duration T_{full}. $\mu_P(T_{\text{full}}, 0) < 0.5$ indicates the steady state expression profile could not be reached within the nuclear cycle duration.
(C) Steep expression profiles (large H) correspond to equilibrium binding models with larger readout errors of the mean activity of the nuclei at the mid-boundary position. The readout error decreases with nc duration. Randomizing parameters of the equilibrium model we plot the lower bound for the readout error of the mean activity of the nuclei, CV_P, defined as the standard variation of the readout f_P divided by its mean, for varying OS number N and steady state-period T. The bounds in (A-B) are calculated numerically from $\sim$ 50000 data points of the solutions of dynamical equations of the equilibrium model (SI section 1.2) with N OS, each corresponding to a randomized kinetic parameter set.

states (see SI - section 5.3). This limit corresponds to very slow promoter switching, $\tau_{\text{active}} \geq \tau_{\text{bind}}$, similarly to conclusions obtained for cell surface receptors [30]. With typically considered parameters for the Bcd-hb system [9, 15, 33], $\tau_{\text{bind}} \sim 4s$ and the $H \sim N$ limit in Eq. 4 corresponds to $\tau_{\text{active}} \gg 4$ s (see SI section 5.3). Even if the currently available estimates for the value τ_{bind} prove inaccurate, the qualitative conclusion about slow promoter switching will remain unchanged.

Given the limited interphase duration in nc 11, $T_{\text{full}} \sim 270s$ (Fig. 2A), randomizing parameters of the equilibrium model (Eq. 2) shows that a steep steady-state expression pattern cannot be established during the interphase: the upper bound for the mean promoter activity level $\mu_P(T_{\text{full}}, 0)$ at the mid-boundary position ($X = 0$) at the end of the interphase of duration T_{full} is less than the steady state value of 0.5 for kinetic parameters giving large Hill coefficients H (Fig. 2B). For long interphases ($T_{\text{full}} \geq 100$ s), all patterns but those close to the

maximum allowed steepness of $H \approx N$ reach steady state. For $H \approx N$, Eq. 4 imposes large τ_{active}, which means there are not enough binding and unbinding events to achieve the steady state expression pattern with $\mu_P(T_{\text{full}}, 0) \sim 0.5$ at the boundary.

Generalizing the model to allow non-equilibrium binding (SI Fig. 1) increases the possible Hill coefficients above $H > N = 6$, but does not alleviate their inaccessibility within the considered nuclear cycles 11-13 (SI Fig. 12). Given the observed steep boundary $H \sim 7$ in nuclear cycles 11-13 [13, 16] and the relatively short interphase duration ($T_{\text{full}} \sim 520s$ in nc 13, see Fig. 2A), it seems unlikely that the steep steady state boundary is reached in early fly development with only the $N = 6$ known Bicoid operator sites of the proximal hb promoter [7, 19]. Nevertheless the steady state results give a best case scenario for readout error estimates so we focus on an equilibrium steady state system in the next section. We then extend the arguments to out-of-equilibrium binding.

B. The single locus readout error at steady-state

Even when the mean promoter dynamics over the nuclear population has reached steady-state, each individual gap gene in each nucleus must independently read the positional information and express mRNA in a way to ensure the transcription pattern's reproducibility. The promoter in each nucleus switches between an active and an inactive state $n(t) = 0, 1$ (Fig. 1C). The reproducibility of the transcriptional readout $f_P = \frac{1}{T} \int_{t=0}^{T} n(t) dt$ at the mid-boundary position in steady state is described by the nuclei-to-nuclei readout error of the mean activity of the nuclei $CV_P = \delta f_P / \langle f_P \rangle$, where the average $\langle \rangle$ is over nuclei at the same position $X = 0$ calculated during the steady state expression window T in a given nuclear cycle (Fig. 1C and E, see SI - section 6).

Randomizing binding parameters in the equilibrium model (Eq. 2) we see that the lower bound for the nuclei-to-nuclei readout error, CV_P, increases with increasing Hill coefficient H and decreases with the nc duration (Fig. 2C). A steep pattern requires slower promoter switching dynamics (Eq. 4), which results in less independent measurements that take part in the single locus readout during each interphase. Therefore, the steeper the pattern, the larger the nuclei-to-nuclei readout error in the expression pattern due to the increased variability in the readouts, f_P, between different nuclei [34]. When the steepness H approaches its upper bound limited by the maximum number of binding sites N, due to a small number of switching events during the interphase, the distribution of readout f_P approaches a Bernoulli distribution with $p = 0.5$ with the relative error always equal to $\sqrt{(1-p)/p} = 1$, regardless of T and N. The decrease in readout error at small steepness depends on the length of the nuclear cycle (Fig. 2C). For very short cycles (i.e. $T < 10$ s), only non steep patterns ($H \leq 2$) are able to significantly reduce the readout errors. For long interphases ($T > 100$ s), significant reduction in readout errors can be achieved with steep patterns ($H \sim 5$), and further decreasing H yields little improvement in reducing the readout error (SI Fig. 4).

In our models, τ_{bind} is the only external time scale in the problem. We assume it is set by diffusion (SI – section 2) and all other timescales (e.g. the time to establish the steady state expression pattern, the value of τ_{active} that will minimize the time to establish the steady state profile) depend on it. If our estimate of $\tau_{\mathrm{bind}} \sim 4s$ is inaccurate and differs by orders of magnitude, then the conclusions about

not being able to establish the steep steady state expression pattern may not hold. However the point of the analysis presented in this section remains valid – steep expression profiles result in large nuclei-to-nuclei variability.

C. Positional resolution

The above analysis uncovers a trade-off between the readout error and steepness of the expression pattern at the boundary: the steeper the boundary, the larger the minimal nuclei-to-nuclei variability, quantified as the readout error (Fig. 2C). Additionally, while long nuclear cycles seem desirable both to obtain the observed steep expression patterns and decrease nuclei-to-nuclei variability, the nuclear cycles 11-13 during which these steep patterns are experimentally observed [13] are very short (Fig. 2A). In light of the experimental facts, steep expression patterns seem like an obstacle to reducing readout errors.

The trade-off between the expression pattern steepness and the nuclei-to-nuclei variability suggests that neither of these features alone can be used as the sole criterion for a reproducible pattern. This observation is not surprising given that these features emerged from looking at the embryo from two different perspectives (Fig. 1A): the expression pattern steepness is perceived from an external observer's perspective when looking at the *whole* embryo at a fixed time (Fig. 1D), while the readout error is calculated by comparing nuclei at a similar position along the AP axis averaged over time (Fig. 1E). These features are likely to be unobtainable to individual nuclei (Fig. 1F), in which the decisions about transcription are made, since they require averaging or comparing the readout of different nuclei.

In order to better understand the readout of reproducible cell fates from the perspective of an individual nucleus in the fly embryo, we use the positional resolution of the expression pattern, ΔX [9, 35], defined as the minimum distance between two nuclei located symmetrically on the two sides of the mid-border position $X = 0$ that have distinct readout levels in steady state (Fig. 3A). Specifically, if F_+ and F_- are the distributions of mRNA concentrations in two nuclei at positions $+\Delta X/2$ and $-\Delta X/2$ (see SI - section 7), we define the positional resolution ΔX such that the probability of a false positive readout is small, $P(F_+ \leq F_-) \leq 0.05$. Positional resolution is a distance measure that we report in length units of % egg length (EL) or nuclei widths,
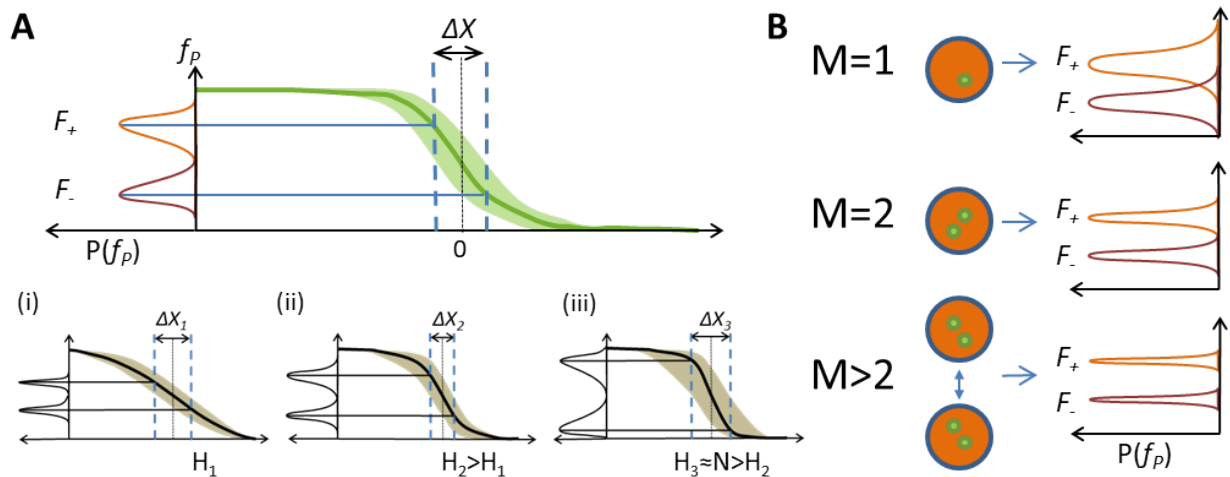
FIG. 3: **The positional resolution of the pattern.**
(A) We use the positional resolution ΔX to describe how well nearby nuclei can readout discernible inputs. F_+ and F_- are the positional readouts in individual nuclei positioned at $\pm \Delta X/2$ from the mid-boundary point ($X = 0$). Positional resolution, ΔX, is the minimal distance between nuclei that make distinct readouts at steady state $P(F_+ \leq F_-) \leq 0.05$. Positional resolution results from a trade-off between the pattern steepness and the readout error: (i-iii) a cartoon representation of the trade-off for a flat pattern (low H, i), pattern of intermediate steepness (ii) and a steep pattern (high H, iii). Both (i) and (iii) have a large value of positional resolution. At low H (i), the readout errors are low but the mean readout values are very similar. At high H (iii), the mean readout values are different but the readout errors are large. The best positional resolution is reached with an intermediate H (ii).
(B) Each nucleus readout is the average of M independent single locus readouts: M can be 1 (there is one copy of the gene as is the case in a heterogeneous gene construct), 2 (there is one gene copy on each chromosome as is the case of the WT embryo) or greater (nuclei at the same position can communicate by diffusion of readout molecules [35, 36] and the readout is the result of spatial averaging). As M increases, the readout error of the nuclei decreases due to spatial averaging.

where one nucleus width corresponds to 2% EL. The width of one nucleus (2% EL) sets a natural resolution scale for the problem – the embryo cannot achieve a better resolution than that of one nucleus. While positional resolution tells us how well a nucleus can distinguish its position from that of other nuclei, it is not a measure of information between the position along the AP axis and Bicoid concentration, such as the previously proposed positional information [37, 38]. The term positional resolution is borrowed from optics, and the higher the resolution the better, since it corresponds to a smaller minimal distance between nuclei that make distinct readouts. To avoid confusion, in the text we refer to the minimal value of the positional resolution ΔX as the best case scenario when nuclei separated by a small distance make discernable readouts.

The trade-off between the pattern steepness and the readout error translates into constraints on the positional resolution. For a flat expression pattern (low H, Fig. 3A, panel (i)), F_+ and F_- have a small difference in their mean value, which makes it hard to differentiate the mRNA concentration in closely positioned nuclei, but the variance around their mean is also small. On the other hand, with a very steep pattern (Fig. 3A, panel (iii)), F_+ and F_- have a big difference in their mean mRNA expression but also an increased variance, due to the increased readout errors in particular nuclei. An intermediate Hill coefficient offers the best positional resolution (Fig. 3A, panel (ii)).

To evaluate the positional resolution for a given pattern steepness H and steady state expression interval T in a given nuclear cycle we randomize all the binding/unbinding parameters for a promoter with $N = 6$ OS – a number inspired by the number of Bicoid binding sites found on the hb promoter [7, 19]. We identify the parameters that give the smallest CV_P to ensure the smallest ΔX. CV_P and ΔX are tightly correlated (SI Fig. 9) but CV_P is faster to evaluate.

For short nuclear cycles (small T), there are hardly any promoter switching events during the readout time window and the readout error $CV_P \sim 1$ for all values of H (Fig. 2C). In this case, the positional resolution is mainly governed by the increase
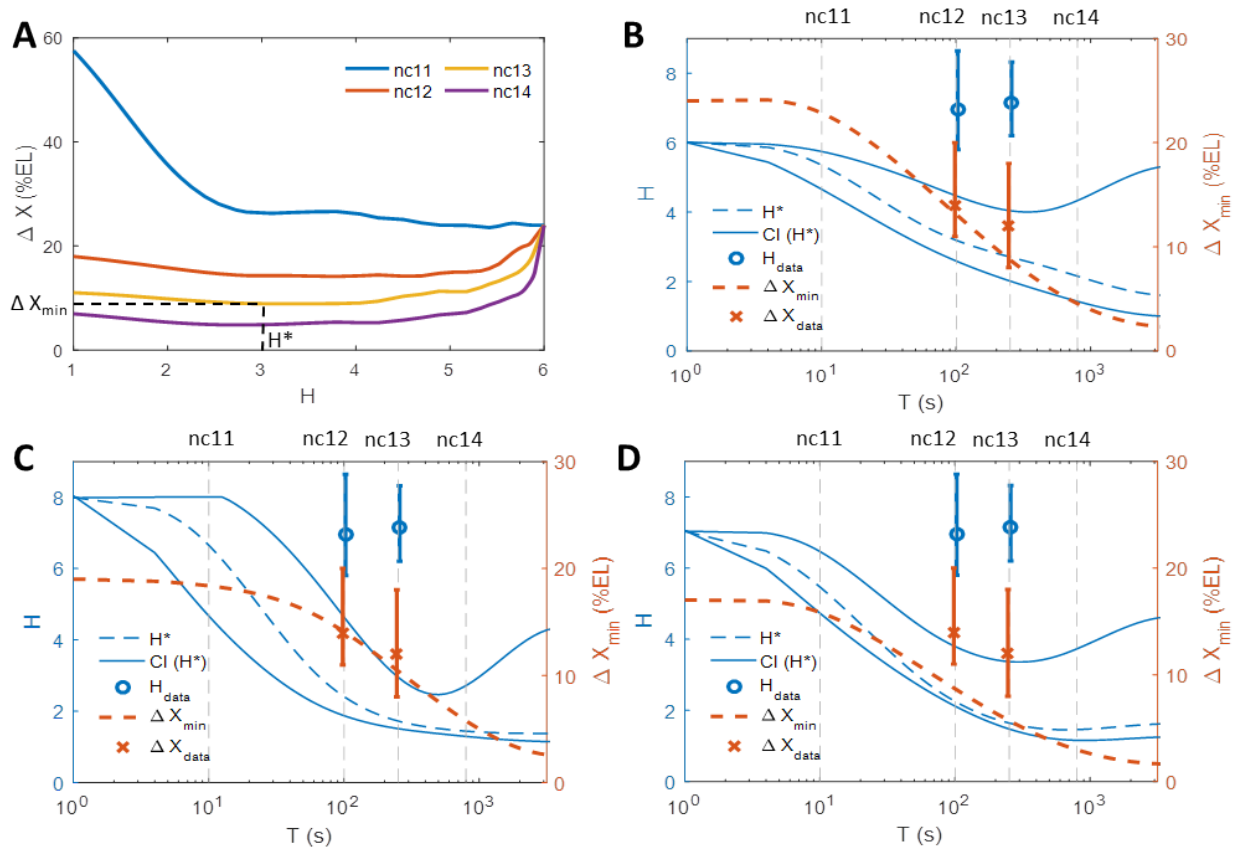
FIG. 4: **The positional resolution of the expression pattern for different nuclear cycles and regulatory models**

(A) Positional resolution calculated from the equilibrium binding site model with randomized kinetic parameters that give different values of the expression profile steepness H for $M = 1$. The colored lines show the results for parameters that give the smallest readout error CV_P from a set of randomized parameters for the steady-state window T in nc 11-14 (Fig. 2A). The curves are smoothed using cubic spline interpolation for better visualization. For each nuclear cycle we find the optimal Hill coefficient H^* that results in a model with the smallest value of positional resolution, $\Delta X_{\min}$.

(B-D) The range of optimal Hill coefficients H^* (dashed blue line) that yield the lowest value of the positional resolution (obtained as described in Fig. 4A) as a function of the steady-state readout duration T for the equilibrium $N = 6$ model (defined in Eq. 2 and SI section 1.2) (B); the hybrid $N = 6$ non-equilibrium model with 3 equilibrium and 3 non-equilibrium OS (defined in section III E and SI section 8) (C); and the two mirror TF gradient model (defined in section III G and SI section 10)(D). Around the optimal Hill coefficients H^* that give the minimal value of the positional resolution ΔX (dashed blue line) we also plot the range of Hill coefficients that come from models resulting in $\Delta X = \Delta X_{\min} \pm 2\%$ EL, allowing for a tolerance interval for the positional resolution (solid blue lines). The curves are smoothed by cubic spline interpolation for better visualization. Also shown is the lowest achievable value of the positional resolution in the numerical randomization experiment for varying T (dashed orange line). The results are obtained assuming a diffusion limited estimate for $\tau_{\text{bind}} = 4s$. The theoretical results for $M = 1$ for all models are compared to the empirical Hill coefficient H_{data} (blue circles with error bars) and positional resolution ΔX_{data} (orange crosses with error bars) extracted from MS2-MCP live imaging data in nc 12 and nc 13 [13] (see SI section 12). The error bars correspond to 95% confidence intervals. In general, only the non-equilibrium model with $N = 6$ is able to produce both Hill coefficients and ΔX values observed in experiments. However, assuming $t_{\text{bind}} = 4s$, even the non-equilibrium model cannot achieve the experimental values of the Hill coefficient during the time T of nc 12-13.

in the difference between F_+ and F_-, with increasing Hill coefficients H, which leads to a decrease in ΔX (Fig. 4A). As T lengthens, the value of the positional resolution ΔX for small Hill coefficients decreases with increasing H, due to the reduced readout error from averaging promoter switching events, until a certain value, $\Delta X_{\min}(T)$. As H approaches N, the readout error increases drastically since $CV_P \to 1$ (Fig. 2C). As a result, the value of the positional resolution ΔX increases and converges to a fixed value $\Delta X_N \approx 24\%$ EL independently of T (see SI section 7).

We asked what values of Hill coefficients give the best ability for close-by nuclei to distinguish their position along the AP axis, and whether these values change with the duration of the nuclear cycle. To this end for each steady state transcription period T, we read-off the minimal value of the positional resolution ΔX predicted by our model, $\Delta X_{\min}(T)$, from Fig. 4A to produce the orange line in Fig. 4B. We also plot the optimal Hill coefficients corresponding to the minimal value of the positional resolution, $H^* = H(\Delta X_{\min})$ as a function of T – the dashed blue line in Fig. 4B. We found that the Hill coefficients H^* that guarantee the best positional resolution decrease with the nuclear cycle duration. Since the embryo need not be performing an optimal positional readout, we found the range of Hill coefficients that allow for a margin of error of about one nucleus (2% of the embryo's length). The choice of 2% of the embryo's length is arbitrary, yet motivated by the observation that close-by nuclei do make different readout and this assumption allows us to explore the properties of the model. The solid blue lines in Fig. 4B denote a confidence interval of H that results in a positional resolution within 2% of the embryo's length of the optimal value.

We see that for short nuclear cycles (up to nc 11), the embryo can best discriminate readouts when producing a very steep pattern (intersect of dashed blue and dashed gray nc 11 line in Fig. 4B). For longer nuclear cycles (12 and 13), a narrow range of moderately steep profiles (H^* between 2 and 5) result in the smallest values of positional resolution (intersect of dashed blue and dashed gray nc 12 and nc 13 line in Fig. 4B). As the steady state transcription period T increases, ΔX becomes very small for expression profiles with a wide range of H and the constraint on H^* is relaxed (blue solid lines for large T in Fig. 4B). In this case a discernible readout owing to small values of positional resolution can be reached even for very flat expression profiles, since time averaging alone can result in reproducible readouts.

To compare the model predictions to experimental data, in Fig. 4B-D we plot the Hill coefficient (blue dot) and positional resolution (orange cross) obtained from the analysis of MS2-MCP imaging of fly embryos in nc 12 and 13 [13]. To avoid variability in the Bcd concentration between embryos, the analysis was performed by aligning 8 embryos in nc 12 and 4 embryos in nc 13 at the point of their half-maximal value of the integral fluorescence intensity. The Hill coefficients are calculated by fitting a sigmoidal curve to the mean normalized fluorescence intensity averaged over nuclei at similar positions as a function of the AP axis from data combined from multiple embryos (see SI section 12 for details). To calculate the positional resolution we take the normalized fluorescence intensity as the readout of each nucleus within a 5% EL bin around $X = 0$ and follow the procedure described above and in SI section 12. The errors bars in Fig. 4B-D for both observables represent the 95% confidence intervals. The experimental positional resolution is $\Delta X_{\text{data}} \sim 14\%$ EL (confidence interval from 11% to 20%) in nc 12 and $\Delta X_{\text{data}} \sim 12\%$ EL (confidence interval from 8% EL to 18% EL) in nc 13. The experimental Hill coefficient value is $H_{\text{data}} = 6.9$ (confidence interval $[5.80, 8.64]$, $p < 0.05$) in nc 12 and $H_{\text{data}} = 7.1$ (confidence interval $[6.20, 8.32]$, $p < 0.05$) in nc 13. The experimental positional resolution in these early nuclear cycles is well predicted by an equilibrium model with $N = 6$ binding sites (orange dashed line and orange dots in Fig. 4B), but the experimental Hill coefficient is larger than the model prediction (blue dashed line and blue dots in Fig. 4B).

D. An effective treatment of spatial averaging

To explore the effect of multiple gene copies on positional resolution we generalize the model with $M = 1$ that describes the readout from a heterogenous gene to $M = 2$, which describes a homogenous gene readout made independently in one nucleus (Fig. 3B). Although the density of nuclei does increase as nuclear cycles progress, assuming that each nuclei is making an independent measurement of the Bcd concentration ($M = 1$ for a heterogenous gene or $M = 2$ for a homogenous gene), the minimal distance between nuclei that make a distinct readout measured in units of length, will not change. However, spatial averaging of the readout concentration changes the positional resolution ΔX. In our model we account for spatial averaging of mRNA in the cytoplasmic space coming from different nuclei [35] in an effective way by assuming that the readout in a

given nucleus is the average of more than two genes ($M > 2$, Fig. 3B).

The results for the mRNA readout in a nucleus coming from a single expressing gene copy ($M = 1$ – a heterozygous fluorescent marker such as in recent MS2-MCP experiments [12]) hold for a readout coming from more gene copies ($M > 1$, Fig. 3B, SI Fig. 8). As expected, averaging over many gene copies further reduces the readout noise and slightly decreases the minimal value of positional resolution (SI Fig. 8). We opt for an effective treatment of spatial averaging at the mRNA level, since the scale of the phenomenon has not yet been quantified in experiments in nc 11-13 and a more detailed model would require making arbitrary assumptions. In general, the strength of the averaging effect is likely to increase with time, as the nuclei density increases and the nuclear cycles get longer. Our model does not capture these time dependent effects because the role of averaging is likely to be limited during the very short time of ~ 2 minutes when the steep expression pattern is established [13].

E. The non-equilibrium model

Comparing the results of the equilibrium binding site model to experimental observations, we note that the steepness values obtained in experiment ($H_{\text{data}} \sim 7$) cannot be reached by an equilibrium model with the identified $N = 6$ Bcd binding sites on the proximal hb promoter. Estrada $et\ al.$ [18] noted that this threshold of $H = N$ can be overcome with a non-equilibrium binding model. We considered a full non-equilibrium model for $N = 3$ (SI Fig. 1) and a hybrid model for $N = 6$ due to the computational complexity of performing a parameter scan of a full $N = 6$ non-equilibrium model. In the hybrid model, the promoter has 3 OS whose interactions with TF are in equilibrium and 3 OS whose interactions with TF are out-of-equilibrium (see SI - section 8).

The boundary steepness within these models can be larger than the number of operator sites ($H \leq 5$ for the $N = 3$ case, SI Fig. 11, and $H \leq 8$ for the hybrid $N = 6$ case, SI Fig. 12). However, the conclusions drawn from the equilibrium model are still valid even for $H > N$. Large Hill coefficients result in larger readout errors (SI Fig. 11 and SI Fig. 12). For the $N = 6$ hybrid model, the value of the positional resolution is minimal for large H only for very short interphase durations, and for longer interphase durations lower Hill coefficients give smaller ΔX (SI Fig. 13). For interphase durations found in the fly embryo, intermediate Hill coefficient val-

ues, $2 \leq H \leq 5$, provide the best positional resolution of ~ 6 to 10 % EL or 6 to 7 nuclei lengths (Fig. 4C), smaller than the observed experimental values of $\sim 14\%$ EL for nc 12 and 12% EL for nc 13 [13] (orange crosses with error bars in Fig. 4C, see SI - section 12).

F. "K-or-more" model

Until now we assumed that the gene is read out only if all the binding sites are occupied. We relax this assumptions and consider the equilibrium "K-or-more" model ($P_{\text{active}} \equiv P_{i \geq K}$, $1 < K < N$), where the gene is transcribed if at least K binding sites are occupied, assuming for simplicity that transcription occurs at the same rate regardless of the promoter state. As in the "all-or-nothing" model, the attainable pattern steepness is also bounded by the number of OS ($H \leq N - \tau_{\text{bind}}/\tau_{\text{active}}$), but to achieve a specific steepness H, the τ_{active} in the "K-or-more" model is $N - 1$ times smaller than that of "all-or-nothing" model. However, since the deactivation process now involves several reversible steps, τ_{active} is also noisier. As a result, the "K-or-more" model has only a slightly faster pattern formation time and slightly lower readout error than the "all-or-nothing" case (SI Fig. 14). In general, the 'K-or-more" setup does not change the conclusions about the parameter regimes where the minimal value of the positional resolution ΔX can be obtained (SI Fig. 15).

G. Transcription pattern formed by additional transcription factor gradients

We also investigated whether two mirrored transcription factor gradients, one anterior activator TF and one posterior repressor TF', could lower the predicted pattern steepness, at the same time keeping low values of positional resolution. While there is no direct evidence for additional regulatory gradients acting in the early nuclear cycles, the idea of an inverse gradient, possibly indirectly due to Caudal, has been suggested [39]. We assume $N = 6$ binding sites for the Anterior-Posterior decreasing gradient (TF) and $L = 6$ binding sites for Posterior-Anterior decreasing (TF') gradient. Transcription is allowed only when the promoter is fully bound by TF and free of TF' and we assume that the interactions of TF and TF' with the promoter are independent (see SI section 10). In the equilibrium model, the pattern can achieve a maximum steepness of $H^* \sim 7$

given the total of 12 binding sites (Fig. 4D). The quantitative conclusions are the same as for the previously considered models (SI Fig 16 and SI Fig. 17) but the minimal value of the positional resolution ($\Delta X \sim 10\%$ EL in nc 12 and nc 13) is smaller than that achieved with a single TF gradient, and smaller than observed experimentally.

Lastly, we investigated the pattern formation when an additional repressor is concentrated in the mid-embryo region (see SI section 10.2). This scenario is motivated by the known pattern of the Capicua (Cic) protein and its potential effect on transcription. In the hb promoter sequence there is one known binding motif for the Cic protein [40]. Since the Cic concentration is relatively constant at the hb pattern boundary ($\sim -5\%$ EL from mid-embryo), Cic does not affect the pattern steepness. We also find that the Cic gradient contributes little to the positional resolution of the hb pattern (SI Fig. 18).

H. A common Hill coefficient for all nuclear cycles

Since the interphase duration varies during the early development phase but the molecular encoding of regulation is unlikely to change, we can use the results of the simplest equilibrium model Fig. 4B to define a value of a Hill coefficient, H_{robust}, that gives the minimal value of the positional resolution in the widest range of steady state transcription periods T (see Fig. 5 inset) as a function of the number of operator sites (N) for different numbers of expressing gene copies (M). For $M = 1$, H_{robust} is slightly greater than $N/2$, resulting in not so steep boundaries (Fig. 5). H_{robust} increases with M but is always smaller than its highest possible value of N allowed by the equilibrium model, even for very large numbers of expressing genes.

The optimal value of the Hill coefficients in nc 12 and 13 for all the considered models, as well as the H_{robust} values, are all between $H \sim 2 - 4$. These values are in very good agreement with *in vitro* experiments that measured the cooperativity of 6 Bcd binding sites on the hb promoter [20, 41] ($H_{\mathrm{data}} \sim 3$).

I. Comparison to experimental data

Comparing the model predictions to the experimental data [13], one can construct an equilibrium model that correctly captures the experimentally ob-
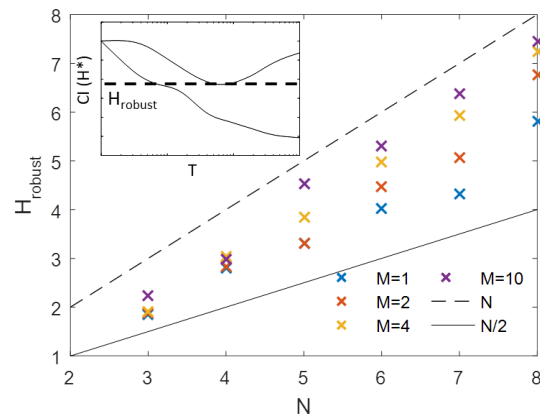

FIG. 5: **The Hill coefficient H_{robust} that gives the best positional resolution for the widest range of steady state transcription periods for the equilibrium model with $N = 6$ binding sites.**
Since the regulatory parameters are unlikely to change between nc 11-13, we look for the model quantified by its Hill coefficient that results in the lowest values of the positional resolution in the largest range of T. Inset: H_{robust} is calculated as the minimum of the upper bound of the confidence interval of the optimal Hill coefficient $CI(H^*)$, where the optimal Hill coefficient H^* is defined in Fig. 4A, and the confidence intervals are the solid blue lines in Fig. 4B. H_{robust} is shown as a function of the OS number (N) and number of readout genes (M). Also plotted for reference are $H = N$ (dashed line) and $H = N/2$ (solid line). H_{robust} for $M = 1$ is much less than the maximum steepness allowed by the model $H = N$ and with a lot of spatial averaging ($M = 10$) H_{robust} approaches N.

served positional resolution, but it is much harder to achieve the readout steepness observed from the endogenous promoter given the currently identified number of binding sites. As has been shown before [18], non-equilibrium models allow for steeper expression profiles. However, increasing the Hill coefficients to $H_{\mathrm{data}} \sim 7$ [13, 16] also increases the minimal obtainable value of the positional resolution within a hybrid non-equilibrium model to $\Delta X \sim 20\%$ EL (~ 10 nuclei widths), slightly above the the experimentally observed value of $\Delta X_{\mathrm{data}} \sim 12\%$ EL in nc 13 (~ 6 nuclei widths) (Fig. 4 C). Unfortunately, from the experimental data it is hard to reliably extract Hill coefficients for nc 11.

Steep boundaries are only possible if the promoter spends most of its time in the fully occupied or fully bound states, which sets boundaries on the switching parameters [30, 42] (SI Fig. 6). We looked for the kinetic parameter set that yields the smallest positional resolution ΔX and, although the values vary

with the interphase duration, we find that a parameter set that results in the experimentally observed $\Delta X_{\text{data}} \sim 12\%$ EL in nc 12 does not change over multiple nuclear cycles of varying duration. This stability throughout the nuclear cycles is consistent with experimental observations that the Bcd interactions with the hb promoter are likely independent of other TF, which suggests the binding rate constant coefficients are independent of the nuclei's positions along the AP axes [43].

Varying the only parameter of the model τ_{bind}, which is set by the 3D diffusion assumption, rescales the steady state transcription period T (see SI Fig. 19). However, this rescaling does not quantitatively change the conclusions of our analysis for the equilibrium models, since only the non-equilibrium model with $N = 6$ binding sites is able to produce boundaries as steep as those observed in the experiments (Fig. 4C). Within a non-equilibrium model longer binding timescales ($\tau_{\text{bind}} = 40s$) than currently estimated within the diffusion approximation ($\tau_{\text{bind}} \sim 4s$) result in a model that reproduces the observed steepness in nc 11-13 (SI Fig. 19 B) but, as discussed above, also results in a much higher minimal value of the positional resolution. Conversely, short binding timescales ($\tau_{\text{bind}} = 0.4s$) allow the model to reach very low values of positional resolution in models with much smaller corresponding Hill coefficients than H_{data} (SI Fig. 19 A).

We also asked what value of the binding rate τ_{bind} in a non-equilibrium model results in both Hill coefficients and positional resolution that is consistent with experimentally observed values. For this, we calculate the positional resolution as a function of τ_{bind} and randomized the remaining set of binding and unbinding parameters to achieve the experimentally observed $H_{\text{data}} \approx 7$ and the lowest value of the positional resolution given the fixed H_{data} constraint (see SI Fig.20). The difference with the analysis in SI Fig. 19 is that now we add an additional constraint on $H = H_{\text{data}}$, so the minimal value of positional resolution is greater than in the results in SI Fig. 19. We find that for small values of $\tau_{\text{bind}} \sim 0.01$s, the mean values of the experimentally observed positional resolution ($\Delta X_{\text{data}} \approx 14\%$ EL in nc12 and $\Delta X_{\text{data}} \approx 12\%$ EL in nc13) are close to the minimal value calculated in the model (SI Fig.20). Taking into account the confidence interval of the experimentally measured positional resolution, the experimental values are very close to the minimal predicted values of positional resolution even for $\tau_{\text{bind}} \sim 0.1s$. We conclude that a hybrid non-equilibrium model with $N = 6$ binding sites can reproduce both the experimentally observed Hill

steepness and positional resolution, if the binding timescales are smaller than currently estimated.

Achieving small $\tau_{\text{bind}} \sim 0.1 - 0.4s$ requires a diffusion coefficient of $D \sim 100\mu m^2/s$, which seems an order of magnitude larger than the current estimates ($D \sim 7.4\mu m^2/s$) [9, 33]. Misestimates in $\tau_{\text{bind}} = 1/(Dac[TF])$ coming from the binding site size a and Bcd concentration $[TF]$ separately are unlikely to be at the origin of such a large difference. Even considering a combined effect of a misestimate in the binding site size, Bcd concentration and the diffusion coefficient, the diffusion coefficient would need to be an order of magnitude larger. However, a different diffusion model, such as a combination of a 1D and 3D TF search for the operator site [44] could help lower the binding timescale. As a result, a non-equilibrium model with a slight modification (additional binding site, additional regulation) and a smaller binding rate does seem a likely candidate for explaining the experimental data.

We compare the readout error $\delta mRNA/\langle mRNA \rangle$ calculated directly from the MS2-MCP experiments in nc 12 and nc 13 (SI Fig. 22). The experimental readout error in nc 12 is $\delta mRNA/\langle mRNA \rangle = 0.82$ and in nc 13 is $\delta mRNA/\langle mRNA \rangle = 0.69$, which are lower than expected from the equilibrium model $CV_P \sim 1$ for the maximum allowed Hill coefficient of $N = 6$, but higher than the $CV_P \sim 0.45$ in nc 12 and $CV_P \sim 0.25$ in nc 13 for the non-equilibrium hybrid model that yields the minimal value of the positional resolution. The higher experimentally observed readout error may be due to the the fact that the living embryo does not saturate the lower bound of positional resolution, as well as additional sources of noise in the experiments that are not considered in this model. These sources of noise include the random arrival times of RNA polymerases [45], non-uniform progression of the polymerases along the DNA [46] or additional modes of regulation that manifest themselves in bursty expression even in the anterior region where Bcd binding should be saturated [25, 47], and possibly experimental noise. To focus on the regulatory architecture, following previous work [48–52], we assumed the mean expression and noise at the promoter level is correlated with the mRNA readout. Exploring the role of these different sources of noise that lead to the observed readout error in conjunction with binding models of different complexity remains a future direction.

The error values reported above are also less than the previously reported $\delta mRNA/\langle mRNA \rangle \sim 1.5$ [25] for nuclei in a 10% EL strip centered at mid-embryo for the same 4 nuclei in nc 13. In the previous analysis the embryos where aligned in the mid-

dle of the embryo (0% EL), which is close to the half-maximal expression point based on the mean probability of the nuclei to transcribe the gene at any point during the interphase. In the current analysis, based on the discussion in the experimental companion paper [13], we align the embryos at their half-maximal expression point of the integral fluorescence intensity, which is typically positioned anterior to the middle of the embryo at $\sim -5\%$ EL. These results suggest that either fluctuations in the Bicoid concentration between embryos influence $\delta mRNA/\langle mRNA \rangle$, or that nuclei that are positioned to the posterior of the mid-boundary point ($X = 0$) contribute more to the readout error, which is likely due to their lower expression probability. We confirm the latter hypothesis by finding for the 4 embryos aligned at $X = 0$ in nc 13 the CV_P in a 5% strip around 0% EL, $CV_P = 1.78$, which is larger than the $CV_P = 0.69$ in the strip centered at -5% EL. However, without experiments that simultaneously measure Bcd concentrations and hb expression, we cannot rule out that fluctuations in the Bicoid concentration also play a role.

To explore the role of a transcriptional repressor in these trade-offs, we also considered the possibility of binding sites for an inversely directed gradient. The choice of a gradient repressor was arbitrary, since the only known mirror gradient in early fly development, Caudal, has no known binding sites in the hb promoter, no known repressor function in fly development and its maternal component has been shown to be non-essential in early fly development [53, 54]. Nevertheless it provided for a simple choice of parameters and was motivated by earlier theoretical ideas [39], and known activator-repressor pairs in other systems [55]. Since we are only looking at a small part of the embryo the precise form of the gradient will not strongly influence our qualitative conclusions, so we opted for the mirror image for simplicity. This two gradient model, even in its equilibrium version, does decrease the positional resolution in short cell cycles while increasing the steepness of the expression profile (Fig. 4D). Again, the exact results of the model do not position the experimental results for the endogenous promoter within the predictions of the model, but for the two TF gradient the minimal value of the positional resolution observed at nc 12 is obtained at earlier nc with $H^* \sim 7$, $\Delta X \sim 16\%$ EL is not far from the experimentally measured value of $\Delta X_{\text{data}} \sim 12\%$ EL in nc 13 (Fig. 4D). Together these results suggest that a repressor gradient working together with Bcd in a non-equilibrium setting, possibly with additional Bcd or Hb binding sites, could explain all

of the experimentally observed results. Following the above results for different binding timescales (SI Fig.20), an equilibrium repressor gradient model with a smaller τ_{bind} is another way to agree the model and the data.

There are also other repressor candidates in the fly development, such as Capicua, which is a known repressor gradient albeit with a different profile [56, 57]. For simplicity, motivated by Capicua, we studied a model with a constant additional repressor gradient in the middle of the embryo. Not surprisingly, due to its symmetry around the boundary, this type of gradient neither increases steepness nor severely modifies the readout error.

IV. DISCUSSION

In order to better understand the trade-off between short cell cycles, steepness, readout error and positional resolution we studied a family of models where transcription is controlled by the binding and unbinding of the Bcd TF to multiple operator sites on the hb promoter: equilibrium binding models with different expression rules, non-equilibrium models and equilibrium models with two TF gradients.

One possible way to reconcile steep profiles with small values of positional resolution are additional unidentified binding sites in the promoter. Currently the minimal hb promoter used in the experiments we are analyzing [13] is known to consist of 6 Bcd binding sites, one proximal and one distal Hb binding site. Of course, it could also include unidentified binding sites. Since we were interested in nc 11-13 – the early cell cycles when the profile is already steep – we did not include the Hb binding sites in our analysis. At that stage of development the zygotic Hb gradient is weak, although there exists a maternal step-like Hb profile with a smaller amplitude than the final zygotic profile [8]. Since these Hunchback gradients have the same direction as Bcd, Hb binding sites would most likely have the same effect as additional Bcd sites so we did not add them to the model promoter. However due to the step-like shape with a boundary in the middle of the emrbyo, maternal hunchback may play a role in establishing the steep profile. The usually characterized minimal hb promoter also includes one to two Zelda binding sites but they either do not change or they decrease the pattern steepness [13]. Nevertheless additional unknown Bcd binding sites would certainly increase steepness, as could Hb binding sites.

The disagreement between the model and the data

is not manifested by the fact that the experimental points do not precisely fall on the theoretical predictions. The fly embryo does not need to function close to the optimal parameter regime and probably it does not. The disagreement arises because the experimentally measured values of these two observables, the Hill coefficient and the positional resolution, cannot be simultaneously obtained within the current regulatory model with the experimentally estimated diffusion limited binding time. In general, within the current models, steep boundaries increase the minimal obtainable value of the positional resolution. Specifically, the results of the model tell us that in the case of the observed steep profiles the best positional resolution that can be achieved has a much larger value than is experimentally measured. Since this is the minimal value of the positional resolution, the experimentally observed value of the positional resolution must be larger. Yet, in experiments we observe much smaller values of the positional resolution. This suggests different modes of regulation, such as described above, or smaller binding timescales than currently estimated (SI Fig. 20). Yet if this process is fast, and early fly development is very fast, the undiscovered modes of regulation have to be simple [43].

Another explanation to consider for the discrepancy between the experimental observations and our current discussion of the model is that the assumption we made about the positional error being minimized in the developing embryo is not valid. However, even if we relax this assumption, the general conclusions do not change: the Bcd-only equilibrium $N = 6$ model is not compatible with the experimentally observed Hill coefficient regardless of this assumption and in the hybrid non-equilibirum model, the predicted positional regulation for the experimentally observed Hill coefficient values within the current regulatory model is larger than the observed positional regulation (SI Fig. 13). Relaxing this assumption does, however, make it even more likely for models with different binding timescales or additional regulators or binding sites can explain both the observables simultaneously.

The observed steep boundaries minimize the positional resolution only for very short cell cycles. Another possible regulatory strategy involves setting up an imprecise boundary with low positional resolution at nc 11 using a steep expression profile. This boundary would further be refined during the following cell cycles, using additional regulatory mechanisms, such as Hb regulation or epigenetic modifications encoding memory in the translational state [9], leading to lower positional regula-

tion. We also demonstrated that if the system starts from an out-of-steady-state condition after mitosis, the interphase duration may not be long enough for steep steady state expression patterns to establish (SI Fig. 3). This may lead the pattern to shift along the AP axis from nuclear cycle to nuclear cycle, as observed in fly development [9].

The "all or nothing" model is clearly a simplifying assumption but we have shown that a "K-or-more" model does not change the quantitative conclusions. In the "K-or-more" model, we further, incorrectly, assume that the transcription rate is the same for all of the promoter states that enable transcription. However, given the generality of our conclusions, introducing intermediate transcription rates would change the precise numerical values of the achievable positional resolution but not the general constraints on steepness and the positional resolution.

As has been pointed out in the context of maximizing information flow between the Bcd gradient and Hb output [38], very steep boundaries decrease the ability of the nuclei to discriminate between similar Bcd concentrations. The optimal expression profiles for minimizing positional resolution are always relatively steep $H > 1$, since large input fluctuations in the posterior end of the embryo coming from small Bcd concentrations limit extremely flat expression profiles. In general, we give a real biological example of the previously identified phenomenon that utlrasensitive systems require extremely slow receptor switching dynamics, which results in increased errors at the single-cell readout level [58]. Other trade-offs imposed by a need for a precise or informative readout have also been explored, including the energy–speed–accuracy constraint that shows that these three quantitates cannot be simultaneously optimized [59] or the cost of optimal information transmission in a finite time [60].

The variability in the expression states of different nuclei in the considered models comes from the binding and unbinding noise of TF to OS. The binding rates are assumed to be diffusion limited, which we implement using the Berg-Purcell bound [26]. In order to concentrate on the trade-off between steepness and positional resolution and simplify the parameter space exploration, we make the simplifying assumption that the binding and unbinding dynamics are uncoupled from diffusion. This approximation means that after an unbinding event the TF diffuses far enough from the OS so that it does not have an increased probability of binding compared to other TF molecules and its rebinding can be considered as an independent event [29]. For the equilibrium model, where all binding sites are the

same, allowing for fast rebinding renormalizes the binding rates depending on the number of available free binding sites [29]. This renormalization would rescale the time axes to shorter times (or shift the time axis to the left on the log scale), but would not qualitatively change the discussed results (see SI Fig. 19). The effects of the full model of coupled binding and diffusion in the non-equilibrium model remain to be investigated in detail. Coupling the search process to the non-equilibrium process is also interesting in light of recent experimental evidence of two Bcd populations, one that spends a long time bound ($\sim 1s$)($< 0.1s$) to the DNA, and the other that spends a short time bound ($< 0.1s$) [61], which could be a manifestation of specific or non-specific rebinding.

We compared the experimentally measured positional resolution and steepness in the MS2-MCP experiments [12, 13, 24] to the $M = 1$ "all or nothing" model, since these experiments look at heterozygous constructs. The developing fly embryo is homozygous and has M=2 genes, and the total resolution of the gene readout that matters for downstream genes should be determined at the protein level. Therefore the overall resolution at the protein level is different than measured by the MS2-MCP system [15].

At the protein level Gregor *et al.* [15] measured a Hill coefficient of $H_{\mathrm{data}}^{\mathrm{protein}} = 5$ in nc 14 and concluded that within the equilibrium limit of $H \leq N$ the known six binding sites are sufficient to achieve this steepness. In this work we consider the steepness of the *mRNA* readout in nc 13 and earlier, which is steeper ($H_{\mathrm{data}} \sim 7$) than the protein boundary at later cycles [13]. Our results therefore do not contradict previous observations [15]. The protein boundary is likely to benefit from averaging of protein concentrations between nuclei [15, 35]. The fast timescale of about 2 minutes for achieving the steep mRNA boundary [13] suggests that the readout mechanism initially produces a steeper boundary, which is then made less steep with time, possibly due to diffusion [35]. While spatial averaging is clearly important for Hb proteins [15, 35], given that the steep expression profile is established in ~ 2 minutes [13], spatial averaging of *hb* mRNA in nc 11-12 probably plays a smaller role.

Inspired by the experiments of Lucas *et al.* [13] we focused on nc 11-13. The *hb* gene is also expressed during later stages of development [62–64]. In nc 14, additionally to the proximal promoter active in nc 11-13, expression of *hb* is also controlled by distal and shadow enhancers [1, 2, 43]. However they are unlikely to play a major role in nc 11-13. Recent studies have also used an optogenetically modifiable

Bcd protein [65] that makes it possible to modify the transcription of Bcd target genes. Combining all these experimental approaches with the knowledge gained both about *hb* mRNA [12, 13, 24] and Hb protein regulation [36] is a much needed future direction.

In summary, we show how trade-offs between steep expression profiles and positional resolution influence the possible regulatory modes of *hb* expression in the short early cell cycles of fly development. We propose a number of possible solutions from non-equilibrium binding, additional regulatory gradients and binding sites, faster binding rates to epigenetic regulation. Additional experiments are needed to discriminate between the proposed scenarios. For example, testing whether the binding of TF to the promoter is equilibrium or non-equilibrium requires analysis of experiments that track TF bound to fluorescent probes that follow their binding and unbinding. Equilibrium dynamics results in time reversible traces – a property that can be evaluated based on such tagged TF data collected using high resolution microscopy.

V.   METHODS

A.   Model of promoter dynamics

The general model of transcription regulation through transcription factor (TF) binding/unbinding to the operator sites (OS) is based on the graph-based framework of biochemical systems [18, 66]. In short, for a promoter with N TF binding sites the model considers all the possible 2^N promoter occupancy states and all transitions between these states that involve the binding and unbinding of one TF. In most treatments of the model we randomize parameters to explore its behavior. The full non-equilibrium model is described in SI - section 1 and solved numerically. Assuming the binding sites are indistinguishable results in the one dimensional equilibrium model in Eq. 2.

B.   Randomization of kinetic parameters

The kinetic rate constants are randomized in $\mathbb{R}^+$ space. Assuming binding is diffusion limited by the Berg-Purcell limit [26], the binding rate constants k_i have an upper bound depending on the OS search time τ_{bind}. Based on measured and typically taken parameters for diffusion, concentration and operator size the we estimate $\tau_{\mathrm{bind}} = 4s$ (SI section 2). How-

ever, since we randomize the parameters, our quantitative conclusions do not depend on the exact values taken for these parameters. For the non-equilibrium model, i ranges from 1 to 2^{N-1} and there are no further constraints on the binding rates. For the equilibrium model, a reaction from P_{i-1} to P_i is the binding of a TF to one of the remaining $N-i+1$ free OS, so the rate constants k_{+i} are bound by $(N-i+1)/\tau_{\mathrm{bind}}$. There are no bounds on the unbinding rate constants k_{-i}, but their values are rescaled *a posteriori* so that the boundary is located in the middle of the embryo ($P(P_{active}, X = 0) = 0.5$, see SI - section 3). The values of the rate constants are sampled to be uniformly distributed on the logarithmic scale, from $10^{-20}\ s^{-1}$ to $10^{20}\ s^{-1}$. The number of randomized configurations tested is on the order of 10^5.

C. Calculating the positional resolution

To find the value of ΔX for a specific kinetic parameter set, we test the condition $P(F_+ \leq F_-) \leq 0.05$ with increasing nuclei distance ΔW. The distribution of F_+ and F_- is taken as the marginal distributions of the gene readout from 500 stochastic simulation runs (SSA) [67, 68] implemented in the SGNS2 simulator [69]. F_- and F_+ are not well-fit by Gaussian distributions, especially for short interphase durations. ΔX is the smallest value of ΔW yielding a tolerable error of $P(F_+ \leq F_-) \leq 0.05$. ΔX and ΔW and nuclei position X can be expressed in units of length relative to the decay length of the TF gradient $\lambda \approx 100\mu m$ [5], which corresponds to $\sim 20\%$ of the embryo length (EL).

D. Experimental data

The data on the dynamics of *hb* pattern are taken from Lucas *et al.* 2018 [13]. In this work, *hb* transcription in nuclear cycle 11 to 13 is monitored using the MS2-MCP RNA tagging system [12, 14]. From the total amount of mRNA produced per nuclei at any given position, we extracted the pattern steepness (H_{data}) and positional resolution (ΔX_{data}) (See SI section 12).

[1] Perry MW, Bothma JP, Luu RD, Levine M. Precision of hunchback expression in the Drosophila embryo. Current Biology. 2012;22(23):2247–2252.

[2] Perry MW, Boettiger AN, Levine M. Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(33):1–12.

[3] Houchmandzadeh B, Wieschaus E, Leibler S. Establishment of developmental precision and proportions in the early Drosophila embryo. Nature. 2002;415(6873):798–802.

[4] Gregor T, Bialek W, de Ruyter van Steveninck RR, Tank DW, Wieschaus E. Diffusion and scaling during early embryonic pattern formation. Proceedings of the National Academy of Sciences. 2005;102(51):18403–18407.

[5] Gregor T, Wieschaus E, McGregor AP, Bialek W, Tank DW. Stability and nuclear dynamics of the Bicoid morphogen gradient. Cell. 2007;130(1):141–152.

[6] Driever W, Nusslein-Volhard C. The Bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. Cell.

1988;54(1):95–104.

[7] Driever W, Thoma G, Nüsslein-Volhard C. Determination of spatial domains of zygotic gene expression in the Drosophila embryo by the affinity of binding sites for the bicoid morphogen. Nature. 1989 aug;340(6232):363–367.

[8] Crauk O, Dostatni N. Bicoid determines sharp and precise target gene expression in the Drosophila embryo. Current Biology. 2005;15(21):1888–1898.

[9] Porcher A, Abu-Arish A, Huart S, Roelens B, Fradin C, Dostatni N. The time to measure positional information: maternal hunchback is required for the synchrony of the Bicoid transcriptional response at the onset of zygotic transcription. Development. 2010;137(16):2795–804.

[10] Jaeger J. The gap gene network. Cellular and Molecular Life Sciences. 2011;68(2):243–274.

[11] Porcher A, Dostatni N. The Bicoid morphogen system. Current Biology. 2010;20(5):249–254.

[12] Lucas T, Ferraro T, Roelens B, Chanes JDLH, Walczak AM, Coppey M, et al. Live imaging of Bicoid-dependent transcription in Drosophila embryos. Current Biology. 2013;23(21):2135–2139.

[13] Lucas T, Tran H, Perez Romero CA, Mathieu

Coppey, Fradin C, Walczak AM, et al. Precision in a rush: trade-offs between positioning and steepness of the hunchback expression pattern. bioRxiv. 2018;305532.

[14] Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM. Localization of ASH1 mRNA particles in living yeast. Molecular Cell. 1998;2(4):437–445.

[15] Gregor T, Tank DW, Wieschaus E, Bialek W. Probing the limits to positional information. Cell. 2007;130(1):153–164.

[16] Xu H, Sepúlveda LA, Figard L, Sokac AM, Golding I. Combining protein and mRNA quantification to decipher transcriptional regulation. Nature methods. 2015;12(8):739–42.

[17] Lopes FJ, Spirov AV, Bisch PM. The role of Bicoid cooperative binding in the patterning of sharp borders in Drosophila melanogaster. Developmental Biology. 2011;72(2):181–204.

[18] Estrada J, Wong F, DePace A, Gunawardena J. Information integration and energy expenditure in gene regulation. Cell. 2016;166(1):234–244.

[19] Ochoa-Espinosa A, Yucel G, Kaplan L, Pare A, Pura N, Oberstein A, et al. The role of binding site cluster strength in Bicoid-dependent patterning in Drosophila. Proceedings of the National Academy of Sciences. 2005;102(14):4960–4965.

[20] Ma X, Yuan D, Diepold K, Scarborough T, Ma J. The Drosophila morphogenetic protein Bicoid binds DNA cooperatively. Development (Cambridge, England). 1996;122(4):1195–206.

[21] Sourjik V, Berg HC. Receptor sensitivity in bacterial chemotaxis. Proceedings of the National Academy of Sciences of the United States of America. 2002;99(1):123–127.

[22] Bray D, Levin MD, Morton-Firth CJ. Receptor clustering as a cellular mechanism to control sensitivity. Nature. 1998 may;393(6680):85–88.

[23] Ozbudak EM, Thattai M, Lim HN, Shraiman BI, Van Oudenaarden A. Multistability in the lactose utilization network of Escherichia coli. Nature. 2004 feb;427(6976):737–740.

[24] Garcia HG, Tikhonov M, Lin A, Gregor T. Quantitative imaging of transcription in living Drosophila embryos links polymerase activity to patterning. Current Biology. 2013;23(21):2140–2145.

[25] Desponds J, Tran H, Ferraro T, Lucas T, Perez Romero C, Guillou A, et al. Precision of readout at the hunchback gene: analyzing short transcription time traces in living fly embryos. PLOS Computational Biology. 2016;12(12):e1005256.

[26] Berg HC, Purcell EM. Physics of chemoreception. Biophysical Journal. 1977;20(2):193–219.

[27] Bialek W, Setayeshgar S. Physical limits to biochemical signaling. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(29):10040–5.

[28] Berezhkovskii AM, Szabo A. Effect of ligand diffusion on occupancy fluctuations of cell-surface recep-

tors. Journal of Chemical Physics. 2013;139(12):1–7.

[29] Kaizu K, De Ronde W, Paijmans J, Takahashi K, Tostevin F, Wolde PRT. The Berg-Purcell limit revisited. Biophysical Journal. 2014;106(4):976–985.

[30] Skoge M, Meir Y, Wingreen NS. Dynamics of cooperativity in chemical sensing among cell-surface receptors. Physical Review Letters. 2011;107(17):1–5.

[31] Bialek W, Setayeshgar S. Cooperativity, sensitivity, and noise in biochemical signaling. Physical Review Letters. 2008;100(25):1–4.

[32] Hill AV. The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. Journal of Physiology. 1910;40:4–7.

[33] Abu-Arish A, Porcher A, Czerwonka A, Dostatni N, Fradin C. High mobility of bicoid captured by fluorescence correlation spectroscopy: Implication for the rapid establishment of its gradient. Biophysical Journal. 2010;99(4):33–35.

[34] ten Wolde PR, Becker NB, Ouldridge TE, Mugler A. Fundamental limits to cellular sensing. Journal of Statistical Physics. 2016;162(5):1395–1424.

[35] Erdmann T, Howard M, Ten Wolde PR. Role of spatial averaging in the precision of gene expression patterns. Physical Review Letters. 2009;103(25):2–5.

[36] Bothma JP, Norstad MR, Alamos S, Garcia HG. LlamaTags: A versatile tool to image transcription factor dynamics in live embryos. Cell. 2018;173(7):1–13.

[37] Dubuis JO, Tkacik G, Wieschaus EF, Gregor T, Bialek W. Positional information, in bits. Proceedings of the National Academy of Sciences. 2013;110(41):16301–16308.

[38] Tkacik G, Dubuis JO, Petkova MD, Gregor T. Positional information, positional error, and readout precision in morphogenesis: A mathematical framework. Genetics. 2015;199(1):39–59.

[39] Sokolowski TR, Erdmann T, ten Wolde PR. Mutual repression enhances the steepness and precision of gene expression boundaries. PLoS Computational Biology. 2012;8(8).

[40] Jiménez G, Guichet A, Ephrussi A, Casanova J. Relief of gene repression by Torso RTK signaling: Role of capicua in Drosophila terminal and dorsoventral patterning. Genes and Development. 2000;14(2):224–231.

[41] Burz DS, Rivera-Pomar R, Jäckle H, Hanes SD. Cooperative DNA-binding by Bicoid provides a mechanism for threshold-dependent gene activation in the Drosophila embryo. EMBO Journal. 1998;17(20):5998–6009.

[42] Sevier SA, Levine H. Properties of cooperatively induced phases in sensing models. Physical Review E. 2015 may;91(5):052707.

[43] Hannon CE, Blythe SA, Wieschaus EF. Concentration dependent chromatin states induced by the bicoid morphogen gradient. eLife. 2017;6:e28275:1–29.

[44] Slutsky M, Mirny LA. Kinetics of protein-DNA interaction: Facilitated target location in sequence-dependent potential. Biophysical Journal. 2004;87(6):4021–4035.

[45] Zoller B, Little SC, Gregor T. Diverse spatial expression patterns emerge from common transcription bursting kinetics.

[46] Coulon A, Ferguson ML, de Turris V, Palangat M, Chow CC, Larson DR. Kinetic competition during the transcription cycle results in stochastic RNA processing. eLife. 2014;3:1–22.

[47] Xu H, Skinner SO, Sokac AM, Golding I. Stochastic kinetics of nascent RNA. Physical Review Letters. 2016;117(12):1–6.

[48] Garcia HG, Sanchez A, Boedicker JQ, Osborne M, Gelles J, Kondev J, et al. Operator sequence alters gene expression independently of transcription factor occupancy in bacteria. Cell Reports. 2012;2(1):150–161.

[49] Jones DL, Brewster RC, Phillips R. Promoter architecture dictates cell-to-cell variability in gene expression. Science. 2014;346(6216):1533–1536.

[50] Oliveira SMD, Häkkinen A, Lloyd-price J, Tran H, Kandavalli V, Ribeiro AS. Temperature-dependent model of multi-step transcription initiation in Escherichia coli based on live single-cell measurements. PLoS computational biology. 2016;p. 1–18.

[51] Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. Gene regulation at the single-cell level. Science. 2005 mar;307(5717):1962–5.

[52] Dunlop MJ, Cox RS, Levine JH, Murray RM, Elowitz MB. Regulatory activity revealed by dynamic correlations in gene expression noise. Nature genetics. 2008;40(12):1493–1498.

[53] Mlodzik M, Gibson G, Gehring WJ. Effects of ectopic expression of caudal during Drosophila development. Development. 1990;109(2):271–277.

[54] Olesnicky EC, Brent AE, Tonnes L, Walker M, Pultz MA, Leaf D, et al. A caudal mRNA gradient controls posterior development in the wasp Nasonia. Development. 2006 sep;133(20):3973–3982.

[55] Zagorski M, Tabata Y, Brandenberg N, Lutolf MP, Tkačik G, Bollenbach T, et al. Decoding of position in the developing neural tube from antiparallel morphogen gradients. Science. 2017;356(6345):1379–1383.

[56] Löhr U, Chung HR, Beller M, Jäckle H. Antagonistic action of Bicoid and the repressor Capicua determines the spatial limits of Drosophila head gene expression domains. Proceedings of the National Academy of Sciences. 2009;106(51):21695–21700.

[57] Löhr U, Chung HR, Beller M, Jäckle H. Bicoid - morphogen function revisited. Fly. 2010;4(3):236–240.

[58] Hu B, Rappel WJ, Levine H. How input noise limits biochemical sensing in ultrasensitive systems. Physical Review E. 2014 sep;90(3):032702.

[59] Wong F, Amir A, Gunawardena J. An energy-speed-accuracy relation in complex networks for biological discrimination. arXiv. 2017;.

[60] Mancini F, Marsili M, Walczak AM. Trade-offs in delayed information transmission in biochemical networks. Journal of Statistical Physics. 2015 apr;162(5):1088–1129.

[61] Mir M, Reimer A, Haines JE, Li XY, Stadler M, Garcia H, et al. Dense bicoid hubs accentuate binding along the morphogen gradient. Genes and Development. 2017 sep;31(17):1784–1794.

[62] Margolis JS, Borowsky ML, Steingrímsson E, Shim CW, Lengyel JA, Posakony JW. Posterior stripe expression of hunchback is driven from two promoters by a common enhancer element. Development. 1995;121(9):3067–3077.

[63] Kambadur R, Koizumi K, Stivers C, Nagle J, Poole SJ, Odenwald WF. Regulation of POU genes by castor and hunchback establishes layered compartments in the Drosophila CNS. Genes & Development. 1998;12(2):246–260.

[64] Wimmer EA, Carleton A, Harjes P, Turner T, Desplan C. bicoid-independent formation of thoracic segments in Drosophila. Science. 2000 mar;287(5462):2476–2479.

[65] Huang A, Amourda C, Zhang S, Tolwinski NS, Saunders TE. Decoding temporal interpretation of the morphogen bicoid in the early drosophila embryo. eLife. 2017;6:1–21.

[66] Gunawardena J. A linear framework for time-scale separation in nonlinear biochemical systems. PLoS ONE. 2012;7(5).

[67] Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. Journal of Computational Physics. 1976;22(4):403–434.

[68] Gillespie DT. Exact stochastic simulation of coupled chemical reactions. Journal of Physical Chemistry. 1977;81(25):2340–2361.

[69] Lloyd-Price J, Gupta A, Ribeiro AS. SGNS2: A compartmentalized stochastic chemical kinetics simulator for dynamic cell populations. Bioinformatics. 2012;28(22):3004–3005.

Precision in a rush: trade-offs between reproducibility and steepness of the hunchback expression pattern

Supporting Information

September 7, 2018

1 Model of transcription regulation

1.1 The binding site model

We use a model of transcription factors (TF) binding/unbinding to operator sites (OS) based on the graph-based method introduced in [1] and implemented in [2]. Within this framework one considers a finite, connected, labeled, directed graph $[\overline{v}, \overline{e}]$, where the vertices $\overline{v}$ describe the promoter micro-states corresponding to the number and arrangement of bound TF on the promoter OS array, and the edges $\overline{e}$ are transitions between micro-states. The edge labels $\overline{k}$ are infinitesimal transition rates for a Markov process.

For a gene whose promoter has N OS, the number of micro-states/vertices is 2^N:

$$\overline{v} = [v_1, v_2 ... v_{2^N}]. \tag{1}$$

Since we assume only one binding/unbinding event can take place at a time, not all vertices are directly connected and the number of edges is $N_{\text{edges}} = N2^N$ [2].

We separate the edges $\overline{e}$ into two sets, corresponding to forward (binding) transitions (e_+) and backward (unbinding) transitions (e_-):

$$e = [e_+, e_-] = [e_{+1}, e_{+2}, ... e_{+N2^{N-1}}, e_{-1}, e_{-2}, ... e_{-N2^{N-1}}], \tag{2}$$

with the corresponding edge labels describing the reaction rate constants of binding and unbinding between the TF and the OS:

$$\overline{k} = [k_+, k_-] = [k_{+1}, k_{+2}, ... k_{+N2^{N-1}}, k_{-1}, k_{-2}, ... k_{-N2^{N-1}}]. \tag{3}$$

An example of the labeled graph is shown in SI Fig. 1 for $N = 3$.

The micro states of the gene are divided into active (when the gene is expressed) and inactive (when the gene is not expressed) states. In this work, we assume that the gene is activated only when the OS are bound by at least K TFs (corresponding to the "K-or-nothing" case in [2], with $K \leq N$). During this active state window RNA polymerases can bind to the target promoter to initiate transcription with a rate that is much faster than the rate of gene activation. Therefore, the mean transcription rate and the mean expression values of a gene can be approximated by the probability of the gene to be active $\langle f(X) \rangle = P(P_{i \geq K})$.
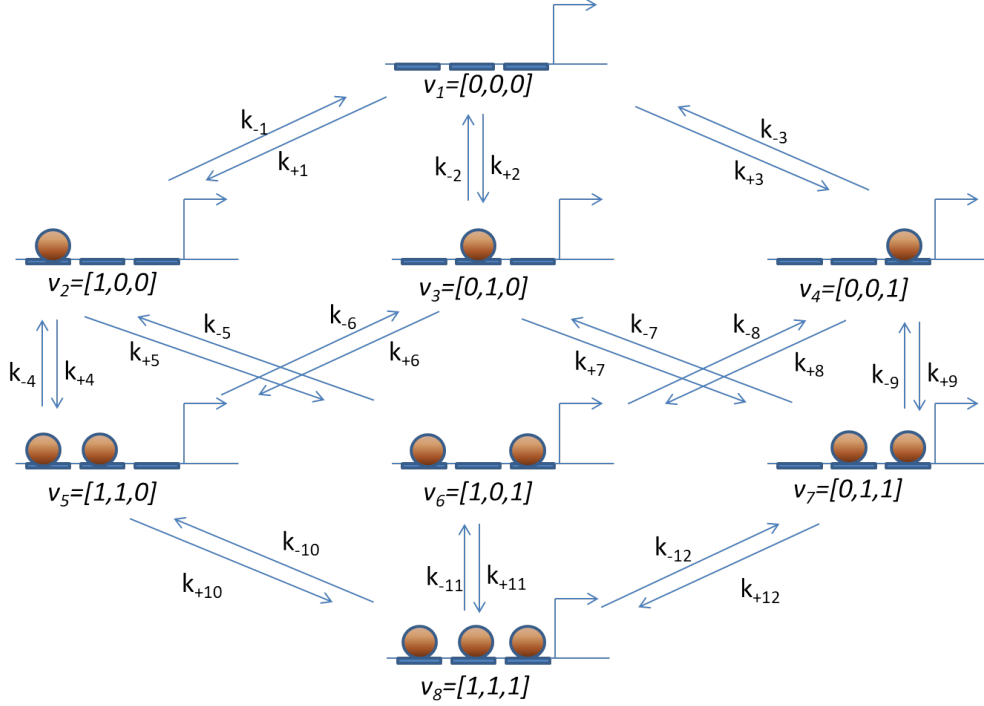
Figure 1: **The labeled graph for the general non-equilibrium binding site model of transcription regulation for $N = 3$ OS**. Each vertex v corresponds to a unique OS array state. The edge labels describe the binding and unbinding reactions, assuming one reaction takes place at a time.

1.2 The thermodynamic equilibrium model

1.2.1 One dimensional model

We assume that the N binding sites are identical, and thus all the micro-states $v = [v_i]_{i=1..N}$ are characterized only by the number of bound TF (Σv_i). In this case the model is reduced to a one dimensional model described by P_i – promoter states with i bound TF molecules:

$$P_0 \xrightleftharpoons[k_{-1}]{k_1[TF]} P_1 \xrightleftharpoons[k_{-2}]{k_2[TF]} P_2... \xrightleftharpoons[k_{-N}]{k_N[TF]} P_N, \tag{4}$$

where the TF-OS interactions are at thermodynamic equilibrium and the detailed balance is satisfied. The binding and unbinding of Bcd molecules to the OS occurs with rate constants k_i and k_{-i}. The maximum value of k_i is dependent on τ_{bind} the time for a free OS to be bound by TF and the number of free remaining operator $[N - i + 1]$: $k_i \leq [N - i + 1]/\tau_{\text{bind}}$. $[TF]$ is the normalized TF concentration, and is equal to 1 at the mid-boundary position $X = 0$ (defined in section 4).

1.2.2 Steady-state solution

The temporal evolution of the probability $P(P_i)$ that the promoter is in state i in the one dimensional model in Eq. 4 is given by:

$$\partial_t P(P_i) = k_i[TF]P(P_{i-1}) + k_{-(i+1)}P(P_{i+1}) - (k_{-i} + k_{i+1}[TF])\,P(P_i). \tag{5}$$

2

The steady state solution is:

$$P(P_i, [TF]) = \frac{\prod_{j=0}^{i} K_j [TF]^i}{\sum_{\ell=0}^{N} \prod_{j=0}^{\ell} K_j [TF]^\ell}, \tag{6}$$

where $K_j = k_j/k_{-j}$ are the equilibrium constants for each transition between two states, and $K_0 = 1$.

We introduce the effective equilibrium constant $\tilde{K}_i = \prod_{j=0}^{i} K_j$ that is proportional to the fraction of time the promoter spends in state P_i:

$$P(P_i, [TF]) = \frac{\tilde{K}_i [TF]^i}{\sum_{j=0}^{N} \tilde{K}_j [TF]^j}. \tag{7}$$

If the TF concentration gradient follows an exponential curve, we can express the TF concentration $[TF]$ in terms of the nuclei's position (Eq. 14) and Eq. 7 becomes:

$$P(P_i, X) = \frac{\tilde{K}_i e^{iX}}{\sum_{j=0}^{N} \tilde{K}_j e^{jX}}. \tag{8}$$

2 Promoter searching time at the boundary position

We estimate the expected time for a binding event between a single binding site and a TF at the mid-boundary position ($[TF] = 1$).

Assuming that the TF can only search for OS by diffusing in the nucleus and that each collision between TF and OS is one successful binding event [3], we estimate $\tau_{bind} = 1/(Dac) \sim 4$ s, using $D \sim 7.4 \mu m^2/s$ – the diffusion coefficient of TF (measured through Bcd-eGFP using FRAP [4, 5]), $[c \sim 11.2/\mu m^3$ [4] – the absolute TF concentration at the mid-boundary position and $a \sim 3nm$ [6] – the size of one operator site for Bicoid.

3 Aligning the pattern boundary position

From the randomized kinetic parameter set $\overline{k} = [k_i, k_{-i}]$, we solve the probability of the gene to be active at any position position X and find the mid-boundary position $X_0(\overline{k})$ such that $P(P_{\text{active}}|\overline{k}, X_0(\overline{k})) = 0.5$. The solution remains the same when we multiply both k_i and k_{-i} to a factor of $e^{X_0(\overline{k})}$:

$$P(P_{\text{active}}|\overline{k}, X) = P(P_{active}|[k_i, k_{-i}], X) \tag{9}$$

$$= P(P_{\text{active}}|[k_i e^{X_0((\overline{k}))}, k_{-i} e^{X_0((\overline{k}))}], X) \tag{10}$$

$$= P(P_{\text{active}}|[k_i, k_{-i} e^{X_0((\overline{k}))}], X - X_0(\overline{k})). \tag{11}$$

The whole pattern can be shifted so that the mid-boundary position is located at $X_0(\overline{k}) = X = 0$:

$$P(P_{\text{active}}([k_i, k_{-i} e^{X_0(\overline{k})}], 0) = P(P_{active}|[k_i, k_{-i}], X_0(\overline{k})) = 0.5. \tag{12}$$

We obtain the new parameter set $\overline{k'} = [k_i, k_{-i} e^{X_0(\overline{k})}]$, which satisfies the model assumptions $X_0(\overline{k'}) = 0$.

4 Boundary steepness and the Hill coefficient

The boundary steepness is a feature that emerges only when looking at the mean transcription readout value $f_P(X)$ of the hunchback gene along the AP axis (see Fig. **??** C and D of the main text) [2, 7, 8, 6]. We assume that the mean expression of the *hunchback* gene is regulated by a single transcription factor (i.e. Bcd), which has normalized concentration [TF], in terms of a Hill function with coefficient H:

$$\langle f_P([TF]) \rangle = \frac{[TF]^H}{1 + [TF]^H}. \tag{13}$$

In the case of Bcd-*hb* system the TF concentration decays exponentially along the embryo length [9] and we can estimate the nuclei's position X from the TF concentration:

$$X = \ln([TF]). \tag{14}$$

X is reported in units of decay length λ of the TF gradient, which is $\sim 100\mu m$ or 20% of the embryo length (EL) [9]. Eq. 13 becomes:

$$\langle f_P(X) \rangle = \frac{e^{H \cdot X}}{1 + e^{H \cdot X}}. \tag{15}$$
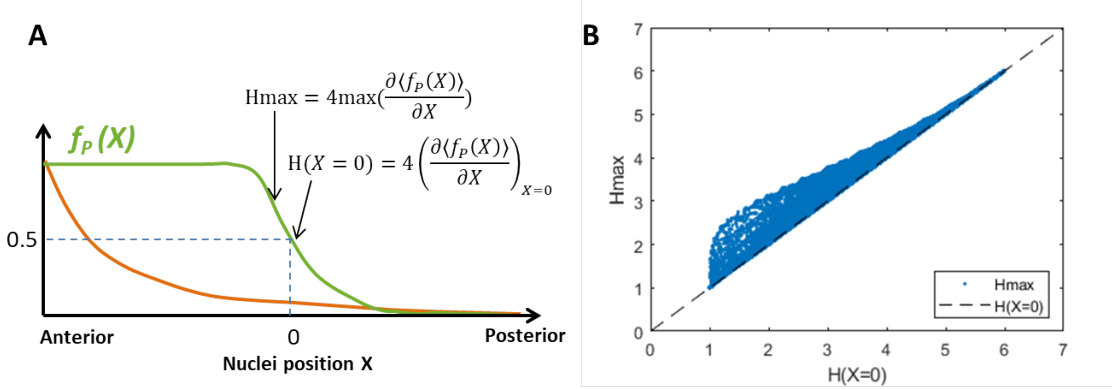
Figure 2: **Quantifying the pattern steepness**.
(A) From the mean readout function $f_P(X)$, the Hill coefficient H can be obtained from either the slope at the mid-boundary position corresponding to half-maximum readout of $\langle f_P(X) \rangle$, ($H(X = 0)$) or at the steepest point ($H_{\max}$) as in [2].
(B) Comparison between the two definitions of steepness $H(X = 0)$ and $H_{\max}$ for the equilibrium regulatory model with $N = 6$ binding sites shows the two values are correlated. The data points are taken from ~ 50000 data points with randomized kinetic parameters.

We define the mid-boundary position, $X = 0$, as the position along the AP axis corresponding to half-maximum of the mean readout function, $\langle f_P(X) \rangle = 0.5$. Note that the expression boundary is not necessarily positioned at the middle of the embryo.

Hill coefficients are typically obtained either by fitting the mean expression function to a sigmoid curve [6, 8, 7] or by comparing the maximum derivative of the mean readout function to that of a sigmoid function [2]. Here, to easily compare different embryos to each other and to analytical predictions, we calculate the Hill coefficient by comparing the slope of the mean readout function at the mid-boundary position ($X = 0$) to the prediction of Eq. 15:

$$H = 4\big(\frac{\partial\langle f_P(X)\rangle}{\partial X}\big)_{X=0}. \tag{16}$$

To see if our definition using the derivative at the half-maximum expression position significantly changes the numerical value of the steepness when calculated at the point of the maximum derivative (SI Fig. 2A), we compare the two values obtained at the steady state of the transcription regulatory model defined in Section 1.2 for different sets of randomized kinetic parameters. The results shown in SI Fig. 2B for $N = 6$ show the two values of the steepness calculated at different positions are tightly correlated, especially in the regime of high steepness. For the remainder of this work, we work with the steepness defined at mid-boundary ($H(X = 0)$) and note that an alternative definition would not change our conclusions.

5 Boundary steepness and promoter switching time for the equilibrium model

5.1 The boundary steepness

We consider the general "K-or-more case", that is the promoter is active when at least K OS are bound by TF, $P(P_{\text{active}}) = \sum_{i=K}^{N} P(P_i)$. When $K = N$, we recover the "all-or-nothing" case, $P(P_{\text{active}}) = P(P_N)$.

At the boundary position $X = 0$ and $P(P_{\text{active}}) = p$ ($0 \le p \le 1$), Eq. 7 simplifies to:

$$P(P_{\text{active}}, X = 0) = \frac{\sum_{i=K}^{N} \tilde{K}_i}{\sum_{j=0}^{N} \tilde{K}_j} = p, \tag{17}$$

which imposes a condition on the effective equilibrium constants:

$$\sum_{i=K}^{N} \tilde{K}_i = \frac{p}{1-p} \sum_{k=0}^{K-1} \tilde{K}_k = p \sum_{j=0}^{N} \tilde{K}_j. \tag{18}$$

The slope of the pattern at mid-boundary position is given by the derivative:

$$\begin{aligned}
\big(\frac{\partial P(P_{\text{active}})}{\partial X}\big)\big|_{X=0} &= \sum_{i=K}^{N} \left(\frac{\partial}{\partial X}\frac{\tilde{K}_i e^{iX}}{\sum_{j=0}^{N}\tilde{K}_j e^{jX}}\right)\big|_{X=0} \\
&= \frac{\sum_{i=K}^{N} i\tilde{K}_i}{\sum_{j=0}^{N}\tilde{K}_j} - \frac{\sum_{i=K}^{N}\tilde{K}_i \sum_{j=0}^{N} j\tilde{K}_j}{(\sum_{j=0}^{N}\tilde{K}_j)^2} \\
&= \frac{\sum_{i=K}^{N} i\tilde{K}_i}{\sum_{j=0}^{N}\tilde{K}_j} - \frac{p\sum_{j=0}^{N} j\tilde{K}_j}{\sum_{j=0}^{N}\tilde{K}_j},
\end{aligned} \tag{19}$$

where in the last step we used Eq. 18.

For clarity, we set the ranges $i = K..N$, $j = 0..N$ and $k = 0..K - 1$. Eq. 19 is then rewritten as:

$$\left(\frac{\partial P(P_{\text{active}})}{\partial X}\right)_{X=0} = \frac{\sum i\tilde{K}_i - p\sum j\tilde{K}_j}{\sum \tilde{K}_j} \tag{20}$$

$$= \frac{(1-p)\sum i\tilde{K}_i - p\sum k\tilde{K}_k}{\sum \tilde{K}_j}$$

$$= \frac{N(1-p)\sum \tilde{K}_i}{\sum \tilde{K}_j} - \frac{(1-p)\sum(N-i)\tilde{K}_i + p\sum k\tilde{K}_k}{\sum \tilde{K}_j}$$

$$= Np(1-p) - \frac{(1-p)\sum(N-i)\tilde{K}_i + p\sum k\tilde{K}_k}{\sum \tilde{K}_j},$$

and the Hill coefficient (Eq. 16) is:

$$H = 4\frac{\partial P(P_{\text{active}})}{\partial X}$$

$$= 4Np(1-p) - 4\frac{(1-p)\sum(N-i)\tilde{K}_i + p\sum k\tilde{K}_k}{\sum \tilde{K}_j}. \tag{21}$$

At the boundary criteria $p = 0.5$ and:

$$H = N - 2\frac{\sum(N-i)\tilde{K}_i + \sum k\tilde{K}_k}{\sum \tilde{K}_j}. \tag{22}$$

In the "all-or-nothing" case ($K = N$), $\sum_{j=0}^{N}\tilde{K}_j = 2\tilde{K}_N$ (Eq. 18), the first term in the nominator disappears and Eq. 22 becomes

$$H = N - \frac{\sum_{k=0}^{N-1} k\tilde{K}_k}{\tilde{K}_N}. \tag{23}$$

5.2 Bounds for pattern steepness

Eq. 22 gives an upper bound of $H \approx N$ at the mid-boundary position, which occurs when

$$\frac{\sum_{i=K}^{N}(N-i)\tilde{K}_i + \sum_{k=0}^{K-1} k\tilde{K}_k}{\sum_{j=0}^{N}\tilde{K}_j/2} << N. \tag{24}$$

When N is not too large (≤ 10), we can rewrite the upper bound condition in Eq. 24:

$$\sum_{i=K}^{N}(N-i)\tilde{K}_i + \sum_{k=0}^{K-1} k\tilde{K}_k \ll \sum_{j=0}^{N}\tilde{K}_j/2 \tag{25}$$

which is equivalent to $\tilde{K}_l \ll \tilde{K}_0 + \tilde{K}_N$ for $l = 1..N - 1$ or $\tilde{K}_0 + \tilde{K}_N \approx \sum_{j=0}^{N}\tilde{K}_j$.

Maximum sharpness ($H = N$) is achieved when $\tilde{K}_0 \approx \tilde{K}_N \approx 0.5\sum_{i=0}^{N}\tilde{K}_i$ – the system spends most of the time in the fully free (P_0) or fully bound states (P_N). In this limit, we have $P(P_0) + P(P_N) \approx 1$, and thus $P(P_{active}) \approx P(P_N)$ regardless of the value of K.

To find the lower bound on H, we consider the difference between H and N from Eq. 22:

$$N - H = \frac{\sum_{i=K}^{N}(N-i)\tilde{K}_i + \sum_{k=0}^{K-1} k\tilde{K}_k}{\sum_{j=0}^{N} \tilde{K}_j/2} \tag{26}$$

$$= \frac{\sum_{i=K}^{N}(N-K)\tilde{K}_i - \sum_{i=K}^{N}(i-K)\tilde{K}_i}{\sum_{i=K}^{N} \tilde{K}_i} + \frac{\sum_{k=0}^{K-1}(K-1)\tilde{K}_k - \sum_{k=0}^{K-1}(K-1-k)\tilde{K}_k}{\sum_{k=0}^{K-1} \tilde{K}_k}$$

$$= N - K - \frac{\sum_{i=K}^{N}(i-K)\tilde{K}_i}{\sum_{i=K}^{N} \tilde{K}_i} + K - 1 - \frac{\sum_{k=0}^{K-1}(K-1-k)\tilde{K}_k}{\sum_{k=0}^{K-1} \tilde{K}_k},$$

$$\leq N - 1$$

Thus $H \geq 1$. $H = 1$ when the sum in Eq. 25 are negligible compared to 1, which happens when $\tilde{K}_K \gg \tilde{K}_{K+1..N}$ and $\tilde{K}_{K-1} \gg \tilde{K}_{0..K-2}$. With these conditions, the promoter spends most of the time in P_{K-1} and P_K.

5.3   Promoter activity time

τ_{active} is the mean duration the promoter is in the active state and the system is at steady state, $\tau_{\text{active}} \sim P(P_{\text{active}})$. We can relate τ_{active} to the average time τ_N the promoter spends in the P_N state where all the operator sites are occupied by TF:

$$\tau_{\text{active}} = \tau_N \frac{P(P_{\text{active}})}{P(P_N)} \tag{27}$$

$$= \frac{1}{k_{-N}} \frac{\sum_{i=K}^{N} \tilde{K}_i}{\tilde{K}_N}$$

$$= \frac{K_N}{k_N} \frac{\sum_{i=K}^{N} \tilde{K}_i}{\tilde{K}_N}$$

$$= \frac{1}{k_N} \frac{\tilde{K}_N}{\tilde{K}_{N-1}} \frac{\sum_{i=K}^{N} \tilde{K}_i}{\tilde{K}_N}$$

$$= \tau_{\text{bind}} \frac{\sum_{j=0}^{N} \tilde{K}_j}{2\tilde{K}_{N-1}},$$

where $\tau_{\text{bind}} = 1/k_N$ is the expected time for a binding event between the remaining free OS of P_{N-1} and a TF at the mid-boundary position ($[TF] = 1$) (as defined in section 2).

Eq. 27 allows us to connect the Hill coefficient in Eq. 23 to τ_{active}. For $K = N$ (the "all-or-nothing" case), using Eq. 27, Eq. 23 becomes:

$$H = N - \frac{\sum_{k=0}^{N-1} k\tilde{K}_k}{\sum_{j=0}^{N} \tilde{K}_j/2} \tag{28}$$

$$= N - \frac{(N-1)\tilde{K}_{N-1} + \sum_{k=0}^{N-2} k\tilde{K}_k}{\sum_{j=0}^{N} \tilde{K}_j/2}$$

$$= N - \frac{(N-1)\tau_{\text{bind}}}{\tau_{\text{active}}} - \frac{\sum_{k=0}^{N-2} k\tilde{K}_k}{\sum_{j=0}^{N} \tilde{K}_j/2}$$

$$\leq N - \frac{(N-1)\tau_{\text{bind}}}{\tau_{\text{active}}},$$

or

$$\tau_{\text{active}} \geq \tau_{\text{bind}} \frac{N-1}{N-H}. \tag{29}$$

This leads to the bound on the Hill coefficient presented in the main text:

$$H \leq N - (N-1)\frac{\tau_{\text{bind}}}{\tau_{\text{active}}}. \tag{30}$$

Given the estimate $\tau_{\text{bind}} \approx 4$s (section 2), $H \sim N$ for $\tau_{\text{active}} \gg 4$ s or $k_{-N} \ll 0.25\ s^{-1}$.
For $K < N$ (the "K-or-more" case), Eq. 22 becomes:

$$
\begin{aligned}
H &= N - \frac{\tilde{K}_{N-1} + \sum_{m=K}^{N-2}(N-m)\tilde{K}_m + \sum_{k=0}^{K-1} k\tilde{K}_k}{\sum_{j=0}^{N} \tilde{K}_j/2} \\
&= N - \frac{\tau_{\text{bind}}}{\tau_{\text{active}}} - \frac{\sum_{m=K}^{N-2}(N-m)\tilde{K}_m + \sum_{k=0}^{K-1} k\tilde{K}_k}{\sum_{j=0}^{N} \tilde{K}_j/2} \\
&\leq N - \frac{\tau_{\text{bind}}}{\tau_{\text{active}}},
\end{aligned}
\tag{31}
$$

or

$$\tau_{\text{active}} \geq \tau_{\text{bind}} \frac{1}{N-H}. \tag{32}$$

The equality in Eq. 28 and Eq. 31 occurs when $\tilde{K}_{1..N-2} \ll \sum_{j=0}^{N} \tilde{K}_j$ – the system spends most of the time in the P_0, P_{N-1} and P_N states.

6 Calculating the mean promoter activity and readout error

In this section we obtain analytical solutions for the time dependent mean promoter activity $(\mu_P(T, 0))$ and readout error $(CV_P(T))$. Those results are expressed in terms of the exponential of the transition rate matrix U of size $N2^N$ for the non-equilibrium model and size $N+1$ for the equilibrium model, defined in Eq. 5. We discuss in what cases the matrix exponentiation can be done analytically or must be done numerically.

The steady state solution for the promoter activity probability vector is given by $U\bar{x} = 0$ and the normalization condition $\sum \bar{x} = 1$. In the equilibrium model, the steady state solution $\bar{x}$ is given by Eq. 8.

6.1 Mean promoter activity out of steady state

We define x_0 as the promoter state probability at the beginning of the interphase $(\sum x_0 = 1)$. The mean promoter activity level at time T is given by:

$$\mu_P(T) = \alpha^{\text{T}} e^{U \cdot T} x_0 \tag{33}$$

where α is a vector of the promoter active states. The ith element of α takes values of either 1 or 0, indicating the ith promoter state is active or inactive respectively.
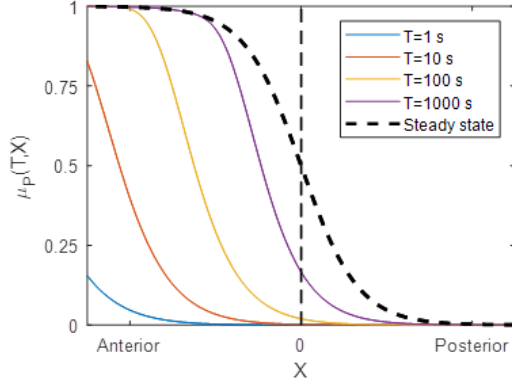
Figure 3: **Example of the pattern formation following mitosis**. The mean promoter activity pattern $\mu_P(T, X)$ as a function of the position in the embryo the end of an interphase of duration T (solid colored lines), for interphases of varying duration. The dashed line shows the steady-state expression pattern.

6.2 The readout error

After the interphase of duration T, we obtain a readout $f_P(T)$ which is the average of the promoter activity trace $n(t)$ over time T:

$$f_P(T) = \frac{1}{T} \int_0^T n(t)dt. \tag{34}$$

At steady state $(x_0 = \overline{x})$, the probability of the gene to be active is a projection of the steady state probability onto the state of the system:

$$\langle f_P \rangle = \alpha^{\mathrm{T}} \times \overline{x}. \tag{35}$$

Let us define x_{fire} in which $x_{\text{fire}}(i) = \alpha(i)\overline{x}(i)$. The second moment of $f_P(T)$ can be found via the autocorrelation function:

$$\langle f_P^2(T) \rangle = \frac{1}{T^2} \int_0^T du \int_0^T dv \; \alpha^{\mathrm{T}} e^{U|u-v|} x_{\text{fire}} \tag{36}$$

$$= \frac{2}{T^2} \int_0^T du \int_0^u dv \; \alpha^{\mathrm{T}} e^{U(u-v)} x_{\text{fire}}$$

$$= \frac{2\alpha^{\mathrm{T}}}{T^2} \int_0^T ds \int_s^T du \; e^{Us} x_{\text{fire}}$$

$$= \frac{2\alpha^{\mathrm{T}}}{T^2} \left[\int_0^T ds(T-s)e^{Us} \right] x_{\text{fire}}.$$

We diagonalize the matrix $U = VDV^{-1}$, where V is the eigenvector matrix and D a diagonal matrix of eigenvalues $[\lambda_1, \lambda_2, ...\lambda_M]$. Eq. 36 becomes

$$\langle f_P^2(T) \rangle = \frac{2\alpha^{\mathrm{T}}}{T^2} V \cdot \text{diag}\big[L_1, L_2, ...L_M\big] \cdot V^{-1} x_{\text{fire}} \tag{37}$$

with

$$L_i = \int_0^T ds(T-s)e^{\lambda_i s}. \tag{38}$$

9
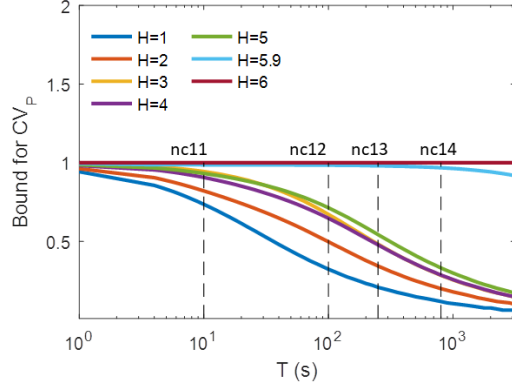
Figure 4: **The readout error of [TF] readout at steady-state.** The smallest values of CV_P as a function of interphase duration T, plotted for varying pattern sharpness H.

Performing the integration, for $\lambda_i = 0$:

$$L_i = \int_0^T ds(T - s) = T^2 - T^2/2 = T^2/2, \tag{39}$$

and for $\lambda_i \neq 0$:

$$
\begin{aligned}
L_i &= \int_0^T ds \cdot T \cdot e^{\lambda_i s} - \int_0^T ds \cdot s \cdot e^{\lambda_i s} \\
&= \int_0^T ds \cdot T e^{\lambda_i s} - 1/\lambda_i \big(s \cdot e^{\lambda_i s} \big|_0^T + \int_0^T ds \cdot T \cdot e^{\lambda_i s} \big) \\
&= T/\lambda_i (e^{\lambda_i T} - 1) - T/\lambda_i e^{\lambda_i T} + 1/\lambda_i^2 (e^{\lambda_i T} - 1) \\
&= 1/\lambda_i^2 (e^{\lambda_i T} - 1) - T/\lambda_i.
\end{aligned}
\tag{40}
$$

The readout error $CV_P(T)$ is calculated as:

$$CV_P(T) = \frac{\delta f_P(T)}{\langle f_P \rangle} = \sqrt{\frac{\langle f_P^2(T) \rangle - \langle f_P \rangle^2}{\langle f_P \rangle^2}} \tag{41}$$

In the special case when H approaches its maximum value, τ_{active} is infinitely long, all eigenvalues λ_i become zero and $L_i = T^2/2$ for all i. In this limiting case:

$$
\begin{aligned}
\langle f_P^2(T) \rangle &= \frac{2\alpha^{\mathrm{T}}}{T^2} V \cdot T^2/2 \cdot V^{-1} \overline{x} \\
&= \alpha^{\mathrm{T}} \overline{x} \\
&= \langle f_P \rangle
\end{aligned}
\tag{42}
$$

Applying Eq. 42 to Eq. 41, the readout error at the mid-boundary position is therefore:

$$CV_P(T) = \sqrt{\frac{\langle f_P \rangle - \langle f_P \rangle^2}{\langle f_P \rangle^2}} = 1 \tag{43}$$

10

6.3 Specific case with 2 binding sites

In the specific case $N = 2$:

$$P_0 \underset{k_{-1}}{\overset{k_1[TF]}{\rightleftharpoons}} P_1 \underset{k_{-2}}{\overset{k_2[TF]}{\rightleftharpoons}} P_2 \tag{44}$$

the matrix U is:

$$U = \begin{bmatrix} -k_1 & k_{-1} & 0 \\ k_1 & -k_2 - k_{-1} & k_{-2} \\ 0 & k_2 & -k_{-2} \end{bmatrix}, \tag{45}$$

where we have set $[TF] = 1$ at the mid-boundary position. The matrix is of size 3×3 and can be diagonalized analytically in the general case. We define the following auxiliary variables:

$$\begin{cases} r = k_1 k_2 + k_{-1} k_2 + k_{-1} k_{-2} \\ w = k_1 + k_2 + k_{-1} + k_{-2} \\ v = k_{-2} - k_1 - k_2 - k_{-1} \\ x = k_2 + k_{-2} - k_1 - k_{-1} \\ d = \sqrt{(k_1 + k_2 + k_{-1} + k_{-2})^2 - 4(k_1 k_2 + k_1 k_{-2} + k_{-1} k_{-2})} \end{cases} \tag{46}$$

and for the all-or-nothing case:

$$\alpha = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{47}$$

the steady state probability is:

$$\langle f_P \rangle = \alpha^{\mathrm{T}} \overline{x} = \alpha^{\mathrm{T}} \begin{pmatrix} k_{-1} k_{-2}/r \\ k_{-2} k 1/r \\ k_1 k_2/r \end{pmatrix} = k_1 k_2/r, \tag{48}$$

and the mean squared of the readout error is:

$$\langle f_P^2(T) \rangle = \frac{2k_1 k_2}{T^2 r} \times \tag{49}$$

$$\left[\frac{T^2}{2r} \left(k_2 k_{-2} + \frac{v^2 - d^2}{4} \right) \right.$$

$$+ \left(\frac{2T}{w - d} - \frac{4}{(w - d)^2} \left(1 + e^{-(w-d)T/2} \right) \right) \frac{k_{-2}(x - d)}{2r - 2w(w - d) + 3(w - d)^2/2})$$

$$\left. + \left(\frac{2T}{w + d} - \frac{4}{(w + d)^2} \left(1 + e^{-(w+d)T/2} \right) \right) \frac{k_{-2}(x + d)}{2r - 2w(w + d) + 3(w + d)^2/2} \right]$$

The analytically calculated readout relative error $CV_P(T) = \sqrt{\langle f_P^2(T) \rangle / \langle f_P \rangle^2 - 1}$ agrees with the numerical calculation for the $N = 2$ equilibrium model (SI Fig. 5).

7 Positional resolution

7.1 Calculation of positional resolution

For each set of parameters $\overline{k}$, integration window T and nuclei distance ΔW, we generate 500 realizations of promoter activity at location $-\Delta W/2$ and $+\Delta W/2$. From each realization, we
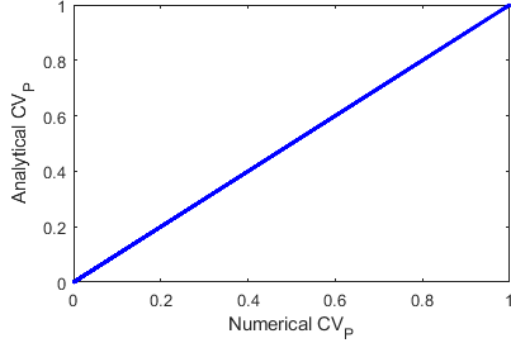
Figure 5: **The analytical and numerical calculation of the relative error $CV_P(T)$ for the equilibrium model with $N = 2$ (SI section 6)**. Small discrepancies result from numerically finding the half-maximum expression point, which due to numerical precision is not exactly at $\langle f_P \rangle = 0.5$.
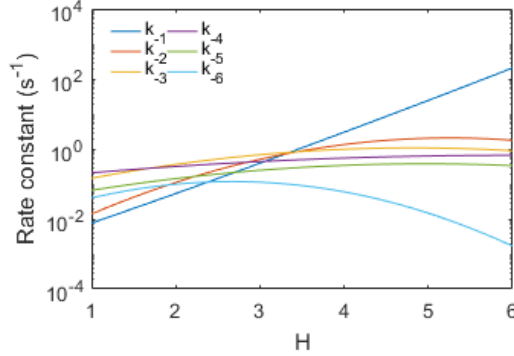

Figure 6: **The unbinding rate constants k_{-i} yielding the highest readout error CV_P in nuclear cycle 12 of the equilibrium model with $N = 6$**. The optimal binding rate constants k_i are at their highest possible values $(N - i + 1)/\tau_{\text{bind}}$.

extract an individual gene readout $f_{-i} = f_P(-\Delta W/2)$ and $f_{+i} = f_P(+\Delta W/2)$, with $i = 1...500$. The distribution of the readout values at the two positions, F_+ and F_-, can be approximated marginally by the sample distribution of f_{+i} and f_{+j} (SI Fig. 7A).

The difference in the activity of two nuclei on opposite sides of the mid-boundary position is:

$$D = F_+ - F_-. \tag{50}$$

When D takes a non-negative value, we have a false negative result suggesting the anterior nucleus is not the anterior region. The probability β of getting such false negative samples is:

$$\beta = P(D \leq 0) = P(F_+ \leq F_-). \tag{51}$$

The value of β for each ΔW can be calculated numerically via the approximated distribution of F_+ and F_-. One observes that β decreases with increasing nuclei distance ΔW (SI Fig. 7B). We set the risk tolerance level $\beta \leq 5\%$ to conclude whether the nuclei distance (ΔW) is large enough for any two nuclei to have different readout values. We define the positional resolution as such a value of ΔW that (SI Fig. 7B):
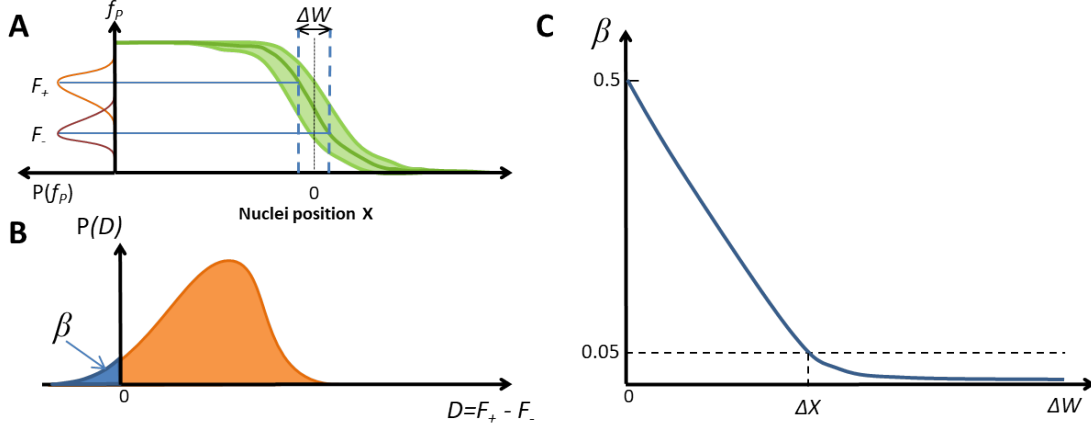
Figure 7: **Finding the positional resolution**.
(A) The distribution of readout F_+ and F_- from two nuclei positioned at a distance of ΔW on opposite sides of the expression boundary.
(B) The coefficient $\beta = P(D \leq 0)$ is the risk of a nucleus wrongly predicting its position.
(C) ΔX is set as the smallest ΔW yielding a tolerable risk ($\beta \leq 5\%$).

$$\Delta X = \min\left(\Delta W | \beta \leq 0.05\right). \tag{52}$$

In practice, to determine the value of ΔX for each parameter set $\bar{k}$, we increase ΔW from $0 \times \lambda$ to $4 \times \lambda$ with an increment of $0.01 \times \lambda$ (λ is the TF gradient's decaying length, which is $\sim 100\mu_m$ or $\sim 20\%$EL), which corresponds to 0% to 80% of the embryo length. For each value of ΔW, the distribution of D and the value of β are computed from stochastic simulations of F_+ and F_- [10, 11]. As β also monotonically decreases with ΔW, ΔX is set as the first value of ΔW that gives $\beta \leq 0.05$ (SI Fig. 7C).

When the nuclei readout is the average of M identical and independent identical single gene readouts ($F_+(j)$ and $F_-(j)$, for $j = 1..M$), the difference in the averaged readout at the two locations $-\Delta W/2$ and $+\Delta W/2$ is:

$$D_M = \frac{1}{M} \sum_{j=1}^{M} (F_+(j) - F_-(j)), \tag{53}$$

and β is calculated as $\beta = P(D_M \leq 0)$.

As M increases, it is expected that the difference in the averaged readout D_M at specific nuclei distance ΔW has reduced variance while maintaining the same mean level. This leads to a smaller risk level β and consequently smaller values of ΔX when compared with $M = 1$ case (SI Fig. 8).

7.2 Correlation between readout error and positional resolution

The correlation between the readout error and positional resolution given the same degree of pattern steepness is demonstrated in the $N = 6$ equilibrium model. We first find the randomized kinetic parameter sets that yield the Hill coefficient $H = 4$. The transcription readout error CV_P given $T = 400s$ varies between 0.12 and 1 (see Section 6). Among these sets, we select 20 parameter sets yielding CV_P linearly spaced between 0.12 and 1 and calculate the positional resolution for each of
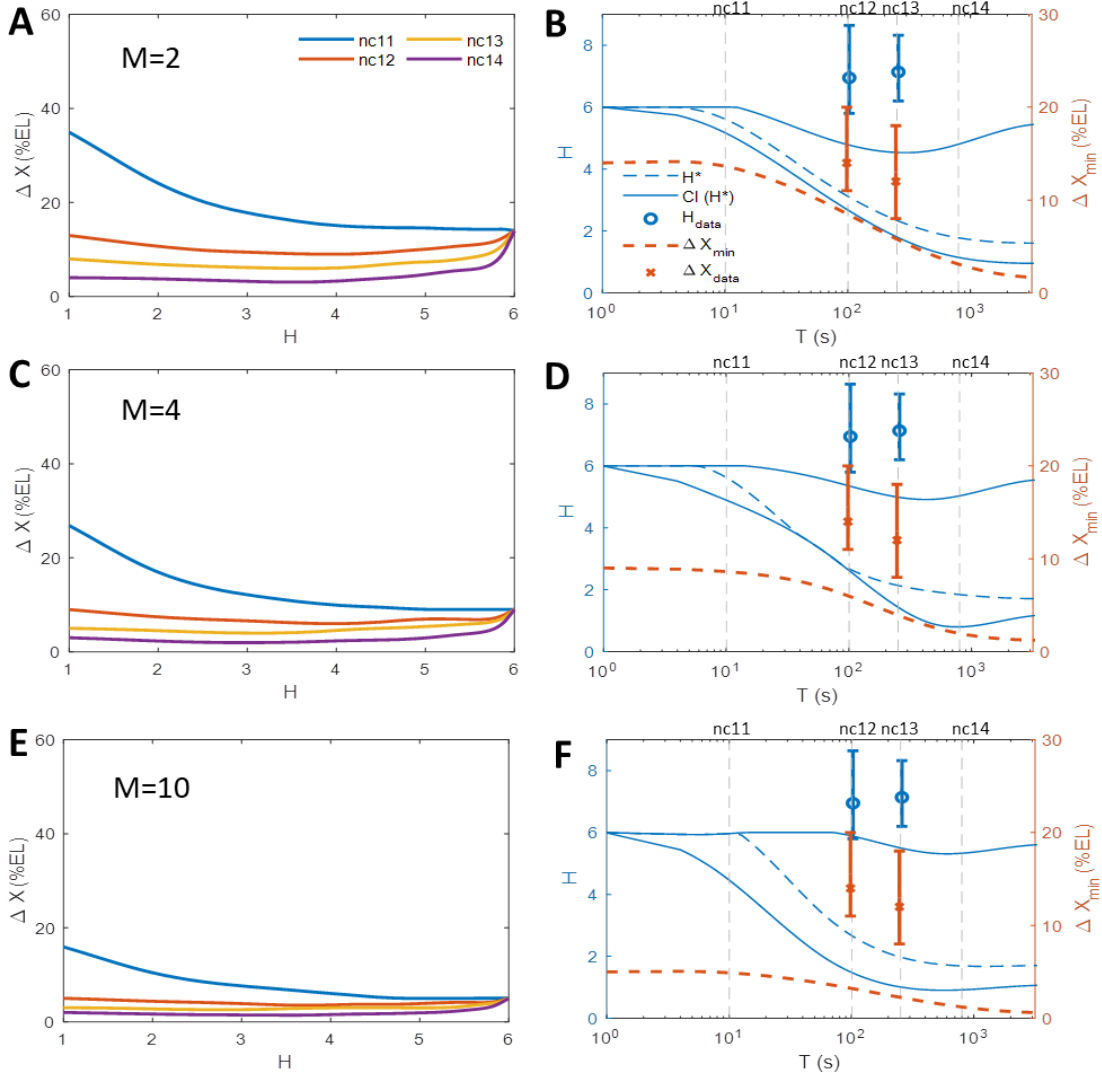
Figure 8: **Positional resolution with varying single gene copy number per nuclei M in the equilibrium model.** Results for the "all-or-nothing" model with $N = 6$ binding sites and M equal 2 (A-B), 4 (C-D) and 10 (E-F).

(A,C,E) Positional resolution calculated from the equilibrium binding site model for varying boundary steepness H.

(B,D,F) The optimal Hill coefficients H^* that gives the minimal positional resolution (dashed blue line), the confidence interval CI(H^*) with 2 %EL tolerance (solid blue lines) and the lowest value of the positional resolution $\Delta X_{\min}$ (orange dashed line), for varying T. The theoretical results are compared to the empirical Hill coefficient H_{data} (blue crosses) and positional resolution ΔX_{data} (orange crosses) extracted from MS2-MCP live imaging data.

the parameter set. The positional resolution (ΔX) as a function of readout error CV_P is shown in SI Fig. 9.

14
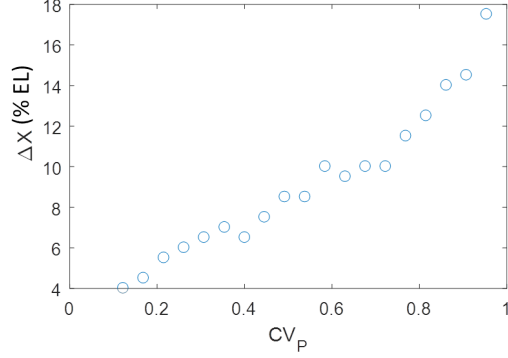
Figure 9: **The correlation between the readout error (CV_P) and positional resolution (ΔX).** Demonstrated with $N = 6$ and $H = 4$ and $T = 400s$. The kinetic parameters are selected so as to generate the same Hill coefficient $H = 4$ and readout error CV_P linearly spaced between its bound 0.12 and 1.

7.3 Positional resolution for a binomial readout

When the nuclear cycle is very short or when the promoter dynamics is very slow, the positional readout value any given position X depends only on the promoter activity state at the beginning of the nuclear cycle. At steady state, this activity state follows a Bernoulli distribution ($CV_P = 1$ as in Fig. **??** of the main manuscript) with a mean value $\langle f_P(X) \rangle$. We assume that the readout pattern can be well fitted with a sigmoid curve with a Hill coefficient H:

$$\langle f_P(X) \rangle = \frac{e^{HX}}{1 + e^{HX}}. \tag{54}$$

For the case $M = 1$ (single gene readout), at the anterior position $\Delta W/2$, the readout F_+ has a chance $f_P(\Delta W/2)$ to be 1. Similarly, at the opposite position $-\Delta W/2$, the readout F_- has a chance $1 - f_P(\Delta W/2)$ to be 1. Thus, the probability that two opposite nuclei falsely determine their position from their readout value is:

$$\begin{aligned} \beta &= P(F_+ \leq F_-) = 1 - P(F_+ > F_-) \\ &= 1 - P(F_+ = 1, F_- = 0) \\ &= 1 - f_P(\Delta W/2)^2. \end{aligned} \tag{55}$$

When we increase ΔW from 0 until β reaches 5%, we find the positional resolution $\Delta X = \Delta W$. Therefore:

$$1 - f_P(\Delta X/2)^2 = 0.05, \tag{56}$$

or

$$\begin{aligned} f_P(\Delta X/2) &= \sqrt{1 - 0.05} \\ &= \frac{e^{H\Delta X/2}}{1 + e^{H\Delta X/2}} \\ &= 1 - \frac{1}{1 + e^{H\Delta X/2}}, \end{aligned} \tag{57}$$

15

which gives

$$\Delta X = \frac{2}{H} ln(\frac{1}{1 - \sqrt{0.95}} - 1) \approx \frac{7.3}{H}. \tag{58}$$

and the value of ΔX in $\%EL$ unit is:

$$\Delta X \approx \frac{7.3}{H} * 20\%EL = \frac{146\%EL}{H} \tag{59}$$

In the case $M > 1$, the positional readout follows a scaled binomial distribution:

$$F_+ \sim \frac{1}{M} B(M, f_P(\Delta W/2)), \tag{60}$$
$$F_- \sim \frac{1}{M} B(M, f_P(-\Delta W/2)),$$

and the value of ΔX is calculated numerically by solving $\beta = P(F_+ \leq F_-) = 0.05$.

7.4 Positional resolution

To calculate the positional resolution of the hb pattern in live fly embryos, for each position along the embryo AP axis, we collect the readout of all nuclei in this position (with a bin width of 5% of the embryo length). We then find the distribution of the difference $P(F_+ - F_-)$ at position $+\Delta W/2$ and $-\Delta W/2$ from the pattern's boundary, with ΔW increasing from 0 %EL. Assuming that this difference follows a normal distribution, we calculate the risk β and its confidence interval (p-value=0.05) (SI Fig. 10). By inspecting when the risk value is tolerable ($\leq$ 5 %), we find $\Delta X \sim 14\%$EL (confidence interval from 11% EL to 20% EL) in nuclear cycle 12 and $\Delta X \sim 12\%$ EL (confidence interval from 8% EL to 18% EL) in nuclear cycle 13.
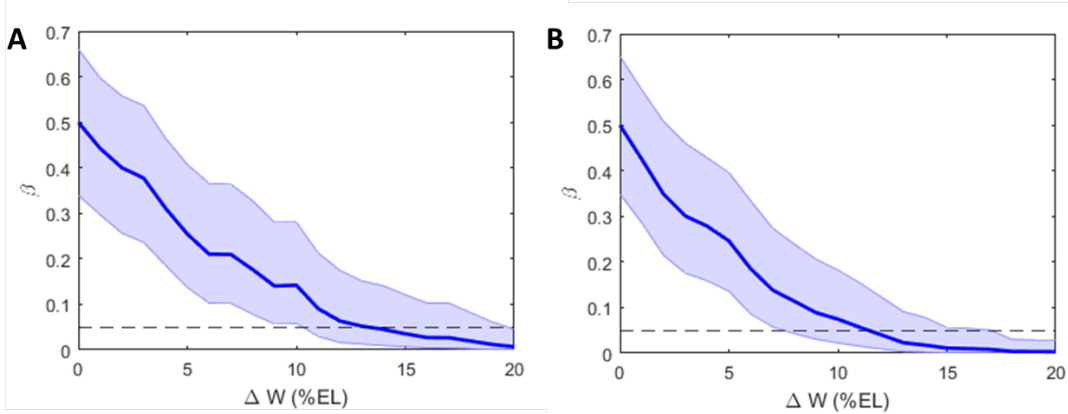
Figure 10: **The risk factor value β as a function of ΔW for the hb proximal promoter (solid line)**, plotted with the confidence interval (shaded) with p-value=0.05. The dashed black line indicates the tolerable risk $\beta = 0.05$. (A) Nuclear cycle 12 (8 embryos). (B) Nuclear cycle 13 (4 embryos).

8 Analysis of the non-equilibrium model

The steady state of the non-equilibrium models can be a limit-cycle instead of simple fix points. Therefore, to assess whether the system has reached steady state we consider both the probability of

the promoter to be active $\mu_P(T,0)$ like we did in steady state and the derivative of this probability over time:

$$\mu_{SS} = |\mu_P(T,0) - 0.5| + |\frac{\partial \mu_P(T,0)}{\partial T}|. \tag{61}$$

If the system has reached steady state at time T, at the mid-embryo position, we expect $\mu_P(T,0)$ to be equal 0.5 and its derivative term to be 0, and thus $\mu_{SS} \approx 0$. If $\mu_{SS} > 0$ the system has not yet reached steady state.

8.1 Full non-equilibrium model with $N = 3$ binding sites

We first investigate the "all-or-nothing" non-equilibrium model with 3 OS ($N = 3$). SI Fig. 11 shows that the model is able to achieve a higher steepness ($H \leq 2N - 1 = 5$) than that with the equilibrium model ($H \leq N = 3$), as described in [2]. Similarly to the equilibrium model we observe a tradeoff between the pattern steepness H, readout error and pattern formation time. In the case of the steepest pattern ($H = 5$), the pattern is not yet formed ($\mu_{SS} = 0.5$) and the noise is at its highest value ($CV_P = 1$).
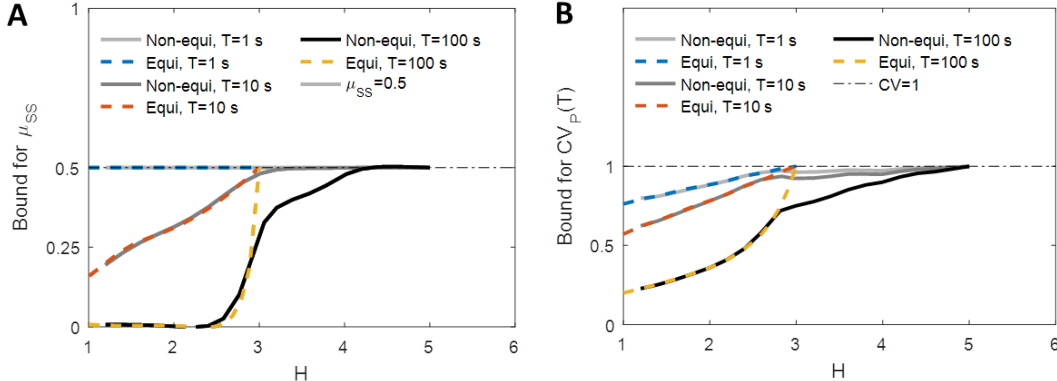
Figure 11: **Readout error of a pure non-equilibrium model for $N = 3$.**
(A) The lower bounds for μ_{SS} from the non-equilibrium model (grey solid line), for varying values of T, computed from 3×10^5 data points. Also shown are the bounds for equilibrium model (colored dashed lines).
(B) The lower bounds for readout error $CV_P(T)$ for the non-equilibrium model (grey solid lines) for varying value of T computed from 3×10^5 data points. Also shown are the bounds for the equilibrium model (colored dashed lines).

8.2 Hybrid non-equilibrium model with $N = 3$ binding sites

We expand the non-equilibrium model to $N = 6$. However, we do not use a full model (as in SI Fig. 1) due to the very large numbers of micro-states ($2^6 = 64$) and possible transitions ($6 \cdot 2^6 = 396$), which makes numerical optimization of the parameters numerically costly. Instead, we opt to use a hybrid model with 2 OS arrays. The first array contains 3 identical OS, the interactions of which with the TF are at equilibrium (as in Eq. 4). The second array contains 3 OS, the interactions of which with the TF are out of equilibrium (as in SI Fig. 1). To include cooperativity between the binding sites and decrease the computational time of the numerical parameter optimization we further assume the dynamics of the two arrays are not independent: TF can only interact with

the first OS array when the second OS array is completely free, and TF can only interact with the second OS array when the first array is fully bound.
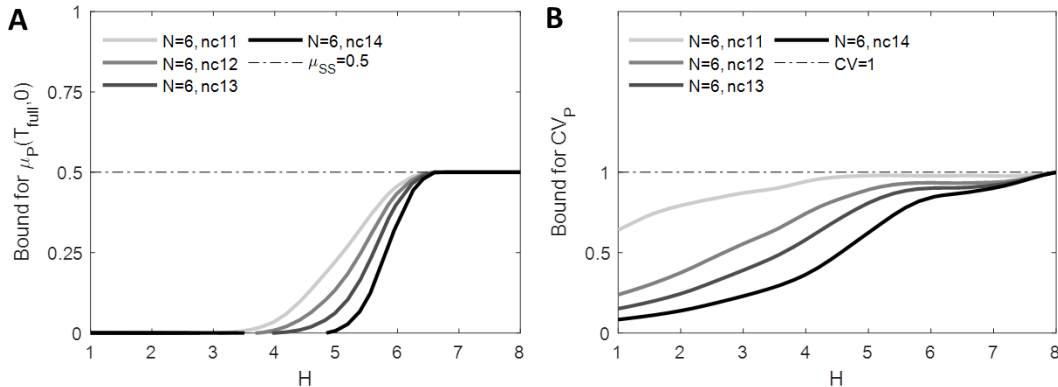

Figure 12: **Readout error of the hybrid non-equilibrium model with $N = 6$, three equilibrium and three non-equilibrium OS**.
(A) The lower bounds for μ_{SS} for a hybrid non-equilibrium model (grey solid lines) for varying values of T computed from 10^6 data points.
(B) The lower bounds for the readout error $CV_P(T)$ for a hybrid non-equilibrium model (grey solid lines), for varying values of T computed from 3×10^5 data points.

The hybrid model is able to achieve a steepness of 8 (SI Fig. 12), as expected from equilibrated activity of 3 OS and non-equilibrated activity of 3 OS. The tradeoff between the pattern steepness H and the readout error and pattern formation time still holds. Note that the hybrid model is not nested in the equilibrium model.

The positional resolution for the hybrid model with varying nuclear cycle is shown in SI Fig. 13 and the optimal steepness with varying interphase duration T is plotted in Fig. **??**C of the main text.
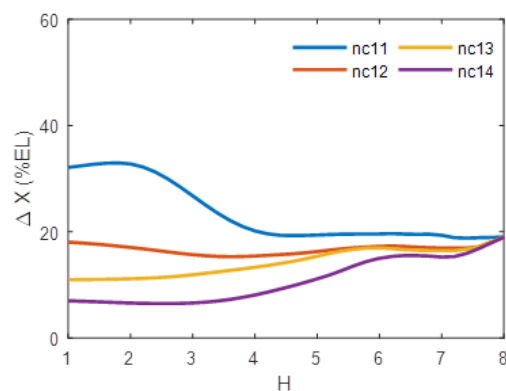

Figure 13: **Positional resolution in the $N = 6$ hybrid non-equilibrium model.** The results are shown for the "all-or-nothing" case with $N = 6$, $M = 1$. Positional resolution calculated for a hybrid non-equilibrium binding site model for varying boundary steepness H.

9 Analysis of the "K-or-more" case

The results concerning the "K-or-more" case is shown in SI Fig. 14, from which qualitatively similar observations as in the "all-or-nothing" case can be drawn.
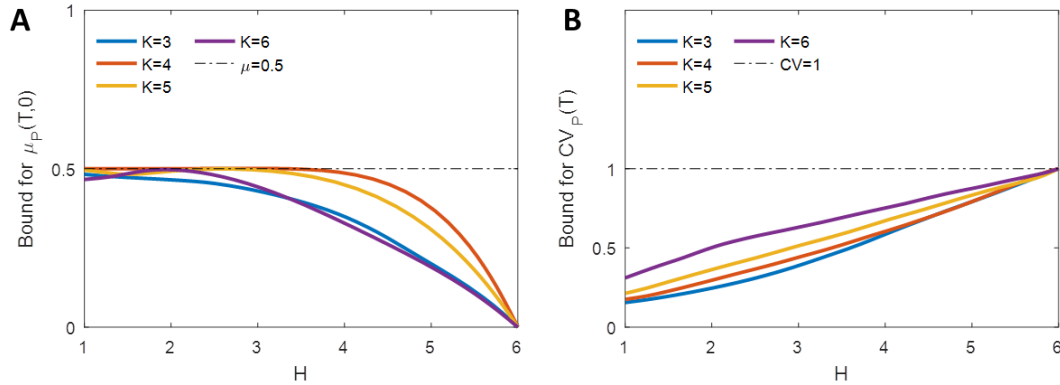

Figure 14: **The "K-or-more" case is qualitatively similar to the "all or nothing" case.** The results are shown with $N = 6$ binding sites and an interphase duration of $T = 500$ s. $K = 6$ corresponds to the "all or nothing" case. Each curve is computed from ~ 20000 data points.
(A) The lower bound for the mean promoter activity level at the boundary position $\mu_P(T, 0)$, for different K values (solid colored lines), as a function of pattern sharpness H. The $\mu = 0.5$ line (dashed line) indicates the steady-state value.
(B) The lower bound for readout error $CV_P(T)$ for different K values (solid colored lines) as a function of pattern sharpness H. Also shown is the upper bound for the noise, $CV = 1$.
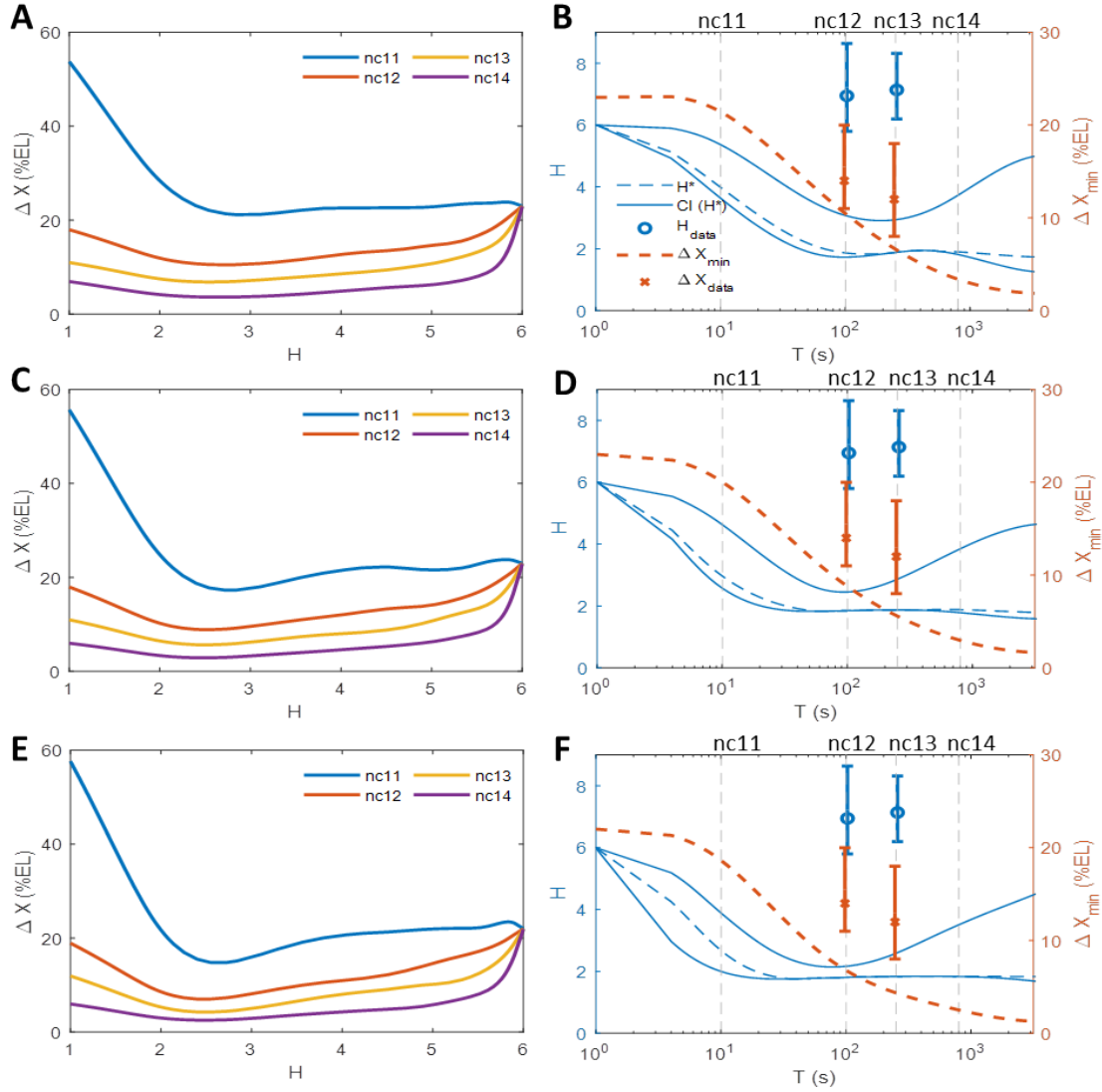
Figure 15: **Positional resolution in the "K-or-more" case, with varying K.** The results are shown with $N = 6$ binding sites and K equal 5 (A-B), 4 (C-D), 3 (E-F). $M = 1$.
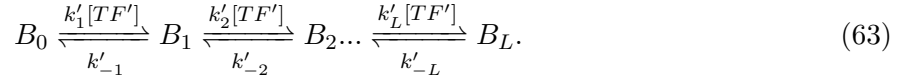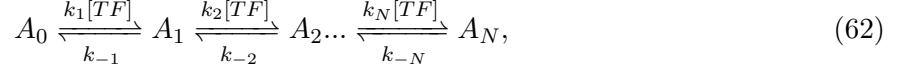
(A,C,E) Positional resolution calculated from the equilibrium binding site model for varying boundary steepness H.

(B,D,F) The optimal Hill coefficients H^* that gives the lowest value of the positional resolution (dashed blue line), the confidence interval CI(H^*) with 2 %EL tolerance (solid blue lines) and the lowest value of the positional resolution $\Delta X_{\min}$ (orange dashed line), for varying T. The theoretical results are compared to the empirical Hill coefficient H_{data} (blue crosses) and positional resolution ΔX_{data} (orange crosses) extracted from MS2-MCP live imaging data.

10   Transcription pattern formed by two transcription factor gradients

We investigate the transcription pattern formation under the independent regulation of two transcription factor gradients: an anterior activator TF (modeled as above) and a repressor TF', which is concentrated at either the posterior (e.g. Cad protein) or mid-embryo (Cic protein).

The transcription factors regulate the target gene via interactions with the activator binding site array A and repressor binding site array B, each with N and L identical binding sites respectively:

$$A_0 \underset{k_{-1}}{\overset{k_1[TF]}{\rightleftharpoons}} A_1 \underset{k_{-2}}{\overset{k_2[TF]}{\rightleftharpoons}} A_2 ... \underset{k_{-N}}{\overset{k_N[TF]}{\rightleftharpoons}} A_N, \tag{62}$$

$$B_0 \underset{k'_{-1}}{\overset{k'_1[TF']}{\rightleftharpoons}} B_1 \underset{k'_{-2}}{\overset{k'_2[TF']}{\rightleftharpoons}} B_2 ... \underset{k'_{-L}}{\overset{k'_L[TF']}{\rightleftharpoons}} B_L. \tag{63}$$

We call α and γ the vectors indicating which states are ON (the ith elements of α and γ respectively indicate whether A_i or B_i is an active or an inactive state). We consider the "all-or-nothing" model for the activator ($\alpha = [00...1]^T$) and a "zero-or-nothing" model for repressor ($\gamma = [10...0]^T$).

In each nuclear cycle of duration T, A and B produce time traces $a(t)$ and $b(t)$. The mean activity levels $A(T)$ and $B(T)$ are:

$$A(T) = \frac{1}{T} \int_{t=0}^{T} \alpha^\mathsf{T} a(t) dt \tag{64}$$

$$B(T) = \frac{1}{T} \int_{t=0}^{T} \gamma^\mathsf{T} b(t) dt.$$

We consider the promoter to be active when both the binding arrays are active:

$$P_{\text{active}} = \alpha^\mathsf{T} a(t) \gamma^\mathsf{T} b(t). \tag{65}$$

The promoter readout is given by:

$$f_P(T) = \frac{1}{T} \int_{t=0}^{T} \alpha^\mathsf{T} a(t) \gamma^\mathsf{T} b(t) dt. \tag{66}$$

At a given position, the two arrays have rate matrices U_a and U_b respectively. We call $\bar{x}_a$ and $\bar{x}_b$ the steady state solution of $U_a x = 0$ and $U_b x = 0$ respectively.

10.1   Scenario 1: posterior repressor

In the first scenario, the repressor has an exponentially decay gradient from the posterior, mirroring the anterior gradient:

$$[TF'] = e^{-X}. \tag{67}$$

We select $N = 6$, $L = 6$.

10.1.1 The pattern steepness

The mean promoter readout is given by:

$$\langle f_P(T) \rangle = \frac{1}{T} \int_{t=0}^{T} \langle \alpha^\mathsf{T} a(t) \gamma^\mathsf{T} b(t) \rangle dt. \tag{68}$$

Given that $a(t)$ and $b(t)$ are independent and assuming the system is at steady-state, we have:

$$\langle f_P(T) \rangle = \frac{1}{T} \int_{t=0}^{T} \langle \alpha^\mathsf{T} a(t) \rangle \langle \gamma^\mathsf{T} b(t) \rangle dt \tag{69}$$
$$= \langle A(T) \rangle \langle B(T) \rangle,$$

and at the promoter activity boundary $(X = 0, \langle P(T) \rangle = 0.5)$ the steepness of the promoter activity pattern is:

$$H_P = 4 \left(\frac{\partial \langle f_P(T) \rangle}{\partial X} \right)_{X_0} \tag{70}$$
$$= 4 \left(\frac{\partial (\langle A(T) \rangle \langle B(T) \rangle)}{\partial X} \right)_{X_0}$$
$$= 4 \langle B(T) \rangle \left(\frac{\partial \langle A(T) \rangle}{\partial X} \right)_{X_0} + 4 \langle A(T) \rangle \left(\frac{\partial \langle B(T) \rangle}{\partial X} \right)_{X_0}$$
$$= \langle B(T) \rangle H_A + \langle A(T) \rangle H_B.$$

Given Eq. 70, we have:

$$H_P \leq 4N \langle B(T) \rangle \langle A(T) \rangle (1 - \langle A(T) \rangle) + 4L \langle A(T) \rangle \langle B(T) \rangle (1 - \langle B(T) \rangle) \tag{71}$$
$$= 4 \langle P(T) \rangle (N + L - N \langle A(T) \rangle - L \langle B(T) \rangle)$$
$$= 2(N + L - N \langle A(T) \rangle - L \langle B(T) \rangle)$$

In the case $N = L$, we have the upper bound for H_P:

$$H_P \leq 4N - 2N(\langle A(T) \rangle + \langle B(T) \rangle) \tag{72}$$
$$\leq 4N - 4N \sqrt{\langle A(T) \rangle \langle B(T) \rangle}$$
$$= 4N \left(1 - \frac{1}{\sqrt{2}} \right) \approx 1.17N.$$

The equality in Eq. 72 occurs when $\langle A(T) \rangle = \langle B(T) \rangle = \frac{1}{\sqrt{2}}$. From Eq. 72, we found that having two independent binding site arrays does not yield significant higher pattern steepness than that achievable with a single array.

10.1.2 Mean promoter activity out-of-steady-state

We call a_0 and b_0 the initial state of the OS arrays A and B. At the end of the interphase of duration T, the mean probability that the promoter is active is:

$$\langle f_P(T) \rangle = \langle \alpha^\mathsf{T} a(T) \rangle \langle \gamma^\mathsf{T} b(T) \rangle \tag{73}$$
$$= (\alpha^\mathsf{T} e^{U_a T} a_0)(\gamma^\mathsf{T} e^{U_b T} b_0).$$

The upper bound for the mean promoter activity level at the end of each nuclear cycle interphase is shown in SI Fig. 16A for $N = L = 6$.

10.1.3 The readout error

Assuming the system is at steady state, the mean square of the readout:

$$\langle(f_P(T))^2\rangle = 1/T^2 \int_{u=0}^{T} du \int_{s=0}^{T} ds \langle p(u)p(s)\rangle \tag{74}$$

$$= 1/T^2 \int_{u=0}^{T} du \int_{s=0}^{T} ds \langle \alpha^{\mathsf{T}} a(u) \alpha^{\mathsf{T}} a(s)\rangle \langle \gamma^{\mathsf{T}} b(u) \gamma^{\mathsf{T}} b(s)\rangle$$

$$= 2/T^2 \int_{s=0}^{T} ds (T-s) \alpha^{\mathsf{T}} e^{U_a|u-s|} (\overline{x}_a.\alpha) \gamma^{\mathsf{T}} e^{U_b|u-s|} (\overline{x}_b.\gamma).$$

is calculated numerically given the transition matrices U_a and U_b and used to calculate the promoter activity readout relative error (the . here corresponds to term by term multiplication of the vectors coordinates).

The lower bound for promoter activity readout error after each nuclear cycle is shown in SI Fig. 16B for $N = L = 6$.

When H approaches its maximal value and τ_{active} goes to infinity, the integrated activity of each transcription factor becomes binomial. If p is the probability of the activator binding array being full bound and and q the probability of the repressor binding array being free, we have $\langle f_P(T)\rangle = pq$ and $CV_P(T) = (1-pq)/pq$. If $pq = 1/2$ we recover $CV_P(T) = 1$. Consistently, the limit of Eq. 74 when all non-zero eigenvalues of U_a and U_b go to $-\infty$ yields the same result.
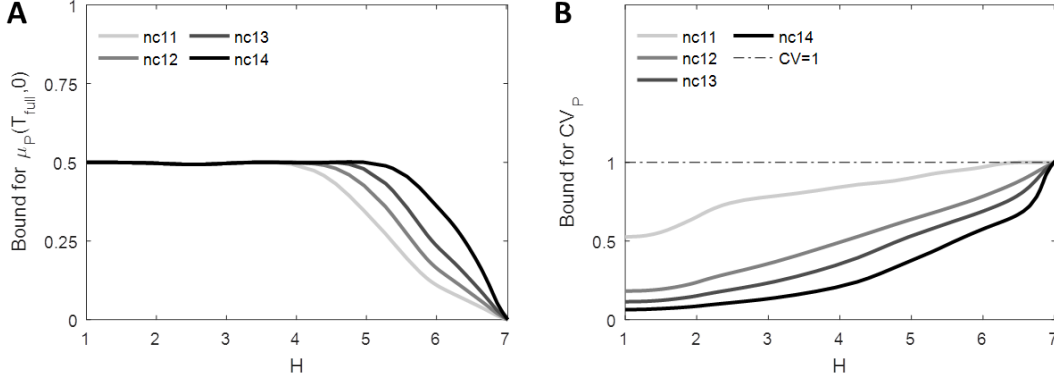

Figure 16: **The trade-off between pattern steepness (H), pattern formation time and readout error in the case of transcription regulation by two transcription factors**. Results for $N = L = 6$. $M = 1$. Each curve is computed from > 20000 data points.
(A) The lower bounds for the mean promoter activity level at the boundary position μ_{SS} for varying nuclear cycles.
(B) The lower bounds for readout error $CV_P(T)$ for varying nuclear cycles. Also plotted is the dashed line $CV = 1$.

10.2 Scenario 2: mid-embryo repressor

In the second scenario, the repressor is concentrated at the boundary position. The repressor gradient is modelled as a Gaussian curve with standard deviation $\sigma = 1.25$ (equivalent to 25 %EL).
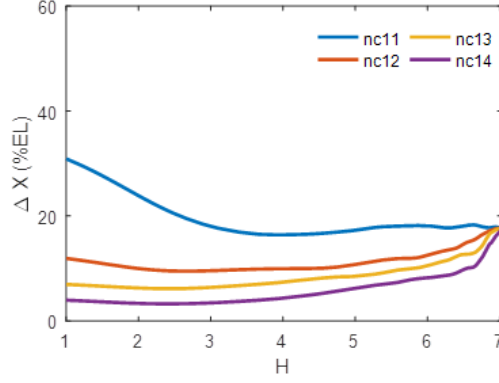
$$[TF'] = e^{\frac{X^2}{\sigma^2}}. \tag{75}$$

Figure 17: **Positional resolution in the case of transcription regulation by two mirror transcription factor gradients**. The results are shown for $N = L = 6$, $M = 1$. Positional resolution calculated from equilibrium binding site model for varying boundary steepness H.

TF can interact with the promoter via $L = 1$ binding site, corresponding to the number of known Cic binding sites found on *hb* promoter [12]. For simplicity, we consider $k_1' = k_{-1}' = 1$.

At the boundary position, given the local flat repressor concentration, the pattern steepness is dependent on the regulation function of only the activator:

$$\frac{\partial \langle f_P(T) \rangle}{\partial X} = \frac{\partial \langle A(T) \rangle \langle B(T) \rangle}{\partial X} = \frac{\partial \langle A(T) \rangle}{\partial X} = H_A \leq N \tag{76}$$

We plot the positional resolution of the readout in SI Fig. 18. The kinetic parameters k_i and k_{-i} are selected so as to minimize the readout error from 6 activator binding sites for varying pattern steepness.
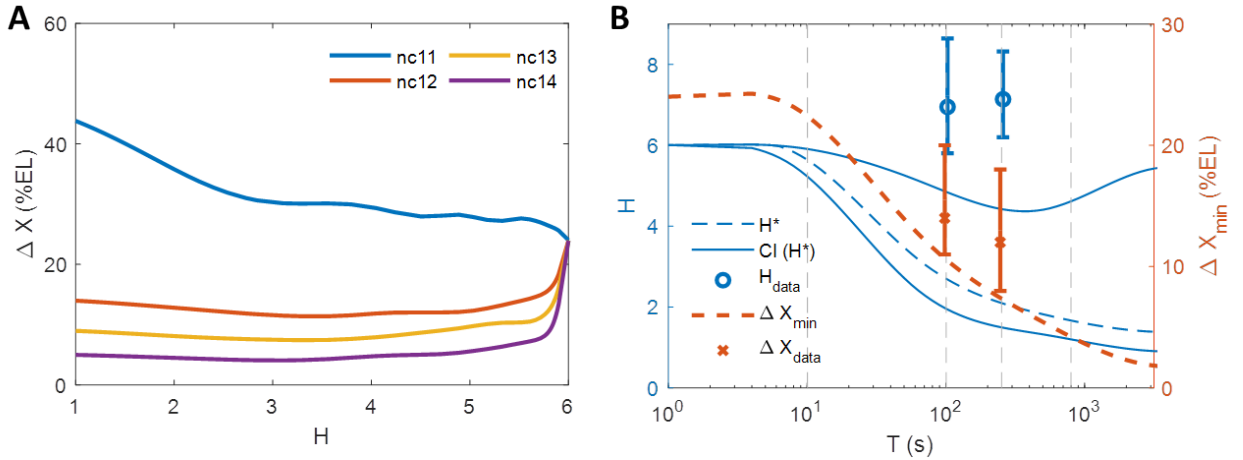

Figure 18: **Positional resolution in the case of transcription regulation by an anterior activator and a mid-embryo repressor**. The results are shown for $N = 6$, $L = 1$, $M = 1$. The positional resolution calculated from the equilibrium binding site model for varying boundary steepness H.
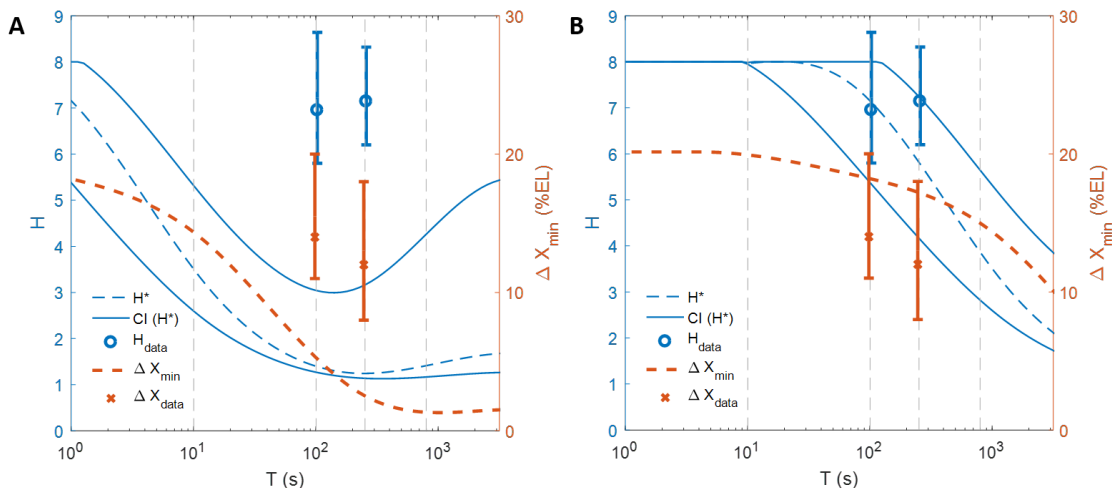
Figure 19: **Optimal positional resolution with different values of** τ_{bind}. (A) $\tau_{\text{bind}} = 0.4s$ (B) $\tau_{\text{bind}} = 40s$. The optimal positional resolution and Hill coefficient are calculated for the hybrid non-equilibrium model $N = 6$, $M = 1$ in the "all-or-nothing" case.

11 τ_{bind} needed to achieve experimentally observed pattern steepness and positional resolution in the hybrid non-equilibrium model

We vary the TF searching time for a single binding site τ_{bind} so as to fix the pattern steepness to the experimentally measured $H_{\text{data}} = 7$ and find the minimal value of the positional resolution ΔX given this constraint. ΔX as a function of τ_{bind} in each cycle is shown in SI Fig. 20. From SI Fig. 20, we find that values close to the experimentally observed positional resolution ($\Delta X_{\text{data}} \sim 14\%$ EL in nc 12 and $\Delta X_{\text{data}} \sim 12\%$ EL in nc 13) and pattern steepness ($H_{\text{data}} = 7$) can be achieved simultaneously with small τ_{bind} ($\tau_{\text{bind}} \sim 1.2s$ in nc12 and $\tau_{\text{bind}} \sim 0.12s$ in nc13).

12 Expression pattern of proximal *hb* promoter in live *Drosophila* embryos

We observe the transcription dynamics of a 700bp *hb* P2 minimal promoter using the RNA-tagging MS2-MCP system [13, 14]. Here, the nascent RNAs in each transcription loci are visualized as bright spots under the confocal microscope, due to the co-localization of fluorescent tagged MS2-GFP molecules [15]. The data for the analysis can be obtained in Lucas *et al.*, 2013 [16].

12.1 The pattern steepness

From each nucleus, we obtain a single gene readout f_P – the total spot intensity observed during the interphase. We fit the readout values along the AP axis with a sigmoid curve using least-mean-square and infer the Hill coefficient (SI Fig. 21). The inferred Hill coefficients in nuclear cycle 12 is from 6.9, with the confidence interval from 5.80 to 8.64 (p-value=0.05). In nuclear cycle 13, the Hill coefficient is 7.1,with the confidence interval from 6.20 to 8.32 (p-value=0.05).
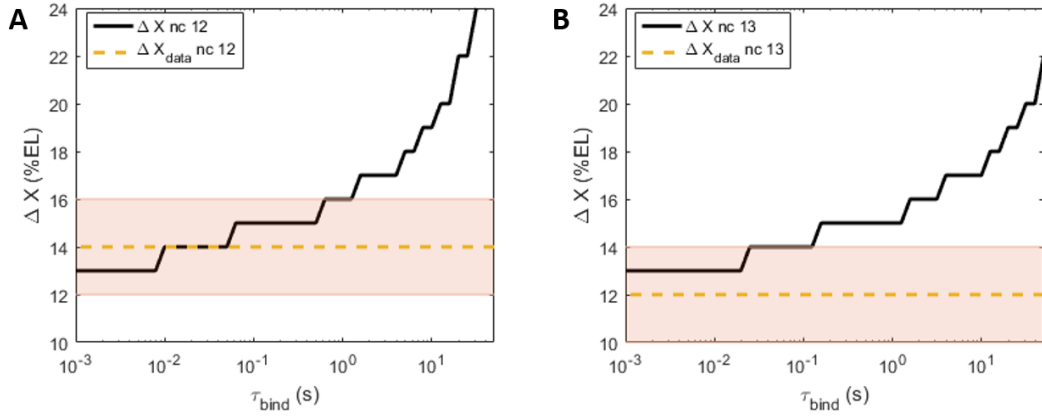
Figure 20: **Positional resolution ΔX of the hybrid non-equilibrium model as a function of TF searching time for a binding site τ_{bind} for nc 12 (A) and nc 13 (B)**. The kinetic parameters are selected to achieve the experimentally observed Hill coefficient $H_{\text{data}} = 7$. Also shown are the observed positional resolution ΔX_{data} in nc 12 (dashed green line in A) and nc 13 (dashed green line in B) and the 95% confidence intervals (shaded stripe in A and B).
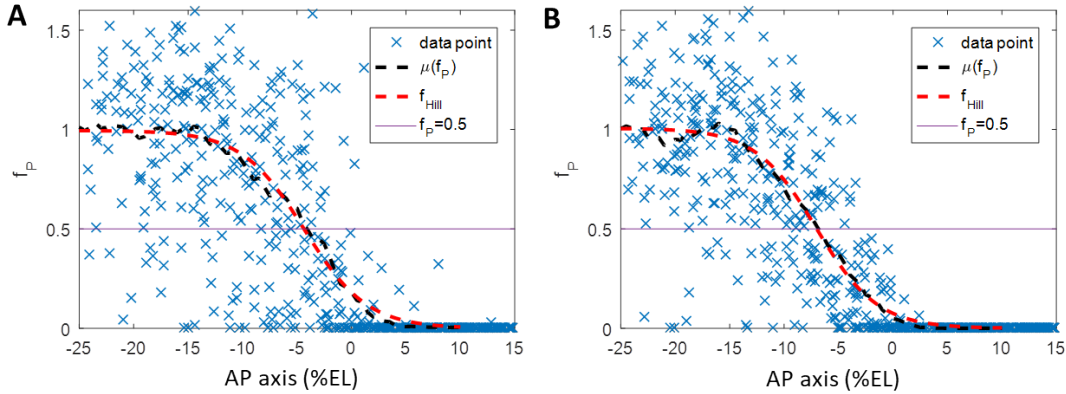

Figure 21: **The transcription readout pattern by hb proximal promoter along AP axis.** normalized fluorescence (blue crosses), the mean readout $\langle f_P \rangle$ (dashed black line), the fitted Hill function (dashed red line) and $f_P = 0.5$ (solid yellow line) as a function of nuclei position. The normalized fluorescence and mean expression curves are normalized by the fitted Hill function's maximum value. (A) Nuclear cycle 12 (8 embryos). (B) Nuclear cycle 13 (4 embryos).

12.2 Transcription readout error

From the fitted sigmoid curve, we identify the hb pattern boundary position at $\sim -5\%$ EL from the middle of the embryo for both nc 12 and nc 13. The readout distributions around the boundary (within $\pm 2.5\%$ EL) are shown in SI Fig. 22. From the distributions, we calculate readout errors CV_P to be ~ 0.82 in nc 12 and ~ 0.69 in nc 13.
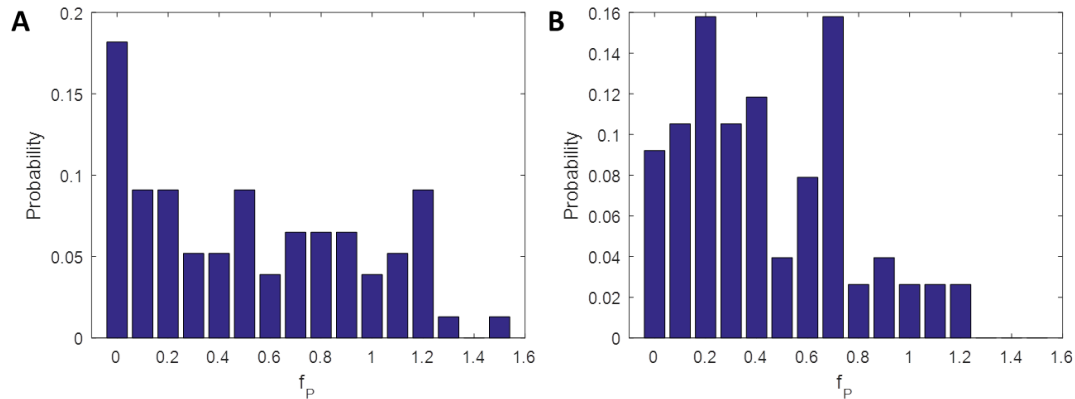
Figure 22: **Distributions of *hb* transcription readout at mid-boundary position** in (A) nuclear cycle 12 (8 embryos) and (B) nuclear cycle 13 (4 embryos).

References

[1] Jeremy Gunawardena. A linear framework for time-scale separation in nonlinear biochemical systems. *PLoS ONE*, 7(5), 2012.

[2] Javier Estrada, Felix Wong, Angela DePace, and Jeremy Gunawardena. Information integration and energy expenditure in gene regulation. *Cell*, 166(1):234–244, 2016.

[3] Howard C Berg and Edward M Purcell. Physics of chemoreception. *Biophysical Journal*, 20(2):193–219, 1977.

[4] Asmahan Abu-Arish, Aude Porcher, Anna Czerwonka, Nathalie Dostatni, and Cecile Fradin. High mobility of bicoid captured by fluorescence correlation spectroscopy: Implication for the rapid establishment of its gradient. *Biophysical Journal*, 99(4):33–35, 2010.

[5] Aude Porcher, Asmahan Abu-Arish, Sébastien Huart, Baptiste Roelens, Cécile Fradin, and Nathalie Dostatni. The time to measure positional information: maternal hunchback is required for the synchrony of the Bicoid transcriptional response at the onset of zygotic transcription. *Development*, 137(16):2795–804, 2010.

[6] Thomas Gregor, David W. Tank, Eric Wieschaus, and William Bialek. Probing the limits to positional information. *Cell*, 130(1):153–164, 2007.

[7] Heng Xu, Leonardo A Sepúlveda, Lauren Figard, Anna Marie Sokac, and Ido Golding. Combining protein and mRNA quantification to decipher transcriptional regulation. *Nature methods*, 12(8):739–42, 2015.

[8] Francisco JP Lopes, Alexander V Spirov, and Paulo M Bisch. The role of Bicoid cooperative binding in the patterning of sharp borders in Drosophila melanogaster. *Developmental Biology*, 72(2):181–204, 2011.

[9] Thomas Gregor, Eric Wieschaus, Alistair P McGregor, William Bialek, and David W Tank. Stability and nuclear dynamics of the Bicoid morphogen gradient. *Cell*, 130(1):141–152, 2007.

[10] Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434, 1976.

[11] Jason Lloyd-Price, Abhishekh Gupta, and Andre Sanches Ribeiro. SGNS2: A compartmentalized stochastic chemical kinetics simulator for dynamic cell populations. *Bioinformatics*, 28(22):3004–3005, 2012.

[12] Ulrike Löhr, Ho-Ryun Chung, Mathias Beller, and Herbert Jäckle. Antagonistic action of Bicoid and the repressor Capicua determines the spatial limits of Drosophila head gene expression domains. *Proceedings of the National Academy of Sciences*, 106(51):21695–21700, 2009.

[13] Tanguy Lucas, Teresa Ferraro, Baptiste Roelens, Jose De Las Heras Chanes, Aleksandra M Walczak, Mathieu Coppey, and Nathalie Dostatni. Live imaging of Bicoid-dependent transcription in Drosophila embryos. *Current Biology*, 23(21):2135–2139, 2013.

[14] Hernan G Garcia, Mikhail Tikhonov, Albert Lin, and Thomas Gregor. Quantitative imaging of transcription in living Drosophila embryos links polymerase activity to patterning. *Current Biology*, 23(21):2140–2145, 2013.

[15] Edouard Bertrand, Pascal Chartrand, Matthias Schaefer, Shailesh M. Shenoy, Robert H. Singer, and Roy M. Long. Localization of ASH1 mRNA particles in living yeast. *Molecular Cell*, 2(4):437–445, 1998.

[16] Tanguy Lucas, Huy Tran, Carmina Angelica Perez Romero, Mathieu Coppey, Cecile Fradin, Aleksandra M Walczak, and Nathalie Dostatni. Precision in a rush: trade-offs between positioning and steepness of the hunchback expression pattern. *bioRxiv*, 305532, 2018.