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

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

[1] *Ecole Normale Superieure, PSL Research University, Paris, France*
[2] *UPMC Univ Paris 06, Sorbonne Universités, Paris, France*
[3] *UMR3664/UMR168/UMR8549, CNRS, Paris, France*
[4] *Institut Curie, PSL Research University, Paris, France*
[5] *McMaster University, Canada*

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, additional gradients or non-equilibirum binding.

I. INTRODUCTION

One question in developmental biology is to understand how cell identity, determined by the expression of specific genes, is reliably established at the correct time and correct location in space. One of the simplest and the best understood developmental examples is the Bicoid-*hunchback* system in *Drosophila melanogaster* (see Fig. 1A). In the fly embryo, the exponentially decaying Bicoid (Bcd) gradient (see Fig. 1B) [1–3] acts as a maternal source of positional information along the embryo's Anterior-Posterior (AP) axis [4]. 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. 1A and B) [1, 5–7]. This Hb pattern later becomes a source of positional information for the formation of other gap gene patterns [8, 9], ensures the reproducibility of developmental patterns.

From real-time monitoring of the *hb* transcription dynamics [10, 11] using the MS2-MCP RNA-tagging system [12], 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 [11], the boundary between regions of high and low *hb* transcription is already steeper than that of the Hb protein concentration profile, even in nc14 [11, 13, 14].

Several studies have proposed that the steep boundary between regions of high and low *hb* expression given the smooth Bcd gradient is due to the cooperativity between the Bcd binding sites (Fig. 1C) [5, 15–18]. This cooperativity gives rise to a strong positive feedback that diversifies gene expression levels given small changes in the input [19–21]. Conventionally, the pattern steepness is represented by the Hill coefficient H, estimated by fitting the regulation function of *hb* expression by Bcd to the functional form:

$$f_{\text{Hill}}([Bcd]) \sim \frac{[Bcd]^H}{Bcd_0^H + [Bcd]^H}, \qquad (1)$$

where Bcd_0 is the Bcd concentration that results in half-maximal *hb* expression, $f_{Hill}(Bcd_0) = 0.5$. It defines the middle of the boundary region separating 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 [16]. Depending on whether this process conserves detailed balance or not, the maximal Hill coefficient is $2N-1$ or N, respectively.

These studies did not address whether such a steep boundary is achievable within the limited time window of 3 to 15 minutes in nuclei 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 [11, 22], suggesting that the *hb* expression pattern needs to be re-established in each cell cycle. Furthermore, the intrinsic noise in chemical processes leads to inherent errors in the concentration readout [14, 23]. This noise results in a lower bound for the ligand concentration readout error that depends on the readout integration time and the diffusion constant of ligand molecules [3, 24–27]. Extending the original work that considered a single or an array of non-interacting receptors [3, 24], other work pointed out that cooperativity from receptor arrays increases the readout noise [28, 29]. Given these effects, it is unclear how the readout accuracy of the Bcd concentration (or nuclei position) changes quantitatively given the highly cooperative readout by the promoter observed as the steep *hb* expression pattern [7, 13, 14] 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

A. A transcription factor binding model of gene regulation

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 [6]. 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 [5, 16]. 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 [16] (Fig. 1C, SI Fig. 2). In this model, the binding rates are modulated by the position-dependent TF concentration and bounded according to the promoter search time of individual TFs. The promoter readout is defined as the mean of the promoter activity level $n(t)$, calculated from the temporal promoter state $n(t)$ over the interphase duration T (see Fig. 1D).

We first focus on a simplified version of the general model of gene regulation for binding of Bcd TF to the N operator sites [16], 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-embryo 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-embryo $X = 0$ and $[TF] = 1$. The binding and unbinding of TF to the promoter occur with rate constants k_i and k_{-i}.

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 [16, 28], where the promoter is active if at least K binding sites are occupied (see SI - section 2). At steady state, we find the probability that the promoter is in the active state given the nucleus position X (see SI - section 2):

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

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

The Hill coefficient [30], calculated as the slope of the expression pattern at the mid-boundary position (see SI - section 1) determines the steepness of the expression pattern $H = N - (\sum_{i=1}^{N-1} i \cdot \tilde{K}_i)/\tilde{K}_N$, which is bounded by 1 and the OS number ($1 \leq H \leq N$), confirming previous results [16]. 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 2).

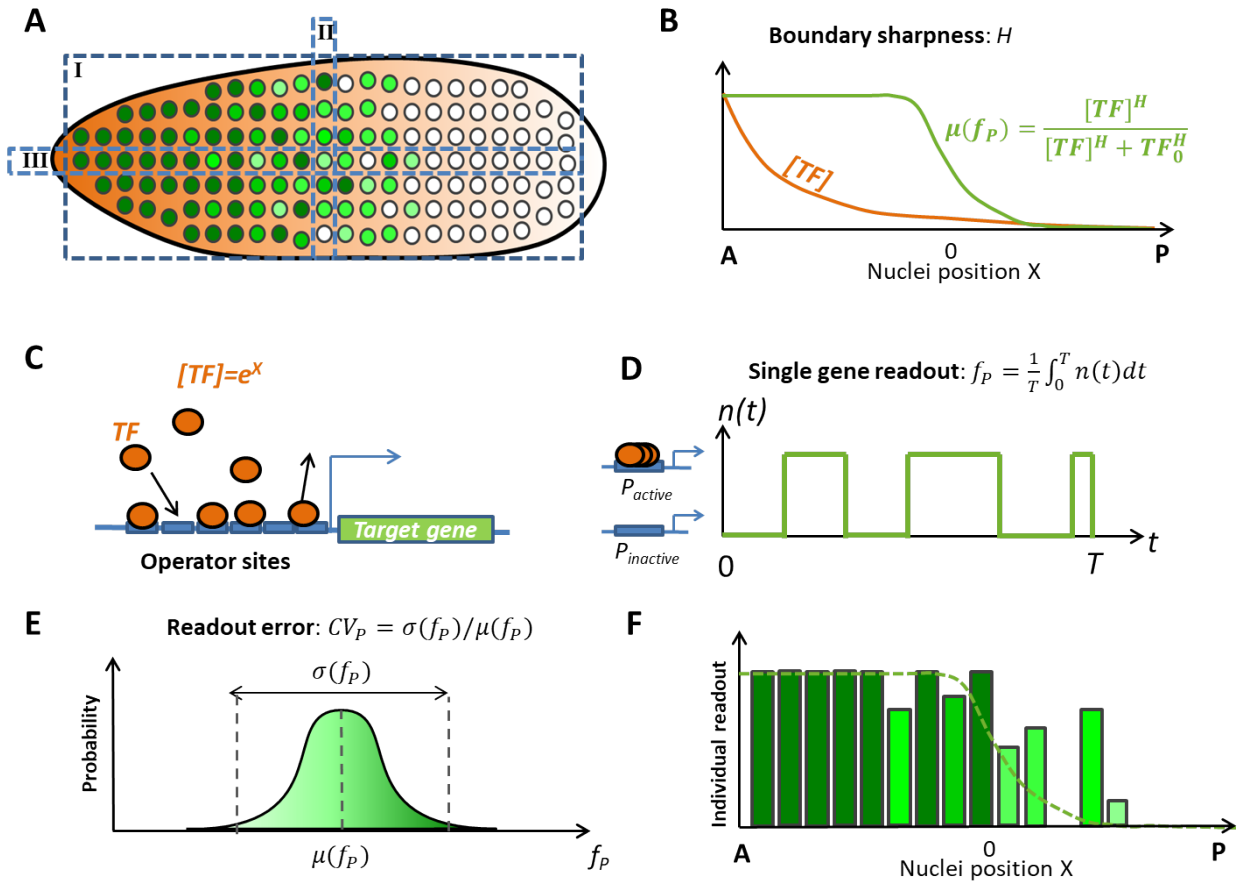

FIG. 1: **Features of the *hb* gene expression pattern.**
(A) The *hb* gene expression pattern from different perspectives: (I) Side view of the whole embryo. (II) Single columns of nuclei at similar position along the AP axis. (III) Single rows of nuclei along the AP axis.
(B) By observing the whole embryo (I), we are able to find the average of the expression pattern f_P of the nuclei, $\mu(f_P)$, as a function of the nuclei's position along the AP axes, from which the boundary steepness (denoted by H) is quantified.
(C) The model of gene regulation via the binding and unbinding of transcription factors (orange) to multiple operator sites (blue) of the promoter.
(D) After each cell cycle, each gene copy produces a readout f_P defined as the averaged promoter activity level during the interphase duration T.
(E) By observing nuclei at similar position X along the AP axis (II), we can calculate the errors CV_P in the single gene readout f_P.
(F) The zoomed-in pattern, extracted by observing a single row of nuclei (III) along the AP axis, is determined by both the averaged pattern and the errors in the readout of this mean value.

B. 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 the interphase of nc 11-13 T_{full} varies between $\sim 100 - 520$ seconds (Fig. 2A). During mitosis, Bcd molecules leave the nuclei and only reenter at the beginning of the interphase [3]. Assuming that at the beginning of the interphase all OS of the *hb* promoter in all nuclei are free, the mean probability $\mu_P(t, X)$ of 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, and with time spreads towards the posterior to reach its steady-state form with the border $(P(P_{\text{active}}, X))$ at mid-boundary position (see SI Fig. 3).

Whether the interphase duration T_{full} in a given cell cycle is long enough for the system to reach steady state depends on the parameters of bind-

A

Nuclear cycle	T_{full}	T
10	<100 s	< 10 s
11	270 s	≤ 10 s
12	320 s	100 s
13	520 s	250 s
14	1060 s	800 s

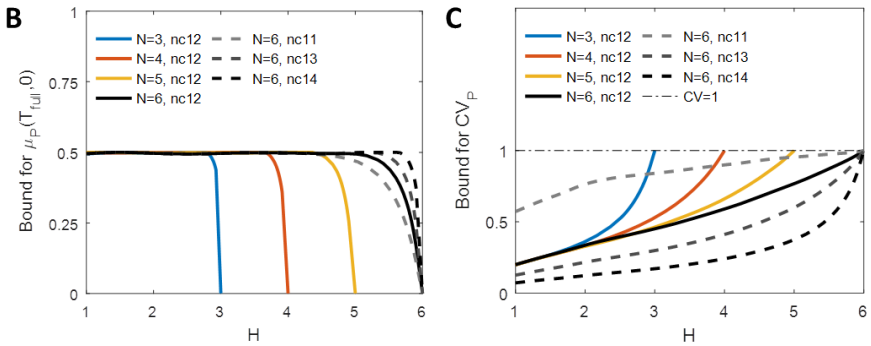

FIG. 2: **The pattern formation time and readout precision as a function of pattern steepness**.
(A) The total transcription window T_{full} and the time window when the transcription rate is at steady state T of *hb* transcription in nuclear cycle 10, 11, 12, 13 and early nuclear cycle 14 (before cellularization) at 25°C obtained from 8 MS2-MCP movies [11]. The short periods of transcription inactivity right before and after mitosis are excluded.
(B) The upper bound for the mean activity of the gene in nuclei positioned at mid-boundary $\mu_P(T_{\text{full}}, 0)$ and the maximum obtained Hill coefficient H for varying OS number N and nuclear cycle with varying T_{full}.
(C) The lower bound for the readout error CV_P for varying OS number N and nuclear cycles of varying T_{full}. The bounds in (A-B) are calculated numerically from $\sim$ 50000 data points of an equilibrium model with N OS, each corresponding to a randomized kinetic parameter set.

ing and unbinding of the TF to the operator sites. Within the "all-or-nothing" model, when the promoter is active ($P_N \equiv P_{\text{active}}$), any unbinding from the promoter inactivates the promoter. The mean time for the promoter to switch from the active state to the inactive state increases with the Hill coefficient $\tau_{\text{active}} = (k_{-N})^{-1} \geq (k_N)^{-1}(N-1)/(N-H)$ (see SI - section 3), which means a steep expression pattern requires the promoter to stay in the active state for a long time or have many OS:

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

where $\tau_{\text{bind}} = (k_N)^{-1}$ is the mean time for a TF to bind to the remaining 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 states (see SI - section 3). This limit corresponds to very slow promoter switching, similarly to conclusions obtained for cell surface receptors [28]. With typically considered parameters for the Bcd-hb system [7, 13, 31] the bound in Eq. 4 corresponds to $\tau_{\text{active}} \gg 4$ s (see SI section 2).

Given the limited interphase duration in nc 11 $T_{\text{full}} < 100s$ (Fig. 2A), a steep steady-state expression pattern at the boundary cannot be established during 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 randomized kinetic parameters giving large H boundaries (Fig. 2B). For long interphases ($T_{\text{full}} \geq 100$ s), all patterns but those close to the OS imposed 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 desired expression pattern with $\mu_P(T_{\text{full}}, 0) \sim 0.5$ at the boundary.

Generalizing the model to allow for non-equilibrium binding (SI Fig. 2) increases the possible Hill coefficients above $H > N = 6$, but does not alleviate their inaccessibility within the nuclear cycles (SI Fig. 12). Given the observed steep boundary $H \sim 7$ in cell cycles 11-13 [11, 14] and the relatively short interphase duration ($T_{\text{full}} < 800$s in nc 13, see Fig. 2B), 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 [5, 17]. Nevertheless the steady state results give a best case scenario for precision estimates so we focus on an equilibrium steady state system in the next section. We then extend the arguments to out-of-equilibrium binding.

C. The single gene readout precision 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 precisely express mRNA 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. 1D). The precision of the gene expression readout at the mid-boundary position in steady state is described by the variance within the boundary strip of the mean activity of the nuclei $f_P = \frac{1}{T}\int_{t=0}^{T_{SS}} n(t)dt$ in a nuclear cycle (Fig. 1D and E, see SI - section 5).

Randomizing binding parameters we see that the lower bound for the readout error CV_P increases with increasing Hill coefficient H and decreases with the nc duration (Fig. 2C). An increased pattern steepness requires slower promoter switching dynamics, which results in less independent measurements that take part in the single gene readout during each interphase. Therefore, the steeper the pattern, the larger the uncertainty in the expression pattern due to the lowered precision in the readout in each nuclei [32]. When the steepness H approaches its upper value limited by the maximum number of binding sites N, due to a small number of switching event during the interphase, the distribution of the average activity 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 increase in readout precision by having a lower steepness depends on the length of the nuclear cycle (Fig. 2B). 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 precision (SI Fig. 4).

III. RESULTS

A. Positional resolution

From the above analysis we see that it takes time to achieve a steep expression boundary and the variance in expression decreases with time. However a steep boundary increases the minimal achievable expression variance. We are left with this trade-off.

The trade-off between the expression pattern steepness and the readout precision means that neither of these features alone can be used as the sole criterion for a reproducible and robust pattern. This observation is not surprising given that these features emerged from looking at the embryo from two different perspectives (Fig. 1): the expression pattern steepness is perceived from an external observer's perspective when looking at the *whole* embryo at a fixed time (Fig. 1B), 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 gene expression 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 [7, 33], 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 at steady state ($P(F_+ \leq F_-) \leq 0.05$) (Fig. 3A). F_+ and F_- are the mRNA concentrations in two nuclei at positions $+\Delta X/2$ and $-\Delta X/2$ (see SI - section 6). The trade-off between the pattern steepness and the readout precision translates into constraints on the positional resolution. For a flat expression pattern (low H, Fig. 3A, top), F_+ and F_- have a small difference in their mean value, which makes the mRNA concentration in closely positioned nuclei hard to differentiate, but the variance around their mean is also small. On the other hand, with a steep pattern (Fig. 3A, bottom), F_+ and F_- have a big difference in their mean mRNA expression but also an increased variance, due to the lowered readout precision.

To evaluate the positional resolution at a given pattern steepness H and interphase duration T 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 [5, 17]. We identify the parameters that give the highest precision (smallest CV_P) to ensure the highest not the smallest ΔX. These two features are tightly correlated but the first one is faster to evaluate.

For short interphase durations (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 in $\Delta F = F_+ - F_-$, with increasing Hill coefficients H and decreasing ΔX with increasing H (Fig. 3C). As T lengthens, the positional resolution ΔX for small Hill coefficients decreases with increasing H due to the reduced readout uncertainty from averaging promoter switching events until a certain value, $\Delta X_{\min}(T)$. As H approaches N, the readout error increases drastically since the uncertainty $CV_P \to 1$ (Fig. 2B). As a result, the positional res-

olution ΔX increases and converges to a fixed value $\Delta X_N \approx 24\%$EL independently of T (see SI section 6).

The minimal positional resolution predicted by our model, $\Delta X_{\min}(T)$, is read off Fig. 3C for each interphase duration T (orange line in Fig. 3D). The optimal Hill coefficients corresponding to the minimal positional resolution, $H^* = H(\Delta X_{\min})$ (dashed blue line in Fig. 3D), decrease with the interphase duration. The solid blue lines denote a confidence interval of H that results in a positional resolution within 2% of the embryo's length (roughly the size of one nucleus) of the optimal value.

We see that for short interphase durations (up to nc 11), the embryo can best discriminate readouts when producing a very steep pattern. For longer nuclear cycles (12 and 13), a narrow range of moderately steep profiles (H between 2 and 5) provide the most discernible readout. As the interphase T increases, ΔX becomes very small for expression profiles with a wide range of H and the constraint on H^* is relaxed. In this case a discernible readout can be reached even for very flat expression profiles, since time averaging alone can result in sufficiently precise readouts.

The results for the mRNA readout measured in a nucleus coming from a single expressing gene copy ($M = 1$ – a heterozygous fluorescent marker such as in recent MS2-MCP experiments [10]) hold for to a readout coming from more gene copies ($M > 1$, Fig. 3B) (SI Fig. 8). $M = 2$ describes the readout from a homozygous gene, whereas $M > 2$ describes the averaging of mRNA in the cytoplasmic space coming from different nuclei [33]. As expected, averaging over many promoter copies further reduces the readout noise and slightly decreases the minimal value of positional resolution.

Since the interphase duration varies during the early development phase, we can use the results of Fig. 3D to define a value of a Hill coefficient, H_{robust}, that gives the minimal positional resolution in the widest range of interphase duration T (see Fig. 4) 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. 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.

B. The non-equilibrium model

Comparing the results of the equilibrium binding site model to experimental observations, we note that the steepness values reached in experiment ($H \sim 7$) cannot be reached by an equilibrium model with the identified $N = 6$ Bcd binding sites on the proximal hb promoter. Estrada *et al.* [16] 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. 2) and a hybrid model for $N = 6$ due to computational complexity. 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). The conclusions drawn from the equilibrium model are still valid even for $H > N$. Large Hill coefficients result in more readout variability. For the $N = 6$ hybrid model, 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 values, $2 \leq H \leq 5$, provide the best discernibility of ~ 12 to 14 %EL or 6 to 7 nuclei lengths, close to the observed experimental values of $\sim 14\%$ EL for nc 12 and 12% EL for nc 13 [11] (orange crosses with error bars in Fig. 3E, see SI - section 7).

C. "K-or-more" model

Until now we assumed that the gene is readout 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 higher readout precision than the "all-or-

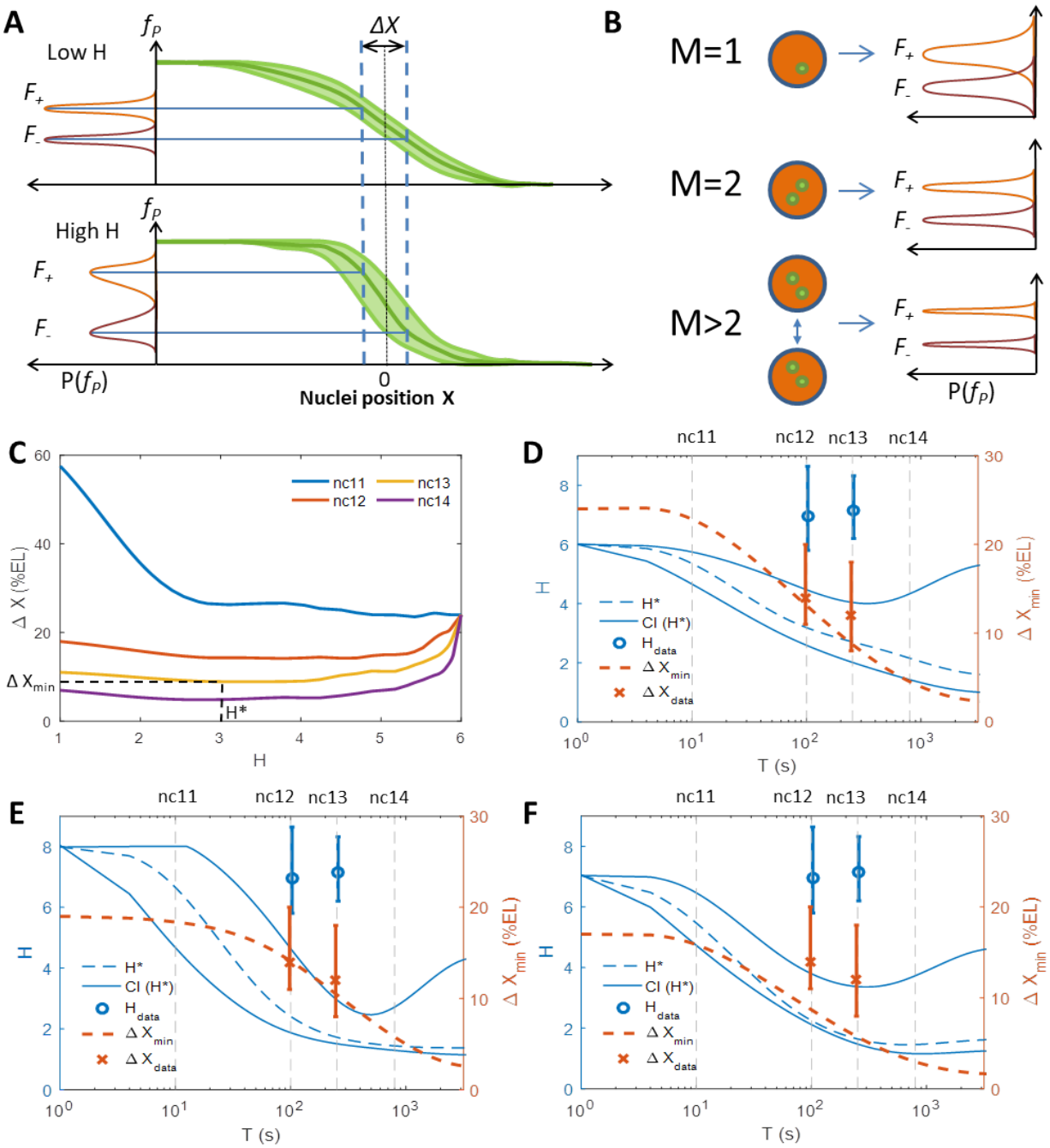

FIG. 3: **The positional resolution of the pattern.**
(A) The positional resolution ΔX results from a trade-off between the pattern steepness and the readout precision: a cartoon representation of the trade-off for a flat pattern (low H, top) and a steep pattern (high H, bottom). F_+ and F_- are the positional readouts in individual nuclei positioned at $\Delta X/2$ from the mid-boundary point ($X = 0$).
(B) Each nuclei readout is the average of M independent single gene readouts: M can be 1 (there is one copy of the gene), 2 (e.g. there is one gene copy on each chromosome pair) or greater (nuclei at the same position can communicate via e.g. diffusion of readout molecules [33]). As M increases, the precision in the nuclei readout increases due to averaging.
(C) Positional resolution calculated from the equilibrium binding site model for varying boundary steepness H for $M = 1$. The colored lines show the results for parameters that give the smallest precision CV_P from a set of randomized parameters at a given steady-state window T in nc 11-14 (Table 1). The curves are smoothed using cubic spline interpolation for better visual.
(D) The range of optimal Hill coefficients H^* that yield the highest positional resolution as a function of the steady-state readout duration T for the equilibrium $N = 6$ model (D), the hybrid $N = 6$ non-equilibrium model with 3 equilibrium and 3 non-equilibrium OS (E) and the two mirror TF gradients model (F). The optimal Hill coefficients H^* that give the minimal positional resolution (dashed blue line) and the 2 %EL tolerance interval (solid blue lines). The curves are smoothed by cubic spline interpolation for better visual. Also shown is the value of the highest positional resolution for varying T (dashed orange line). The theoretical results for $M = 1$ for all models are compared to the empirical Hill coefficient H_{data} (blue crosses with error bars) and positional resolution ΔX_{data} (orange crosses with error bars) extracted from MS2-MCP live imaging data[11].

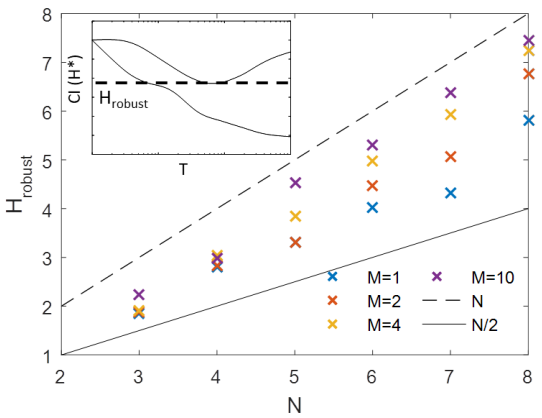

FIG. 4: **The robust optimal Hill coefficient H_{robust} that gives the best positional resolution for the widest range of interphase durations.**
Inset: H_{robust} is calculated as the minimum of the upper bound of the confidence interval $CI(H^*)$. H_{robust} is shown as a function of the OS number (N) and number of readout genes (M). Also plotted for reference are $H = N$ (dashed line) and $H = N/2$ (solid line).

nothing" case (SI Fig. 14). In general, the 'K-or-more" setup does not change the conclusions about the parameter regimes where minimal positional resolution can be obtained (SI Fig. 15).

D. Transcription pattern formed by two 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. While there is no direct evidence for additional regulatory gradients acting in the early cell cycles, the idea of an inverse gradient, possibly indirectly due to Caudal, has been suggested [34]. We assume $N = 6$ binding sites for the Anterior-Posterior decreasing gradient (TF) and $L = 6$ binding sites for Posterior-Anterior decreasing (TF') gradient. Gene expression 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. In the equilibrium model, the pattern can achieve a maximum steepness of $H \sim 7$ given the total of 12 binding sites (Fig. 3F). The quantitative conclusions are the same as for the previously considered models (SI Fig 16 and SI Fig. 17) but the minimal positional resolution (~ 10%EL in nc 12 and nc 13) is smaller 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. This scenario is motivated by the known pattern of the Capicua (Cic) protein and its potential effect on gene expression. In the *hb* promoter sequence there is one known binding motif for the Cic protein [35]. Since the Cic concentration is relatively constant at the *hb* pattern boundary (~ -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).

IV. DISCUSSION

As an external observer looking at the whole embryo, we perceive the boundary steepness as the slope of the gene expression pattern at the border between regions of high and low *hb* expression (Fig. 1B) [13, 14, 16]. Analyzing data from live embryos using the MS2-MCP reporter, when we zoom into individual transcription loci in this boundary region, we observe large variability in the expression state of particular nuclei at similar positions (Fig. 1E, $\delta mRNA/\langle mRNA\rangle \sim 1.5$ [23]). As a result, the pattern appears to be both steep and imprecise at the same time. A simple model of gene regulation identifies a clear trade-off between the boundary expression pattern steepness and the readout precision that impacts the nuclei's ability to make decision about their position with respect to the boundary in early fly development – the quantity that matters for future cellular differentiation [7, 20, 24, 25, 33, 36, 37]. Within the simplest regulatory model there is too little time during the early nuclear cycles to produce a steep border pattern and, within the same regulatory setup, obtain high positional resolution or precision. Conversely, the experimentally observed positional resolution is attainable in the allocated time but it does not correspond to a steep expression pattern within the simplest model of regulation by only Bicoid.

In order to better understand the trade-off between short cell cycles, steepness and precision we studied a family of models where gene expression 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 an equilibiurm model of two TF gradients. Comparing the model predictions to the experimental data, one can construct an equilibrium model that correctly captures

the experimentally observed 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 [16], non-equilibrium models allow for steeper expression profiles. However reaching the experimentally determined Hill coefficients of $H \sim 7$ [11, 14] increases the minimal obtainable positional resolution within a Bcd only binding model with 6 binding sites to $\Delta X \sim 20\%$ EL (corresponding to $\sim$ 10 nuclei widths), slightly above the the experimentally observed value of $\Delta X \sim 12\%$EL in nc 13, corresponding to $\sim$ 6 nuclei widths (Fig. 3 E). Within the hybrid model with 3 equilibrium and 3 non-equilibrium binding sites, the positional resolution can be decreased to observable values in nc 13 and nc 14 for $H < 5$ and $\Delta X \sim 20\%$EL can also be reached with $H \sim 7$ for long cell cycles (SI Fig. 13). However the minimal positional resolution goes quickly up at large H, resulting in small tolerance intervals and a very small steepness value that minimizes the positional resolution in the largest time interval, $H_{\text{robust}} = 2.75$, compared to $H_{\text{robust}} = 3.86$ for the "all-or-nothing" equilibrium model and $H_{\text{robust}} = 3.27$ for the model with two opposing TF gradients. Unfortunately, from the experimental data it is hard to reliably extract Hill coefficients for nc 11. Nevertheless a full non-equilibrium model, which is computationally challenging to explore, could increase the values of the Hill coefficients that are compatible with the observed positional resolution at longer cycles.

One possible way to reconcile steep profiles with small positional information are additional unidentified binding sites in the promoter. Currently the minimal *hb* promoter used in the experiments we are analyzing [11] 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 mostly 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 smaller amplitude than the final zygotic profile. 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 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 decrease the pattern steepness [11]. Nevertheless additional unknown Bcd binding would certainly increase steepness, as could Hb binding sites.

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 [18, 38] ($H \sim 3$).

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 [39, 40]. Nevertheless it provided for a simple choice of parameters and was motivated by earlier theoretical ideas [34], and known activator-repressor pairs in other systems [41]. 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. 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 positional resolution observed at nc 12 is obtained at earlier nc with $H \sim 7$. 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.

There are also other repressor candidates in the fly development, such as Capicua, which is a known repressor gradient albeit with a different profile [42, 43]. 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 precision.

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 fundamental disagreement comes

from the general observation that the steep boundaries do not increase the minimal 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 is much higher than the observed value. Since this is the minimal resolution, the experimentally observed resolution must be greater. This suggests different modes of regulation, such as described above. Yet if a process is fast, and early fly development is very fast, it has to be simple [44].

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 [7], leading to lower positional regulation. 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 is observed in fly development [7].

The positional resolution calculated in the model at the level of promoter activation corresponds in the most noisy case to $CV_P \sim 1$, which is smaller than the variability of mRNA concentrations observed in the MS2-MCP experiments, $\delta mRNA/\langle mRNA \rangle \sim 1.5$ [23].The additional noise in the experiments comes from random arrival times of the polymerase [45], nonuniform progression of the polymerase 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 [23, 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 variance in conjunction with binding models of different complexity remains a future direction.

As has been pointed out in the context of maximizing information flow between the Bcd gradient and Hb output [36], 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 [53]. 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 [54] or the cost of optimal information transmission in a finite time [55].

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 [28, 56] (SI Fig. 6). We looked for the kinetic parameter set that yield 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 $\Delta X \sim 12\%$ 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 rates are independent of the nuclei's positions along the AP axes [44].

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 [24]. 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 [27]. 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 [27]. This renormalization would rescale the time axes to shorter times, but would not qualitatively change the discussed results. 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 ($< 0.1s$) [57], 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 [10, 11, 22] 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 [13]. 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 result in 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.

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 to epigenetic regulation. Additional experiments are needed to discriminate between the proposed scenarios.

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 [16, 58]. 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 2 and solved numerically. Assuming the binding sites are indistinguishable results in the one dimensional equilibrium model in Eq. 2.

B. Randomization of kinetic parameters

To explore all the possible wiring schemes of the model with N OS, we opted to randomize all the kinetic rate constants in $\mathbb{R}^+$ space. Assuming binding is diffusion limited by the Berg-Purcell limit [24], the binding rate constants k_i have an upper bound depending on the OS search time τ_{bind}. 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 $(N-i+1)/\tau_{\text{bind}}$. 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 4). 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) [59, 60] implemented in the SGNS2 simulator [61]. F_- and F_+ are not well-fit by Gaussian distributions, especially for short interphase durations. ΔX is the smallest value of ΔW yielding an 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 $\lambda \approx 100\mu m$ [3], which corresponds to $\sim 20\%$ of the embryo length (EL).

[1] Bahram Houchmandzadeh, Eric Wieschaus, and Stanislas Leibler. Establishment of developmental precision and proportions in the early Drosophila

embryo. *Nature*, 415(6873):798–802, 2002.

[2] Thomas Gregor, William Bialek, Rob R. de Ruyter van Steveninck, David W. Tank, and Eric Wieschaus. Diffusion and scaling during early embryonic pattern formation. *Proceedings of the National Academy of Sciences*, 102(51):18403–18407, 2005.

[3] 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.

[4] Wolfgang Driever and Christiane Nusslein-Volhard. The Bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. *Cell*, 54(1):95–104, 1988.

[5] Wolfgang Driever, Gudrun Thoma, and Christiane Nüsslein-Volhard. Determination of spatial domains of zygotic gene expression in the Drosophila embryo by the affinity of binding sites for the bicoid morphogen. *Nature*, 340(6232):363–367, aug 1989.

[6] Olivier Crauk and Nathalie Dostatni. Bicoid determines sharp and precise target gene expression in the Drosophila embryo. *Current Biology*, 15(21):1888–1898, 2005.

[7] 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.

[8] Johannes Jaeger. The gap gene network. *Cellular and Molecular Life Sciences*, 68(2):243–274, 2011.

[9] Aude Porcher and Nathalie Dostatni. The Bicoid morphogen system. *Current Biology*, 20(5):249–254, 2010.

[10] 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.

[11] Tanguy Lucas, Huy Tran, Carmina Perez Romero, Mathieu Coppey, Cecile Fradin, Aleksandra M Walczak, and Nathalie Dostatni. In preparation.

[12] 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.

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

[14] 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.

[15] 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.

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

[17] Amanda Ochoa-Espinosa, Gozde Yucel, Leah Kaplan, Adam Pare, Noel Pura, Adam Oberstein, Dmitri Papatsenko, and Stephen Small. The role of binding site cluster strength in Bicoid-dependent patterning in Drosophila. *Proceedings of the National Academy of Sciences*, 102(14):4960–4965, 2005.

[18] Xiuguang Ma, Dong Yuan, Katharina Diepold, Tom Scarborough, and Jun Ma. The Drosophila morphogenetic protein Bicoid binds DNA cooperatively. *Development (Cambridge, England)*, 122(4):1195–206, 1996.

[19] Victor Sourjik and Howard C Berg. Receptor sensitivity in bacterial chemotaxis. *Proceedings of the National Academy of Sciences of the United States of America*, 99(1):123–127, 2002.

[20] Dennis Bray, Matthew D Levin, and Carl J Morton-Firth. Receptor clustering as a cellular mechanism to control sensitivity. *Nature*, 393(6680):85–88, may 1998.

[21] Ertugrul M Ozbudak, Mukund Thattai, Han N Lim, Boris I Shraiman, and Alexander Van Oudenaarden. Multistability in the lactose utilization network of Escherichia coli. *Nature*, 427(6976):737–740, feb 2004.

[22] 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.

[23] Jonathan Desponds, Huy Tran, Teresa Ferraro, Tanguy Lucas, Carmina Perez Romero, Aurelien Guillou, Cecile Fradin, Mathieu Coppey, Nathalie Dostatni, and Aleksandra M Walczak. Precision of readout at the hunchback gene: analyzing short transcription time traces in living fly embryos. *PLOS Computational Biology*, 12(12):e1005256, 2016.

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

[25] William Bialek and Sima Setayeshgar. Physical limits to biochemical signaling. *Proceedings of the National Academy of Sciences of the United States of America*, 102(29):10040–5, 2005.

[26] Alexander M Berezhkovskii and Attila Szabo. Effect of ligand diffusion on occupancy fluctuations of cell-surface receptors. *Journal of Chemical Physics*, 139(12):1–7, 2013.

[27] Kazunari Kaizu, Wiet De Ronde, Joris Paijmans, Koichi Takahashi, Filipe Tostevin, and Pieter Rein Ten Wolde. The Berg-Purcell limit revisited. *Biophysical Journal*, 106(4):976–985, 2014.

[28] Monica Skoge, Yigal Meir, and Ned S Wingreen. Dynamics of cooperativity in chemical sensing among cell-surface receptors. *Physical Review Letters*, 107(17):1–5, 2011.

[29] William Bialek and Sima Setayeshgar. Cooperativity, sensitivity, and noise in biochemical signaling. *Physical Review Letters*, 100(25):1–4, 2008.

[30] Archibald V Hill. The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *Journal of Physiology*, 40:4–7, 1910.

[31] 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.

[32] Pieter Rein ten Wolde, Nils B. Becker, Thomas E. Ouldridge, and Andrew Mugler. Fundamental limits to cellular sensing. *Journal of Statistical Physics*, 162(5):1395–1424, 2016.

[33] Thorsten Erdmann, Martin Howard, and Pieter Rein Ten Wolde. Role of spatial averaging in the precision of gene expression patterns. *Physical Review Letters*, 103(25):2–5, 2009.

[34] Thomas R Sokolowski, Thorsten Erdmann, and Pieter Rein ten Wolde. Mutual repression enhances the steepness and precision of gene expression boundaries. *PLoS Computational Biology*, 8(8), 2012.

[35] Gerardo Jiménez, Antoine Guichet, Anne Ephrussi, and Jordi Casanova. Relief of gene repression by Torso RTK signaling: Role of capicua in Drosophila terminal and dorsoventral patterning. *Genes and Development*, 14(2):224–231, 2000.

[36] Gasper Tkacik, Julien O Dubuis, Mariela D Petkova, and Thomas Gregor. Positional information, positional error, and readout precision in morphogenesis: A mathematical framework. *Genetics*, 199(1):39–59, 2015.

[37] Julien O Dubuis, Gasper Tkacik, Eric F Wieschaus, Thomas Gregor, and William Bialek. Positional information, in bits. *Proceedings of the National Academy of Sciences*, 110(41):16301–16308, 2013.

[38] David S Burz, Rolande Rivera-Pomar, Herbert Jäckle, and Steven D Hanes. Cooperative DNA-binding by Bicoid provides a mechanism for threshold-dependent gene activation in the Drosophila embryo. *EMBO Journal*, 17(20):5998–6009, 1998.

[39] Marek Mlodzik, Greg Gibson, and Walter J Gehring. Effects of ectopic expression of caudal during Drosophila development. *Development*, 109(2):271–277, 1990.

[40] Eugenia C Olesnicky, Ava E Brent, Lori Tonnes, Megan Walker, Mary Anne Pultz, David Leaf, and Claude Desplan. A caudal mRNA gradient controls posterior development in the wasp Nasonia. *Development*, 133(20):3973–3982, sep 2006.

[41] Marcin Zagorski, Yoji Tabata, Nathalie Brandenberg, Matthias P. Lutolf, Gašper Tkačik, Tobias Bollenbach, James Briscoe, and Anna Kicheva. Decoding of position in the developing neural tube from antiparallel morphogen gradients. *Science*, 356(6345):1379–1383, 2017.

[42] 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.

[43] Ulrike Löhr, Ho-Ryun Chung, Mathias Beller, and Herbert Jäckle. Bicoid - morphogen function revisited. *Fly*, 4(3):236–240, 2010.

[44] Colleen E Hannon, Shelby A Blythe, and Eric F Wieschaus. Concentration dependent chromatin states induced by the bicoid morphogen gradient. *eLife*, 6:e28275:1–29, 2017.

[45] Benjamin Zoller, Shawn C Little, and Thomas Gregor. Diverse spatial expression patterns emerge from common transcription bursting kinetics. *arXiv*, dec 2017.

[46] Antoine Coulon, Matthew L Ferguson, Valeria de Turris, Murali Palangat, Carson C Chow, and Daniel R Larson. Kinetic competition during the transcription cycle results in stochastic RNA processing. *eLife*, 3:1–22, 2014.

[47] Heng Xu, Samuel O Skinner, Anna Marie Sokac, and Ido Golding. Stochastic kinetics of nascent RNA. *Physical Review Letters*, 117(12):1–6, 2016.

[48] Hernan G Garcia, Alvaro Sanchez, James Q Boedicker, Melisa Osborne, Jeff Gelles, Jane Kondev, and Rob Phillips. Operator sequence alters gene expression independently of transcription factor occupancy in bacteria. *Cell Reports*, 2(1):150–161, 2012.

[49] Daniel L Jones, Robert C Brewster, and Rob Phillips. Promoter architecture dictates cell-to-cell variability in gene expression. *Science*, 346(6216):1533–1536, 2014.

[50] Samuel M D Oliveira, Antti Häkkinen, Jason Lloyd-price, Huy Tran, Vinodh Kandavalli, and Andre S Ribeiro. Temperature-dependent model of multi-step transcription initiation in Escherichia coli based on live single-cell measurements. *PLoS computational biology*, pages 1–18, 2016.

[51] Nitzan Rosenfeld, Jonathan W Young, Uri Alon, Peter S Swain, and Michael B Elowitz. Gene regulation at the single-cell level. *Science*, 307(5717):1962–5, mar 2005.

[52] Mary J Dunlop, Robert Sidney Cox, Joseph H Levine, Richard M Murray, and Michael B Elowitz. Regulatory activity revealed by dynamic correlations in gene expression noise. *Nature genetics*, 40(12):1493–1498, 2008.

[53] Bo Hu, Wouter-Jan Rappel, and Herbert Levine. How input noise limits biochemical sensing in ultrasensitive systems. *Physical Review E*, 90(3):032702, sep 2014.

[54] Felix Wong, Ariel Amir, and Jeremy Gunawardena. An energy-speed-accuracy relation in complex networks for biological discrimination. *arXiv*, 2017.

[55] Francesca Mancini, Matteo Marsili, and Aleksandra M Walczak. Trade-offs in delayed information transmission in biochemical networks. *Journal of Statistical Physics*, 162(5):1088–1129, apr 2015.

[56] Stuart A Sevier and Herbert Levine. Properties of cooperatively induced phases in sensing models. *Physical Review E*, 91(5):052707, may 2015.

[57] Mustafa Mir, Armando Reimer, Jenna E Haines, Xiao Yong Li, Michael Stadler, Hernan Garcia, Michael B Eisen, and Xavier Darzacq. Dense bicoid hubs accentuate binding along the morphogen gradient. *Genes and Development*, 31(17):1784–1794, sep 2017.

[58] Jeremy Gunawardena. A linear framework for timescale separation in nonlinear biochemical systems. *PLoS ONE*, 7(5), 2012.

[59] 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.

[60] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[61] 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.

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

Supporting Information

April 18, 2018

1 Boundary steepness and the Hill coefficient

The boundary steepness is a feature that emerges only when looking at the mean gene expression values $f_P(X)$ of the hunchback gene along the AP axis (see Fig. 1 C and D of the main text) [1, 2, 3, 4]. We assume that the 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:

$$f([TF]) = \frac{[TF]^H}{1 + [TF]^H}, \tag{1}$$

In the case of Bcd-*hb* system the TF concentration decays exponentially along the embryo length [5] and we can estimate the nuclei's position X from the TF concentration:

$$X = \ln([TF]). \tag{2}$$

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) [5]. Eq. 1 becomes:

$$f(X) = \frac{e^{H \cdot X}}{1 + e^{H \cdot X}}. \tag{3}$$

We define the mid-boundary position, $X = 0$, as the position along the AP axis corresponding to half-maximum of the mean expression function, $f_P(X) = 0.5$. 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 [4, 3, 2] or by comparing the maximum derivative of the mean expression function to that of a sigmoid function [1]. 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 expression function at the mid-boundary position ($X = 0$) to the prediction of Eq. 3):

$$H = 4\left(\frac{\partial f_P(X)}{\partial X}\right)_{X=0}. \tag{4}$$

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. 1A), we compare the two values obtained at the steady state of the transcription regulatory model defined in Section 2 for different sets of randomized kinetic parameters. The results shown in SI Fig. 1B for $N = 6$ show the two values of the steepness calculated at different positions

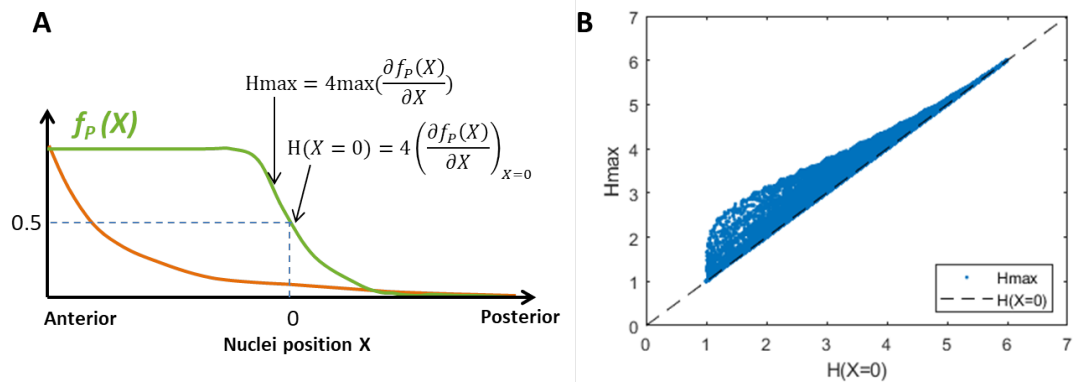

Figure 1: **Quantifying the pattern steepness**.
(A) From the mean expression function $f_P(X)$, the Hill coefficient H can be obtained from either the slope at the mid-boundary position corresponding to half-maximum expression of $f_P(X)$, ($H(X = 0)$) or at the steepest point ($H_{\max}$) as in [1].
(B) Comparison between the two definitions of steepness $H(X = 0)$ and $H_{\max}$ for the 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.

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.

2 Model of transcription regulation

2.1 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 [6] and implemented in [1]. 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:

$$\overline{v} = [v_1, v_2 ... v_{2^N}]. \tag{5}$$

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 $N_{\text{edges}} = N2^N$ [1].

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

$$e = [e_+, e_-] = [e_{+1}, e_{+2}, ...e_{+N2^{N-1}}, e_{-1}, e_{-2}, ...e_{-N2^{N-1}}], \tag{6}$$

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

$$\overline{k} = [k_+, k_-] = [k_{+1}, k_{+2}, ...k_{+N2^{N-1}}, k_{-1}, k_{-2}, ...k_{-N2^{N-1}}]. \tag{7}$$

An example of the labeled graph is shown in SI Fig. 2 for $N = 3$.

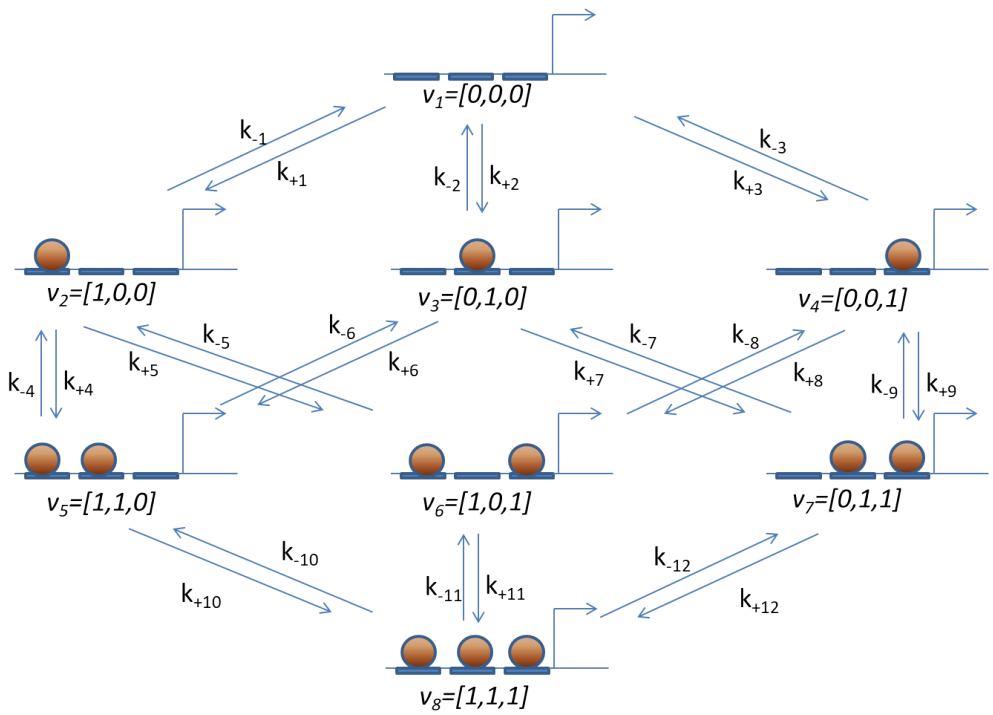

Figure 2: **The labeled graph for the general non-equilibrium binding site model of transcription regulation for $N = 3$ OS**. Each vertex v corresponds to a unique OS array state. The edge labels describe the binding and unbinding reactions, assuming one reaction takes place at a time.

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 [1], 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 at any given time can be approximated by the probability of the gene to be active, $f(X) = P(X)$.

2.2 The thermodynamic equilibrium model

2.2.1 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 – a promoter states with i bound TF molecules:

$$P_0 \underset{k_{-1}}{\overset{k_1[TF]}{\rightleftharpoons}} P_1 \underset{k_{-2}}{\overset{k_2[TF]}{\rightleftharpoons}} P_2 ... \underset{k_{-N}}{\overset{k_N[TF]}{\rightleftharpoons}} P_N, \tag{8}$$

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 occur 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 1 at mid-boundary position.

2.2.2 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. 8 is given by:

$$\partial_t P(P_i) = k_i[TF]P(P_{i-1}) + k_{-(i+1)}P(P_{i+1}) - (k_{-i} + k_{i+1}[TF])\,P(P_i). \tag{9}$$

The steady state solution is:

$$P(P_i, [TF]) = \frac{\prod_{j=0}^{i} K_j [TF]^i}{\sum_{\ell=0}^{N} \prod_{j=0}^{\ell} K_j [TF]^\ell}, \tag{10}$$

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:

$$P(P_i, [TF]) = \frac{\tilde{K}_i [TF]^i}{\sum_{j=0}^{N} \tilde{K}_j [TF]^j}. \tag{11}$$

If the TF concentration gradient follows an exponential curve, we can express the TF concentration $[TF]$ in terms of the nuclei's position (Eq. 2) and Eq. 11 becomes:

$$P(P_i, X) = \frac{\tilde{K}_i e^{iX}}{\sum_{j=0}^{N} \tilde{K}_j e^{jX}}. \tag{12}$$

3 Boundary steepness and promoter switching time for the equilibrium model

3.1 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 $P(P_{\text{active}}) = p$ $(0 \leq p \leq 1)$, Eq. 11 is simplified to:

$$P(P_{\text{active}}, X = 0) = \frac{\sum_{i=K}^{N} \tilde{K}_i}{\sum_{j=0}^{N} \tilde{K}_j} = p, \tag{13}$$

which imposes a condition on the effective equilibrium constants:

$$\sum_{i=K}^{N} \tilde{K}_i = \frac{p}{1-p} \sum_{k=0}^{K-1} \tilde{K}_k = p \sum_{j=0}^{N} \tilde{K}_j. \tag{14}$$

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

$$
\begin{aligned}
\left(\frac{\partial P(P_{\text{active}})}{\partial X}\right)\Big|_{X=0} &= \sum_{i=K}^{N}\left(\frac{\partial}{\partial X}\frac{\tilde{K}_i e^{iX}}{\sum_{j=0}^{N}\tilde{K}_j e^{jX}}\right)\Big|_{X=0} \\
&= \frac{\sum_{i=K}^{N} i\tilde{K}_i}{\sum_{j=0}^{N}\tilde{K}_j} - \frac{\sum_{i=K}^{N}\tilde{K}_i \sum_{j=0}^{N} j\tilde{K}_j}{(\sum_{j=0}^{N}\tilde{K}_j)^2} \\
&= \frac{\sum_{i=K}^{N} i\tilde{K}_i}{\sum_{j=0}^{N}\tilde{K}_j} - \frac{p\sum_{j=0}^{N} j\tilde{K}_j}{\sum_{j=0}^{N}\tilde{K}_j},
\end{aligned}
\tag{15}
$$

where in the last step we used Eq. 14.

For clarity, we set the ranges $i = K..N$, $j = 0..N$ and $k = 0..K-1$. Eq. 15 is then rewritten as:

$$
\begin{aligned}
\left(\frac{\partial P(P_{\text{active}})}{\partial X}\right)_{X=0} &= \frac{\sum i\tilde{K}_i - p\sum j\tilde{K}_j}{\sum \tilde{K}_j} \\
&= \frac{(1-p)\sum i\tilde{K}_i - p\sum k\tilde{K}_k}{\sum \tilde{K}_j} \\
&= \frac{N(1-p)\sum \tilde{K}_i}{\sum \tilde{K}_j} - \frac{(1-p)\sum (N-i)\tilde{K}_i + p\sum k\tilde{K}_k}{\sum \tilde{K}_j} \\
&= Np(1-p) - \frac{(1-p)\sum (N-i)\tilde{K}_i + p\sum k\tilde{K}_k}{\sum \tilde{K}_j},
\end{aligned}
\tag{16}
$$

and the Hill coefficient (Eq. 4) is:

$$
\begin{aligned}
H &= 4\frac{\partial P(P_{\text{active}})}{\partial X} \\
&= 4Np(1-p) - 4\frac{(1-p)\sum (N-i)\tilde{K}_i + p\sum k\tilde{K}_k}{\sum \tilde{K}_j}.
\end{aligned}
\tag{17}
$$

At the boundary criteria $p = 0.5$ and:

$$
H = N - 2\frac{\sum (N-i)\tilde{K}_i + \sum k\tilde{K}_k}{\sum \tilde{K}_j}. \tag{18}
$$

In the "all-or-nothing" case ($K = N$), $\sum_{j=0}^{N}\tilde{K}_j = 2\tilde{K}_N$ (Eq. 14), the first term in the nominator disappears and Eq. 18 becomes

$$
H = N - \frac{\sum_{k=0}^{N-1} k\tilde{K}_k}{\tilde{K}_N}. \tag{19}
$$

3.2 Bounds for pattern steepness

Eq. 18 gives an upper bound of $H \approx N$ at the mid-boundary position, which occurs when

$$
\frac{\sum_{i=K}^{N}(N-i)\tilde{K}_i + \sum_{k=0}^{K-1} k\tilde{K}_k}{\sum_{j=0}^{N}\tilde{K}_j/2} << N. \tag{20}
$$

When N is not too large (≤ 10), we can rewrite the bound condition:

$$\sum_{i=K}^{N}(N-i)\tilde{K}_i + \sum_{k=0}^{K-1} k\tilde{K}_k \ll \sum_{j=0}^{N}\tilde{K}_j/2 \tag{21}$$

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).

These conditions correspond to the case when the promoter spends most of its time in P_0 or P_N ($P(P_0) + P(P_N) \approx 1$), 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. 18:

$$\begin{aligned} N - H &= \frac{\sum_{i=K}^{N}(N-i)\tilde{K}_i + \sum_{k=0}^{K-1} k\tilde{K}_k}{\sum_{j=0}^{N}\tilde{K}_j/2} \\ &= \frac{\sum_{i=K}^{N}(N-K)\tilde{K}_i - \sum_{i=K}^{N}(i-K)\tilde{K}_i}{\sum_{i=K}^{N}\tilde{K}_i} + \frac{\sum_{k=0}^{K-1}(K-1)\tilde{K}_k - \sum_{k=0}^{K-1}(K-1-k)\tilde{K}_k}{\sum_{k=0}^{K-1}\tilde{K}_k} \\ &= N - K - \frac{\sum_{i=K}^{N}(i-K)\tilde{K}_i}{\sum_{i=K}^{N}\tilde{K}_i} + K - 1 - \frac{\sum_{k=0}^{K-1}(K-1-k)\tilde{K}_k}{\sum_{k=0}^{K-1}\tilde{K}_k}, \\ &\leq N - 1 \end{aligned} \tag{22}$$

Thus $H \geq 1$. $H = 1$ when the sum in Eq. 21 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.

3.3 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:

$$\begin{aligned} \tau_{\text{active}} &= \tau_N \frac{P(P_{\text{active}})}{P(P_N)} \\ &= \frac{1}{k_{-N}} \frac{\sum_{i=K}^{N}\tilde{K}_i}{\tilde{K}_N} \\ &= \frac{K_N}{k_N} \frac{\sum_{i=K}^{N}\tilde{K}_i}{\tilde{K}_N} \\ &= \frac{1}{k_N} \frac{\tilde{K}_N}{\tilde{K}_{N-1}} \frac{\sum_{i=K}^{N}\tilde{K}_i}{\tilde{K}_N} \\ &= \tau_{\text{bind}} \frac{\sum_{j=0}^{N}\tilde{K}_j}{2\tilde{K}_{N-1}}, \end{aligned} \tag{23}$$

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 ($[TF] = 1$).

Eq. 23 allows us to connect the Hill coefficient in Eq. 19 to τ_{active}. For $K = N$ (the "all-or-nothing" case), using Eq. 23, Eq. 19 becomes:

$$\begin{aligned} H &= N - \frac{\sum_{k=0}^{N-1} k\tilde{K}_k}{\sum_{j=0}^{N} \tilde{K}_j/2} \\ &= N - \frac{(N-1)\tilde{K}_{N-1} + \sum_{k=0}^{N-2} k\tilde{K}_k}{\sum_{j=0}^{N} \tilde{K}_j/2} \\ &= N - \frac{(N-1)\tau_{\text{bind}}}{\tau_{\text{active}}} - \frac{\sum_{k=0}^{N-2} k\tilde{K}_k}{\sum_{j=0}^{N} \tilde{K}_j/2} \\ &\leq N - \frac{(N-1)\tau_{\text{bind}}}{\tau_{\text{active}}}, \end{aligned} \tag{24}$$

or

$$\tau_{\text{active}} \geq \tau_{\text{bind}} \frac{N-1}{N-H}. \tag{25}$$

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

$$H \leq N - (N-1)\frac{\tau_{\text{bind}}}{\tau_{\text{active}}}. \tag{26}$$

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 [7], 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 [8, 9]), $c \sim 11.2/\mu m^3$ [8] – the absolute TF concentration at the mid-boundary position and $a \sim 3nm$ [4] – the size of one operator site for Bicoid. Given this estimate, $H \sim N$ for $\tau_{\text{active}} \gg 4$ s or $k_{-N} \ll 0.25$ s^{-1}.

For $K < N$ (the "K-or-more" case), Eq. 18 becomes:

$$\begin{aligned} H &= N - \frac{\tilde{K}_{N-1} + \sum_{m=K}^{N-2}(N-m)\tilde{K}_m + \sum_{k=0}^{K-1} k\tilde{K}_k}{\sum_{j=0}^{N} \tilde{K}_j/2} \\ &= N - \frac{\tau_{\text{bind}}}{\tau_{\text{active}}} - \frac{\sum_{m=K}^{N-2}(N-m)\tilde{K}_m + \sum_{k=0}^{K-1} k\tilde{K}_k}{\sum_{j=0}^{N} \tilde{K}_j/2} \\ &\leq N - \frac{\tau_{\text{bind}}}{\tau_{\text{active}}}, \end{aligned} \tag{27}$$

or

$$\tau_{\text{active}} \geq \tau_{\text{bind}} \frac{1}{N-H}. \tag{28}$$

The equality in Eq. 24 and Eq. 27 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.

4 Aligning the pattern boundary position

We rescale the randomized unbinding rate values k_{-i} to $\overline{k}$ so that the mid-boundary of the expression pattern is located at the mid-embryo position, $P(P_{\text{active}}, X = 0) = 0.5$. We do this to eliminate additional variability coming from different TF binding times to the promoter (τ_{bind}) at different Bcd concentrations and compare all the results at one position.

By solving for the probability of the gene to be active given a kinetic parameter set $\overline{k}$ at position X, $P_{active}(\overline{k}, X)$ at steady state (Eq. 11), we find the mid-boundary position $X_0(\overline{k})$ such that $P_{\text{active}}(\overline{k}, X_0(\overline{k})) = 0.5$. Multiplying all the kinetic parameters in $\overline{k}$ by $e^{X_0((\overline{k}))}$ and rescaling the relative position does not alter the system's steady-state solution:

$$P_{\text{active}}(\overline{k}, X) = P_{active}([k_+, k_-], X) \tag{29}$$

$$= P_{\text{active}}([k_+ e^{X_0((\overline{k}))}, k_- e^{X_0((\overline{k}))}], X) \tag{30}$$

$$= P_{\text{active}}([k_+, k_- e^{X_0((\overline{k}))}], X - X_0(\overline{k})), \tag{31}$$

and the whole pattern can be shifted so that the mid-boundary position is located at $X = 0$:

$$P_{\text{active}}([k_+, k_- e^{X_0(\overline{k})}], 0) = P_{active}([k_+, k_-], X_0(\overline{k})) = 0.5, \tag{32}$$

where the new parameter set is $\overline{k'} = [k_+, k_- e^{X_0(\overline{k})}]$.

5 Calculating mean promoter activity level and readout precision

In this section we obtain analytical solutions for the time dependent mean promoter activity ($\mu_P(T, 0)$) and readout precision (CV_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. 9. 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 $U\overline{x} = 0$ and the normalization condition $\sum \overline{x} = 1$. In the equilibrium model, the steady state solution $\overline{x}$ is given by Eq. 12.

5.1 Mean promoter activity level from 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:

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

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.

5.2 The readout precision

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:

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

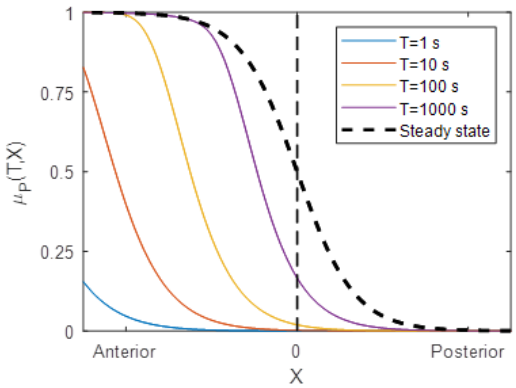

Figure 3: **Example of the pattern formation following mitosis:**. The mean expression level pattern $\mu_P(T, X)$ as a function of the position in the embryo the end of an interphase of duration T (solid colored lines), for interphases of varying duration. The dashed line shows the steady-state expression pattern.

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 at time :

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

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:

$$\begin{aligned}
\langle f_P^2(T) \rangle &= \frac{1}{T^2} \int_0^T du \int_0^T dv \; \alpha^{\mathrm{T}} e^{U|u-v|} x_{\text{fire}} \\
&= \frac{2}{T^2} \int_0^T du \int_0^u dv \; \alpha^{\mathrm{T}} e^{U(u-v)} x_{\text{fire}} \\
&= \frac{2\alpha^{\mathrm{T}}}{T^2} \int_0^T ds \int_s^T du \; e^{Us} x_{\text{fire}} \\
&= \frac{2\alpha^{\mathrm{T}}}{T^2} \left[\int_0^T ds (T-s) e^{Us} \right] x_{\text{fire}}.
\end{aligned} \tag{36}$$

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

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

with

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

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

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

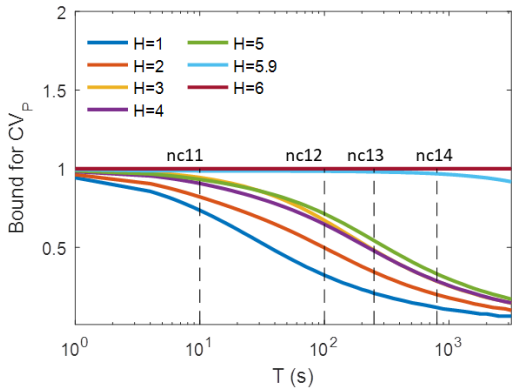

Figure 4: **The precision of [TF] readout at steady-state**: The smallest values of CV_P as a function of interphase duration T, plotted for varying pattern sharpness H.

and for $\lambda_i \neq 0$:

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

In the 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:

$$\begin{aligned} \langle f_P^2(T) \rangle &= \frac{2\alpha^{\mathrm{T}}}{T^2} V \cdot T^2/2 \cdot V^{-1} \overline{x} \\ &= \alpha^{\mathrm{T}} \overline{x} \\ &= \langle f_P(T) \rangle \end{aligned} \tag{41}$$

The readout error at the mid-boundary position is therefore:

$$CV_P(T) = \sqrt{\frac{\langle f_P(T) \rangle - \langle f_P(T) \rangle^2}{\langle f_P(T) \rangle^2}} = 1 \tag{42}$$

5.3 Specific case with 2 binding sites

In the specific case $N = 2$:

$$P_0 \underset{k_{-1}}{\overset{k_1[TF]}{\rightleftharpoons}} P_1 \underset{k_{-2}}{\overset{k_2[TF]}{\rightleftharpoons}} P_2 \tag{43}$$

the matrix U is:

$$U = \begin{bmatrix} -k_1 & k_{-1} & 0 \\ k_1 & -k_2 - k_{-1} & k_{-2} \\ 0 & k_2 & -k_{-2} \end{bmatrix}, \tag{44}$$

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

$$\begin{cases} r = k_1k_2 + k_{-1}k_2 + k_{-1}k_{-2} \\ w = k_1 + k_2 + k_{-1} + k_{-2} \\ v = k_{-2} - k_1 - k_2 - k_{-1} \\ x = k_2 + k_{-2} - k_1 - k_{-1} \\ d = \sqrt{(k_1 + k_2 + k_{-1} + k_{-2})^2 - 4(k_1k_2 + k_1k_{-2} + k_{-1}k_{-2})} \end{cases} \tag{45}$$

and for the all-or-nothing case:

$$\alpha = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{46}$$

the steady state probability is:

$$\langle f_P(T) \rangle = \alpha^{\mathrm{T}} \overline{x} = \alpha^{\mathrm{T}} \begin{pmatrix} k_{-1}k_{-2}/r \\ k_{-2}k1/r \\ k_1k_2/r \end{pmatrix} = k_1k_2/r, \tag{47}$$

and the mean squared of the readout error is:

$$\begin{aligned} \langle f_P^2(T) \rangle = & \frac{2k_1k_2}{T^2 r} \times \\ & \Big[\frac{T^2}{2r} \big(k_2k_{-2} + \frac{v^2 - d^2}{4}\big) \\ & + \big(\frac{2T}{w-d} - \frac{4}{(w-d)^2}\big(1 + e^{-(w-d)T/2}\big)\big) \frac{k_{-2}(x-d)}{2r - 2w(w-d) + 3(w-d)^2/2}\big) \\ & + \big(\frac{2T}{w+d} - \frac{4}{(w+d)^2}\big(1 + e^{-(w+d)T/2}\big) \frac{k_{-2}(x+d)}{2r - 2w(w+d) + 3(w+d)^2/2} \Big] \end{aligned} \tag{48}$$

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

6 Positional resolution

6.1 Calculation of positional resolution

For each set of parameters $\overline{k}$, interphase duration 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 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 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:

$$D = F_+ - F_-. \tag{49}$$

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 s:

$$\beta = P(D \leq 0) = P(F_+ \leq F_-). \tag{50}$$

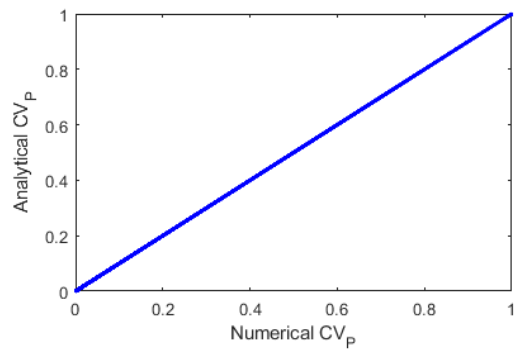

Figure 5: **The analytical and numerical calculation of the relative error $CV_P(T)$ for the equilibrium model with $N = 2$ (SI section 5)**. Small discrepancies result from numerically finding the half-maximum expression point, which due to numerical precision is not exactly at $f_P = 0.5$.

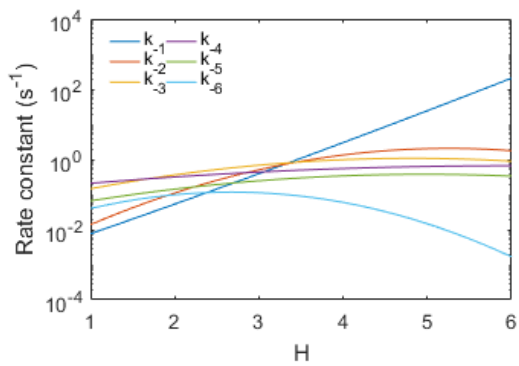

Figure 6: **The unbinding rate constants k_{-i} yielding the highest precision CV_P in nuclear cycle 12 of the equilibrium model with N=6**. The optimal binding rate constants k_i are at their highest possible values $(N - i + 1)/\tau_{\text{bind}}$.

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):

$$\Delta X = \min\left(\Delta W | \beta \leq 0.05\right). \tag{51}$$

In practice, to determine the value of ΔX for each parameter set $\overline{k}$, we increase ΔW from 0 λ to 4 λ with an increment of 0.01 λ (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

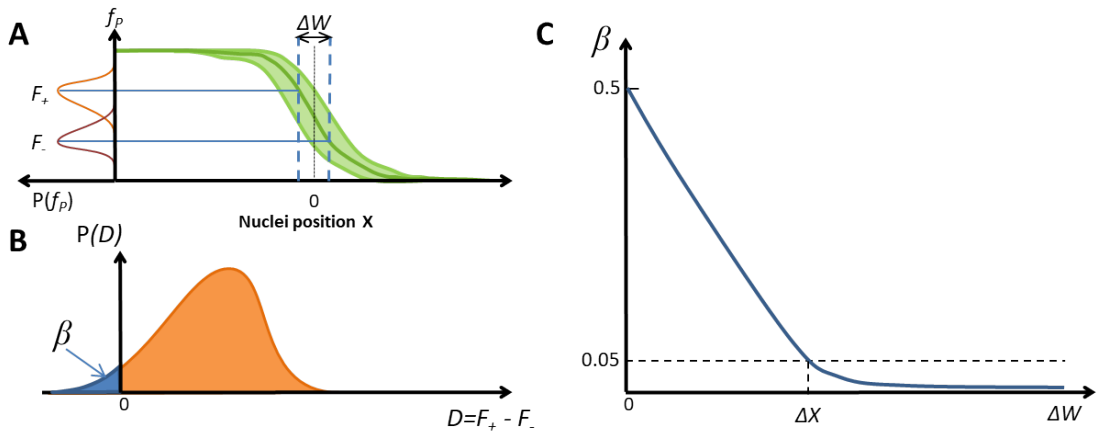

Figure 7: **Finding the positional resolution**.
(A) The distribution of readout F_+ and F_- from two nuclei positioned at a distance of ΔW on opposite sides of the expression boundary.
(B) The coefficient $\beta = P(D \leq 0)$ is the risk of a nucleus wrongly predicting its position.
(C) ΔX is set as the smallest ΔW yielding a tolerable risk ($\beta \leq 5\%$).

locations $-\Delta W/2$ and $+\Delta W/2$ is:

$$D_M = \frac{1}{M}\sum_{j=1}^{M}(F_+(j) - F_-(j)), \tag{52}$$

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).

6.2 Positional resolution for 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. 2 of the main manuscript) with a mean value $f_P(X)$. We assume that the readout pattern can be well fit with a sigmoid curve with a Hill coefficient H:

$$f_P(X) = \frac{e^{HX}}{1 + e^{HX}}. \tag{53}$$

For the case $M = 1$ (single gene readout), at the anterior position $\Delta W/2$, the positional readout F_+ has a chance $f_P(\Delta W/2)$ to be 1. Similarly, at the opposite position $-\Delta W/2$, the positional 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:

$$\begin{aligned} \beta &= P(F_+ \leq F_-) = 1 - P(F_+ > F_-) \\ &= 1 - P(F_+ = 1, F_- = 0) \\ &= 1 - f_P(\Delta W/2)^2. \end{aligned} \tag{54}$$

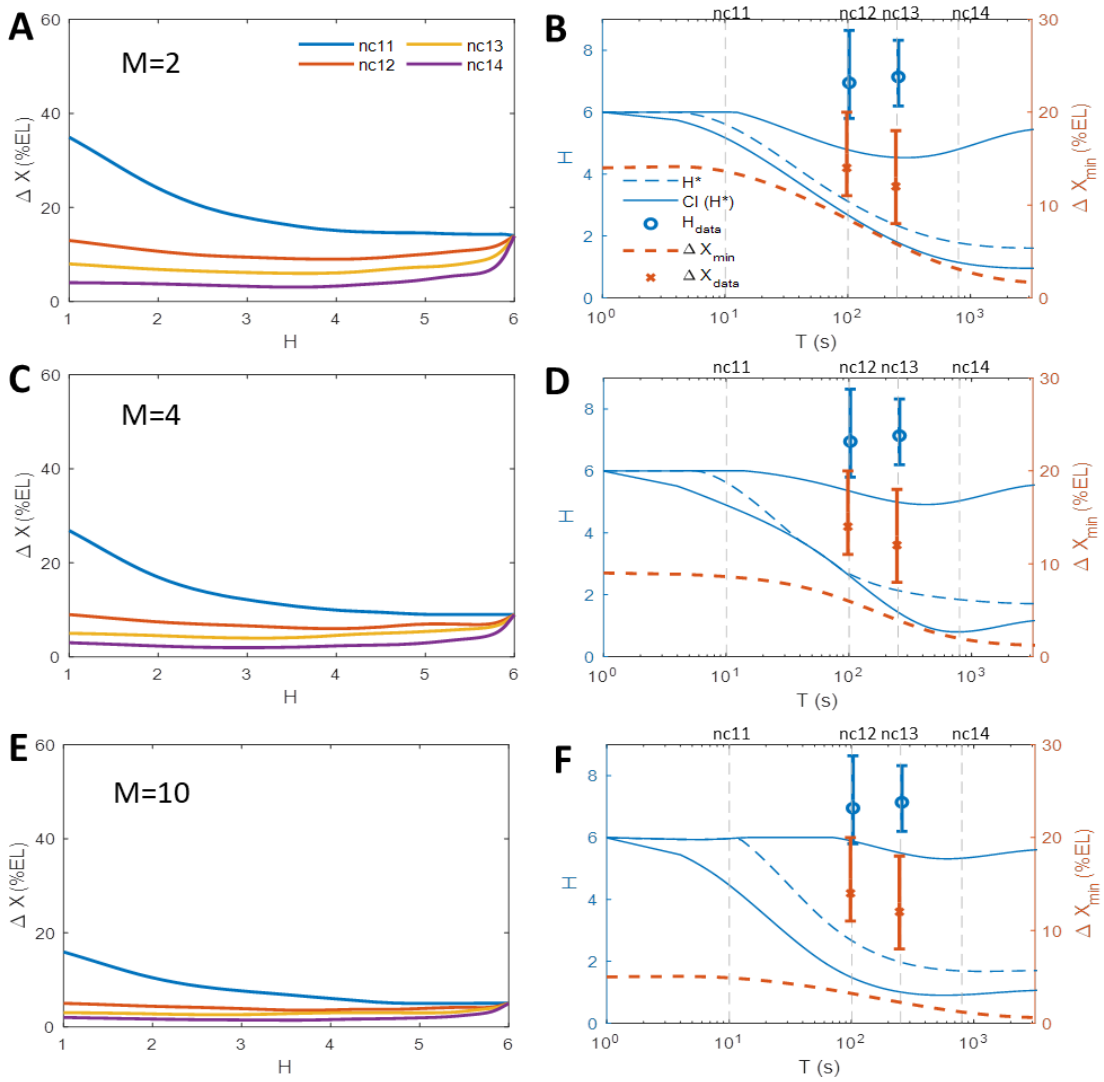

Figure 8: **Positional resolution with varying single gene copy number per nuclei M in the equilibrium model.**. Results for the "all-or-nothing" model with $N = 6$ binding sites and M equal 2 (A-B), 4 (C-D) and 10 (E-F).
(A,C,E) Positional resolution calculated from the equilibrium binding site model 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 value of the highest positional resolution ΔX_{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.

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

$$1 - f_P(\Delta X/2)^2 = 0.05, \tag{55}$$

or

$$f_P(\Delta X/2) = \sqrt{1-0.05} \tag{56}$$
$$= \frac{e^{H\Delta X/2}}{1+e^{H\Delta X/2}}$$
$$= 1 - \frac{1}{1+e^{H\Delta X/2}},$$

which gives

$$\Delta X = \frac{2}{H} ln(\frac{1}{1-\sqrt{0.95}} - 1) \approx \frac{7.3}{H}. \tag{57}$$

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

$$\Delta X \approx \frac{7.3}{H} * 20\%EL = \frac{146\%EL}{H} \tag{58}$$

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

$$F_+ \sim \frac{1}{M} B(M, f_P(\Delta W/2)), \tag{59}$$
$$F_- \sim \frac{1}{M} B(M, f_P(-\Delta W/2)),$$

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

7 Expression pattern of *hb* promoter variants in live *Drosophila* embryos

7.1 Proximal *hb* promoter

To calculate the positional resolution of the *hb* pattern in live fly embryos, we observe the transcription dynamics of a 700bp *hb* proximal promoter using the RNA-tagging MS2-MCP system [12, 13]. 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 [14].

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 on the pattern's Hill coefficient (SI Fig. 9). 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).

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 $\Delta X \sim 14\%$ (confidence interval from 11% to 20%) in nuclear cycle 12 and $\Delta X \sim 12\%$ (confidence interval from 8% to 18%) in nuclear cycle 13.

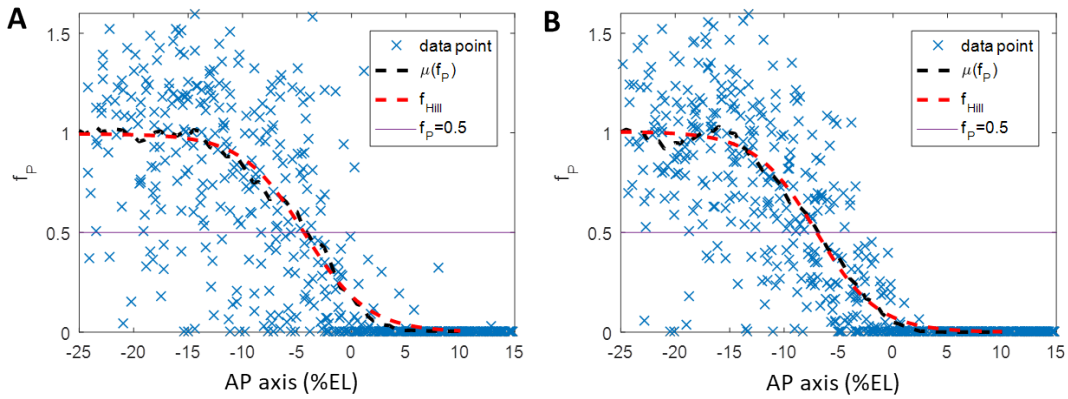

Figure 9: **The gene expression readout pattern by *hb* proximal promoter along AP axis:** normalized fluorescence (blue crosses), the mean readout $\mu(f_P)$ (dashed black line), the fitted Hill function (dashed red line) and $f_P = 0.5$ (solid yellow line) as a function of nuclei position. The normalized 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).

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:

$$\mu_{SS} = |\mu_P(T, 0) - 0.5| + |\frac{\partial \mu_P(T, 0)}{\partial T}|. \tag{60}$$

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.

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 [1]. Similarly to the equilibrium model we observe a tradeoff between the pattern steepness H, readout precision 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 ($CV_P = 1$).

We expand the non-equilibrium model to $N = 6$. However, we do not use a full model (as in SI Fig. 2) 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. 8). The second array contains 3 OS, the interactions of which with the TF are out of equilibrium (as in SI Fig. 2). 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.

The hybrid model is able to achieve a steepness of 8 (SI Fig. 12), as expected from equilibrated

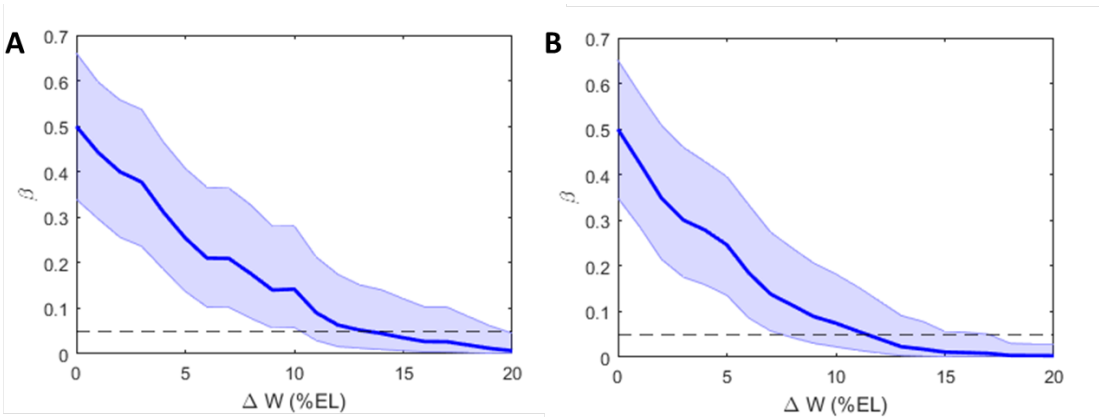

Figure 10: **The risk factor value β as a function of ΔW for the *hb* proximal promoter (solid line)**, plotted with the confidence interval (shaded) with p-value=0.05. The dashed black line indicates the tolerable risk $\beta = 0.05$. (A) Nuclear cycle 12 (8 embryos). (B) Nuclear cycle 13 (4 embryos).

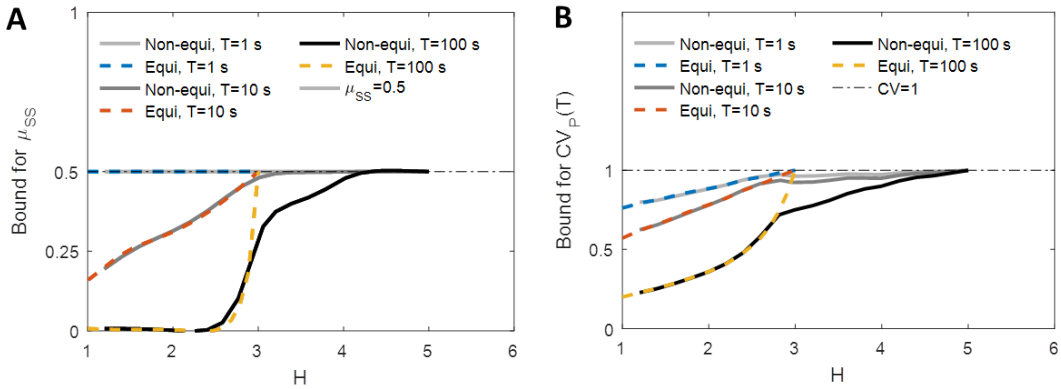

Figure 11: **Readout precision of a pure non-equilibrium model for $N = 3$.**
(A) The lower bounds for μ_{SS} from the non-equilibrium model (grey solid line), for varying values of T, computed from 3×10^5 data points. Also shown are the bounds for equilibrium model (colored dashed lines).
(B) The lower bounds for readout precision $CV_P(T)$ 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).

activity of 3 OS and non-equilibrated activity of 3 OS. The tradeoff between the pattern steepness H and the readout precision and pattern formation time still hold. 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. 13A and the optimal steepness with varying interphase duration T is plotted in Fig. 3E of the main text.

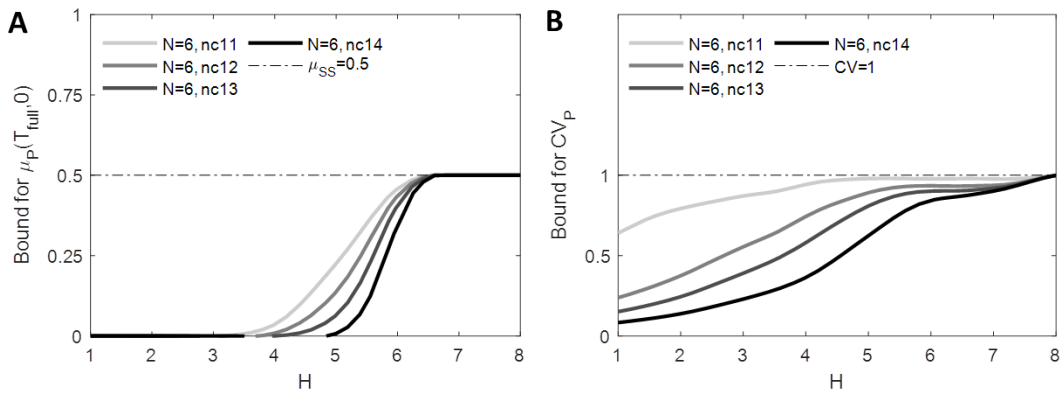

Figure 12: **Readout precision of the hybrid non-equilibrium model with $N = 6$, three equilibrium and three non-equilibrium OS**.
(A) The lower bounds for μ_{SS} for a hybrid non-equilibrium model (grey solid lines) for varying values of T computed from 10^6 data points.
(B) The lower bounds for the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

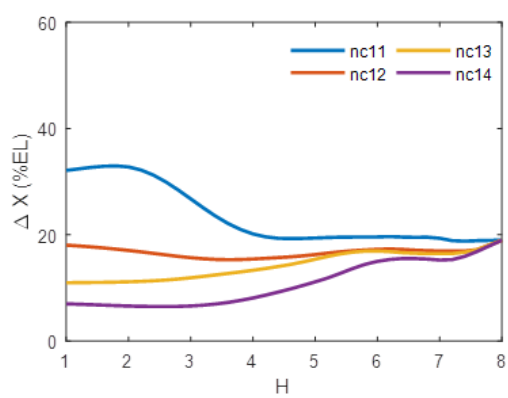

Figure 13: **Positional resolution in the $N = 6$ hybrid non-equilibrium model.** The results are shown for the "all-or-nothing" case with $N = 6$, $M = 1$. Positional resolution calculated for a hybrid non-equilibrium binding site model for varying boundary steepness H.

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.

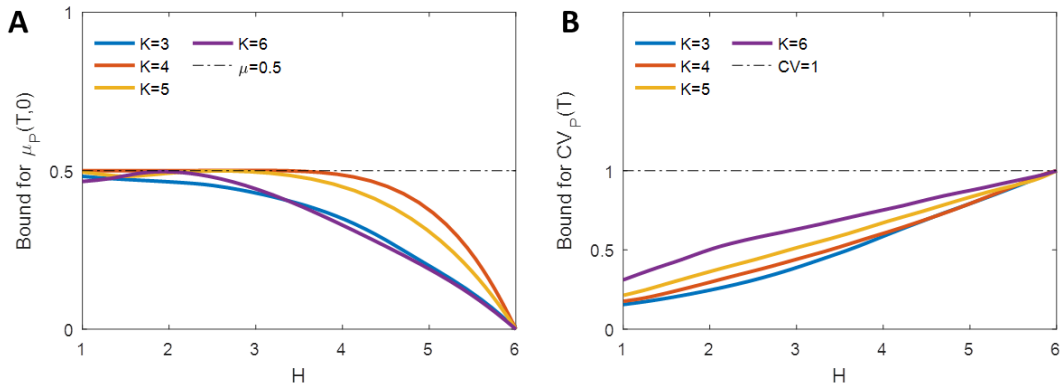

Figure 14: **The "K-or-more" case is qualitatively similar to the "all or nothing" case.** The results are shown with $N = 6$ binding sites and an interphase duration of $T = 500$ s. $K = 6$ corresponds to the "all or nothing" case. Each curve is computed from ~ 20000 data points.
(A) The lower bound 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, $CV = 1$.

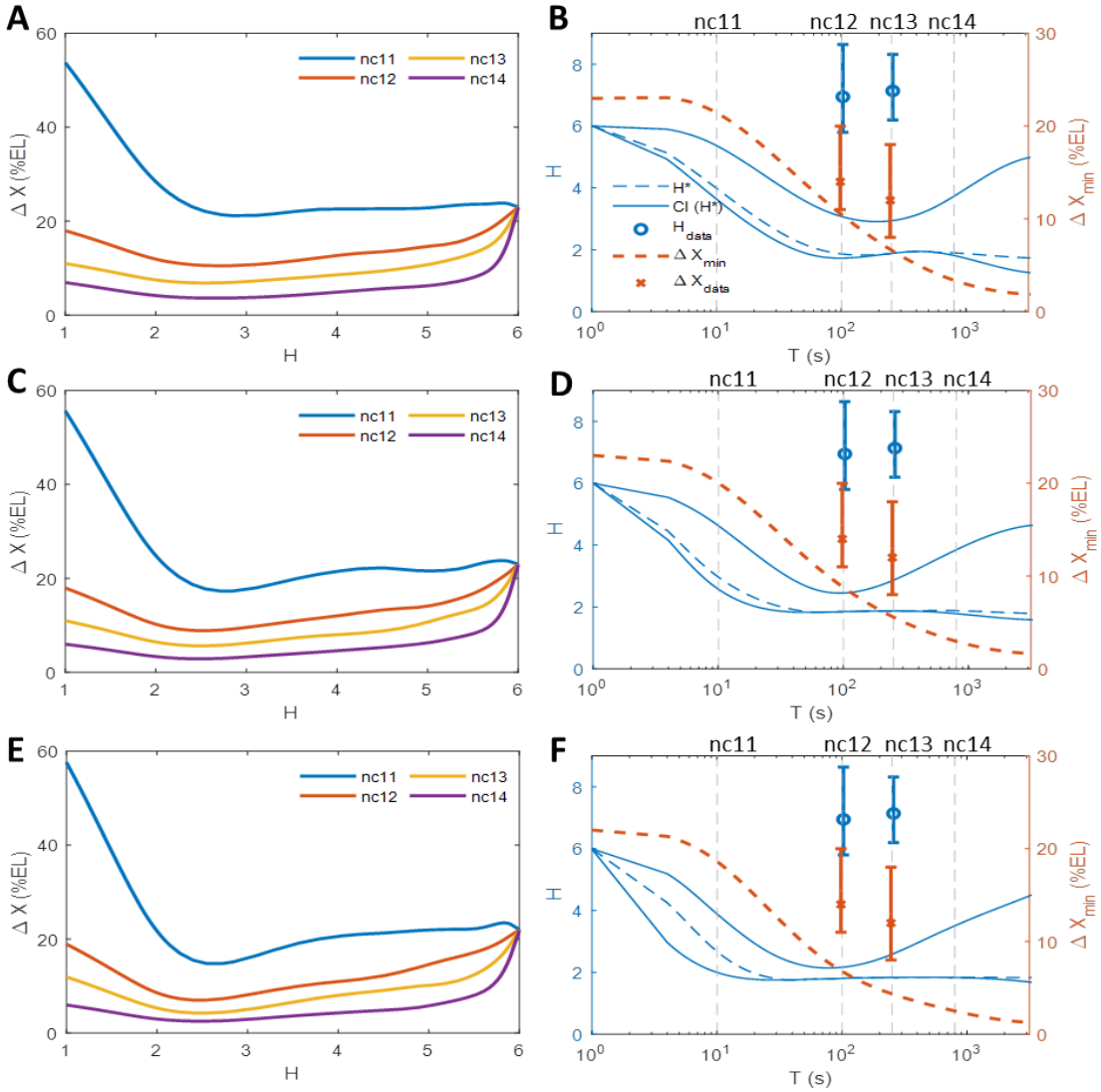

Figure 15: **Positional resolution in the "K-or-more" case, with varying K.** The results are shown with $N = 6$ binding sites andADS K equal 5 (A-B), 4 (C-D), 3 (E-F). $M = 1$.
(A,C,E) Positional resolution calculated from the equilibrium binding site model for varying boundary steepness H.
(B,D,F) The optimal Hill coefficients H^* that gives the minimal positional resolution (dashed blue line), the confidence interval CI(H^*) with 2 %EL tolerance (solid blue lines) and the value of the highest positional resolution ΔX_{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 gene expression 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:

$$A_0 \underset{k_{-1}}{\overset{k_1[TF]}{\rightleftharpoons}} A_1 \underset{k_{-2}}{\overset{k_2[TF]}{\rightleftharpoons}} A_2 ... \underset{k_{-N}}{\overset{k_N[TF]}{\rightleftharpoons}} A_N, \tag{61}$$

$$B_0 \underset{k'_{-1}}{\overset{k'_1[TF']}{\rightleftharpoons}} B_1 \underset{k'_{-2}}{\overset{k'_2[TF']}{\rightleftharpoons}} B_2 ... \underset{k'_{-L}}{\overset{k'_L[TF']}{\rightleftharpoons}} B_L, \tag{62}$$

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:

$$A(T) = \frac{1}{T}\int_{t=0}^{T} \alpha^\intercal a(t)dt \tag{63}$$
$$B(T) = \frac{1}{T}\int_{t=0}^{T} \gamma^\intercal b(t)dt.$$

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

$$P_{\text{active}} = \alpha^\intercal a(t)\gamma^\intercal b(t). \tag{64}$$

The promoter readout is given by:

$$P(T) = \frac{1}{T}\int_{t=0}^{T} \alpha^\intercal a(t)\gamma^\intercal b(t)dt. \tag{65}$$

At a given position, the two arrays have rate matrices U_a and U_b respectively. We call $\overline{x}_a$ and $\overline{x}_b$ 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:

$$[TF'] = e^{-X}. \tag{66}$$

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

10.1.1 The pattern steepness

The mean promoter readout is given by:

$$\langle P(T)\rangle = \frac{1}{T}\int_{t=0}^{T}\langle \alpha^{\intercal} a(t)\gamma^{\intercal} b(t)\rangle dt. \tag{67}$$

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

$$\begin{aligned}\langle P(T)\rangle &= \frac{1}{T}\int_{t=0}^{T}\langle \alpha^{\intercal} a(t)\rangle\langle \gamma^{\intercal} b(t)\rangle dt \\ &= \langle A(T)\rangle\langle B(T)\rangle,\end{aligned} \tag{68}$$

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

$$\begin{aligned}H_P &= 4(\frac{\partial\langle P(T)\rangle}{\partial X})_{X_0} \\ &= 4(\frac{\partial(\langle A(T)\rangle\langle B(T)\rangle)}{\partial X})_{X_0} \\ &= 4\langle B(T)\rangle(\frac{\partial\langle A(T)\rangle}{\partial X})_{X_0} + 4\langle A(T)\rangle(\frac{\partial\langle B(T)\rangle}{\partial X})_{X_0} \\ &= \langle B(T)\rangle H_A + \langle A(T)\rangle H_B.\end{aligned} \tag{69}$$

Given Eq. 69, we have:

$$\begin{aligned}H_P &\leq 4N\langle B(T)\rangle\langle A(T)\rangle(1-\langle A(T)\rangle) + 4L\langle A(T)\rangle\langle B(T)\rangle(1-\langle B(T)\rangle) \\ &= 4\langle P(T)\rangle(N + L - N\langle A(T)\rangle - L\langle B(T)\rangle) \\ &= 2(N + L - N\langle A(T)\rangle - L\langle B(T)\rangle)\end{aligned} \tag{70}$$

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

$$\begin{aligned}H_P &\leq 4N - 2N(\langle A(T)\rangle + \langle B(T)\rangle) \\ &\leq 4N - 4N\sqrt{\langle A(T)\rangle\langle B(T)\rangle} \\ &= 4N(1-\frac{1}{\sqrt{2}}) \approx 1.17N.\end{aligned} \tag{71}$$

The equality in Eq. 71 occurs when $\langle A(T)\rangle = \langle B(T)\rangle = \frac{1}{\sqrt{2}}$. From Eq. 71, 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 from 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:

$$\begin{aligned}\langle p(T)\rangle &= \langle \alpha^{\intercal} a(t)\rangle\langle \gamma^{\intercal} b(t)\rangle \\ &= (\alpha^{\intercal} e^{U_a T} a_0)(\gamma^{\intercal} e^{U_b T} b_0).\end{aligned} \tag{72}$$

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:

$$\begin{aligned}\langle P(T)^2\rangle &= 1/T^2 \int_{u=0}^{T} du \int_{s=0}^{T} ds \langle p(u)p(s)\rangle \\ &= 1/T^2 \int_{u=0}^{T} du \int_{s=0}^{T} ds \langle \alpha^\intercal a(u)\alpha^\intercal a(s)\rangle \langle \gamma^\intercal b(u)\gamma^\intercal b(s)\rangle \\ &= 2/T^2 \int_{s=0}^{T} ds (T-s)\alpha^\intercal e^{U_a|u-s|}\left(\overline{x}_a.\alpha\right)\gamma^\intercal e^{U_b|u-s|}\left(\overline{x}_b.\gamma\right).\end{aligned} \tag{73}$$

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 $N = L = 6$.

Here again, 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 $\langle P(T)\rangle = pq$ and $CV_P(T) = (1-pq)/pq$. If $pq = 1/2$ we recover $CV_P(T) = 1$. Consistently, the limit of Eq. 73 when all non-zero eigenvalues of U_a and U_b go to $-\infty$ yields the same result.

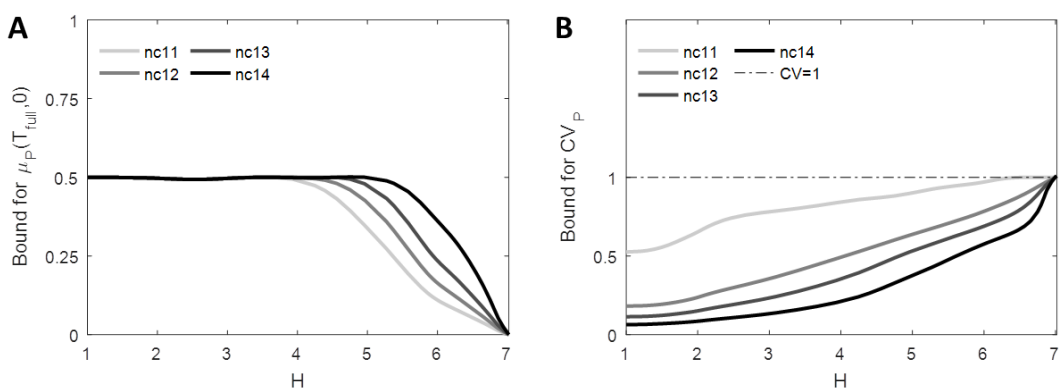

Figure 16: **The trade-off between pattern steepness (H), pattern formation time and readout error in the case of transcription regulation by two transcription factors**. Results for $N = L = 6$. $M = 1$. Each curve is computed from > 20000 data points.
(A) The lower bounds for the mean promoter activity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

$$[TF'] = e^{\frac{x^2}{\sigma^2}}. \tag{74}$$

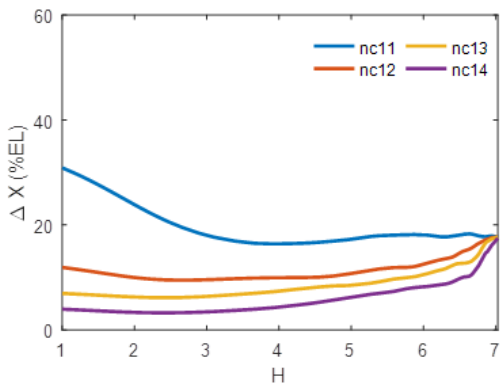

Figure 17: **Positional resolution in the case of transcription regulation by two mirror transcription factor gradients**. The results are shown for $N = L = 6$, $M = 1$. Positional resolution calculated from equilibrium binding site model for varying boundary steepness H.

TF can interact with the promoter via $L = 1$ binding site, corresponding to the number of known Cic binding sites found on *hb* promoter [15]. 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:

$$\frac{\partial\langle P(T)\rangle}{\partial X} = \frac{\partial\langle A(T)\rangle\langle B(T)\rangle}{\partial X} = \frac{\partial\langle A(T)\rangle}{\partial X} = H_A \leq N \tag{75}$$

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.

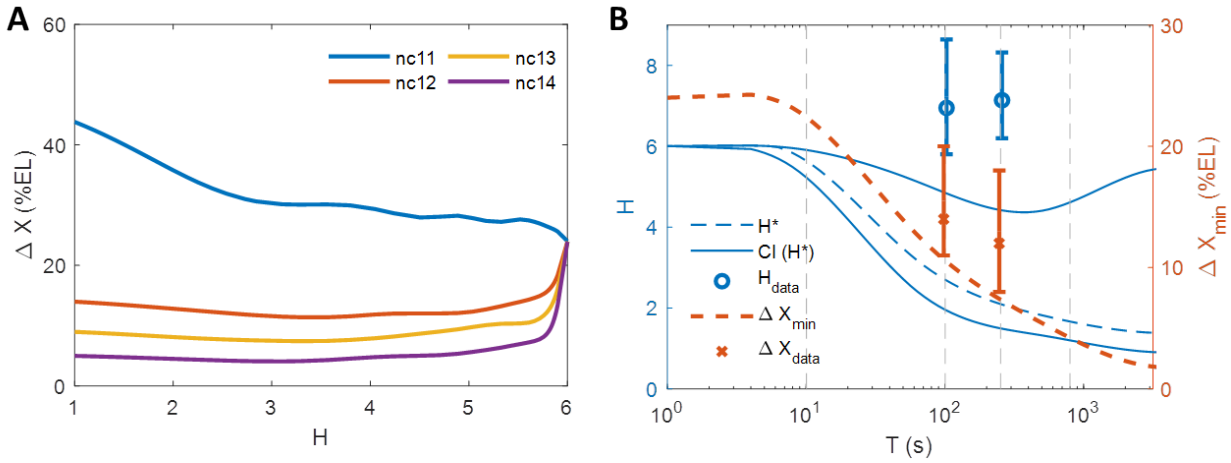

Figure 18: **Positional resolution in the case of transcription regulation by a anterior activator and a mid-embryo repressor**. The results are shown for $N = 6$, $L = 1$, $M = 1$. The positional resolution calculated from the equilibrium binding site model for varying boundary steepness H.

References

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

[2] 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.

[3] 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.

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

[5] 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.

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

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

[8] 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.

[9] 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.

[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] 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.

[13] 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.

[14] 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.

[15] 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.