PNAS
www.pnas.org

Supplementary Information for

Analytical distributions for detailed models of stochastic gene expression in eukaryotic cells

Zhixing Cao and Ramon Grima

Corresponding Author Ramon Grima.
E-mail: ramon.grima@ed.ac.uk

This PDF file includes:

Supporting Information Text

Contents

1. Master equation of Model M5

As stated in Point 1 in Model Specification and shown experimentally in Ref. (1), gene-copy independence means that each mother and daughter copy pair constitutes an independent stochastic subsystem. Therefore, it is enough to study one of the two subsystems.

Given the reactions associated with one of the subsystems of $M5$ (illustrated in Fig. 1D of main text), we can write the chemical master equation (CME) describing the stochastic dynamics

$$\begin{aligned}\frac{\mathrm{d}\mathbf{P}_n(n_\mathrm{N},n_\mathrm{M},t)}{\mathrm{d}t} &= \underbrace{\mathbf{T}(t)[\mathbf{P}_n(n_\mathrm{N}-1,n_\mathrm{M},t)-\mathbf{P}_n(n_\mathrm{N},n_\mathrm{M},t)]}_{\text{transcription of nascent mRNA}} + \underbrace{k[(n_\mathrm{N}+1)\mathbf{P}_n(n_\mathrm{N}+1,n_\mathrm{M}-1,t)-n_\mathrm{N}\mathbf{P}_n(n_\mathrm{N},n_\mathrm{M},t)]}_{\text{mRNA maturation}} \\ &+ \underbrace{d[(n_\mathrm{M}+1)\mathbf{P}_n(n_\mathrm{N},n_\mathrm{M}+1,t)-n_\mathrm{M}\mathbf{P}_n(n_\mathrm{N},n_\mathrm{M},t)]}_{\text{degradation of mature mRNA}} + \underbrace{\mathbf{\Phi}(t)\mathbf{P}_n(n_\mathrm{N},n_\mathrm{M},t)}_{\text{gene state switching}},\end{aligned}$$

where

$$\mathbf{P}_n(n_\mathrm{N},n_\mathrm{M},t) = [P_n^{00}(n_\mathrm{N},n_\mathrm{M},t) \quad P_n^{10}(n_\mathrm{N},n_\mathrm{M},t) \quad P_n^{01}(n_\mathrm{N},n_\mathrm{M},t) \quad P_n^{11}(n_\mathrm{N},n_\mathrm{M},t)]^\top,$$

where n_N and n_M denote the numbers of nascent mRNA and mature mRNA respectively. The first and second binary digit in superscripts label the gene state of the mother copy and its daughter copy produced during replication: 1 for OFF state and 0 for ON state. Hence, $P_n^{00}(n_\mathrm{M},n_\mathrm{N},t)$ is the probability of having two actively transcribing genes. The subscript n stands for the cell cycle number or equivalently the cellular generation number. Note that t is the cell age which varies between 0 and t_d, the time at which cell division occurs. Transcription details are encoded in the diagonal matrix

$$\mathbf{T}(t) = \mathrm{diag}[2\rho_\mathrm{u}(t), \quad \rho_\mathrm{u}(t), \quad \rho_\mathrm{u}(t), \quad 0]$$

where $\rho_\mathrm{u}(t) = \rho_0 \mathrm{e}^{\alpha t}$ (Point 4 in Model Specification), while the details of gene switching are encoded in the matrix

$$\mathbf{\Phi}(t) = \begin{cases} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\sigma_\mathrm{b} & \sigma_\mathrm{u}^- \\ 0 & 0 & \sigma_\mathrm{b} & -\sigma_\mathrm{u}^- \end{bmatrix} & t \in [0, t_\mathrm{r}), \\ \begin{bmatrix} -2\sigma_\mathrm{b} & \sigma_\mathrm{u}^+ & \sigma_\mathrm{u}^+ & 0 \\ \sigma_\mathrm{b} & -\sigma_\mathrm{b}-\sigma_\mathrm{u}^+ & 0 & \sigma_\mathrm{u}^+ \\ \sigma_\mathrm{b} & 0 & -\sigma_\mathrm{b}-\sigma_\mathrm{u}^+ & \sigma_\mathrm{u}^+ \\ 0 & \sigma_\mathrm{b} & \sigma_\mathrm{b} & -2\sigma_\mathrm{u}^+ \end{bmatrix} & t \in [t_\mathrm{r}, t_\mathrm{d}), \end{cases}$$

where the change from σ_u^- to σ_u^+ at replication time t_r is a reflection of Point 6 in Model Specification, whereas the mRNA maturation, degradation of mature mRNA and gene state switching capture the processes described in Points 2, Point 3 and Point 1 in Model Specification, respectively.

Next we describe the non-reactive processes in the model, i.e. replication and cell division. At the replication time t_r, the pre-replication and post-replication probabilities are linked by the equation

$$\mathbf{P}_n(n_\mathrm{N},n_\mathrm{M},t_\mathrm{r}^+) = \Pi\mathbf{P}_n(n_\mathrm{N},n_\mathrm{M},t_\mathrm{r}^-)$$

where Π represents the transition probability undergoing the gene replication process. The variables t_r^+ and t_r^- stand for the time before and after gene replication respectively.

The following two choices of Π represent the two possible conditions for gene states during replication as mentioned in Model Specification Point 5: (i) daughter copy inherits gene state from the mother copy, and (ii) all copies are reset to OFF state upon replication,

$$\Pi = \begin{cases} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} & \text{case (i)}, \\ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} & \text{case (ii)}. \end{cases}$$

At cell age (t_d) a cell divides into two daughter cells (Point 7 in Model Specification), and the generation number increases by one. The pre-division and post-division probabilities are linked by the equations

$$\begin{cases} P_{n+1}^{11}(n_\mathrm{N},n_\mathrm{M},0) = \sum\limits_{n'_\mathrm{N},n'_\mathrm{M}} B(n_\mathrm{N}|n'_\mathrm{N})B(n_\mathrm{M}|n'_\mathrm{M})P_n^{11}(n'_\mathrm{N},n'_\mathrm{M},t_\mathrm{d}) + \sum\limits_{n'_\mathrm{N},n'_\mathrm{M}} B(n_\mathrm{N}|n'_\mathrm{N})B(n_\mathrm{M}|n'_\mathrm{M})P_n^{10}(n'_\mathrm{N},n'_\mathrm{M},t_\mathrm{d}), \\ P_{n+1}^{01}(n_\mathrm{N},n_\mathrm{M},0) = \sum\limits_{n'_\mathrm{N},n'_\mathrm{M}} B(n_\mathrm{N}|n'_\mathrm{N})B(n_\mathrm{M}|n'_\mathrm{M})P_n^{01}(n'_\mathrm{N},n'_\mathrm{M},t_\mathrm{d}) + \sum\limits_{n'_\mathrm{N},n'_\mathrm{M}} B(n_\mathrm{N}|n'_\mathrm{N})B(n_\mathrm{M}|n'_\mathrm{M})P_n^{00}(n'_\mathrm{N},n'_\mathrm{M},t_\mathrm{d}), \\ P_{n+1}^{10}(n_\mathrm{N},n_\mathrm{M},0) = P_{n+1}^{00}(n_\mathrm{N},n_\mathrm{M},0) = 0. \end{cases} \quad [1]$$

These equations model the random allocation of nascent and mature mRNA into two daughter cells according to a binomial distribution. Since only the mother copy is tracked, the probabilities with daughter copy being at an ON gene state (the second digit in subscript being 0, i.e., P^{10}_{n+1} and P^{00}_{n+1}) are set to zero. The binomial distribution kernel is

$$B(n_{\rm N}|n'_{\rm N}) = \left(\begin{array}{c} n'_{\rm N} \\ n_{\rm N} \end{array} \right) 2^{-n'_{\rm N}}.$$

At any time, the joint probability of $n_{\rm N}$ nascent mRNA and $n_{\rm M}$ mature mRNA $\bar{P}(n_{\rm N}, n_{\rm M}, t)$ of the whole of the two subsystems can be computed via a convolution

$$\bar{P}_n(n_{\rm N}, n_{\rm M}, t) = \sum_{i,j} P_n(n_{\rm N}-i, n_{\rm M}-j, t) P_n(i, j, t) \quad \text{and} \quad P_n(n_{\rm N}, n_{\rm M}, t) = P^{11}_n(n_{\rm N}, n_{\rm M}, t) + P^{01}_n(n_{\rm N}, n_{\rm M}, t) + P^{10}_n(n_{\rm N}, n_{\rm M}, t) + P^{00}_n(n_{\rm N}, n_{\rm M}, t).$$

This result follows from the independence of the two subsystems and the fact that they are identical.

All stochastic simulations using the finite state projection algorithm and the stochastic simulation algorithm reported in the main text simulate the processes described in this section. The master equation is too complex to solve exactly and hence approximations are needed which we describe from Sections 3 onwards.

2. Estimation of experimental parameter values

2.1. Table 1 in Main Text. The data processing for each line in this table is summarized as follows:

- **Line 1:** The values of $\rho_{\rm u}$, $\sigma_{\rm b}$, $\sigma_{\rm u}$ are extracted from the bottom tabular of Supplemental Table 3 of Ref. (2). As noted in the latter paper, d is a fixed value taken from (3); this gives a half life of the mRNA of POL1 equal to 10 min which is equivalent to $d = \ln 2/\tau_{1/2} = 0.0693\ \text{min}^{-1}$.
- **Line 2:** Same as Line 1 except that the half life of mRNA corresponding to PDR5 is 14 min (3) which means $d = 0.0495\ \text{min}^{-1}$.
- **Line 3:** The degradation rate d is extracted from Table F in Supplementary file 1 of Ref. (1), being $0.14\ \text{hr}^{-1} = 0.0023\ \text{min}^{-1}$. The other three parameters are extracted from Table A in Supplementary file 2 of Ref. (1).
- **Line 4:** Same as Line 3. The degradation rate d is $0.13\ \text{hr}^{-1} = 0.0022\ \text{min}^{-1}$.
- **Line 5:** The values of $\sigma_{\rm b}$ and $\sigma_{\rm u}$ are extracted from Fig. 3g of (4), while the values of d and $\rho_{\rm u}$ are obtained from Lines 7 and 11 in the left-column text (p. 4) of (4), respectively.
- **Line 6:** The degradation rate d is in Line 8 under Eq. (2) in Section Mathematical modeling of c-Fos serum induction data in SI of (5), which is from literature Ref. (6). The half life of c-Fos mRNA is 15 min, indicating $d = \ln 2/15 = 0.462\ \text{min}^{-1}$. The other kinetic parameters are from Line 2 in Table S1 of the SI of Ref. (5).
- **Lines 7-10:** All the four kinetic parameters are quoted from Table S1 of Ref. (7) (Genes Acly, Actb, Srebf1, Insr1 fed PP (periportal)), and converted from units hr^{-1} to min^{-1}.
- **Line 11:** The values of $\rho_{\rm u}$, $\sigma_{\rm b}$ and $\sigma_{\rm u}$ of 6935 genes are imported from Table S1 of Ref. (8), which have all been normalized by decay rate d. The half life time of mRNA (3575 out of 6935 genes) are found in the file slam_seq.csv in (https://github.com/sandberg-lab/txburst/tree/master/data), from which the decay rates d are calculated. With the decay rates d, the absolute values of $\rho_{\rm u}$, $\sigma_{\rm b}$ and $\sigma_{\rm u}$ can be computed as well. The mean of the absolute values of $\rho_{\rm u}$, $\sigma_{\rm b}$, $\sigma_{\rm u}$ and d over 3575 genes are summarized in Line 11 of Table 1. The burst size, fraction ON time and timescale ratio δ are computed for each gene, and their mean values over 3575 genes are listed in Table 1.
- **Line 12:** The decay rate normalized data are from Table S6 of Ref. (8). The decay rate is absent so that we compute the mean of $\rho_{\rm u}$, $\sigma_{\rm b}$ and $\sigma_{\rm u}$ with respect to degradation rate d.
- **Line 13:** The mean values reported are calculated from Table S1, which were obtained from the authors of Ref. (9).

2.2. Estimation of $t_{\rm d}$ (division time) and $t_{\rm r}$ (replication time). Experimental values are collected in Table S2 and discussed below.

- **Line 1:** The cell cycle time is from the line under Fig. 2 in Ref. (10), while the replication time is calculated from the first line in Table 2 in the same paper.
- **Line 2:** The value is indicated in Fig. 5 of Ref. (2).
- **Line 3:** Both values are from (https://bionumbers.hms.harvard.edu/files/Cell%20cycle%20times.pdf).
- **Line 4:** The cell cycle and replication time are from Section 9 ($\tau_{\rm DIV}$, p. 20) and Fig. 3B of Ref. (1).

Table S1. Data from Suter et. al. (9) which was used to compute values in Line 13 of Table 1 Main text

Cell type (Gene)	σ_u (min^{-1})	σ_b (min^{-1})	ρ_u (min^{-1})	d (min^{-1})	Burst size (ρ_u/σ_b)	Fraction ON time ($\sigma_u/(\sigma_u+\sigma_b)$)	Timescale δ
Mouse fibroblasts (DBP_FRT)	0.0061	0.067	1.60	0.0065	23.9	0.084	10.93
Mouse fibroblasts (Per2_het)	0.0029	0.261	1.14	0.0093	4.4	0.011	89.83
Mouse fibroblasts (Per2_hom)	0.0028	0.149	1.28	0.0093	8.6	0.018	53.11
Mouse fibroblasts (GT:Glutaminase)	0.0251	0.256	2.49	0.0153	9.7	0.089	10.21
Mouse fibroblasts (GT:Serpine1)	0.0346	0.296	1.00	0.0168	3.4	0.105	8.55
Mouse fibroblasts (GT:Prl2C2)	0.0265	0.129	3.82	0.0123	29.7	0.171	4.86
Mouse fibroblasts (GT:Sh3kbp1)	0.0344	0.231	1.86	0.0183	8.1	0.130	6.71
Mouse fibroblasts (GT:Ctgf)	0.0077	0.128	4.69	0.0055	36.6	0.057	16.66
Mouse fibroblasts (H1)	0.0085	0.174	3.61	0.0102	20.7	0.046	20.51
Mouse fibroblasts (H2)	0.0109	0.148	4.65	0.0102	31.5	0.069	13.57
Mouse fibroblasts (1M1C)	0.0073	0.099	1.13	0.0102	11.5	0.069	13.49
Mouse fibroblasts (1M2C)	0.0124	0.126	3.82	0.0102	30.3	0.090	10.15
Mouse fibroblasts (2M1C)	0.0087	0.164	1.60	0.0102	9.8	0.050	18.86
Mouse fibroblasts (2M2C)	0.0119	0.144	4.30	0.0102	29.9	0.076	12.08
Mouse fibroblasts (3M1C)	0.0094	0.137	0.86	0.0102	6.3	0.064	14.54
Mouse fibroblasts (3M2C)	0.0089	0.157	1.74	0.0102	11.1	0.054	17.66

Table S2. Cell cycle and gene replication time in literature

Cell type	Cell cycle (min)	Replication time (G1+S) (min)	Reference
Yeast	150	55	(10)
Yeast	85	/	(2)
Mouse fibroblasts (NIH3T3)	1140	900	(11)
Mouse embryonic stem cell	780	600 (Oct4) / 400 (Nanog)	(1)

2.3. Estimation of σ_u^+. The values of σ_u^+ for Oct4 and Nanog are calculated from α in Table A of Supplementary file 2 of (1), being $0.63 \times 9.2 \times 10^{-3} = 5.8 \times 10^{-3}$ and $0.71 \times 1.9 \times 10^{-3} = 1.3 \times 10^{-3}$.

2.4. Estimation of k. Experimental values are collected in Table S3 and discussed below.

- **Line 1:** The value of k is converted from t_{prod} in Line 2 of Table S1 in Ref (5) ($k = 1/t_{\text{prod}}$).
- **Lines 2-3:** The values of k come from Table A of Supplementary file 2 of Ref (1).

Table S3. mRNA maturation and decay times in literature

Cell type (Gene)	k (min^{-1})	d (min^{-1})	Reference
Human osteosarcoma (c-Fos)	1.25	0.0462	(5)
Mouse embryonic stem cells (Nanog)	0.13	0.0022	(1)
Mouse embryonic stem cells (Oct4)	0.29	0.0023	(1)

Another source for the ratio k/d is Extended Data Fig. 2f of Ref. (12), wherein the mode of the ratio is 8.

3. Calculating an approximate marginal distribution of nascent mRNA

3.1. Analytical marginal distribution of nascent mRNA in a telegraph model (no cell division, replication or dosage compensation). Table 1 (main text) shows that $\sigma_b \gg \sigma_u$ for many eukaryotic genes, i.e. most genes spend most of their time in the OFF state. In this section we will use this observation to calculate the approximate distribution of nascent mRNA; for now our calculation will ignore cell division and replication (hence we simply denote the gene activation rate as σ_u and we do not track the cellular generation time). Note that since the production and degradation rates of nascent mRNA are not dependent on mature mRNA there is no description of the latter in the master equations that follow.

The master equation for the distribution of nascent mRNA numbers (of a single copy gene system) is given by

$$\begin{cases} \dfrac{\mathrm{d}P_0(n_N)}{\mathrm{d}t} = \rho_u(t)[P_0(n_N-1) - P_0(n_N)] + k[(n_N+1)P_0(n_N+1) - n_N P_0(n_N)] - \sigma_b P_0(n_N) + \sigma_u P_1(n_N), \\ \dfrac{\mathrm{d}P_1(n_N)}{\mathrm{d}t} = k[(n_N+1)P_1(n_N+1) - n_N P_1(n_N)] + \sigma_b P_0(n_N) - \sigma_u P_1(n_N), \end{cases} \quad [2]$$

where σ_u is a piecewise constant function in time, and $P_0(n_N)$ and $P_1(n_N)$ are the probability of n_N nascent mRNAs when the gene is ON and OFF respectively. The argument t is suppressed for simplicity. Note that the probability of n_N nascent mRNAs is given by the sum $P(n_N) = P_0(n_N) + P_1(n_N)$. Defining the generating functions for the two probabilities as

$$G_0(z_1) = \sum_{n_N=0}^{\infty} z_1^{n_N} P_0(n_N), \quad G_1(z_1) = \sum_{n_N=0}^{\infty} z_1^{n_N} P_1(n_N),$$

we can rewrite Eq. [2] as a set of two coupled partial differential equations

$$\begin{cases} \partial_t G_0(z_1) = \rho_u(t)(z_1-1)G_0(z_1) + k(1-z_1)\partial_{z_1} G_0(z_1) - \sigma_b G_0(z_1) + \sigma_u G_1(z_1), \\ \partial_t G_1(z_1) = k(1-z_1)\partial_{z_1} G_1(z_1) + \sigma_b G_0(z_1) - \sigma_u G_1(z_1). \end{cases} \quad [3]$$

Further letting $u = z_1 - 1$, one can simplify Eq. [3] as

$$\begin{cases} \partial_t G_0 = \rho_u(t) u G_0 - k u \partial_u G_0 - \sigma_b G_0 + \sigma_u G_1, \\ \partial_t G_1 = -k u \partial_u G_1 + \sigma_b G_0 - \sigma_u G_1. \end{cases} \quad [4]$$

Summing the two equations in Eq. [4] and letting $G = G_0 + G_1$, we have

$$\begin{cases} \partial_t G_0 + k u \partial_u G_0 = \rho_u(t) u G_0 - \sigma_b G_0 + \sigma_u G_1, \\ \partial_t G + k u \partial_u G = \rho_u(t) u G_0. \end{cases}$$

By means of the method of characteristics, we have the following set of ordinary differential equations (ODEs)

$$\begin{aligned} &\frac{\mathrm{d}t}{\mathrm{d}s} = 1 && \Rightarrow && t = s, \\ &\frac{\mathrm{d}u}{\mathrm{d}s} = ku && \Rightarrow && u = u_0 \mathrm{e}^{ks}, \\ &\frac{\mathrm{d}G_0}{\mathrm{d}s} = \rho_u(t) u G_0 - \sigma_b G_0 + \sigma_u G_1 && && [5] \\ &\frac{\mathrm{d}G}{\mathrm{d}s} = \rho_u(t) u G_0, && && [6] \end{aligned}$$

with $u_0 = u|_{s=0}$. If we define $\delta = \sigma_b/\sigma_u$ and divide both sides of Eqs. [5] and [6] by σ_u then we obtain

$$\begin{cases} \dfrac{\mathrm{d}G_0}{\mathrm{d}\tilde{s}} = \tilde{\rho}_u(t) u G_0 - \delta G_0 + G_1 \\ \dfrac{\mathrm{d}G}{\mathrm{d}\tilde{s}} = \tilde{\rho}_u(t) u G_0, \end{cases} \quad [7]$$

where $\tilde{\rho}_u(t) = \rho_u(t)/\sigma_u$ and $\tilde{s} = \sigma_u s$. Then, we further divide both equations in Eq. [7] by δ and denote $\epsilon = 1/\delta$ to obtain

$$\begin{cases} \epsilon \dfrac{\mathrm{d}G_0}{\mathrm{d}\tilde{s}} = b_t u G_0 - G_0 + \epsilon G_1 \\ \epsilon \dfrac{\mathrm{d}G}{\mathrm{d}\tilde{s}} = b_t u G_0, \end{cases} \quad [8]$$

with b_t being $\rho_u(t)/\sigma_b$ (the mean nascent mRNA burst size). Since from Table 1 in the main text, we know that $\sigma_b \gg \sigma_u$ it then follows that ϵ is a small positive real number, and hence as per perturbation theory, we postulate that G_0 and G have a series expansion in ϵ of the form

$$G_0 = G_0^{(0)} + \epsilon G_0^{(1)} + \mathcal{O}(\epsilon^2), \quad G = G^{(0)} + \epsilon G^{(1)} + \mathcal{O}(\epsilon^2).$$

Matching the coefficients of different orders in ϵ in Eq. [8], the following set of equations ensue

$$\text{Order } \epsilon^0: \qquad b_t u G_0^{(0)} - G_0^{(0)} = 0 \qquad \Rightarrow \qquad G_0^{(0)} = 0,$$

$$\text{Order } \epsilon^1: \qquad \begin{cases} \dfrac{\mathrm{d}G_0^{(0)}}{\mathrm{d}\tilde{s}} = b_t u G_0^{(1)} - G_0^{(1)} + G^{(0)} - G_0^{(0)} \\ \qquad \Rightarrow b_t u G_0^{(1)} - G_0^{(1)} + G^{(0)} = 0, \\ \dfrac{\mathrm{d}G^{(0)}}{\mathrm{d}\tilde{s}} = b_t u G_0^{(1)}, \end{cases}$$

all of which reduce to

$$\frac{\mathrm{d}G^{(0)}}{\mathrm{d}\tilde{s}} = -\frac{b_t u}{b_t u - 1} G^{(0)} = -\frac{\rho_\mathrm{u}(t) u}{\rho_\mathrm{u}(t) u - \sigma_\mathrm{b}} G^{(0)}.$$

Combining with $u = u_0 \mathrm{e}^{ks}$ and $\rho_\mathrm{u}(t) = \rho_0 \mathrm{e}^{\alpha t}$, it is equivalent to

$$\frac{\mathrm{d}G^{(0)}}{\mathrm{d}s} = -\frac{\rho_0 \mathrm{e}^{\alpha s} \sigma_\mathrm{u} u_0 \mathrm{e}^{ks}}{\rho_0 \mathrm{e}^{\alpha s} \sigma_\mathrm{u} u_0 \mathrm{e}^{ks} - \sigma_\mathrm{b}} G^{(0)}.$$

Note that here we have used the relation $\rho_\mathrm{u}(t) = \rho_0 \mathrm{e}^{\alpha t}$ which models growth-dependent transcription as explained in Section 1. The solution immediately follows and is given by

$$G^{(0)} = C(u_0)(\rho_0 u_0 \mathrm{e}^{(\alpha+k)s} - \sigma_\mathrm{b})^{-\frac{\sigma_\mathrm{u}}{\alpha+k}}, \qquad [9]$$

with $C(u_0)$ being a function of u_0 to be determined from initial condition. Suppose that the initial condition for this process is $g(u) = G^{(0)}|_{t=0}$, which is known a priori. For instance, the initial marginal distribution of i nascent mRNA molecules is $P(i) = p_i$, and then $g(u) = \sum_i p_i (u+1)^i$. Letting s be equal to 0 (or equivalently $t = 0$), it follows $u = u_0$ and $g(u) = g(u_0)$, and we can establish the following relation

$$g(u_0) = C(u_0)(\rho_0 u_0 - \sigma_\mathrm{b})^{-\frac{\sigma_\mathrm{u}}{\alpha+k}},$$

from which we can solve for $C(u_0)$ to obtain

$$C(u_0) = g(u_0)(\rho_0 u_0 - \sigma_\mathrm{b})^{\frac{\sigma_\mathrm{u}}{\alpha+k}}.$$

Substituting the latter in Eq. [9] and replacing $u_0 = u\mathrm{e}^{-kt}$, we can calculate the leading-order solution of G from Eq. [9] as

$$\boxed{G(u,t) = g(u\mathrm{e}^{-kt}) \left(\frac{\rho_0 \mathrm{e}^{-kt} u - \sigma_\mathrm{b}}{\rho_0 \mathrm{e}^{\alpha t} u - \sigma_\mathrm{b}} \right)^{\frac{\sigma_\mathrm{u}}{\alpha+k}}}. \qquad [10]$$

In the special case $\alpha = 0$ (no growth-dependent transcription), in the steady-state limit of long times, Eq. [10] reduces to

$$G(z_1) = \left(\frac{\sigma_\mathrm{b}}{\sigma_\mathrm{b} - \rho_0 (z_1 - 1)} \right)^{\frac{\sigma_\mathrm{u}}{k}}$$

which implies the probability distribution of nascent mRNA is negative binomial and of the form NB $\left(\frac{\sigma_\mathrm{u}}{k}, \frac{\rho_0}{\rho_0 + \sigma_\mathrm{b}} \right)$, a result previously reported in (13).

3.2. Reduced model of nascent mRNA dynamics (no cell division, replication or dosage compensation). Next we show that the results of the previous section can be obtained from a reduced model which does not have an explicit gene state description. This will be useful in later sections when we derive results for mature mRNA and protein.

We start by assuming that nascent mRNA is actively transcribed in bursts whose size are distributed according to a negative binomial distribution with mean burst size $\beta_t = \beta_0 \mathrm{e}^{\alpha t}$. This implies the time-dependent propensity function describing nascent mRNA production is $f(t,n) = \frac{f_0 (\beta_0 \mathrm{e}^{\alpha t})^n}{(1+\beta_0 \mathrm{e}^{\alpha t})^{n+1}}$ where f_0 and β_0 are constants to be determined. In other words, we consider the reaction system

$$\varnothing \xrightarrow{f(t,n)} nN, \qquad N \xrightarrow{k} \varnothing. \qquad [11]$$

The master equation describing the stochastic dynamics of this system is given by

$$\partial_t P(n_\mathrm{N}, t) = \sum_{i=0}^{\infty} \frac{f_0 (\beta_0 \mathrm{e}^{\alpha t})^i}{(1+\beta_0 \mathrm{e}^{\alpha t})^{i+1}} [P(n_\mathrm{N} - i, t) - P(n_\mathrm{N}, t)] + k[(n_\mathrm{N}+1) P(n_\mathrm{N}+1, t) - n_\mathrm{N} P(n_\mathrm{N}, t)],$$

and the corresponding generating function equation is

$$\partial_t G(u,t) + ku\partial_u G(u,t) = \frac{f_0\beta_0 e^{\alpha t}u}{1-\beta_0 e^{\alpha t}u} G(u,t).$$

This can be solved by the method of characteristics which leads to the following set of ODEs

$$\begin{aligned}
\frac{dt}{ds} &= 1 \qquad &\Rightarrow \qquad & t = s, \\
\frac{du}{ds} &= ku \qquad &\Rightarrow \qquad & u = u_0 e^{ks}, \\
\frac{dG}{ds} &= \frac{f_0\beta_0 e^{\alpha t}u}{1-\beta_0 e^{\alpha t}u} G.
\end{aligned}$$

which simplifies to

$$\frac{dG}{ds} = \frac{f_0\beta_0 u_0 e^{(\alpha+k)s}}{1-\beta_0 u_0 e^{(\alpha+k)s}} G, \tag{12}$$

The solution of this ODE is given by

$$G = C(u_0)(\beta_0 u_0 e^{(\alpha+k)s} - 1)^{-\frac{f_0}{\alpha+k}}.$$

Hence following the same reasoning as in the previous section, given an initial condition on generating functions $g(u)$, the exact solution to [11] is

$$G = g(ue^{-kt}) \left(\frac{1-\beta_0 u e^{-kt}}{1-\beta_0 u e^{\alpha t}} \right)^{\frac{f_0}{\alpha+k}}. \tag{13}$$

Interestingly, comparing Eq. [10] and Eq. [13], one can conclude that the reduced model of bursty nascent mRNA [11] has the same probability distribution of nascent mRNA as the telegraph model in the limit of large σ_b/σ_u, provided we choose the free constants to be $\beta_0 = b_0 = \rho_0/\sigma_b$ and $f_0 = \sigma_u$.

3.3. Approximate marginal distribution of nascent mRNA in model M5 (including cell division, replication and dosage compensation). We now extend the reduced model of the previous section to take into account the rest of the processes present in model M5 namely cell division, replication and dosage compensation. The advantage of using the reduced model rather than a telegraph model is the fact that we can avoid discussing the change of gene state during gene replication, i.e. the choice of Π matrix.

The number of nascent mRNAs at a particular cell age t of a given cell cycle n is contributed by two processes: (i) the decay (maturation) of nascent mRNAs inherited from the previous cycle, and (ii) the production of new nascent mRNAs in cell cycle n. The two processes are independent across cell generations. Therefore, it further allows us to establish a compact relation on generating functions of nascent mRNA

$$G_n(u,t) = \underbrace{G_n(ue^{-kt},0)}_{\text{death process}} \underbrace{G_n^{\text{new}}(u,t)^2}_{\text{new born mRNA}} \qquad \forall t \in [0,t_d), \tag{14}$$

where $G_n(u,t)$ is the generating function corresponding to the marginal distribution of nascent mRNA at cell age t and cell cycle n, $G_n(u,0)$ is the initial condition at the beginning of cell cycle n, and $G_n^{\text{new}}(u,t)$ is the generating function for new born nascent mRNA, which will be detailed later. Noted that the power of 2 in the right hand side of Eq. [14] arises from the diploidy, the number of independent subsystems.

Since in the limit of $\sigma_b \gg \sigma_u$, the mapping of the telegraph model onto the effective model Eq. [11] dispenses with a discussion on gene states when gene replicates, all gene copies are independent and so are the species associated with them. Hence the generating function $G_n^{\text{new}}(u,t)$ can be further decomposed into two components $G_A(u,t)$ and $G_B(u,t)$, which are contributed by mother and daughter copy respectively. Consequently Eq. [14] can be generally written as

$$\boxed{G_n(u,t) = G_n(ue^{-kt},0)G_A^2(u,t)G_B^2(u,t) \quad \forall t \in [0,t_d).} \tag{15}$$

Next we introduce cell division which results in a binomial partitioning of the number of nascent mRNA at time t_d. Since we only considering nascent mRNA and since in our reduced model there are no explicit gene states then it follows that Eq. [1] reduces to the simpler form

$$P_{n+1}(n_N,0) = \sum_{n'_N} B(n_N,n'_N)P_n(n'_N,t_d) = \sum_{n'_N} \binom{n'_N}{n_N} 2^{-n'_N} P_n(n'_N,t_d),$$

From the definition of generating function $G_{n+1}(z_1,0)=\sum_{n_\mathrm{N}} z_1^{n_\mathrm{N}} P_{n+1}(n_\mathrm{N},0)$, it follows that

$$\begin{aligned}
G_{n+1}(z_1,0) =& \sum_{n_\mathrm{N}}\sum_{n'_\mathrm{N}} \begin{pmatrix} n'_\mathrm{N} \\ n_\mathrm{N} \end{pmatrix} 2^{-n'_\mathrm{N}} P_n(n'_\mathrm{N},t_\mathrm{d}) z_1^{n_\mathrm{N}} \\
=& \sum_{n'_\mathrm{N}} P_n(n'_\mathrm{N},t_\mathrm{d}) \left(\frac{1}{2}\right)^{n'_\mathrm{N}} \sum_{n_\mathrm{N}} \begin{pmatrix} n'_\mathrm{N} \\ n_\mathrm{N} \end{pmatrix} (z_1)^{n_\mathrm{N}} \times 1^{n'_\mathrm{N}-n_\mathrm{N}} \\
=& \sum_{n'_\mathrm{N}} P_n(n'_\mathrm{N},t_\mathrm{d}) \left(\frac{1}{2}\right)^{n'_\mathrm{N}} (z_1+1)^{n'_\mathrm{N}} \\
=& G_n\left(\frac{z_1+1}{2},t_\mathrm{d}\right).
\end{aligned}$$

The second equality comes from swapping the summation operator, the third equality is by the binomial theorem, and the fourth is yielded by the definition of generating function. Note that this implies that binomial partitioning is akin to the operation

$$z_1 \mapsto \frac{z_1+1}{2}.$$

Since $u=z_1-1$, it follows that the equivalent mapping on u is

$$u=z_1-1 \mapsto \frac{z_1+1}{2}-1=\frac{u}{2}.$$

Therefore, we have a simpler expression for cell division in terms of variable u

$$G_{n+1}(u,0)=G_n\left(\frac{u}{2},t_\mathrm{d}\right). \tag{16}$$

Combining Eq. [15] and Eq. [16] together we obtain a recursive relation between the generating functions across two successive cell cycles

$$\boxed{G_{n+1}(u,0)=G_n\left(\frac{u}{2},t_\mathrm{d}\right)=G_n(\eta u,0)G_\mathrm{A}^2\left(\frac{u}{2},t_\mathrm{d}\right)G_\mathrm{B}^2\left(\frac{u}{2},t_\mathrm{d}\right)} \tag{17}$$

where $\eta=\mathrm{e}^{-kt_\mathrm{d}}/2$.

To complete the time-dependent solution we need to derive the explicit forms of G_A and G_B for nascent mRNA. We remind the reader that these are the generating functions describing the dynamics of nascent mRNA which is produced in the present cell cycle (taking into account transcription, maturation and dosage compensation) due to the mother and daughter copy. By using Eq. [13], it follows that the generating functions for mother and daughter copy $G_\mathrm{A}(u,t)$ and $G_\mathrm{B}(u,t)$ of nascent mRNA are piece-wisely defined as

$$\boxed{G_\mathrm{A}(u,t)=\begin{cases} \left(\dfrac{\rho_0 u\mathrm{e}^{-kt}-\sigma_\mathrm{b}}{\rho_0 u\mathrm{e}^{\alpha t}-\sigma_\mathrm{b}}\right)^{\frac{\sigma_\mathrm{u}^-}{\alpha+k}} & t\in[0,t_\mathrm{r}), \\ \left(\dfrac{\rho_0 u\mathrm{e}^{-kt}-\sigma_\mathrm{b}}{\rho_0 u\mathrm{e}^{-k(t-t_\mathrm{r})+\alpha t_\mathrm{r}}-\sigma_\mathrm{b}}\right)^{\frac{\sigma_\mathrm{u}^-}{\alpha+k}} \left(\dfrac{\rho_1 u\mathrm{e}^{-k(t-t_\mathrm{r})}-\sigma_\mathrm{b}}{\rho_1 u\mathrm{e}^{\alpha(t-t_\mathrm{r})}-\sigma_\mathrm{b}}\right)^{\frac{\sigma_\mathrm{u}^+}{\alpha+k}} & t\in[t_\mathrm{r},t_\mathrm{d}), \end{cases}} \tag{18a}$$

$$\boxed{G_\mathrm{B}(u,t)=\begin{cases} 1 & t\in[0,t_\mathrm{r}), \\ \left(\dfrac{\rho_1 u\mathrm{e}^{-k(t-t_\mathrm{r})}-\sigma_\mathrm{b}}{\rho_1 u\mathrm{e}^{\alpha(t-t_\mathrm{r})}-\sigma_\mathrm{b}}\right)^{\frac{\sigma_\mathrm{u}^+}{\alpha+k}} & t\in[t_\mathrm{r},t_\mathrm{d}). \end{cases}} \tag{18b}$$

The first part $t\in[0,t_\mathrm{r})$ of $G_\mathrm{A}(u,t)$ describes the stochastic dynamics of nascent mRNA born in the pre-replication time. Note that the initial condition is zero and $\sigma_\mathrm{u}=\sigma_\mathrm{u}^-$. The second part $t\in[t_\mathrm{r},t_\mathrm{d})$ of $G_\mathrm{A}(u,t)$ describes that the stochastic dynamics of nascent mRNA born in the post-replication time. This is given by Eq. [13] with g replaced by the initial condition which is the generating function at replication time (from the expression for $t\in[0,t_\mathrm{r})$); also note that $\sigma_\mathrm{u}=\sigma_\mathrm{u}^+$ (due to dosage compensation), ρ_0 is replaced by $\rho_1=\rho_0\mathrm{e}^{\alpha t_r}$ since this is the transcription rate at replication time, and time t is replaced by $t-t_\mathrm{r}$. Since there is no transcription activity in the pre-replication time for the daughter copy, the generating function $G_\mathrm{B}(u,t)$ is trivially equal to 1 for $t\in[0,t_\mathrm{r})$. The second part $t\in[t_\mathrm{r},t_\mathrm{d})$ of $G_\mathrm{B}(u,t)$ can be found similarly as for $G_\mathrm{A}(u,t)$.

Hence summarizing Eqs. [15], [17], [18a] and [18b] together explicitly define the approximate marginal distribution for nascent mRNA for all cell ages and all generations in our model M5.

Indeed, the probability distributions for Eqs. [18a] and [18b] can be found exactly. For simplicity, we only illustrate the technique for $G_{\rm B}(u,t)$. Denoting $\psi_0=\rho_1 e^{-k(t-t_{\rm r})}/\sigma_{\rm b}$, $\psi_1=\rho_1 e^{\alpha(t-t_{\rm r})}/\sigma_{\rm b}$, $\psi_2=\sigma_{\rm u}^+/(\alpha+k)$ and $q=\psi_1/(\psi_1+1)$, the generating function $G_{\rm B}$ can be decomposed as

$$G_{\rm B}(z_1,t)=\underbrace{(\psi_0+1-\psi_0 z_1)^{\psi_2}}_{\text{binomial}}\underbrace{\left(\frac{1-q}{1-qz_1}\right)^{\psi_2}}_{\text{negative binomial}}.$$

Since the generating function is a z-transform of probability distribution, the multiplication in z-domain is equivalent to a convolution in probability domain, hence suggesting that the probability of $G_{\rm B}(u,t)$ can be computed by parts:

$$P_{\rm B}(n)=\sum_{i=0}^{n}P_{\rm B0}(i)P_{\rm B1}(n-i)$$

where $P_{\rm B0}(i)$ corresponds to the expansion of binomial part (similar to a binomial distribution but with non-integer trials and one negative coefficient)

$$P_{\rm B0}(i)=\frac{\Gamma(\psi_2+1)}{\Gamma(i+1)\Gamma(\psi_2-i+1)}(-\psi_0)^i(\psi_0+1)^{\psi_2-i}$$

and $P_{\rm B1}(i)$ corresponds to the negative binomial distribution

$$P_{\rm B1}(i)=\frac{\Gamma(i+\psi_2)}{\Gamma(i+1)\Gamma(\psi_2)}(1-q)^{\psi_2}q^i.$$

4. Approximate marginal distribution of mature mRNA in model M5 (including cell division, replication and dosage compensation)

The derivation for mature mRNA follows very similarly to that for nascent mRNA. First we construct the generating function describing mature mRNA dynamics in the absence of cell division, replication and dosage compensation (similar to what we did in Section 3.2). The extension of effective reaction system [11] to include mature mRNA is

$$\varnothing\xrightarrow{f(t,n)}nN,\qquad N\xrightarrow{k}M\xrightarrow{d}\varnothing \tag{19}$$

Note that the effective bursty nascent mRNA production is due to the implicit assumption of $\sigma_{\rm b}$ much larger than $\sigma_{\rm u}$. For each gene copy, the master equation describing the stochastic dynamics is

$$\begin{aligned}\partial_t P(n_{\rm N},n_{\rm M},t)=&\sum_{i=0}^{\infty}\frac{f_0(\beta_0 e^{\alpha t})^i}{(1+\beta_0 e^{\alpha t})^{i+1}}[P(n_{\rm N}-i,n_{\rm M},t)-P(n_{\rm N},n_{\rm M},t)]+k[(n_{\rm N}+1)P(n_{\rm N}+1,n_{\rm M}-1,t)-n_{\rm N}P(n_{\rm N},n_{\rm M},t)]\\&+d[(n_{\rm M}+1)P(n_{\rm N},n_{\rm M}+1,t)-n_{\rm M}P(n_{\rm N},n_{\rm M},t)],\end{aligned}$$

and the corresponding generating function equation is given by

$$\partial_t G(u,v,t)+k(u-v)\partial_u G(u,v,t)+dv\partial_v G(u,v,t)=\frac{f_0\beta_0 e^{\alpha t}u}{1-\beta_0 e^{\alpha t}u}G(u,v,t), \tag{20}$$

with $f_0=\sigma_{\rm u}$ and $\beta_0=\rho_0/\sigma_{\rm b}$ as shown in Section 3.2. Applying the method of characteristics on Eq. [20], we have the following set of ODEs

$$\begin{aligned}\frac{{\rm d}t}{{\rm d}s}&=1 &\Rightarrow\quad t&=s,\\ \frac{{\rm d}u}{{\rm d}s}&=k(u-v),&&\\ \frac{{\rm d}v}{{\rm d}s}&=dv &\Rightarrow\quad v&=v_0e^{ds},\\ \frac{{\rm d}G}{{\rm d}s}&=\frac{f_0\beta_0 e^{\alpha t}u}{1-\beta_0 e^{\alpha t}u}G.&&\end{aligned}$$

with $v_0=v|_{s=0}$. Next we assume that nascent mRNA is fast degraded such that it becomes an intermediate and can be eliminated from the equations above (14). This assumption is realistic since $k\gg d$ holds for the experimental data that we collected from various sources (see Section 2 Table S3). In particular for mouse embryonic stem cells, the timescale of nascent mRNA ($1/k$) is of the order of a few minutes while the timescale of mature mRNA ($1/d$) is several hours. The fast equilibrium assumption implies that ${\rm d}u/{\rm d}s$ is small, equivalently implying $u=v$. Hence, using the latter result, the set of ODEs reduces

$$\frac{{\rm d}G}{{\rm d}s}=\frac{f_0\beta_0 v_0 e^{(\alpha+d)s}}{1-\beta_0 v_0 e^{(\alpha+d)s}}G. \tag{21}$$

Comparing Eq. [21] with Eq. [12] we see that the two are precisely the same if we replace u_0 by v_0 and k by d. Since Eq. [12] corresponds to the reaction scheme [11], it hence follows that Eq. [21] corresponds to the effective reaction scheme

$$\varnothing \xrightarrow{f(t,n)} nM, \qquad M \xrightarrow{d} \varnothing.$$

Hence under the assumption that nascent mRNA is fast, the intermediate steps of production of nascent mRNA and its transformation into mature mRNA, are reduced to the direct production of mature mRNA. In other words, since nascent mRNA is short lived, its dynamics does not affect that of mature mRNA. A special case of the fast nascent mRNA assumption presented here, namely where the transcription rate is independent of time, is indeed implicit in standard models of gene expression which have only an explicit description of mature mRNA dynamics.

By arguments akin to those used previously in solving for nascent mRNA, we immediately have the solution for the generating function corresponding to the marginal distribution of mature mRNA

$$\boxed{G(v,t) = g(v\mathrm{e}^{-dt}) \left(\frac{\rho_0 \mathrm{e}^{-dt} v - \sigma_\mathrm{b}}{\rho_0 \mathrm{e}^{\alpha t} v - \sigma_\mathrm{b}} \right)^{\frac{\sigma_\mathrm{u}}{\alpha + d}}}. \qquad [22]$$

Note that Eqs. [13] and [22] are the same if we replace k by d. Indeed it is straightforward to show that all the equations derived in Section 3.3 for nascent mRNA follow for mature mRNA provided that we replace k by d. Hence we have obtained the approximate marginal distribution for mature mRNA for all cell ages and all generations in our model M5.

5. Derivation of the cyclo-stationary moments

Based on the analytic solution given by Eqs. [15] and [16], we now derive the stationary solutions for the mean and variance of fluctuations of mRNA in the limit of infinite cell cycles, i.e. the moments in the cyclo-stationary limit. Note that in this limit the moments at a particular cell age t are independent of the generation (15). Prior to proceeding to the main results, we define the following notations to simplify the presentation:

$$n_{\mathrm{A},t} = \partial_u G_\mathrm{A}(u,t)|_{u=0}, \qquad n_{\mathrm{B},t} = \partial_u G_\mathrm{B}(u,t)|_{u=0},$$
$$\sigma^2_{\mathrm{A},t} = \partial^2_{uu} G_\mathrm{A}(u,t)|_{u=0} + n_{\mathrm{A},t} - n^2_{\mathrm{A},t}, \qquad \sigma^2_{\mathrm{B},t} = \partial^2_{uu} G_\mathrm{B}(u,t)|_{u=0} + n_{\mathrm{B},t} - n^2_{\mathrm{B},t},$$

The first line corresponds to the mean of the molecule numbers of mother copy and daughter copy while the second line gives the corresponding variance of molecule number fluctuations.

The first step is to find the cyclo-stationary initial condition (at cell age $t = 0$) for the mean. To this end, we take the derivative on Eq. [17] with respect to u, evaluate it at $u = 0$ and then enforce the cyclo-stationary condition ($\langle n_\mathrm{N}\rangle_{n+1,0} = \langle n_\mathrm{N}\rangle_{n,0} = \overline{\langle n_\mathrm{N}\rangle}_0$) to obtain

$$\overline{\langle n_\mathrm{N}\rangle}_0 = \eta \overline{\langle n_\mathrm{N}\rangle}_0 + n_{\mathrm{A},t_\mathrm{d}} + n_{\mathrm{B},t_\mathrm{d}},$$

with operator $\bar{\cdot}$ indicating "cyclo-stationary", hence

$$\boxed{\overline{\langle n_\mathrm{N}\rangle}_0 = \frac{n_{\mathrm{A},t_\mathrm{d}} + n_{\mathrm{B},t_\mathrm{d}}}{1-\eta}}. \qquad [23]$$

Taking the second-order derivative with respect to u on both sides of Eq. [17], we obtain

$$\begin{aligned}\overline{\sigma}^2_{n_\mathrm{N},0} + \overline{\langle n_\mathrm{N}\rangle}^2_0 - \overline{\langle n_\mathrm{N}\rangle}_0 =& \frac{1}{2}\Big[4\eta \overline{\langle n_\mathrm{N}\rangle}_0 n_{\mathrm{B},t_\mathrm{d}} + 4 n_{\mathrm{A},t_\mathrm{d}} n_{\mathrm{B},t_\mathrm{d}} + n^2_{\mathrm{A},t_\mathrm{d}} + 2\eta^2 (\overline{\sigma}^2_{n_\mathrm{N},0} + \overline{\langle n_\mathrm{N}\rangle}^2_0 - \overline{\langle n_\mathrm{N}\rangle}_0) \\ &+ 4\eta \overline{\langle n_\mathrm{N}\rangle}_0 n_{\mathrm{A},t_\mathrm{d}} + (\sigma^2_{\mathrm{A},t_\mathrm{d}} - n_{\mathrm{A},t_\mathrm{d}} + n^2_{\mathrm{A},t_\mathrm{d}}) + n^2_{\mathrm{B},t_\mathrm{d}} + (\sigma^2_{\mathrm{B},t_\mathrm{d}} - n_{\mathrm{B},t_\mathrm{d}} + n^2_{\mathrm{B},t_\mathrm{d}})\Big]\end{aligned}$$

which when simplified leads to the cyclo-stationary initial condition (at cell age $t = 0$) for the variance

$$\boxed{\overline{\sigma}^2_{n_\mathrm{N},0} = \frac{(1+2\eta)(n_{\mathrm{A},t_\mathrm{d}} + n_{\mathrm{B},t_\mathrm{d}}) + \sigma^2_{\mathrm{A},t_\mathrm{d}} + \sigma^2_{\mathrm{B},t_\mathrm{d}}}{2-2\eta^2}}. \qquad [24]$$

Now we find the full time-dependent solution for the cyclo-stationary mean and variance. Taking the derivative of Eq. [15] with respect to u and evaluating at zero, we have

$$\boxed{\overline{\langle n_\mathrm{N}\rangle}_t = \xi \overline{\langle n_\mathrm{N}\rangle}_0 + 2n_{\mathrm{A},t} + 2n_{\mathrm{B},t},} \qquad [25]$$

with $\xi = \mathrm{e}^{-kt}$. Taking second-order derivative of Eq. [15] and evaluating at zero, leads to

$$\begin{aligned}\overline{\sigma}^2_{n_\mathrm{N},t} + \overline{\langle n_\mathrm{N}\rangle}^2_t - \overline{\langle n_\mathrm{N}\rangle}_t =& 4\xi n_{\mathrm{B},t} \overline{\langle n_\mathrm{N}\rangle}_0 + 8 n_{\mathrm{A},t} n_{\mathrm{B},t} + 4\xi n_{\mathrm{A},t} \overline{\langle n_\mathrm{N}\rangle}_0 + \xi^2 \left(\overline{\sigma}^2_{n_\mathrm{N},0} + \overline{\langle n_\mathrm{N}\rangle}^2_0 - \overline{\langle n_\mathrm{N}\rangle}_0\right) \\ &+ 2n^2_{\mathrm{A},t} + 2(\sigma^2_{\mathrm{A},t} - n_{\mathrm{A},t} + n^2_{\mathrm{A},t}) + 2n^2_{\mathrm{B},t} + 2(\sigma^2_{\mathrm{B},t} - n_{\mathrm{B},t} + n^2_{\mathrm{B},t}),\end{aligned}$$

which admits the solution

$$\overline{\sigma^2_{n_{\rm N},t}} = \underbrace{\xi^2\overline{\sigma^2_{n_{\rm N},0}} + \xi(1-\xi)\overline{\langle n_{\rm N}\rangle_0}}_{\text{cell division}} + \underbrace{2\sigma^2_{{\rm A},t}}_{\text{intrinsic noise}} + \underbrace{2\sigma^2_{{\rm B},t}}_{\text{gene replication}}. \quad [26]$$

In summary, Eqs. [23] and [25] define the cyclo-stationary mean of nascent mRNA while Eqs. [24] and [26] define the cyclo-stationary variance of nascent mRNA. By the results of Section 4, it follows that the same expressions hold for mature mRNA if k is replaced by d, and $n_{\rm N}$ by $n_{\rm M}$.

6. Including cell cycle variability

Next we discuss how the variability of cell cycle affects the distributions of mRNA numbers in a population of cells. Previously we have assumed a synchronized population of cells with fixed cell cycle time (Fig. S3A). Now we assume that the cell cycle times $t_{\rm d}$ are random variables independently drawn from an Erlang distribution (of which exponential distribution is a special case), which is consistent with experimental findings (16); this is illustrated in Fig. S3B. The first pronounced observation is that there exists a steady state in the normal sense for the population of cells instead of cyclo-stationary states as shown in Fig. S4A, which arises from the shuffling effect introduced by the cell cycle variability. Since different cells may be in different generations, the generating function of the mRNA numbers at time T is denoted as $G_{\rm v}(u,T)$ (subscript v for cell cycle variability). Hence we are interested in finding an analytic expression in the limit of infinite time, i.e., $\lim_{T\to\infty} G_{\rm v}(u,T)$. Note that the second argument of $G_{\rm v}(u,T)$ represents the absolute time, whereas the second argument of generation function of the n-th generation $G_n(u,t)$ is the cell age.

Let $T_{\rm d}$ be a series of cell cycle times $\{t_{\rm d1},\cdots,t_{{\rm d}i},\cdots\}$, where each $t_{{\rm d}i}$ is i.i.d. sampled from a given cell cycle distribution (denoted as $p_{\rm v}(t_{{\rm d}i})$) such as the Erlang distribution. For a time of interest T, the cell generation n is then defined by the unique integer satisfying $\sum_{i=1}^{n-1} t_{t_{{\rm d}i}} \le T < \sum_{i=1}^{n} t_{t_{{\rm d}i}}$. Hence, generally the generation number n is a random number and a function of a given realization $T_{\rm d}$ and time of interest T. Following that, the cell age of time T becomes

$$t = T - \sum_{i=1}^{n-1} t_{{\rm d}i}. \quad [27]$$

We assume that the cell cycle distribution $p_{\rm v}(t_{{\rm d}i})$ is zero for $t_{{\rm d}i} < \underline{t_{\rm d}}$. The assumption is practically valid as we note from Fig. 3A in the main text that the probability of cell cycle time less than 500 min is almost none. We also assume that gene replication occurs in the middle of a cell cycle, i.e. $t_{{\rm r}i} = t_{{\rm d}i}/2$. This assumption is only for simplicity and can be lifted for generalization. It is further assumed that $\mathrm{e}^{-\varepsilon d\underline{t_{\rm d}}} \ll 1$ for some small $\varepsilon \in [0,1/2)$, which is the case for the mRNA degradation rates presented in Table S1 and the cell cycle distributin Fig. 3A for mouse fibroblast cells. For simplicity, we also do not consider growth-dependent transcription ($\alpha = 0$). The time symbols and their interpretations are summarized in Table S4.

We consider a special deterministic sequence $\bar{T}_{\rm d}$, wherein each element is the mean of the cell cycle distribution $\langle t_{\rm d}\rangle$. The following proof entails three parts: (i) we shall show that the difference between $G_n(u,t_{{\rm d}n}|T_{\rm d})$ and $G_n(u,\langle t_{\rm d}\rangle|\bar{T}_{\rm d})$ is of the order of $\exp(-\varepsilon d\underline{t_{\rm d}})$ for $n\to\infty$ (equivalently $T\to\infty$). (ii) the cell cycle progression $(t/t_{{\rm d}n})$ distribution is shown to be uniform when the cell cycle time distribution is Erlang. (iii) the last step requires finding the generation function at the observation time T in the presence of cell cycle variability.

6.1. Finding generating function solution upon cell division for large generations. By iterating Eqs. [15] and [17] from 1 to n, we get that

$$G_n(u,t_{{\rm d}n}|T_{\rm d}) = \prod_{i=1}^{n} G^2_{\rm A}(\eta_i u,t_{{\rm d}i}) G^2_{\rm B}(\eta_i u,t_{{\rm d}i}), \quad [28]$$

while $G_1(u,0|T_{\rm d}) = 1$ for any $T_{\rm d}$. The decay factor can be written as $\eta_i = \prod_{j=i+1}^{n}\exp(-dt_{{\rm d}j})/2^{n-i}$, and $\eta_n = 1$. Since $\exp(-\varepsilon d\underline{t_{\rm d}}) \ll 1$, then we have $\eta_i \ll 1$ for $i < n$.

Table S4. Summary for all time symbols used in the proof.

Time symbol	Interpretation
$t_{{\rm d}i}$	Cell cycle time for generation i of a cell, randomly sampled form a given cell cycle distribution
$\underline{t_{\rm d}}$	Lower bound of cell cycle time
$\langle t_{\rm d}\rangle$	Mean cell cycle time
$T_{\rm d}$	A series of cell cycle times $\{t_{\rm d1},\cdots,t_{{\rm d}i},\cdots\}$
$\bar{T}_{\rm d}$	Deterministic cell cycle time sequence $\{\langle t_{\rm d}\rangle,\cdots,\langle t_{\rm d}\rangle,\cdots\}$
T	Observation time
t	Cell age, $t = T - \sum_{i=1}^{n-1} t_{{\rm d}i}$
n	Cell generation

Taking logarithms on both sides of Eq. [28] and using Eqs. [18a] and [18b] we obtain

$$\begin{aligned}\ln G_n(u,t_{\mathrm{d}i}|T_{\mathrm{d}}) =& 2\sum_{i=1}^{n}\ln G_{\mathrm{A}}(\eta_i u,t_{\mathrm{d}i})+\ln G_{\mathrm{B}}(\eta_i u,t_{\mathrm{d}i})\\ =&\frac{2}{d}\sum_{i=1}^{n}\sigma_{\mathrm{u}}^{-}\left[\ln(1-b_0\mathrm{e}^{-dt_{\mathrm{d}i}}\eta_i u)-\ln(1-b_0\mathrm{e}^{-dt_{\mathrm{d}i}/2}\eta_i u)\right]\\ &+2\sigma_{\mathrm{u}}^{+}\left[\ln(1-b_0\mathrm{e}^{-dt_{\mathrm{d}i}/2}\eta_i u)-\ln(1-b_0\eta_i u)\right],\end{aligned}$$

with $b_0=\rho_0/\sigma_{\mathrm{b}}$. Note that $u\in[-1,0]$ and $\ln(1+x)\approx x$ for small positive x. Hence, it can be further shown that

$$\ln G_n(u,t_{\mathrm{d}i}|T_{\mathrm{d}})=-\frac{4\sigma_{\mathrm{u}}^{+}}{d}\ln(1-b_0u)+\mathcal{O}(\mathrm{e}^{-\varepsilon d\underline{t}_{\mathrm{d}}}).$$

Note that the leading order term of $\ln G_n(u,t_{\mathrm{d}i}|T_{\mathrm{d}})$ is independent of $t_{\mathrm{d}i}$ and T_{d}. Thus, it allows us to do the following approximation

$$G_n(u,t_{\mathrm{d}n}|T_{\mathrm{d}})\approx G_n(u,\langle t_{\mathrm{d}}\rangle|\bar{T}_{\mathrm{d}}), \tag{29}$$

for $n\to\infty$ provided that $\exp(-\varepsilon d\underline{t}_{\mathrm{d}})$ is small. The physical intuition behind this conclusion is that the distribution of mRNA numbers can only be a weak function of cell cycle duration due to the fast mRNA equilibration. Using Eqs. [16] and [29], it follows that $G_n(u,0|T_{\mathrm{d}})\approx G_n(u,0|\bar{T}_{\mathrm{d}})$ for large n. Equivalently,

$$G_n(u,0|T_{\mathrm{d}})\approx\bar{G}(u,0|\bar{T}_{\mathrm{d}}), \tag{30}$$

which is the cyclo-stationary (deterministic) solution of cell age $t=0$ and cell cycle time $\langle t_{\mathrm{d}}\rangle$.

6.2. Finding the cell cycle progression distribution in the steady state. The next step is to find the distribution of cell age t for different realizations T_{d}. More precisely, we will show that the distribution of cell cycle progression $t/t_{\mathrm{d}n}$ is uniform. It is known that the Erlang waiting time distribution (for cell cycle time) admits a multi-stage decomposition whereby the waiting time of each stage is exponential. Specifically, we divide a cell cycle into s stages and let the switching rates to a next stage be equal to a. The probability of finding a cell in the ith stage at time τ is denoted as $p_i(\tau)$. The temporal evolution of the vector of probabilities $P(\tau)=[p_1(\tau),\cdots,p_s(\tau)]^{\top}$ is then given by

$$\frac{\mathrm{d}}{\mathrm{d}\tau}P(\tau)=AP(\tau),$$

with

$$A=\begin{bmatrix}-a & 0 & \cdots & a\\ a & -a & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & -a\end{bmatrix}.$$

It is easy to conclude that

$$P(\infty)=[1/s,\cdots,1/s]^{\top}. \tag{31}$$

6.3. Finding the generating function at the observation time T. By the definition of $G_{\mathrm{v}}(u,T)$, we have by using Eq. [15] that for large T

$$G_{\mathrm{v}}(u,T)=\langle G_n(u\mathrm{e}^{-dt},0|T_{\mathrm{d}})G_{\mathrm{A}}^2(u,t|T_{\mathrm{d}})G_{\mathrm{B}}^2(u,t|T_{\mathrm{d}})\rangle.$$

Further combining with Eq. [30], we have

$$G_{\mathrm{v}}(u,T)\approx\langle\bar{G}(u\mathrm{e}^{-dt},0|\bar{T}_{\mathrm{d}})G_{\mathrm{A}}^2(u,t|T_{\mathrm{d}})G_{\mathrm{B}}^2(u,t|T_{\mathrm{d}})\rangle. \tag{32}$$

The approximation in Eq. [32] indicates the elimination of one random variable n for large T, and we still have one random variable – cell age t as per Eq. [27]. Then, Eq. [31] suggests the following discretized approximation

$$G_{\mathrm{v}}(u,T)\approx\sum_{i=1}^{s}p_i(\infty)\langle\bar{G}(u\mathrm{e}^{-di\Delta},0|\bar{T}_{\mathrm{d}})G_{\mathrm{A}}^2(u,i\Delta|t_{\mathrm{d}})G_{\mathrm{B}}^2(u,i\Delta|t_{\mathrm{d}})\rangle. \tag{33}$$

The new variable $\Delta=t_{\mathrm{d}}/s$ is still random since t_{d} is stochastic. We change G_{A} and G_{B} conditional on T_{d} to t_{d} because they are generating functions only for one cycle initiated with zero mRNA. We also let

$$\hat{G}(u,t|t_{\mathrm{d}})=\bar{G}(u\mathrm{e}^{-dt},0|\bar{T}_{\mathrm{d}})G_{\mathrm{A}}^2(u,t|t_{\mathrm{d}})G_{\mathrm{B}}^2(u,t|t_{\mathrm{d}})$$

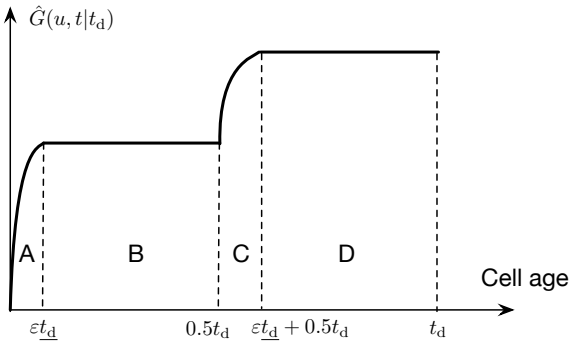

Fig. S1. Illustration of integral interval partition

for brevity, and Eq. [33] can then be rewritten as

$$G_\mathrm{v}(u,T)\approx\sum_{i=1}^{s}p_i(\infty)\langle\hat{G}(u,i\Delta|t_\mathrm{d})\rangle. \tag{34}$$

The continuous analog of Eq. [34] is

$$G_\mathrm{v}(u,T)\approx\left\langle\frac{1}{t_\mathrm{d}}\int_0^{t_\mathrm{d}}\hat{G}(u,t|t_\mathrm{d})\mathrm{d}t\right\rangle=\int_0^\infty\frac{1}{t_\mathrm{d}}\int_0^{t_\mathrm{d}}\hat{G}(u,t|t_\mathrm{d})\mathrm{d}t\mathrm{d}p_\mathrm{v}(t_\mathrm{d})=\underbrace{\int_0^{\underline{t_\mathrm{d}}}\frac{1}{t_\mathrm{d}}\int_0^{t_\mathrm{d}}\hat{G}(u,t|t_\mathrm{d})\mathrm{d}t\mathrm{d}p_\mathrm{v}(t_\mathrm{d})}_{\mathcal{A}}+\underbrace{\int_{\underline{t_\mathrm{d}}}^\infty\frac{1}{t_\mathrm{d}}\int_0^{t_\mathrm{d}}\hat{G}(u,t|t_\mathrm{d})\mathrm{d}t\mathrm{d}p_\mathrm{v}(t_\mathrm{d})}_{\mathcal{B}}.$$

The part $\mathcal{A}$ is zero since $p_\mathrm{v}(t_\mathrm{d})=0$ for $t_\mathrm{d}\le\underline{t_\mathrm{d}}$. We further analyse the inner integral in part $\mathcal{B}$

$$\frac{1}{t_\mathrm{d}}\int_0^{t_\mathrm{d}}\hat{G}(u,t|t_\mathrm{d})\mathrm{d}t=\frac{1}{t_\mathrm{d}}\int_{t\in A\cup B\cup C\cup D}\hat{G}(u,t|t_\mathrm{d})\mathrm{d}t$$

where $A=[0,\varepsilon\underline{t_\mathrm{d}}]$, $B=[\varepsilon\underline{t_\mathrm{d}},t_\mathrm{d}/2]$, $C=[t_\mathrm{d}/2,t_\mathrm{d}/2+\varepsilon\underline{t_\mathrm{d}}]$ and $D=[t_\mathrm{d}/2+\varepsilon\underline{t_\mathrm{d}},\underline{t_\mathrm{d}}]$, which stands for the pre-replication transition period, the pre-replication steady state, the post-replication transition period and the post-replication steady state (see Fig. S1).

Then, we have the following three observations:

(i) When $t\in B\cup C\cup D$, $\exp(-\varepsilon d\underline{t_\mathrm{d}})\ll 1$ gives that $\bar{G}(ue^{-dt},0|\bar{T}_\mathrm{d})\approx 1$, representing the physical insight that steady state is independent of initial conditions.

(ii) According to Eq. [18a], for $t\in B$, $G_\mathrm{A}(u,t|t_\mathrm{d})\approx G_\mathrm{A}^-(u,\infty)=\left(\frac{\sigma_\mathrm{b}}{\sigma_\mathrm{b}-\rho_0 u}\right)^{\sigma_\mathrm{u}^-/d}$. Similary for $t\in D$, $G_\mathrm{A}(u,t|t_\mathrm{d})\approx G_\mathrm{A}^+(u,\infty)=\left(\frac{\sigma_\mathrm{b}}{\sigma_\mathrm{b}-\rho_0 u}\right)^{\sigma_\mathrm{u}^+/d}$.

(iii) According to Eq. [18b], for $t\in A\cup B$, $G_\mathrm{B}(u,t|t_\mathrm{d})=1$, while $G_\mathrm{B}(u,t|t_\mathrm{d})\approx G_\mathrm{A}^+(u,\infty)$ for $t\in D$.

The three observations tell us that $\hat{G}(u,t|t_\mathrm{d})$ is approximately indepentent of t and t_d for $t\in B$ and $t\in D$. Thus, we have

$$\begin{cases}\dfrac{1}{t_\mathrm{d}}\displaystyle\int_{t\in B}\hat{G}(u,t|t_\mathrm{d})\mathrm{d}t\approx\frac{t_\mathrm{d}/2-\varepsilon\underline{t_\mathrm{d}}}{t_\mathrm{d}}[G_\mathrm{A}^-(u,\infty)]^2=\frac{1}{2}[G_\mathrm{A}^-(u,\infty)]^2+\mathcal{O}(\varepsilon),\\ \dfrac{1}{t_\mathrm{d}}\displaystyle\int_{t\in D}\hat{G}(u,t|t_\mathrm{d})\mathrm{d}t\approx\frac{t_\mathrm{d}/2-\varepsilon\underline{t_\mathrm{d}}}{t_\mathrm{d}}[G_\mathrm{A}^+(u,\infty)]^4=\frac{1}{2}[G_\mathrm{A}^+(u,\infty)]^4+\mathcal{O}(\varepsilon).\end{cases} \tag{35}$$

For $t\in A$ and $t\in C$, there exists $\xi_\mathrm{A}\in A$ and $\xi_\mathrm{C}\in C$ such that

$$\frac{1}{t_\mathrm{d}}\int_{t\in A}\hat{G}(u,t|t_\mathrm{d})\mathrm{d}t=\frac{\varepsilon\underline{t_\mathrm{d}}}{t_\mathrm{d}}\hat{G}(u,\xi_\mathrm{A})=\mathcal{O}(\varepsilon)\qquad\text{and}\qquad\frac{1}{t_\mathrm{d}}\int_{t\in C}\hat{G}(u,t|t_\mathrm{d})\mathrm{d}t=\frac{\varepsilon\underline{t_\mathrm{d}}}{t_\mathrm{d}}\hat{G}(u,\xi_\mathrm{C})=\mathcal{O}(\varepsilon), \tag{36}$$

according to the mean-value theorem.

Next we propose a mean-field approximation to $\mathcal{B}$ and have that

$$\mathcal{C}=\int_{\underline{t_\mathrm{d}}}^\infty\mathrm{d}p_\mathrm{v}(t_\mathrm{d})\int_0^{\langle t_\mathrm{d}\rangle}\frac{1}{\langle t_\mathrm{d}\rangle}\hat{G}(u,t|\langle t_\mathrm{d}\rangle)\mathrm{d}t.$$

Analogously, we can partition the interval $[0,\langle t_\mathrm{d}\rangle]$ into A, B, C and D four sets (by replacing t_d therein with $\langle t_\mathrm{d}\rangle$). Hence, Eqs. [35] and [36] immediately become valid for $\mathcal{C}$ (again by replacing t_d therein with $\langle t_\mathrm{d}\rangle$). Therefore, one can conclude that the differences between $\mathcal{B}$ and $\mathcal{C}$ is of the order of $\mathcal{O}(\varepsilon)$.

Further, using the definitions of $\bar{G}(u,t)$ and $\hat{G}(u,t|\langle t_{\rm d}\rangle)$, one can derive that

$$\mathcal{C} = \int_0^\infty \mathrm{d}p_{\rm v}(t_{\rm d}) \int_0^{\langle t_{\rm d}\rangle} \frac{1}{\langle t_{\rm d}\rangle} \bar{G}(u,t)\mathrm{d}t = \frac{1}{\langle t_{\rm d}\rangle} \int_0^{\langle t_{\rm d}\rangle} \bar{G}(u,t)\mathrm{d}t.$$

Hence summarizing, we have shown that an approximate analytic solution to $G_{\rm v}(u,T)$ for large T is

$$\lim_{T\to\infty} G_{\rm v}(u,T) \approx \frac{1}{\langle t_{\rm d}\rangle} \int_0^{\langle t_{\rm d}\rangle} \bar{G}(u,t)\mathrm{d}t, \qquad [37]$$

provided the condition $\mathrm{e}^{-\varepsilon d \underline{t}_{\rm d}} \ll 1$ for some $\varepsilon \in [0, 1/2]$,

This means that the transcriptional dynamics of the stochastic gene expression model with cell cycle variability (Fig. S3B) is approximated by a model for an asynchronous cell population but with identical cell cycle time (see Fig. S3C). Note that the cyclo-stationary generating function $\bar{G}(u,t)$ in the integral above is the generating function of the effective negative binomial distribution constructed from the cyclo-stationary mean and variance, as described in the main text. Simulations indicate that the approximation given by Eq. [37] is accurate for various choices of cell cycle distributions (see Fig. S4B,C). Furthermore we found that Eq. [37] agrees very well with SSA predictions using experimentally determined parameters and cell cycle distribution for 16 mouse fibroblast genes (See Fig. 3A,B in main text and Fig. S5). Surprisingly, the solution Eq. [37] is still accurate for cell cycle time distribution being exponential (see Fig. S4), which suggests that the condition $\exp(-\varepsilon d \underline{t}_{\rm d}) \ll 1$ may be relaxed to $\exp(-\varepsilon d\langle t_{\rm d}\rangle) \ll 1$.

6.4. Comparison to the exact solution for the mean mRNA numbers. In support of the proof above, by using the method presented in Ref. (17), we next provide an exact calculation for the mean mRNA numbers under the condition that the cycle time distribution is exponential and gene dosage compensation is perfect ($\sigma_{\rm u}^+ = 0.5\sigma_{\rm u}^-$). Under the aforementioned condition, the temporal evolution of the mean within a cell cycle becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle n_{\rm M}\rangle_t = 2\kappa - d\langle n_{\rm M}\rangle_t,$$

which is exactly Eq. [35] in Ref. (17) with $\kappa = \frac{\rho_0\sigma_{\rm u}^-}{\sigma_{\rm b}+\sigma_{\rm u}} \approx \frac{\rho_0\sigma_{\rm u}^-}{\sigma_{\rm b}}$ since $\sigma_{\rm b} \gg \sigma_{\rm u}$. The number 2 on the right hand side arises from the presence of two gene copies, and perfect gene dosage compensation so that four gene copies behave like two on the mean level. Hence, we can calculate the mean at steady state exactly by means of Eq. [37] in Ref. (17), which reads

$$\langle n_{\rm M}\rangle_{\rm ss} = \frac{2\kappa}{d} - \frac{\kappa}{d^2\langle t_{\rm d}\rangle} \frac{1-\langle \mathrm{e}^{-d t_{\rm d}}\rangle}{1-\langle \mathrm{e}^{-d t_{\rm d}}\rangle/2}.$$

Here $\langle \mathrm{e}^{-d t_{\rm d}}\rangle$ can be computed from Eq. [13] in Ref. (17) and is equal to

$$\langle \mathrm{e}^{-d t_{\rm d}}\rangle = \int_0^\infty \frac{1}{\langle t_{\rm d}\rangle} \mathrm{e}^{-\frac{t}{\langle t_{\rm d}\rangle}} \mathrm{e}^{-dt}\mathrm{d}t = \frac{1}{1+d\langle t_{\rm d}\rangle}.$$

Hence we have,

$$\langle n_{\rm M}\rangle_{\rm ss} = \frac{4\kappa\langle t_{\rm d}\rangle}{1+2d\langle t_{\rm d}\rangle}.$$

On the other hand, using Eqs. [23] and [25], the cyclo-stationary mean is given by

$$\overline{\langle n_{\rm M}\rangle}_t = \frac{2\kappa}{d} \cdot \frac{2-\mathrm{e}^{-d\langle t_{\rm d}\rangle} - \mathrm{e}^{-dt-d\langle t_{\rm d}\rangle}}{2-\mathrm{e}^{-d\langle t_{\rm d}\rangle}}.$$

Then, the average of the cyclo-stationary mean can be calculated as

$$\langle n_{\rm M}\rangle_{\rm avg} = \frac{1}{\langle t_{\rm d}\rangle} \int_0^{\langle t_{\rm d}\rangle} \overline{\langle n_{\rm M}\rangle}_t \mathrm{d}t = \frac{2\kappa}{d} \cdot \frac{-\mathrm{e}^{-d\langle t_{\rm d}\rangle} + (2\mathrm{e}^{d\langle t_{\rm d}\rangle}-1)^{-1} + d\langle t_{\rm d}\rangle}{d\langle t_{\rm d}\rangle} \approx \frac{2\kappa}{d}$$

Therefore, the approximation error is

$$|\langle n_{\rm M}\rangle_{\rm ss} - \langle n_{\rm M}\rangle_{\rm avg}| \approx 2\kappa \left| \frac{2\langle t_{\rm d}\rangle}{1+2d\langle t_{\rm d}\rangle} - \frac{1}{d} \right| \propto \frac{1}{d\langle t_{\rm d}\rangle},$$

which confirms that the error in the approximation Eq. [37] is approximately proportional to the ratio of the mRNA lifetime and the average cell cycle time.

6.5. Adder mechanism. Previously we have taken a phenomenological approach to modelling cell cycle time variability. Here we consider the more specific case where this variability is due to stochasticity in the cell growth rates and a form of cell size control. The model in the main text uses the "timer mechanism" to determine when to divide a cell into two; specifically, when the cell age t reaches $t_{\rm d}$, a cell divides into two. It has been argued (18) that the timer mechanism cannot provide size homeostasis. We here focus on an alternative mechanism, the adder, which has received the most attention in the literature. Let us denote the volume of a cell upon birth and at cell age t as V_0 and V_t respectively, such that the incremental volume is $\delta V = V_t - V_0$. The adder mechanism assumes that a cell divides when the incremental volume δV exceeds a constant threshold ΔV, independent of the cell volume at birth. As per the data shown in Ref. (19), the volume growth is almost linear. Hence, we establish that given a particular generation we have

$$V_0 + r_{\rm grow} t_{\rm d} = V_{\rm f}, \qquad [38]$$

where $r_{\rm grow}$ is the growth rate (constant for a given generation) and $V_{\rm f}$ is the volume of a cell at mitosis. Equivalently, one can also rewrite Eq. [38] as

$$t_{\rm d} = \frac{\Delta V}{r_{\rm grow}}.$$

Hence if the growth rate is variable across different generations then the cell cycle length shows a similar variability. For example if the growth rate distribution follows an inverse gamma distribution (Fig. S6A) then the cell cycle length is gamma distributed (Fig. S6B). Again we assume the transcription rate is independent of cell volume i.e. $\alpha = 0$, and gene replication occurs in the middle of the cell cycle, i.e. $t_{\rm r} = t_{\rm d}/2$. Modifying the SSA to take into account the adder mechanism of size control and an inverse gamma distribution of growth rates, we can obtain from stochastic simulations the distribution of mature mRNA numbers across the population of cells (blue dots in Fig. S6C). This agrees very well with the distribution of mature mRNA estimated from Eq. [37] with the mean cell cycle time $\langle t_{\rm d} \rangle$ given by the mean of the gamma distribution in Fig. S6B.

7. Derivation of the error in the negative binomial approximation of the telegraph model

A popular effective model of our full model M5 is the telegraph model

$$G \xrightarrow{\rho_{\rm u}} G + M, \qquad M \xrightarrow{\hat{d}} \varnothing, \qquad G \underset{\sigma_{\rm u}}{\overset{\sigma_{\rm b}}{\rightleftharpoons}} G^* \qquad [39]$$

Note that $\hat{d} = d + \ln 2/t_{\rm d}$, where d is mature mRNA degradation and $\ln 2/t_{\rm d}$ approximates the dilution effect of cell division (20). The mean and variance on M of the telegraph model at steady state are

$$\langle n \rangle_{\rm tele} = \frac{\hat{\rho}_{\rm u}\hat{\sigma}_{\rm u}}{\hat{\sigma}_{\rm u} + \hat{\sigma}_{\rm b}} \qquad \text{and} \qquad \sigma^2_{\rm tele} = \frac{\hat{\rho}_{\rm u}\hat{\sigma}_{\rm u}(\hat{\sigma}_{\rm b}^2 + \hat{\sigma}_{\rm u} + \hat{\sigma}_{\rm u}^2 + \hat{\sigma}_{\rm b}(1 + \hat{\rho}_{\rm u} + 2\hat{\sigma}_{\rm u}))}{(\hat{\sigma}_{\rm b} + \hat{\sigma}_{\rm u})^2(1 + \hat{\sigma}_{\rm b} + \hat{\sigma}_{\rm u})},$$

where $\hat{\rho}_{\rm u} = \rho_{\rm u}/\hat{d}, \hat{\sigma}_{\rm u} = \sigma_{\rm u}/\hat{d}, \hat{\sigma}_{\rm b} = \sigma_{\rm b}/\hat{d}$. The negative binomial approximation $\mathrm{NB}(r, p)$ to the telegraph model is then obtained by equating the first two moments of this distribution to the exact first two moments of the telegraph model, namely,

$$r = \frac{\langle n \rangle^2_{\rm tele}}{\sigma^2_{\rm tele} - \langle n \rangle_{\rm tele}}, \qquad p = 1 - \frac{\langle n \rangle_{\rm tele}}{\sigma^2_{\rm tele}}.$$

Given r and p, it is easy to find the non-central 3$^{\rm rd}$ moment of the negative binomial distribution as a function of r and p which we denote as $\langle n^3 \rangle_{\rm NB}$. Similarly, the non-central 3$^{\rm rd}$ moment of the telegraph model [39] can be also found and denoted as $\langle n^3 \rangle_{\rm tele}$. Then, the absolute error between the 3$^{\rm rd}$ moments of the telegraph model and of its negative binomial approximation is defined as

$$\Theta = |\langle n^3 \rangle_{\rm NB} - \langle n^3 \rangle_{\rm tele}| = \frac{2\hat{\rho}_{\rm u}^3\hat{\sigma}_{\rm b}\hat{\sigma}_{\rm u}(1 + \hat{\sigma}_{\rm u})}{(\hat{\sigma}_{\rm b} + \hat{\sigma}_{\rm u})^2(1 + \hat{\sigma}_{\rm b} + \hat{\sigma}_{\rm u})^2(2 + \hat{\sigma}_{\rm b} + \hat{\sigma}_{\rm u})}.$$

In the main text we show that this expression serves as a simple but accurate index of how well the negative binomial can approximate the time-dependent distribution of model M5.

8. Approximate marginal distribution of protein for model M5

Next we extend the effective system [19] to include protein translation and degradation

$$\varnothing \xrightarrow{f(n)} nN, \quad N \xrightarrow{k} M \xrightarrow{d} \varnothing, \quad M \xrightarrow{\lambda} M + P, \quad P \xrightarrow{d_{\rm p}} \varnothing.$$

Note that the bursts of nascent mRNA were earlier derived by assuming that the gene spends most of its time in the OFF state. We shall specifically consider the case of non growth-dependent transcription, i.e. $\alpha = 0$ for the following results. For each gene copy, the chemical master equation reads

$$\begin{aligned}\frac{\mathrm{d}P(n_\mathrm{N}, n_\mathrm{M}, n_\mathrm{P})}{\mathrm{d}t} = &\sum_{i=0}^{\infty} \frac{f_0 \beta_0^i}{(1+\beta_0)^{i+1}} [P(n_\mathrm{N}-i, n_\mathrm{M}, n_\mathrm{P}) - P(n_\mathrm{N}, n_\mathrm{M}, n_\mathrm{P})] + k[(n_\mathrm{N}+1)P(n_\mathrm{N}+1, n_\mathrm{M}-1, n_\mathrm{P}) - n_\mathrm{N} P(n_\mathrm{N}, n_\mathrm{M}, n_\mathrm{P})] \\ &+ d[(n_\mathrm{M}+1)P(n_\mathrm{N}, n_\mathrm{M}+1, n_\mathrm{P}) - n_\mathrm{M} P(n_\mathrm{N}, n_\mathrm{M}, n_\mathrm{P})] + \lambda[n_\mathrm{M} P_0(n_\mathrm{N}, n_\mathrm{M}, n_\mathrm{P}-1) - n_\mathrm{M} P_0(n_\mathrm{N}, n_\mathrm{M}, n_\mathrm{P})] \\ &+ d_\mathrm{p}[(n_\mathrm{P}+1)P(n_\mathrm{N}, n_\mathrm{M}, n_\mathrm{P}+1) - n_\mathrm{P} P(n_\mathrm{N}, n_\mathrm{M}, n_\mathrm{P})],\end{aligned} \tag{40}$$

By assuming generating functions of the form

$$G = \sum_{n_\mathrm{N}=0}^{\infty} \sum_{n_\mathrm{M}=0}^{\infty} \sum_{n_\mathrm{P}=0}^{\infty} z_1^{n_\mathrm{N}} z_2^{n_\mathrm{M}} z_3^{n_\mathrm{P}} P(n_\mathrm{N}, n_\mathrm{M}, n_\mathrm{P})$$

the chemical master equation Eq. [40] can also be equivalently represented as

$$\partial_t G = \frac{f_0 \beta_0 (z_1 - 1)}{1 - \beta_0 (z_1 - 1)} G + k(z_2 - z_1)\partial_{z_1} G - d(z_2 - 1)\partial_{z_2} G + \lambda z_2 (z_3 - 1)\partial_{z_2} G - d_\mathrm{p}(z_3 - 1)\partial_{z_3} G,$$

where the arguments z_1, z_2, z_3 and t are suppressed for simplicity. Using $u = z_1 - 1$, $v = z_2 - 1$ and $w = z_3 - 1$, we have

$$\partial_t G = \frac{f_0 \beta_0 u}{1 - \beta_0 u} G + k(v - u)\partial_u G - dv\partial_v G + \lambda(v+1)w\partial_v G - d_\mathrm{p} w \partial_w G. \tag{41}$$

By means of the method of characteristics, we have the following set of ODEs

$$\begin{aligned}&\frac{\mathrm{d}t}{\mathrm{d}s} = 1 && \Rightarrow && t = s, \\ &\frac{\mathrm{d}u}{\mathrm{d}s} = k(u - v) \\ &\frac{\mathrm{d}v}{\mathrm{d}s} = [dv - \lambda(v+1)w] \\ &\frac{\mathrm{d}w}{\mathrm{d}s} = d_\mathrm{p} w && \Rightarrow && w = w_0 \mathrm{e}^{d_\mathrm{p} s}, \\ &\frac{\mathrm{d}G}{\mathrm{d}s} = \frac{f_0 \beta_0 u}{1 - \beta_0 u} G,\end{aligned}$$

with $w_0 = w|_{s=0}$.

The first step towards solving Eq. [41] for the marginal distribution of protein is to eliminate the variables of both nascent and mature mRNAs as they are fast intermediates. This assumption is motivated by the large experimentally measured values of k/d and d/d_p for many genes (see main text for discussion). This of course suggests that $\mathrm{d}u/\mathrm{d}s$ and $\mathrm{d}v/\mathrm{d}s$ are small. It hence follows that

$$u = v, \qquad dv - \lambda(v+1)w = 0,$$

which together with $w = w_0 \exp(d_\mathrm{p} s)$ yield

$$u = \frac{\lambda w_0 \mathrm{e}^{d_\mathrm{p} s}}{d - \lambda w_0 \mathrm{e}^{d_\mathrm{p} s}}.$$

Hence, the set of ODEs reduce to

$$\frac{\mathrm{d}G}{\mathrm{d}s} = -\frac{\rho_\mathrm{u} \sigma_\mathrm{u} \lambda w_0 \mathrm{e}^{d_\mathrm{p} s}}{(\rho_\mathrm{u} + \sigma_\mathrm{b})\lambda w_0 \mathrm{e}^{d_\mathrm{p} s} - \sigma_\mathrm{b} d} G$$

by using the definitions of $f_0 = \sigma_\mathrm{u}$ and $\beta_0 = \rho_0/\sigma_\mathrm{b}$ explicitly. Assuming the initial condition $G|_{t=0,\, u=0,\, v=0} = g(w)$, the analytic solution for protein marginal distribution can be immediately obtained as

$$G = g(w\mathrm{e}^{-d_\mathrm{p} t}) \left[\frac{b(\rho_\mathrm{u} + \sigma_\mathrm{b}) w \mathrm{e}^{-d_\mathrm{p} t} - \sigma_\mathrm{b}}{b(\rho_\mathrm{u} + \sigma_\mathrm{b}) w - \sigma_\mathrm{b}} \right]^{\frac{\rho_\mathrm{u} \sigma_\mathrm{u}}{d_\mathrm{p}(\rho_\mathrm{u} + \sigma_\mathrm{b})}}, \tag{42}$$

where $b = \lambda/d$ is the mean protein burst size (mean number of proteins produced during the lifetime of mature mRNA).

It is straightforward to show that Eqs. [15] and [17] also hold for the case of proteins. By the same reasoning as in Section 3.3 and using Eq. [42] one finds the expressions for $G_{\rm A}$ and $G_{\rm B}$ are given by

$$G_{\rm A}(w,t) = \begin{cases} \left[\dfrac{b(\rho_{\rm u}+\sigma_{\rm b})we^{-d_{\rm p}t}-\sigma_{\rm b}}{b(\rho_{\rm u}+\sigma_{\rm b})w-\sigma_{\rm b}}\right]^{\frac{\rho_{\rm u}\sigma_{\rm u}^{-}}{d_{\rm p}(\rho_{\rm u}+\sigma_{\rm b})}}, & t\in[0,t_{\rm r}), \\ \left[\dfrac{b(\rho_{\rm u}+\sigma_{\rm b})we^{-d_{\rm p}t}-\sigma_{\rm b}}{b(\rho_{\rm u}+\sigma_{\rm b})we^{-d_{\rm p}(t-t_{\rm r})}-\sigma_{\rm b}}\right]^{\frac{\rho_{\rm u}\sigma_{\rm u}^{-}}{d_{\rm p}(\rho_{\rm u}+\sigma_{\rm b})}} \left[\dfrac{b(\rho_{\rm u}+\sigma_{\rm b})we^{-d_{\rm p}(t-t_{\rm r})}-\sigma_{\rm b}}{b(\rho_{\rm u}+\sigma_{\rm b})w-\sigma_{\rm b}}\right]^{\frac{\rho_{\rm u}\sigma_{\rm u}^{+}}{d_{\rm p}(\rho_{\rm u}+\sigma_{\rm b})}} & t\in[t_{\rm r},t_{\rm d}), \end{cases} \quad [43a]$$

$$G_{\rm B}(w,t) = \begin{cases} 1 & t\in[0,t_{\rm r}), \\ \left[\dfrac{b(\rho_{\rm u}+\sigma_{\rm b})we^{-d_{\rm p}(t-t_{\rm r})}-\sigma_{\rm b}}{b(\rho_{\rm u}+\sigma_{\rm b})w-\sigma_{\rm b}}\right]^{\frac{\rho_{\rm u}\sigma_{\rm u}^{+}}{d_{\rm p}(\rho_{\rm u}+\sigma_{\rm b})}} & t\in[t_{\rm r},t_{\rm d}). \end{cases} \quad [43b]$$

Hence summarising, Eqs. [15], [16], [43a] and [43b] together define the temporal evolution of the protein marginal distribution when the gene is mostly in the OFF state and protein timescales are significantly slower than those of nascent and mature mRNA.

9. Including gene-protein interactions (feedback loop)

The stochastic model including feedback dynamics will be analyzed by the recent developed method – *linear mapping approximation (LMA)* (21). The underlying idea is to map a gene regulatory system with bimolecular interactions (known as a nonlinear system) onto a system involving only effective first-order reactions (known as linear system). If the exact time-dependent solution of the linear system is known then the LMA provides a computational recipe by which an approximate time-dependent solution of the nonlinear system can be obtained. Specifically the recipe involves: (i) finding the mapped linear system by removing protein from the reversible protein-gene interactions in the nonlinear system; (ii) finding the effective binding rates of the linear system by solving the moment equations of its master equation. Note that the mapping between the constants of the linear and nonlinear systems involves the use of a conditional mean-field approximation; (iii) calculating the time-average of these effective rates; (iv) finding the time-dependent probability distribution solution of the master equation of the linear system assuming constant rates; (v) replacing the constant rates in (iv) by the time-average of the effective rates found in (iii). Note that steps (iii)-(v) are equivalent to the first term of the Magnus expansion of the master equation of the linear system.

9.1. Illustrative Example. We will first illustrate the computational recipe of the LMA by using an example – an auto-regulatory feedback loop adapted from Ref. (21). Specifically, the reaction scheme is given by:

$$G \xrightarrow{\rho_{\rm u}} G+P,\ \ G^* \xrightarrow{\rho_{\rm b}} G^*+P,\ \ G+P \underset{\sigma_{\rm u}}{\overset{\sigma_{\rm b}}{\rightleftharpoons}} G^*,\ \ P \xrightarrow{d_{\rm p}} \varnothing. \quad [44]$$

The first step is to find the linear network corresponding to Eq. [44] which is obtained by replacing the bimolecular reaction therein with a first-order reaction leading to

$$G \xrightarrow{\rho_{\rm u}} G+P,\ \ G^* \xrightarrow{\rho_{\rm b}} G^*+P,\ \ G \underset{\sigma_{\rm u}}{\overset{\tilde{\sigma}_{\rm b}}{\rightleftharpoons}} G^*,\ \ P \xrightarrow{d_{\rm p}} \varnothing. \quad [45]$$

Here $\tilde{\sigma}_{\rm b}$ is the effective reaction rate which is dependent (in some way still to be found) on the rate constants of the non-linear network. Next we write down the moment equations of the linear network Eq. [45] which are given by

$$\mathcal{M}: \begin{cases} \partial_t\langle n_{\rm g}\rangle = -\tilde{\sigma}_{\rm b}\langle n_{\rm g}\rangle + \sigma_{\rm u}(1-\langle n_{\rm g}\rangle), \\ \partial_t\langle n_{\rm P}\rangle = \rho_{\rm u}\langle n_{\rm g}\rangle + \rho_{\rm b}(1-\langle n_{\rm g}\rangle) - d_{\rm p}\langle n_{\rm P}\rangle, \\ \partial_t\langle n_{\rm P}n_{\rm g}\rangle = \rho_{\rm u}\langle n_{\rm g}\rangle + \sigma_{\rm u}\langle n_{\rm P}\rangle - (d_{\rm p}+\sigma_{\rm u}+\tilde{\sigma}_{\rm b})\langle n_{\rm P}n_{\rm g}\rangle. \end{cases} \quad [46]$$

The moments can be solved from Eq. [46] together with the effective rate parametrization

$$\tilde{\sigma}_{\rm b} = \sigma_{\rm b}\langle n_{\rm P}|n_{\rm g}=1\rangle = \sigma_{\rm b}\frac{\langle n_{\rm P}n_{\rm g}\rangle}{\langle n_{\rm g}\rangle}, \quad [47]$$

where the second step holds because $n_{\rm g}$ is Boolean. The quantity $\langle n_{\rm P}|n_{\rm g}=1\rangle$ stands for the mean number of protein conditional on gene state G. Note that Eq. [47] implements a conditional mean-field approximation to the propensity of the bimolecular reaction in reaction scheme Eq. [44]. A sufficient condition for this approximation to hold is that the number of proteins conditioned on state G is large enough such that the fluctuations of this quantity about the mean are small. This condition is however not necessary: for instance, the approximation is still accurate when the modes of the protein distribution conditioned on each state

(ρ_u/d and ρ_b/d) are very close to each other since in this case, gene switching doesn't play a significant role in determining the protein distribution.

By solving Eqs. [46] and [47] simultaneously, we have the solution for $\tilde{\sigma}_\mathrm{b}$ from $t = 0$ to any time of interest T. The time-dependent solution to the protein distribution of the linear reaction scheme Eq. [45] is reported in Ref. (22) and also Eqs. (13)(14) in Ref. (21). The solution depends on rate constants ρ_u, ρ_b, d_p, σ_u, $\tilde{\sigma}_\mathrm{b}$ and the time of interest T, and is denoted as $\mathcal{S}(\rho_\mathrm{u}, \rho_\mathrm{b}, \sigma_\mathrm{u}, d_\mathrm{p}, \tilde{\sigma}_\mathrm{b}, T)$. Note that the solution assumes that all the reaction rates are constant; in other words, the solution $\mathcal{S}(\rho_\mathrm{u}, \rho_\mathrm{b}, \sigma_\mathrm{u}, d_\mathrm{p}, \tilde{\sigma}_\mathrm{b}, T)$ is valid if $\tilde{\sigma}_\mathrm{b}$ is a constant. To use the solution, we apply the time-averaging approximation

$$\underline{\sigma}_\mathrm{b} = \frac{1}{T}\int_0^T \tilde{\sigma}_\mathrm{b} \mathrm{d}t, \qquad [48]$$

and the protein distribution of the feedback loop Eq. [44] is approximately given by $\mathcal{S}(\rho_\mathrm{u}, \rho_\mathrm{b}, \sigma_\mathrm{u}, d_\mathrm{p}, \underline{\sigma}_\mathrm{b}, T)$. It can be shown that implementing the time-averaging approximation Eq. [48] corresponds to the first term of the Magnus expansion of the time-dependent solution of the master equation of the linear network Eq. [45] with time-varying rate constant $\tilde{\sigma}_\mathrm{b}$. The approximation error of the LMA is dominated by the error stemming from the time-averaging approximation which increases with the value of σ_b; a more detailed discussion of the approximation error can be found in Ref. (21).

9.2. Application to Model M5. Next we will outline each step of the computational recipe and show how it can be used to obtain an approximate time-dependent solution of an auto-regulatory feedback loop version of M5 (Fig. 1C in main text). This means replacing the reactive components of the model (centre of Fig. 1C in main text) by the following scheme:

$$\begin{aligned} &G_0 \xrightarrow{\rho_\mathrm{u}} G_0 + N, \quad G_0^* \xrightarrow{\rho_\mathrm{b}} G_0^* + N, \quad G_0 + P \underset{\sigma_\mathrm{u}}{\overset{\sigma_\mathrm{b}}{\rightleftharpoons}} G_0^*, \quad N \xrightarrow{k} M, \quad M \xrightarrow{\lambda} M + P, \\ &G_1 \xrightarrow{\rho_\mathrm{u}} G_1 + N, \quad G_1^* \xrightarrow{\rho_\mathrm{b}} G_1^* + N, \quad G_1 + P \underset{\sigma_\mathrm{u}}{\overset{\sigma_\mathrm{b}}{\rightleftharpoons}} G_1^*, \quad M \xrightarrow{d} \varnothing, \quad P \xrightarrow{d_\mathrm{p}} \varnothing, \end{aligned}$$

Note that G_0 and G_1 are the two gene copies prior to replication. The reactions with rate ρ_b stand for the leakage of gene expression and are introduced for generality. Protein P binds to both gene copies with rate σ_b and unbinds with rate σ_u. This interaction is bimolecular and the heart of the feedback loop. For simplicity, transcription rates ρ_u and ρ_b are assumed to be independent of cell volume ($\alpha = 0$).

Step (i). The linear system corresponding to the nonlinear feedback loop above is obtained by ignoring the protein in the bimolecular reactions

$$\begin{aligned} &G_0 \xrightarrow{\rho_\mathrm{u}} G_0 + N, \quad G_0^* \xrightarrow{\rho_\mathrm{b}} G_0^* + N, \quad G_0 \underset{\sigma_\mathrm{u}}{\overset{\tilde{\sigma}_{\mathrm{b}0}}{\rightleftharpoons}} G_0^*, \quad N \xrightarrow{k} M, \quad M \xrightarrow{\lambda} M + P, \\ &G_1 \xrightarrow{\rho_\mathrm{u}} G_1 + N, \quad G_1^* \xrightarrow{\rho_\mathrm{b}} G_1^* + N, \quad G_1 \underset{\sigma_\mathrm{u}}{\overset{\tilde{\sigma}_{\mathrm{b}1}}{\rightleftharpoons}} G_1^*, \quad M \xrightarrow{d} \varnothing, \quad P \xrightarrow{d_\mathrm{p}} \varnothing, \end{aligned} \qquad [49]$$

as illustrated in Fig. S8. Here $\tilde{\sigma}_{\mathrm{b}0}$ and $\tilde{\sigma}_{\mathrm{b}1}$ are the effective binding rates to be determined from moment predictions, which will be detailed later. If the mRNA maturation process is fast and mature mRNA is unstable, Eq. [49] further reduces to

$$\begin{aligned} &G_0 \xrightarrow{\rho_\mathrm{u}} G_0 + iP, \quad G_0^* \xrightarrow{\rho_\mathrm{b}} G_0^* + iP, \quad G_0 \underset{\sigma_\mathrm{u}}{\overset{\tilde{\sigma}_{\mathrm{b}0}}{\rightleftharpoons}} G_0^*, \quad P \xrightarrow{d_\mathrm{p}} \varnothing, \\ &G_1 \xrightarrow{\rho_\mathrm{u}} G_1 + iP, \quad G_1^* \xrightarrow{\rho_\mathrm{b}} G_1^* + iP, \quad G_1 \underset{\sigma_\mathrm{u}}{\overset{\tilde{\sigma}_{\mathrm{b}1}}{\rightleftharpoons}} G_1^*, \end{aligned} \qquad [50]$$

where i is a random variable and the corresponding protein production reactions occur with probability $\frac{b^i}{(1+b)^{i+1}}$ and $b = \lambda/d$.

Step(ii). The moments of the master equation of the linear system [50] prior to gene replication is given by a set of coupled differential equations

$$\begin{cases} \partial_t \langle n_\mathrm{P} \rangle_n = \rho_\mathrm{u} b(\langle n_{\mathrm{g}0} \rangle_n + \langle n_{\mathrm{g}1} \rangle_n) + \rho_\mathrm{b} b(2 - \langle n_{\mathrm{g}0} \rangle_n - \langle n_{\mathrm{g}1} \rangle_n) - d_\mathrm{p} \langle n_\mathrm{P} \rangle_n, \\ \partial_t \langle n_{\mathrm{g}0} \rangle_n = -\tilde{\sigma}_{\mathrm{b}0} \langle n_{\mathrm{g}0} \rangle_n + \sigma_\mathrm{u}(1 - \langle n_{\mathrm{g}0} \rangle_n), \\ \partial_t \langle n_{\mathrm{g}1} \rangle_n = -\tilde{\sigma}_{\mathrm{b}1} \langle n_{\mathrm{g}1} \rangle_n + \sigma_\mathrm{u}(1 - \langle n_{\mathrm{g}1} \rangle_n), \\ \partial_t \langle n_\mathrm{P} n_{\mathrm{g}0} \rangle_n = (\rho_\mathrm{u} + \rho_\mathrm{b}) b \langle n_{\mathrm{g}0} \rangle_n + (\rho_\mathrm{u} - \rho_\mathrm{b}) b \langle n_{\mathrm{g}0} n_{\mathrm{g}1} \rangle_n - (d_\mathrm{p} + \sigma_\mathrm{u} + \tilde{\sigma}_{\mathrm{b}0}) \langle n_\mathrm{P} n_{\mathrm{g}0} \rangle_n, \\ \partial_t \langle n_\mathrm{P} n_{\mathrm{g}1} \rangle_n = (\rho_\mathrm{u} + \rho_\mathrm{b}) b \langle n_{\mathrm{g}1} \rangle_n + (\rho_\mathrm{u} - \rho_\mathrm{b}) b \langle n_{\mathrm{g}0} n_{\mathrm{g}1} \rangle_n - (d_\mathrm{p} + \sigma_\mathrm{u} + \tilde{\sigma}_{\mathrm{b}1}) \langle n_\mathrm{P} n_{\mathrm{g}1} \rangle_n, \\ \partial_t \langle n_{\mathrm{g}0} n_{\mathrm{g}1} \rangle_n = \sigma_\mathrm{u} \langle n_{\mathrm{g}0} \rangle_n + \sigma_\mathrm{u} \langle n_{\mathrm{g}1} \rangle_n - (2\sigma_\mathrm{u} + \tilde{\sigma}_{\mathrm{b}0} + \tilde{\sigma}_{\mathrm{b}1}) \langle n_{\mathrm{g}0} n_{\mathrm{g}1} \rangle_n. \end{cases} \qquad [51]$$

where $\langle . \rangle_n$ is the mean in generation n. According to the LMA, the effective binding rates of the linear system are functions of time and of the parameters of the non-linear system. They are are functions of the moments of the linear system and are given by

$$\tilde{\sigma}_{\mathrm{b}0} = \sigma_\mathrm{b} \frac{\langle n_\mathrm{P} | n_{\mathrm{g}0} = 1 \rangle_n}{\Omega(t)} = \sigma_\mathrm{b} \frac{\langle n_\mathrm{P} n_{\mathrm{g}0} \rangle_n}{\Omega(t) \langle n_{\mathrm{g}0} \rangle_n} \quad \text{and} \quad \tilde{\sigma}_{\mathrm{b}1} = \sigma_\mathrm{b} \frac{\langle n_\mathrm{P} | n_{\mathrm{g}1} = 1 \rangle_n}{\Omega(t)} = \sigma_\mathrm{b} \frac{\langle n_\mathrm{P} n_{\mathrm{g}1} \rangle_n}{\Omega(t) \langle n_{\mathrm{g}1} \rangle_n}. \qquad [52]$$

Here n_P is the number of molecules of protein P, n_{g0}, n_{g1} are Boolean variables taking the value of 1 if the promoter is in state G_0 or G_1 and the value 0 if it is in state G_0^* or G_1^*, respectively. The cell volume $\Omega(t)$ is a function of cell age t, and changes linearly as $\Omega(t) = 1 + t/t_d$. The last step in Eq. [52] holds because n_{g0} and n_{g1} are Boolean. Due to the symmetry of n_{g0} and n_{g1}, Eq. [51] can be written more compactly

$$\begin{cases} \partial_t \langle n_P \rangle_n = 2\rho_u b \langle n_{g0} \rangle_n + 2\rho_b b(1 - \langle n_{g0} \rangle_n) - d_p \langle n_P \rangle_n, \\ \partial_t \langle n_{g0} \rangle_n = -\tilde{\sigma}_{b0} \langle n_{g0} \rangle_n + \sigma_u (1 - \langle n_{g0} \rangle_n), \\ \partial_t \langle n_P n_{g0} \rangle_n = (\rho_u + \rho_b) b \langle n_{g0} \rangle_n + (\rho_u - \rho_b) b \langle n_{g0} n_{g1} \rangle_n - (d_p + \sigma_u + \tilde{\sigma}_{b0}) \langle n_P n_{g0} \rangle_n, \\ \partial_t \langle n_{g0} n_{g1} \rangle_n = 2\sigma_u \langle n_{g0} \rangle_n - 2(\sigma_u + \tilde{\sigma}_{b0}) \langle n_{g0} n_{g1} \rangle_n. \end{cases} \quad [53]$$

Results analogous to Eq. [53] can be derived for the post-replication period, and both can be compactly written as

$$\boxed{\begin{cases} \partial_t \langle n_P \rangle_n = 2\theta\rho_u b \langle n_{g0} \rangle_n + 2\theta\rho_b b(1 - \langle n_{g0} \rangle_n) - d_p \langle n_P \rangle_n, \\ \partial_t \langle n_{g0} \rangle_n = -\tilde{\sigma}_{b0} \langle n_{g0} \rangle_n + \sigma_u (1 - \langle n_{g0} \rangle_n), \\ \partial_t \langle n_P n_{g0} \rangle_n = [\rho_u + (2\theta - 1)\rho_b] b \langle n_{g0} \rangle_n + (2\theta - 1)(\rho_u - \rho_b) b \langle n_{g0} n_{g1} \rangle_n - (d_p + \sigma_u + \tilde{\sigma}_{b0}) \langle n_P n_{g0} \rangle_n, \\ \partial_t \langle n_{g0} n_{g1} \rangle_n = 2\sigma_u \langle n_{g0} \rangle_n - 2(\sigma_u + \tilde{\sigma}_{b0}) \langle n_{g0} n_{g1} \rangle_n, \end{cases}} \quad [54]$$

where $\theta = 1$ if it is in pre-replication period $t \in [0, t_r)$ and $\theta = 2$ if it is in post-replication period $t \in [t_r, t_d)$. Due to cell division, the moments of two successive generations are linked by

$$\begin{cases} \langle n_P \rangle_{n+1}|_{t=0} = \frac{1}{2} \langle n_P \rangle_n|_{t=t_d}, \\ \langle n_{g0} \rangle_{n+1}|_{t=0} = \langle n_{g0} \rangle_n|_{t=t_d}, \\ \langle n_P n_{g0} \rangle_{n+1}|_{t=0} = \frac{1}{2} \langle n_P n_{g0} \rangle_n|_{t=t_d}, \\ \langle n_{g0} n_{g1} \rangle_{n+1}|_{t=0} = \langle n_{g0} n_{g1} \rangle_n|_{t=t_d}. \end{cases} \quad [55]$$

Hence all the moments up to second order (of the linear system) for any cell cycle n and any cell age t can be obtained by solving together Eqs. [52], [54] and [55]. Substituting these moments in Eq. [54] leads to the effective rates of the linear system.

Step (iii). The next step is to find the time-average of the effective rates

$$\underline{\sigma}_{b0} = \underline{\sigma}_{b1} = \begin{cases} \sigma_b \frac{1}{t} \int_0^t \frac{\langle n_P n_{g0} \rangle_n|_\tau}{\Omega(t) \langle n_{g0} \rangle_n|_\tau} d\tau & t \in [0, t_r), \\ \sigma_b \frac{1}{t - t_r} \int_{t_r}^{t - t_r} \frac{\langle n_P n_{g0} \rangle_n|_\tau}{\Omega(t) \langle n_{g0} \rangle_n|_\tau} d\tau & t \in [t_r, t_d). \end{cases}$$

Given the model is defined piecewisely, one need to find the values of $\underline{\sigma}_{b0}$ for all the times t_r and t_d of all generations and the final time of interest. For example, if one is interested in distributions of mRNA and protein numbers at cell cycle n and cell age t, then all the values of $\underline{\sigma}_{b0}|_{i,j}$ for cell age $j \in \{t_r, t_d\}$ and cell cycle $i < n$ and $\underline{\sigma}_{b0}|_{n,t}$ need to be found as well.

Steps (iv) and (v). Next one finds the time-dependent solution of the master equation of the linear system (assuming constant parameter values) followed by replacing the bimolecular rates in the linear system by the time-averaged effective rates derived earlier. The time-dependent solution of the master equation of the linear system with constant parameters is essentially what we have done in previous sections, except that it needs minor modification to account for gene expression leakage. Specifically the generating function for mature mRNA is given by Eq. [15] (with k replaced by d) and G_A and G_B specified as

$$G_A(u,t) = \begin{cases} \left(\frac{\rho_\Delta u e^{-dt} - \underline{\sigma}_{b0}}{\rho_\Delta u - \underline{\sigma}_{b0}} \right)^{\frac{\sigma_u^-}{d}} \exp\left[\frac{\rho_b}{d} (1 - e^{-dt}) \right] & t \in [0, t_r), \\ \left(\frac{\rho_\Delta u e^{-dt} - \underline{\sigma}_{b0}}{\rho_\Delta u e^{-d(t-t_r)} - \underline{\sigma}_{b0}} \right)^{\frac{\sigma_u^-}{d}} \left(\frac{\rho_\Delta u e^{-d(t-t_r)} - \underline{\sigma}_{b0}}{\rho_\Delta u - \underline{\sigma}_{b0}} \right)^{\frac{\sigma_u^+}{d}} \exp\left[\frac{\rho_b}{d} (1 - e^{-dt}) \right] & t \in [t_r, t_d), \end{cases} \quad [56a]$$

$$G_B(u,t) = \begin{cases} 1 & t \in [0, t_r), \\ \left(\frac{\rho_\Delta u e^{-d(t-t_r)} - \underline{\sigma}_{b0}}{\rho_\Delta u - \underline{\sigma}_{b0}} \right)^{\frac{\sigma_u^+}{d}} \exp\left[\frac{\rho_b}{d} (1 - e^{-d(t-t_r)}) \right] & t \in [t_r, t_d), \end{cases} \quad [56b]$$

Similarly it can be shown that for the protein distributions, the generating function is given by Eq. [15] (with k replaced by $d_{\rm p}$) and $G_{\rm A}$ and $G_{\rm B}$ specified as

$$G_{\rm A}(w,t)=\begin{cases}\left(\frac{bw{\rm e}^{-d_{\rm p}t}-1}{bw-1}\right)^{\frac{\rho_{\rm b}}{d_{\rm p}}}\left[\frac{b(\rho_\Delta+\underline{\sigma}_{\rm b0})w{\rm e}^{-d_{\rm p}t}-\underline{\sigma}_{\rm b0}}{b(\rho_\Delta+\underline{\sigma}_{\rm b0})w-\underline{\sigma}_{\rm b0}}\right]^{\frac{\rho_\Delta\sigma_{\rm u}^-}{d_{\rm p}(\rho_\Delta+\underline{\sigma}_{\rm b0})}}, & t\in[0,t_{\rm r}),\\ \left(\frac{bw{\rm e}^{-d_{\rm p}t}-1}{bw{\rm e}^{-d_{\rm p}(t-t_{\rm r})}-1}\right)^{\frac{\rho_{\rm b}}{d_{\rm p}}}\left[\frac{b(\rho_\Delta+\underline{\sigma}_{\rm b0})w{\rm e}^{-d_{\rm p}t}-\underline{\sigma}_{\rm b0}}{b(\rho_\Delta+\underline{\sigma}_{\rm b0})w{\rm e}^{-d_{\rm p}(t-t_{\rm r})}-\underline{\sigma}_{\rm b0}}\right]^{\frac{\rho_\Delta\sigma_{\rm u}^-}{d_{\rm p}(\rho_\Delta+\underline{\sigma}_{\rm b0})}}\left(\frac{bw{\rm e}^{-d_{\rm p}(t-t_{\rm r})}-1}{bw-1}\right)^{\frac{\rho_{\rm b}}{d_{\rm p}}}\left[\frac{b(\rho_\Delta+\underline{\sigma}_{\rm b0})w{\rm e}^{-d_{\rm p}(t-t_{\rm r})}-\underline{\sigma}_{\rm b0}}{b(\rho_\Delta+\underline{\sigma}_{\rm b0})w-\underline{\sigma}_{\rm b0}}\right]^{\frac{\rho_\Delta\sigma_{\rm u}^+}{d_{\rm p}(\rho_\Delta+\underline{\sigma}_{\rm b0})}} & t\in[t_{\rm r},t_{\rm d}),\end{cases} \tag{57a}$$

$$G_{\rm B}(w,t)=\begin{cases}1 & t\in[0,t_{\rm r}),\\ \left(\frac{bw{\rm e}^{-d_{\rm p}(t-t_{\rm r})}-1}{bw-1}\right)^{\frac{\rho_{\rm b}}{d_{\rm p}}}\left[\frac{b(\rho_\Delta+\underline{\sigma}_{\rm b0})w{\rm e}^{-d_{\rm p}(t-t_{\rm r})}-\underline{\sigma}_{\rm b0}}{b(\rho_\Delta+\underline{\sigma}_{\rm b0})w-\underline{\sigma}_{\rm b0}}\right]^{\frac{\rho_\Delta\sigma_{\rm u}^+}{d_{\rm p}(\rho_\Delta+\underline{\sigma}_{\rm b0})}} & t\in[t_{\rm r},t_{\rm d}).\end{cases} \tag{57b}$$

The parameter ρ_Δ in Eqs. [56a], [56b], [57a] and [57b] is equal to $\rho_{\rm u}-\rho_{\rm b}$.

10. Including time-dependent transcription rate

Here we discusses how to extend the solution Eq. [10] to handle time-dependent transcription rate. First, we denote the second term of the right hand side of Eq. [10] as

$$g_{\rm tv}(u,\rho_{\rm u},t)=\left(\frac{\rho_{\rm u}{\rm e}^{-kt}u-\sigma_{\rm b}}{\rho_{\rm u}u-\sigma_{\rm b}}\right)^{\frac{\sigma_{\rm u}}{k}},$$

by setting α to 0. A general form of the transcription rate $\rho_{\rm u}(t)$ can be discretized into a piecewise constant function in time, i.e., $\{\rho_1,\rho_2,\cdots,\rho_n\}$ with Δt being the time interval. In view of Eq. [10] and assuming the initial condition $G(u,0)=1$ for brevity, the solution at time $T=(n-1)\Delta t+t$ ($t\leq\Delta t$) is then well approximated by

$$G(u,T)\approx\prod_{i=1}^{n-1}g_{\rm tv}(u{\rm e}^{-k((n-1-i)\Delta t+t)},\rho_i,\Delta t)\times g_{\rm tv}(u,\rho_n,t) \tag{58}$$

given sufficiently small Δt. The solution Eq. [58] can be used to assemble $G_{\rm A}$ and $G_{\rm B}$ in Eq. [2] in the main text.

11. Technical details about Fig 3

Specifically we chose the values of $\rho_{\rm u}$, d, $\sigma_{\rm u}^-$ and $\sigma_{\rm b}$ from 1051 genes of CAST allele data of mouse embryonic stem cells (Table S3 of Ref. (8)); since the estimation in this paper did not take into account dosage compensation, $\sigma_{\rm u}^-$ is equalled to the the rate of switching from OFF to ON reported in the paper. The rest of the parameters, α, $t_{\rm d}$, $t_{\rm r}$, $\sigma_{\rm u}^+$, were permuted over the 4 dimensional lattice constituted by $\{0,10^{-3}\}$, 600 to 1000 in steps of 100, $0.3t_{\rm d}$ to $0.8t_{\rm d}$ in steps of $0.1t_{\rm d}$ and $0.3\sigma_{\rm u}^-$ to $0.8\sigma_{\rm u}^-$ in steps of $0.1\sigma_{\rm u}^-$, respectively. The values of $t_{\rm d}$ are in a range centered on the value of 780 mins reported in (1), the values of $t_{\rm r}$ span the cell cycle duration, and the value of $\sigma_{\rm u}^+$ is a fraction of $\sigma_{\rm u}^-$ as required by dosage compensation. Note that $\alpha=0$ corresponds to the case of no growth-dependent transcription while $\alpha=10^{-3}$ (with $t_{\rm d}=780$ mins) corresponds to a cell with growth-dependent transcription and whose volume before cell division is approximately twice that at birth.

12. Technical details about Fig 5

Panel A. The kinetic parameters are: $\rho_{\rm u}=1.8$ min^{-1}, $\sigma_{\rm b}=0.9$ min^{-1}, $\sigma_{\rm u}^-=3\times10^{-3}$ min^{-1}, $\sigma_{\rm u}^+=0.71\sigma_{\rm u}^-$, $t_{\rm r}=800$ min, $t_{\rm d}=1560$ min, $d_{\rm p}=3\times10^{-4}$ min^{-1}, $k=10$ min^{-1}, $\lambda=1.2d$, $\alpha=0$. Specifically, the cell cycle duration $t_{\rm d}$ is selected to be 26 hours close to the data reported for NIH3T3 in the SI of (23), the gene replication occurs roughly in the middle of cell cycle. According to (23), the median of protein half life (∼17 hr) is close to the cell cycle, while the half life of mRNA is 1.7 hr (ratio =10), being reasonably within the range of experimental data. The transcription rate $\rho_{\rm u}$ and gene inactivation rate $\sigma_{\rm b}$ are close to those reported in (4). The gene activation rate $\sigma_{\rm u}^-$ is also within the experimental range, for example close to Per2_het gene in (9). The translation rate is selected to give a low protein burst size which constrains the protein numbers within a computationally friendly region.

Panel B. $d_{\rm p}$ is 10 times larger than that in Panel A, and d, λ are scaled (10 times larger compared to Panel A) accordingly. All the other parameters remain the same as Panel A.

References

1. Skinner SO, et al. (2016) Single-cell analysis of transcription kinetics across the cell cycle. *Elife* 5:e12175.
2. Zenklusen D, Larson DR, Singer RH (2008) Single-RNA counting reveals alternative modes of gene expression in yeast counting reveals alternative modes of gene expression in yeast. *Nat. Struct. Mol. Biol.* 15(12):1263.
3. Holstege FC, et al. (1998) Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95(5):717–728.
4. Ochiai H, Sugawara T, Sakuma T, Yamamoto T (2014) Stochastic promoter activation affects nanog expression variability in mouse embryonic stem cells. *Sci. Rep.* 4:7125.
5. Senecal A, et al. (2014) Transcription factors modulate c-Fos transcriptional bursts. *Cell. Rep.* 8(1):75–83.
6. Shyu AB, Greenberg ME, Belasco JG (1989) The c-Fos transcript is targeted for rapid decay by two distinct mrna degradation pathways. *Genes. Dev.* 3(1):60–72.
7. Halpern KB, et al. (2015) Bursty gene expression in the intact mammalian liver. *Mol. Cell.* 58(1):147–156.
8. Larsson AJ, et al. (2019) Genomic encoding of transcriptional burst kinetics. *Nature* 565(7738):251.
9. Suter DM, et al. (2011) Mammalian genes are transcribed with widely different bursting kinetics. *Science* 332(6028):472–474.
10. Sveiczer A, Csikasz-Nagy A, Gyorffy B, Tyson JJ, Novak B (2000) Modeling the fission yeast cell cycle: Quantized cycle times in wee1- cdc25δ mutant cells. *Proc. Natl Acad. Sci. USA* 97(14):7865–7870.
11. Baserga R (1985) *The biology of cell reproduction*. (Harvard University Press).
12. La Manno G, et al. (2018) RNA velocity of single cells. *Nature* 560(7719):494.
13. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S (2006) Stochastic mRNA synthesis in mammalian cells. *PLoS. Biol.* 4(10):e309.
14. Cappelletti D, Wiuf C (2016) Elimination of intermediate species in multiscale stochastic reaction networks. *Ann. Appl. Probab.* 26(5):2915–2958.
15. Berg OG (1978) A model for the statistical fluctuations of protein numbers in a microbial population. *J. Theor. Biol.* 71(4):587–603.
16. Yates CA, Ford MJ, Mort RL (2017) A multi-stage representation of cell proliferation as a markov process. *Bull. Math. Biol.* 79(12):2905–2928.
17. Soltani M, Singh A (2019) Moment analysis of linear time-varying dynamical systems with renewal transitions. *SIAM J. Control Optim.* 57(4):2660–2685.
18. Amir A (2014) Cell size regulation in bacteria. *Physical Review Letters* 112(20):208102.
19. Cadart C, et al. (2018) Size control in mammalian cells involves modulation of both growth rate and cell cycle duration. *Nat. Commun.* 9(1):3275.
20. Friedman N, Cai L, Xie XS (2006) Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Phys. Rev. Lett.* 97(16):168302.
21. Cao Z, Grima R (2018) Linear mapping approximation of gene regulatory networks with stochastic dynamics. *Nat. Commun* 9(1):3305.
22. Iyer-Biswas S, Hayot F, Jayaprakash C (2009) Stochasticity of gene products from transcriptional pulsing. *Phys. Rev. E* 79(3):031911.
23. Schwanhäusser B, et al. (2011) Global quantification of mammalian gene expression control. *Nature* 473(7347):337.

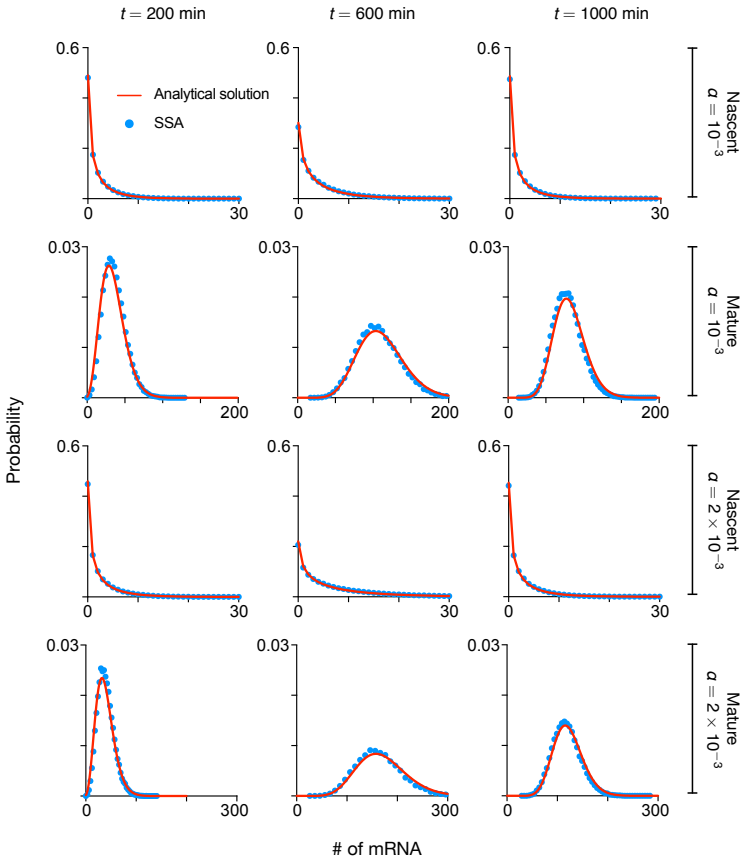

Fig. S2. Approximate analytic solution accurately predicts marginal distributions of nascent and mature mRNA for non-zero α in time. For details see the caption of Fig. 1 of main text. Note that all parameters except α are as specified in Fig. 1 of main text.

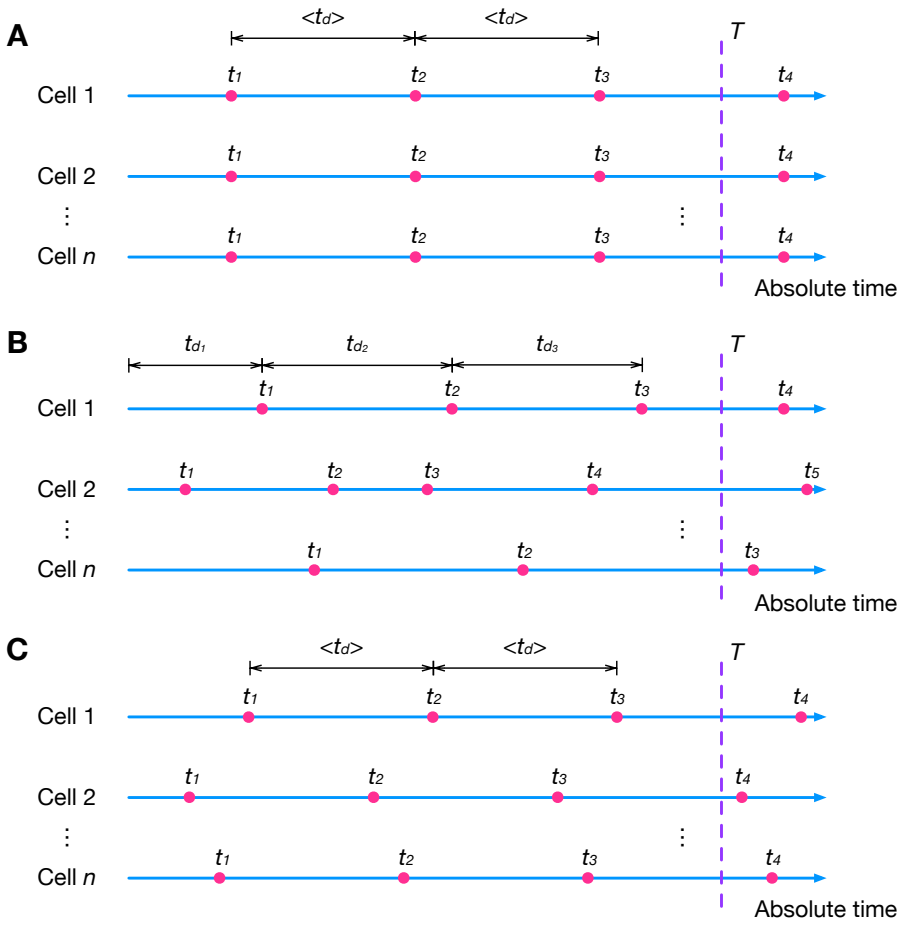

Fig. S3. Illustration of different single cell scenarios: (A) synchronized cells with constant cell cycle length; (B) unsynchronized cells with variable cell cycle length; (C) unsynchronized with constant cell cycle length. Note that t_i is the time at which generation i starts, T is the measurement time, and t_{d_i} is a random cell cycle duration sampled from a distribution with mean $\langle t_d \rangle$.

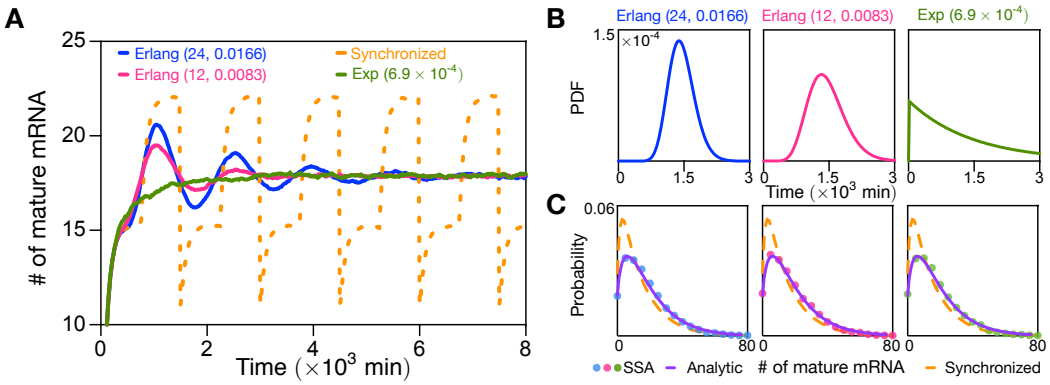

Fig. S4. Stochastic gene expression model with cell cycle time variability. (A) Plot of the mean mature mRNA number versus time calculated from the SSA for synchronized cells with constant cell cycle length and for cells with different cell cycle time distributions shown in (B). The Erlang distributions are given by Erlang(h_1,h_2) which is defined as $P(t_\mathrm{d}) = h_2^{h_1} t_\mathrm{d}^{h_1-1} e^{-h_2 t_\mathrm{d}} / (h_1-1)!$; all the Erlang distributions have the same mean of about 24 hours which is the average cell cycle time for mouse fibroblasts (see Fig. 3A in main text). Note that the mean mature mRNA numbers reach the same steady state value though different cell cycle time distributions are used. (C) Steady-state distributions of mature mRNA numbers calculated using the SSA for the cell cycle length distributions in (B) are found to be well described by the approximation Eq. [37]. The steady-state distribution of mature mRNA for synchronized cells with fixed cell cycle length at $t = 4600$ min is shown for comparison. The kinetic parameters used in the simulations are those of the gene 1M1C (see Table S1); the other parameters are $k = 1$ min^{-1}, $\sigma_\mathrm{u}^+ = 0.71\sigma_\mathrm{u}^-$ and $\alpha = 0$.

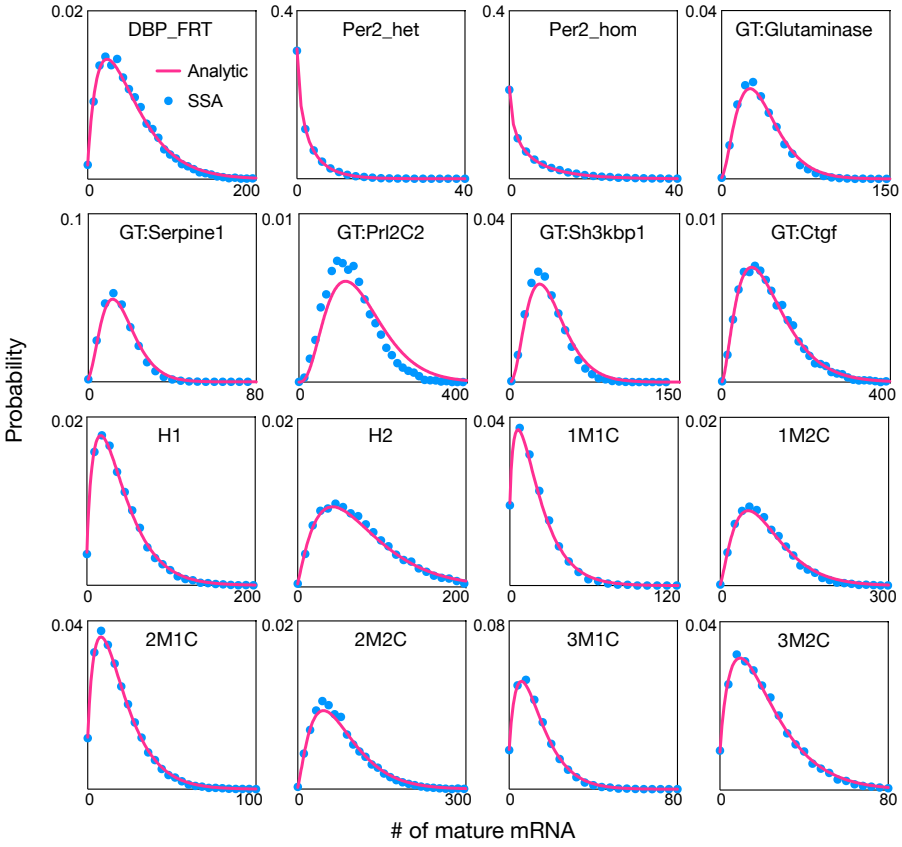

Fig. S5. Accuracy of the approximation given by Eq. [37]. Steady-state mature mRNA distributions from this approximation are compared with steady-state SSA predictions for 16 mouse genes whose kinetic data is reported in Table S1 and with cell cycle length distribution given by Fig. 3A in main text. The other parameters are: $k = 1$ min^{-1}, $\sigma_u^+ = 0.71\sigma_u^-$ and $\alpha = 0$.

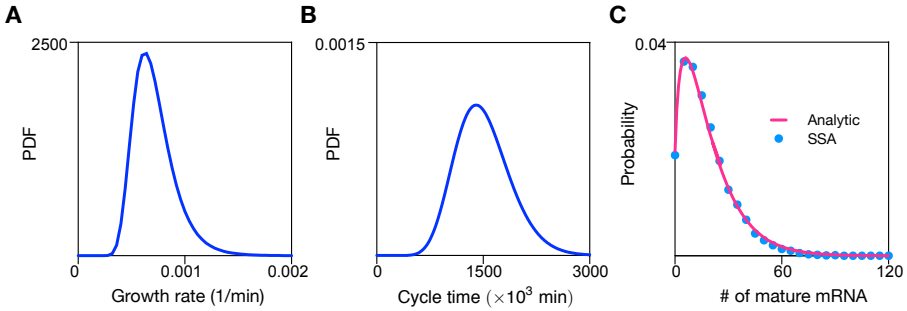

Fig. S6. Stochastic gene expression model with adder size-control mechanism. (A) shows that the distribution of growth rate of each generation is an inverse gamma distribution, i.e., InverseGamma(h_1, h_2) whose PDF is defined as $P(r_{\text{grow}}) = \left(\frac{h_2}{r_{\text{grow}}}\right)^{h_1} \exp\left(-\frac{h_2}{r_{\text{grow}}}\right) / r_{\text{grow}}\Gamma(h_1)$. Here $h_1 = 15$ and $h_2 = 0.01$. The growth rate is defined as the normalized incremental volume (=1) divided by cell cycle time, i.e., cell cycle time $\times$ growth rate $=$ incremental volume $= 1$. (B) shows that the corresponding distribution of cell cycle time is a gamma distribution, i.e., Gamma(h_1, h_2) defined by the PDF $P(t_{\text{d}}) = \left(\frac{t_{\text{d}}}{h_2}\right)^{h_1} \exp\left(-\frac{t_{\text{d}}}{h_2}\right) / t_{\text{d}}\Gamma(h_1)$. Here $h_1 = 15$ and $h_2 = 100$. (C) shows that the analytic solution Eq. [37] agrees well with the steady-state distribution computed using the SSA of our model with adder mechanism. The kinetic parameters for (C) is that of gene 1M1C reported in Table S1. The other parameters are: $k = 1$ min^{-1}, $\sigma_{\text{u}}^{+} = 0.71\sigma_{\text{u}}^{-}$ and $\alpha = 0$.

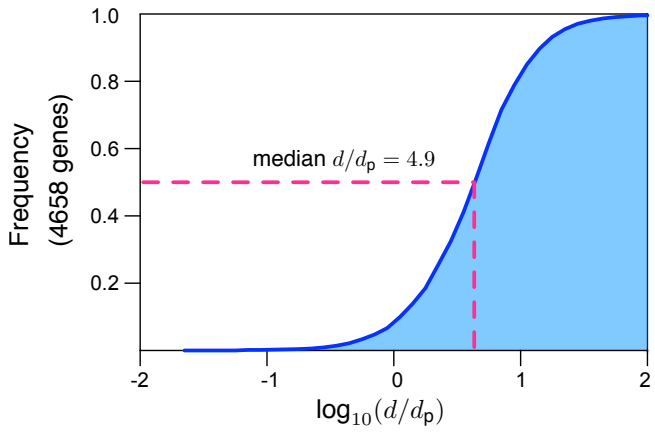

Fig. S7. Cumulative frequency distribution for the ratio of mRNA and protein decay rates for NIH3T3 mouse fibroblasts estimated in Ref. (23). The protein and mRNA half lives are from Supplemental Table 3 of Ref. (23). The ratio d/d_p is calculated by dividing the protein half life by the mRNA half life after excluding non-value pairs. Half of all proteins decay at least five times slower than mRNA.

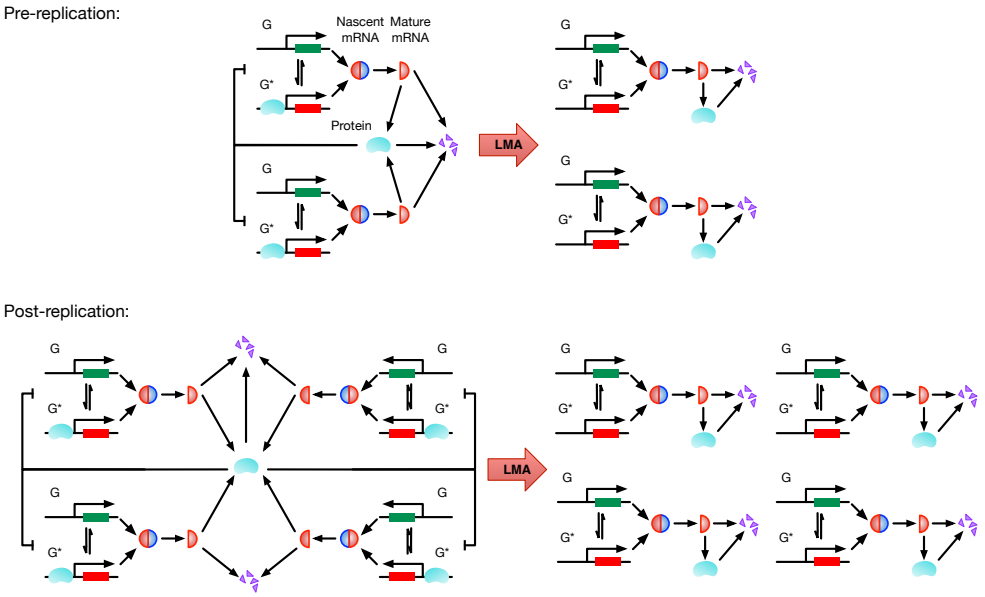

Fig. S8. Linear mapping approximation decouples the multiple-gene auto-regulatory system into several independent linear reaction systems.