PRX LIFE **3**, 043002 (2025)

Cyclo-Stationary Distributions of mRNA and Protein Counts for Random Cell Division Times

Syed Yunus Ali [1,*] Aditya Saran [1] Ashok Prasad [2,†] Abhyudai Singh [3,‡] and Dibyendu Das [1,4,§]

[1] *Department of Physics, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India*
[2] *School of Biomedical and Chemical Engineering, Colorado State University, Fort Collins, Colorado 80521, USA*
[3] *Department of Electrical and Computer Engineering, University of Delaware, Newark, Delaware 19716, USA*
[4] *Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany*

(Received 6 June 2025; accepted 12 September 2025; published 7 October 2025)

There is a long history of using experimental and computational approaches to study noise in single-cell levels of mRNA and proteins. The noise originates from myriad factors: intrinsic processes of gene expression, partitioning errors during division, and extrinsic effects, such as random cell-cycle times. Although theoretical methods are well developed to analytically understand the statistics of copy numbers for fixed or Erlang distributed cell cycle times, the general problem of random division times is still open. For any random (but uncorrelated) division time distribution, we present a method to address this challenging problem and obtain exact series representations of the copy number distributions in the cyclo-stationary state. We provide explicit cell age-specific and age-averaged results, and analyze the relative contribution to noise from intrinsic and extrinsic sources. Our analytical approach will aid the analysis of single-cell expression data and help in disentangling the impact of variability in division times.

DOI: 10.1103/ysj3-gkrr

I. INTRODUCTION

Advances in technologies of single-cell RNA sequencing and single-molecule fluorescence *in situ* hybridization to quantify mRNA levels and fluorescent proteomic imaging, mass cytometry, and mass spectrometry to quantify protein levels in individual cells have unmasked tremendous intercellular variability within isogenic populations over the last two decades [1–7]. Understanding the different sources of stochasticity that drive this variability is key to the analysis of single-cell transcriptomic and proteomics data and using stochasticity as a tool to infer complex regulatory networks [8–10]. Stochastic expression has been implicated in diverse emerging medical problems, such as cancer drug resistance [11–14], microbial persistence, and replication of human viruses [15,16]. Advancing analytical tools for understanding and modeling these inherent noise mechanisms can directly impact controlling cell-to-cell variation for therapeutic benefit.

The protein and mRNA copy numbers in cells are determined by a series of coupled stochastic chemical processes, leading to the above mentioned significant cell-to-cell variability. The transcriptional and translational noise arise due to multiple factors—genes switching between transcriptionally active and inactive states, rapid decay of short-lived mRNA leaving behind long-lived proteins making them appear in bursts, and other factors like RNA splicing and post-transcriptional regulation by micro-RNAs [17–23]. Theoretical studies of simple models (ignoring certain complexities) have obtained analytical moments and probability distributions of mRNA and protein count in the steady state [19,24–29], as well as for transient perturbations [30] and along the cell cycle [31].

In addition to intrinsic noise in gene expression specific to every gene, there are extrinsic factors affecting all genes. Two significant contributors to cell-to-cell variability in copy number, on which we focus in this paper, are noise incorporated through variable cell-cycle times [32–38] and random partitioning of copy numbers to daughter cells after cell division [39–42]. The noise associated with cell division has also been considered by various theoretical studies on copy number fluctuations [31,43–50], discussed in more detail below. In this paper we revisit this problem.

Within a cell cycle with finite division time, steady states are not attained for mRNA and protein counts. Yet after many successive cycles of division, a *cyclo-stationary* state is reached when time-independent distributions are attained for every cell "age"—the age is zero at birth and maximum just before division. The statistics of copy numbers in the cyclo-stationary state have been of central interest in the literature mentioned above.

What decides the instant of cell division still remains an intriguing question. Cell division has been argued to be triggered by a "timer," a "sizer," or an "adder" mechanism. In the "timer" scenario, it is assumed that cell division happens after fixed times T [51]. This is a common

---

[*] Contact author: syed_24@iitb.ac.in
[†] Contact author: ashok.prasad@colostate.edu
[‡] Contact author: absingh@udel.edu
[§] Contact author: dibyendudas@iitb.ac.in

assumption and has been extensively used in the theoretical literature [31,39,43,44,48,49,52]. Yet quite generally, cell division times are known to be random and also dependent on the cell size to ensure size homeostasis [32–36,53–58]. In the "sizer" view, division occurs when a cell size threshold is crossed [51,59], while in the "adder" view, it happens when the additional cell size growth from the initial birth size crosses a threshold [60–62].

In the literature, "threshold crossing" has been modeled in different ways. One approach splits the cell cycle into a sequence of N steps with exponentially distributed time intervals, leading to cell division and copy number partitioning at the end of the Nth step, which therefore acts like a threshold. The N steps are not actual cell cycle phases, but a model assumption. The total cell division time follows an Erlang or hypoexponential distribution in such cases [37,47,48,50,63–66]. In contrast, another view treats cell division as a *first-passage time problem* [67,68] (i.e., first threshold crossing of key regulatory protein). For an auto-catalytic growth process this leads to a Beta-exponential distribution of division times [69,70]. Some recent works have assumed stochastically fluctuating thresholds instead of a fixed threshold, and proposed some empirically relevant distributions of cell cycle times [71,72].

The variability of cell cycle times is thus a fact howsoever diverse may be the cause, yet theoretical works in the past have not treated this aspect in full generality. One study considered random cell division times and obtained exact moments of the copy numbers, but nevertheless assumed deterministic growth kinetics and deterministic partitioning [73]. A large body of analytical work has completely ignored the randomness in cell cycle times. The coefficient of variation was studied in [39] for fixed division times, comparing the relative role of gene expression noise versus binomial partitioning. Under the same assumption of fixed times, the generating functions for the distributions of mRNA and proteins in the cyclo-stationary state were derived for various models [31,43,44,48,49,74]. Specifically, constitutive bursty protein production [48], a two-stage model for mRNA synthesis with active and inactive transcription states [49], and a three-stage model for protein synthesis [31] were studied. The exact generating functions were derived for age-specific and age-averaged cases, and in the presence or absence of gene duplication. The desired probability distributions of copy number were then obtained through numerical derivatives of the generating functions.

Random cell cycle times were treated in another set of works, but for a special class of distributions: Erlang, mixed Erlang, and hypoexponential. These distributions permit an alternate representation of the cell cycle by a Markov chain of N stages each with an exponentially distributed lifetime. This technical simplicity, along with further assumption of steady state in each stage, led to exact moments and cumulants [47,63,64]. Under the same assumptions, the generating function for "age-averaged" protein distributions, for bursty synthesis without degradation, were derived in [48]. The complexities of volume-dependent gene expression, gene duplication, and dosage compensation were also incorporated in two of these papers [48,50]. The power spectra of the copy number autocorrelation function in the cyclo-stationary state were studied in [65].

Although the literature discussed above has made valuable contributions to our theoretical understanding of the problem and compared with some experimental data, it is evident that an analytical approach is lacking to treat arbitrary random division times which may arise in experiments. Even the only case widely studied, namely, Erlang (and hypoexponential), used an assumption of steady state for each of the constitutive stages of the process, and thereby was limited to obtain the "age-averaged" distributions of the copy number. Cell age-specific distributions are preferable as age-averaged ones may be derived from those, but the other way around is not possible. Given this, and the fact that other empirically relevant distributions have been reported [69,71,72], there is sufficient motivation to take a fresh look at the problem in this paper.

We adapt the framework of generating functions used for solving the cyclo-stationary distributions of copy numbers to the case of random division times, and first indicate why going beyond the case of fixed cell cycle times has remained technically challenging. We then develop a method to tackle this problem mathematically and evaluate the cyclo-stationary distributions as certain *analytically exact series* (Sec. II). Instead of enumerating the derivatives of the generating functions [31,48,49], the method now requires summing the relevant series directly. It works for any random (uncorrelated) cell cycle time distribution: the Erlang distribution can now be plugged in directly without splitting it into steps, just like Beta exponential, lognormal, or other empirically relevant distributions (Secs. III and VI D) as we show in the paper. We also show that for fixed division times, the cyclo-stationary mRNA distribution is exactly Poisson (Sec. IV). We provide explicit results for the basic gene expression model of mRNA synthesis and bursty protein production, with both having finite degradation rates [27,75] (Secs. IV and V). The cyclo-stationary distributions we obtain are age-specific: at cell birth, before division, or any time in between (Sec. VI A). We show that the age-averaged results may be obtained using suitable age frequency functions (Sec. VI B). Exact noise (CV^2) formulas and separate contributions to it from intrinsic and extrinsic sources are calculated; the part due to random cell cycle times depending on its distribution can be as high as 50% (Sec. VI C). The skewness shows a nonmonotonicity with mean division times, which is more pronounced with higher variability of cell cycle times, and thus can serve as a diagnostic of the variability itself (Sec. V). The impact of correlations in successive cycle times is studied in Sec. VI E.

II. THE METHOD TO TREAT RANDOM CELL CYCLE TIMES AND CYCLO-STATIONARY DISTRIBUTIONS AT CELL BIRTH AND BEFORE DIVISION

In this section, we present a general framework within which cyclo-stationary distributions of mRNAs or proteins may be analytically derived, for *random* cell cycle times. Let us denote the integer copy number of either mRNA or protein by $y(t)$. Later we will replace $y(t) = m(t)$ for mRNA and $y(t) = n(t)$ for protein. As shown schematically in Fig. 1, $y(t)$ grows from a value $y_{+,i-1}$ at the beginning of the ith cell

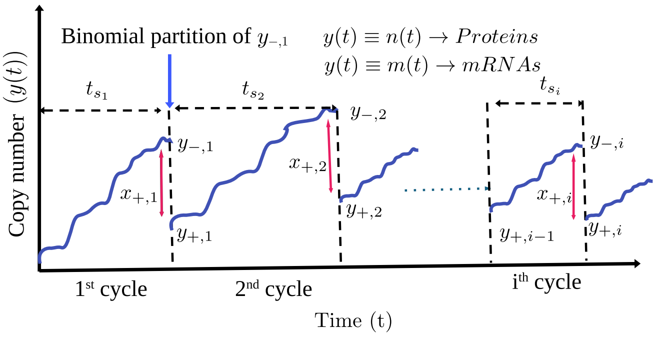

FIG. 1. Schematic diagram of time evolving copy number, of either mRNA $[m(t)]$ or proteins $[n(t)]$ in successive generations, interrupted by cell divisions when the number $y_{-,i}$ binomially divides to $x_{+,i}$ and $y_{+,i}$ in the two daughter cells. The durations $\{t_{s_i}\}$ of the cell cycles are random, and drawn from a probability distribution $g(t_s)$. After several generations, $i \gg 1$, the copy numbers attained cyclo-stationary distributions $P_+^{ss}(m_+)$ and $P_+^{ss}(n_+)$, which are studied in the text.

cycle to a value of $y_{-,i}$ at its end. Then as the cell divides, $y_{-,i}$ binomially partitions between two daughter cells with copy numbers $y_{+,i}$ and $x_{+,i} = y_{-,i} - y_{+,i}$ [76]. This process repeats over several generations through repeated cell divisions as depicted in Fig. 1, i.e., $i = 1, 2, \ldots$. All the details of gene expression within a cell cycle through transcription and translation are described by the probability distribution $p(y_{+,i} + x_{+,i}, t_{s_i} | y_{+,i-1})$. Here t_{s_i} denotes the duration of the ith cell cycle. Using the binomial distribution $B(N_0, p, x) = \binom{N_0}{x} p^x (1-p)^{N_0 - x}$ with $p = 1/2$, we may relate the distributions of copy numbers $P_+^i(y_{+,i}, t_{s_i})$ and $P_-^i(y_{-,i}, t_{s_i})$ just after and just before the ith cell division, respectively, to the distribution $P_+^{i-1}(y_{+,i-1}, t_{s_{i-1}})$ after the $(i-1)$th cycle as follows:

$$P_+^i(y_{+,i}, t_{s_i}) = \sum_{y_{+,i-1}} \sum_{x_{+,i}} B\left(y_{+,i} + x_{+,i}, \frac{1}{2}, x_{+,i}\right) \times p(y_{+,i} + x_{+,i}, t_{s_i} | y_{+,i-1}) P_+^{i-1}(y_{+,i-1}, t_{s_{i-1}}), \tag{1}$$

$$P_-^i(y_{-,i}, t_{s_i}) = \sum_{y_{+,i-1}} p(y_{-,i}, t_{s_i} | y_{+,i-1}) P_+^{i-1}(y_{+,i-1}, t_{s_{i-1}}). \tag{2}$$

Cyclo-stationary state is attained for $i \gg 1$, when the above two distributions approach steady (cycle independent) forms $P_+^{ss}(y_+)$ (for new born cells) and $P_-^{ss}(y_-)$ (for most mature cells before division). Our aim in this section is to solve for these.

To reemphasize, there are three sources of stochasticity: gene expression controlling the evolution of $y(t)$ within every cycle, the random binomial partitioning at every division step, and the random cell cycle times $\{t_{s_i}\}$. Let the distribution of these division times be $g(t_s)$, that the successive cell division time function $g_2(t_{s_i}, t_{s_{i-1}}) = g(t_{s_i}) g(t_{s_{i-1}})$. In such cases, starting from Eqs. (1) and (2), it may be shown [see the Supplemental Material (SM) Sec. IA [77]] that the cyclo-stationary copy number distributions $P_\pm^{ss}(y_\pm) = \int_0^\infty dt_s g(t_{s_i}) P_+^i(y_{+,i}, t_{s_i})$ satisfy

$$P_+^{ss}(y_+) = \sum_{y'_+} \sum_{x_+} \int_0^\infty dt_s g(t_s) B\left(y_+ + x_+, \frac{1}{2}, x_+\right) \times p(y_+ + x_+, t_s | y'_+) P_+^{ss}(y'_+), \tag{3}$$

$$P_-^{ss}(y_-) = \sum_{y'_+} \int_0^\infty dt_s g(t_s) p(y_-, t_s | y'_+, t_s) P_+^{ss}(y'_+). \tag{4}$$

Here subscripts i has been dropped by setting $t_{s_i} = t_s$ and $y_{\pm,i} = y_\pm$ to indicate the history independence. Next we need information of the model of gene expression. For a concrete study, we suppose the evolving copy number distribution $p(y, t | y'_+)$ has a generating function $F(q, t | y'_+) = \sum_{y=0}^\infty q^y p(y, t | y'_+)$ of the form

$$F(q, t | y'_+) = \mathscr{H}(q - 1, \gamma_y t) \times [1 + (q - 1) e^{-\gamma_y t}]^{y'_+}, \tag{5}$$

which is indeed the case for gene expression models of mRNA and protein, studied below [see Eqs. (20) and (25)]. The initial count y'_+-dependent factor is expected in any model with the degradation rate γ_y, while the function $\mathscr{H}(\cdot)$ is specific to the process (e.g., see [49]). Given Eq. (5), it may then be shown (SM Sec. IB [77]) that the generating functions $F_\pm(q) = \sum_{y_\pm = 0}^\infty q^{y_\pm} P_\pm^{ss}(y_\pm)$ of the cyclo-stationary distributions follow:

$$F_+(q) = \int_0^\infty dt_s g(t_s) \mathscr{H}\left(\frac{q-1}{2}, \gamma_y t_s\right) F_+\left(1 + \frac{q-1}{2} e^{-\gamma_y t_s}\right), \tag{6}$$

$$F_-(q) = F_+(2q - 1). \tag{7}$$

For fixed cell cycle time T, i.e., $g(t_s) = \delta(t_s - T)$, Eq. (6) solves exactly for the generating function (see SM Sec. IC)

$$F_+(q) = \prod_{k=1}^\infty \mathscr{H}\left(\frac{(q-1)}{2} \left(\frac{e^{-\gamma_y T}}{2}\right)^{k-1}, \gamma_y T\right), \tag{8}$$

reminiscent of earlier works [43,48,49]. Consequently, the steady-state probability is obtained through its derivatives: $P_+^{ss}(y_+) = \frac{1}{y_+!} \frac{\partial^{y_+}}{\partial q^{y_+}} F_+(q)|_{q=0}$. In contrast, for a random t_s, with a general distribution $g(t_s)$, the F_+ in the right-hand side of Eq. (6) on repeated iteration leads to nested integrals as shown in Eq. (21) of SM Sec. IC [77], which are generally intractable. Thus, the function $F_+(q)$ in general seems hard to find analytically.

We bring a useful insight to this challenging problem by noting that the mathematical structure of the problem at hand is very similar to the one arising in the process of synaptic vesicle fusion and release across chemical synapses on cyclic stimulation by action potentials [78–81]. The size of the ready-release pool of synaptic vesicles in the presynaptic neuron evolves as the copy number in Fig. 1. The arrival of an action potential at the presynaptic terminal causes sudden vesicle release and reduction of the pool size, just like the reduction in copy number by partitioning during cell division. The number of vesicles fused and released is binomially distributed, resembling the binomial partitioning of the copy number to daughter cells. The interspike intervals between the arrival of action potentials are like random cell cycle times. In the study of statistics of the quantal content of synaptic vesicle release, a similar equation as Eq. (6) arises [80], and we follow an idea found useful in that context.

Note that in Eq. (6), the argument q of the generating function F_+ on the left side, maps to another argument $q' = 1 + \frac{q-1}{2}e^{-\gamma_s t}$ on the right side. A fixed point of this map is $q' = q = 1$. Hence, a useful way to proceed analytically is to do an alternate series expansion of $F_+(q)$ about $q = 1$ as follows:

$$F_+(q) = \sum_{j=0}^{\infty} \frac{(q-1)^j}{j!} F_+^{(j)}(1). \tag{9}$$

Substituting the above on both sides of Eq. (6), and equating the coefficients of the power series in $(q-1)$, an exact recursion relation between the coefficients $F_+^{(j)}(1)$ of the following general form may be obtained:

$$F_+^{(k)}(1) = \sum_{j=0}^{k} c_{k,j} F_+^{(j)}(1). \tag{10}$$

Explicit versions of the above equation for mRNA and protein appear in Eqs. (21) and (26), respectively, and their detailed derivations are given in SM Secs. II B and III B [77]. The information of the distribution $g(t_s)$ gets embedded in the coefficients $c_{k,j}$; see, e.g., $\Psi_{k,j}$ and L_k below Eqs. (21) and (26), respectively. Thus, the key to solving the cyclo-stationary distributions of the copy number for random cell cycle times is to evaluate the coefficients $F_+^{(j)}(1)$ using Eq. (10). Using those, as shown in SM Sec. I D [77], the final distributions are given by

$$P_+^{ss}(y_+) = \sum_{k=y_+}^{\infty} \binom{k}{y_+} \frac{(-1)^{k-y_+}}{k!} F_+^{k}(1), \tag{11}$$

$$P_-^{ss}(y_-) = \sum_{k=y_-}^{\infty} \binom{k}{y_-} \frac{(-1)^{k-y_-}2^k}{k!} F_+^{(k)}(1). \tag{12}$$

Although we have so far discussed copy numbers of the new-born (y_+) and most mature cells before division (y_-), the calculations above may be extended to obtain the cyclo-stationary distribution $P^{ss}(y, \tau)$ at any intermediate cell age τ as shown in Sec. VI E. The age-averaged distributions $\overline{P^{ss}(y)}$ may further be obtained, given appropriate weights of cell age.

The first three cumulants associated with $P_+^{(ss)}(y_+)$ are given exactly in terms of the same coefficients $F_+^{(j)}(1)$ as (see SM Sec. I E [77])

$$\kappa_1 = \langle y_+ \rangle = F_+^{(1)}(1), \tag{13}$$

$$\kappa_2 = \langle (y_+ - \kappa_1)^2 \rangle = F_+^{(1)}(1) + F_+^{(2)}(1) - \kappa_1^2, \tag{14}$$

$$\begin{aligned} \kappa_3 = \langle (y_+ - \kappa_1)^3 \rangle &= 3F_+^{(2)}(1) + F_+^{(3)}(1) + F_+^{(1)}(1) \\ &\quad - 3\kappa_1\kappa_2 - \kappa_1^3. \end{aligned} \tag{15}$$

Note that below we will study the standard measures of fluctuations, namely, $CV^2 = \frac{\kappa_2}{\kappa_1^2}$ and skewness $= \frac{\kappa_3}{\kappa_2^{3/2}}$ for $P_+^{(ss)}(y_+)$ using Eqs. (13)–(15).

In summary, while for constant $t_s = T$ the exact generating function $F_+(q)$ [like in Eq. (8)] may be found and inverted (through derivatives) to obtain the $P_+^{ss}(y_+)$, here we have shown that for more realistic random t_s, the distribution $P_+^{ss}(y_+)$ may be found directly as series sums [Eqs. (11)], once the crucial coefficients $F_+^{(j)}(1)$ are evaluated through an exact recursion formula like Eq. (10).

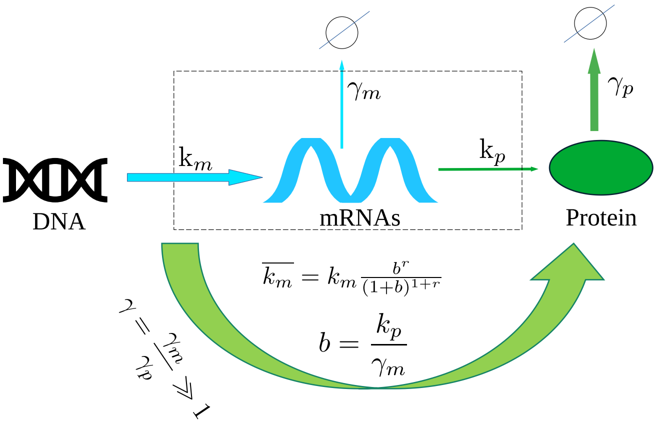

FIG. 2. A schematic figure showing transcriptional production of mRNAs from DNA at rate k_m and their translation to protein at rate k_p. They degrade at rates γ_m and γ_p, respectively. In the limit of slow protein decay, $\gamma = \gamma_m/\gamma_p \gg 1$, the protein production is bursty with an effective rate $\overline{k_m}$ with average burst size b.

III. THE MODELS AND DISTRIBUTIONS STUDIED

Gene expression: We consider the basic model of constitutive gene expression (see Fig. 2), in which mRNAs are produced ($m \to m+1$) at a rate k_m from the DNA template, and they degrade ($m \to m-1$) at a rate γ_m. Proteins are produced from mRNAs at a rate k_p, and they degrade ($n \to n-1$) at a rate γ_p. In this work we focus on cells which have very slow protein degradation compared to the mRNA (i.e., $\gamma = \frac{\gamma_m}{\gamma_p} \gg 1$). This scenario is common in yeast and bacteria, and one may treat the production of proteins effectively in bursts, ignoring the intermediate creation of mRNAs [20,27,75,82]. Hence in the protein production model we study, protein copy number $n \to n + r$ in a burst, with the increment r distributed geometrically as $\frac{b^r}{(1+b)^{1+r}}$. The mean burst size $b = \frac{k_p}{\gamma_m}$ and the rate of production is indicated in Fig. 2. The degradation rate γ_p is taken to be finite throughout this work.

Note that all our mathematical results for the mRNA in this paper can be used as it is for proteins that have nonbursty production ($n \to n+1$) and degradation, at constant rates.

Division time distributions: Although our results would apply to any $g(t_s)$, for concrete comparison, we study few distributions below with the same $\langle t_s \rangle = T$ but different CV^2. The base line case is constant $t_s = T$ with $g(t_s) = \delta(t_s - T)$ has $CV^2 = 0$. The opposite extreme is the exponential distribution $g(t_s) = \frac{1}{T}\exp(-t_s/T)$ with $\langle t_s \rangle = T$ and $CV^2 = 1$. The third one is the Erlang distribution

$$g(t_s) = \frac{\lambda^N t_s^{N-1} \exp(-\lambda t_s)}{(N-1)!} \tag{16}$$

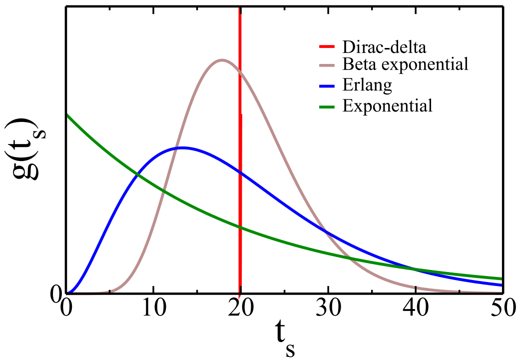

FIG. 3. The four cell cycle time distributions $g(t_s)$ versus t_s, used in the text: Dirac delta, Beta exponential with $X = 20$ and $n_0 = X/2$, Erlang with $N = 3$, and exponential. Other parameters are chosen so that they all have the same average $\langle t_s\rangle = T = 20$ min. Their CV^2 values are 0, 0.104, 0.33, and, 1, respectively.

with $\langle t_s\rangle = T = \frac{N}{\lambda}$ and $CV^2 = 1/N$. Note that the Erlang interpolates between the exponential ($N = 1$) and Dirac delta ($N \to \infty$). It is a popular distribution studied in many earlier works [37,48,64,73], but analyzed as an effective N step Markov process with exponentially distributed waiting times $1/\lambda$ [47,48]. We would however be using it directly (like [73]) with the given form of $g(t_s)$ in Eq. (16).

We would also study distributions of time t_s, where division arises due to some threshold crossing. The Beta exponential distribution

$$g(t_s) = \frac{\beta \exp(-\beta n_0 t_s)[1-\exp(-\beta t_s)]^{X-1-n_0}}{B[n_0, X-n_0]} \tag{17}$$

describes cell division times t_s when a *fixed threshold* X is crossed for the first time by either cell size or protein biomass in an autocatalytic growth process from an initial amount n_0 at rate β. It has been applied to cell division times in the bacteria *Caulobacter crescentus* [69,70,83]. Below we choose the initial quantity $n_0 = \frac{X}{2}$ (i.e., half of the threshold), and the values of threshold X and growth rate β, to have the mean cell cycle time $\langle t_s\rangle = T = \frac{1}{\beta}\prod_{s=n_0}^{X-1}(\frac{1}{s})$ as desired in Fig. 3. The CV^2 of the Beta exponential distribution works out to be lower than the Erlang in Fig. 3, for our chosen parameters.

The quantity which crosses the threshold X to give rise to the Beta exponential distribution [Eq. (17)], grows stochastically with its mean growing exponentially as $\sim\exp(\beta t)$. Instead, if a deterministic growth of cell size is considered as $\sim\exp(\beta t)$, but the growth rate β has a Gaussian variation $\mathcal{N}(\beta_0, \sigma_\beta^2)$ within a population of cells, then for $n_0 = X/2$ the effective division time distribution is

$$g(t_s) = \frac{\ln 2}{t_s^2\sqrt{2\pi\sigma_\beta^2}} \exp\left(-\frac{(\beta_0 t_s - \ln 2)^2}{2(t_s\sigma_\beta)^2}\right). \tag{18}$$

Recently, the possibility of a fluctuating (not fixed) threshold X has been considered [71,72], which is crossed by an exponentially growing cell size. With a Gaussian fluctuation of $\ln(X/n_0)$ having variance σ_X^2, the following distribution was obtained [72] and compared with experiments:

$$g(t_s) = \frac{\beta_0\sigma_X^2 + \sigma_\beta^2\, t_s \ln 2}{\sqrt{2\pi}\left(\sigma_X^2 + (\sigma_\beta\, t_s)^2\right)^{3/2}} \times \exp\left(-\frac{(\beta_0 t_s - \ln 2)^2}{2\left(\sigma_X^2 + (t_s\sigma_\beta)^2\right)}\right). \tag{19}$$

In Sec. IV we will compare the exact cyclo-stationary distributions for Eqs. (18) and (19) with different σ_X (threshold width).

IV. CYCLO-STATIONARY DISTRIBUTIONS OF mRNA

The master equation for the stochastic kinetics of mRNA number $m(t)$ within a cell cycle starting from m'_+ at the beginning of the cycle, is provided in SM Sec. II A [77], and is solved to obtain the generating function $F(q,t|m'_+) = \sum_{m=0}^{\infty} q^m p(m,t|m'_+)$ having a form the same as Eq. (5):

$$F(q,t|m'_+) = e^{\frac{k_m}{\gamma_m}[1-e^{-\gamma_m t}](q-1)}[1+(q-1)e^{-\gamma_m t}]^{m'_+}. \tag{20}$$

Thus, here the function $\mathscr{H} = e^{\frac{k_m}{\gamma_m}[1-e^{-\gamma_m t}](q-1)}$. Hence the generating functions for the cyclo-stationary distributions of mRNA $F_\pm(q) = \sum_{m_\pm=0}^{\infty} q^{m_\pm} P_\pm^{ss}(m_\pm)$ satisfy Eqs. (6) and (7) with $y_\pm = m_\pm$. We show in SM Sec. II B [77] that the series expansion of $F_+(q)$ about $q = 1$ [similar as Eq. (9)] leads to the following recursion relation among the coefficients $F_+^{(j)}(1)$ [of the form as Eq. (10)]:

$$F_+^{(k)}(1) = \frac{1}{2^k}\sum_{j=0}^{k}\binom{k}{j}\left(\frac{k_m}{\gamma_m}\right)^{k-j}\Psi_{k,j}F_+^{(j)}(1), \tag{21}$$

where $\Psi_{k,j} = \int_0^\infty dt_s g(t_s) e^{-j\gamma_m t_s}(1-e^{-\gamma_m t_s})^{k-j}$.

The recursive Eq. (21) can be exactly solved for $F_+^{(k)}(1)$ [see Eq. (39) in SM Sec. II B [77]]. Hence through Eqs. (11) and (12), cyclo-stationary distributions $P_\pm^{ss}(m_\pm)$ are formally solved. The expression for $F_+^{(k)}(1)$ is a bit cumbersome, involving sum over subsets of integers. For the following two $g(t_s)$, simpler closed forms are obtained:

(i) For $g(t_s) = \delta(t_s - T)$, it may be shown (see details in SM Sec. II C [77]) that $F_+^{(k)}(1) = d^k$ with $d = \frac{k_m}{\gamma_m}\frac{1-e^{-\gamma_m T}}{2-e^{-\gamma_m T}}$. The corresponding cyclo-stationary distributions [from Eqs. (11) and (12)] are Poisson:

$$P_\pm^{ss}(m_\pm) = \frac{(d^\pm)^{m_\pm}}{m_\pm!}\exp(-d^\pm) \tag{22}$$

with $d^+ = d$ and $d^- = 2d$.

(ii) For the exponential distribution $g(t_s) = 1/T \exp(-t_s/T)$, $F_+^{(k)}(1) = (\frac{k_m}{2})^k k!/\prod_i(\frac{1}{T} + \gamma_m i - \frac{1}{2^i T})$ (see SM Sec. II D [77]) and Eqs. (11) and (12) lead to

$$P_\pm^{ss}(m_\pm) = \sum_{k=m_\pm}^{\infty}\binom{k}{m_\pm}(-1)^{k-m_\pm}\frac{(k_m^\pm)^k}{\prod_{i=1}^{k}\left(\frac{1}{T}+\gamma_m i - \frac{1}{2^i T}\right)}, \tag{23}$$

with $k_m^+ = k_m/2$ and $k_m^- = k_m$.

As noted above, all the analytical expressions for mRNA in this section also apply to nonbursty protein kinetics.

Summarizing from above, the recursion formula (21) was solved by observing patterns in successive coefficients, and the result further simplified for the Dirac delta and exponential

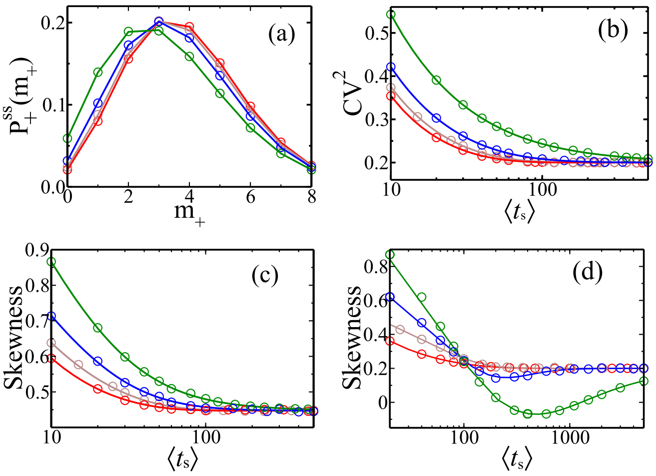

FIG. 4. (a) Cyclo-stationary distribution $P^{ss}_+(m_+)$ of mRNA for the four $g(t_s)$ shown in Fig. 3 (corresponding colors being the same). The corresponding variation of CV^2 (b) and skewness (c) with varying $\langle t_s\rangle$ are shown. Here $k_m = 0.5$, and $\gamma_m = 0.05$. (d) Skewness for $k_m = 0.5$, and $\gamma_m = 0.01$. The solid lines follow analytical formulas, while filled symbols represent KMC simulation data (using 10×10^7 histories).

$g(t_s)$ (SM Sec. II [77]). In general, for other distributions (including Erlang and Beta exponential), simple formulas are hard to derive. So for practical applications, one needs to numerically estimate $F^{(k)}_+(1)$ from Eq. (21) and obtain the $P^{ss}_\pm(m\pm)$ from analytical series in Eqs. (11) and (12). It needs some care as the coefficients $F^{(k)}_+(1)$ grow exponentially large with k and the series for $P^{ss}_\pm(m_\pm)$ typically converge very slowly. In Sec. IV of SM [77] we discuss our numerical protocol to estimate accurate $F^{(k)}_+(1)$ by storing logarithms of terms which are large numbers, and using very high precision in *Mathematica* for calculations. With those coefficients the series of $P^{ss}_\pm(m_\pm)$ converge with reasonable number of terms.

In Fig. 4(a), the $P^{ss}_+(m_+)$ obtained are shown (in solid lines) for the four distributions from Fig. 3. They are validated by independent data from Gillespie simulations [84] of these models (see SM Sec. IV [77]).

The differences of the curves in Fig. 4(a) reflect the differences in the "extrinsic" factor (namely, the division time statistics). Note that not just a few moments but the full distributions $g(t_s)$ contribute through $\Psi_{k,j}$ and $F^{(j)}_+(1)$ to different $P^{ss}_+(m_+)$.

Further insight on cyclo-stationary fluctuations comes from study of second- and third-order cumulants. Since the necessary quantities $F^{(1)}_+(1)$, $F^{(2)}_+(1)$, and $F^{(3)}_+(1)$ appearing in Eqs. (13)–(15) may be exactly solved in terms of $\Psi_{k,j}$ [see Eqs. (36), (37), and (38) in Sec. IIB of SM [77]], one may study the CV^2 and skewness for any $g(t_s)$. The explicit formula for

$$CV^2 = -1 + \frac{1-\frac{1}{2}\Psi_{1,1}}{\Psi_{1,0}}\left[\frac{2\gamma_m}{k_m} + \frac{\Psi_{2,1}}{1-\frac{1}{2^2}\Psi_{2,2}}\right] + \frac{\Psi_{1,0}^2\Psi_{2,0}}{\left(1-\frac{1}{2^2}\Psi_{2,2}\right)\left(1-\frac{1}{2}\Psi_{1,1}\right)^2}. \qquad (24)$$

In Fig. 4(b) we see that the CV^2 of $P^{ss}_+(m_+)$ vary monotonically with mean division times $\langle t_s\rangle$, and for all distributions approach the asymptotic value of $2\gamma_m/k_m$ [$= 0.2$ in Fig. 4(b)] for large $\langle t_s\rangle$. This follows immediately for the Dirac-delta distribution from Eq. (24), as $\Psi_{k,j} = e^{-j\gamma_m T}(1-e^{-\gamma_m T})^{k-j} \to 0$ for any $j \neq 0$, and $\to 1$ for $j = 0$, at large T. For other $g(t_s)$, the times t_s around the mean dominate the integral of $\Psi_{k,j}$ at large $\langle t_s\rangle$, implying similar asymptotic values. In Fig. 4(b) we observe that slower is the decrease of CV^2 of mRNA count, when higher is the CV^2 of $g(t_s)$ (see the hierarchy in Fig. 3), as slower are the corresponding approaches of $\Psi_{k,j}$ to the asymptotic values. The CV^2 can never be nonmonotonic with $\langle t_s\rangle$, since the contributing terms $1/F^{(1)}_+(1)$, $F^{(2)}_+(1)/[F^{(1)}_+(1)]^2$ add with the same sign and each monotonically decreases.

In Fig. 4(c) we see a similar asymptotic approach of skewness to a value $(2\gamma_m/k_m)^{1/2}$ ($= 0.447$ in the figure), for any $g(t_s)$. This follows from the formula of skewness in Eq. (57) in SM [77], based on the arguments given above that $\Psi_{k,j} \to 0$ for $j \neq 0$, and $\to 1$ for $j = 0$ at large $\langle t_s\rangle$ for any $g(t_s)$. A striking feature of the skewness formula is that it may exhibit nonmonotonic behavior if the degradation rate (i.e., γ_m) is sufficiently small. Observe that although curves in Fig. 4(c) are monotonic for $\gamma_m = 0.05$, they are nonmonotonic in Fig. 4(d) for smaller $\gamma_m = 0.01$ before asymptotically flattening. The reason is that, in contrast to CV^2, in the skewness formula, some of the terms dependent on $g(t_s)$ have a plus sign while some have a minus sign. Hence if the group of terms with the minus sign are relatively slower in attaining their asymptotic value in comparison to the group of terms with a positive sign, the value of skewness can get depressed and then again rise as a function of $\langle t_s\rangle$ [as in Fig. 4(d)]; this effect will be magnified if γ_m is small and CV^2 of $g(t_s)$ is high, both delaying the asymptotics. We will show below this interesting feature in skewness is also present for the skewness of protein counts, due to similar reasons.

We have also studied the analytical probability distributions $P^{ss}_-(m_-)$ of mRNA count just before cell division and compared with simulation data; see Fig. 1 in SM [77] for four different $g(t_s)$.

V. CYCLO-STATIONARY DISTRIBUTIONS OF PROTEIN

In Sec. III of SM [77], the master equation for the bursty production and degradation kinetics (see Fig. 2) of protein number $n(t)$, starting from n'_+ at the beginning of the cycle, is shown. The corresponding probability distribution $p(n,t|n'_+)$ has a generating function $F(q,t|n'_+) = \sum_{n=0}^\infty q^n\, p(n,t|n'_+)$ [27]:

$$F(q,t|n'_+) = \left(\frac{1-b(q-1)e^{-\gamma_p t_s}}{1-b(q-1)}\right)^a \times \left(1+(q-1)e^{-\gamma_p t_s}\right)^{n'_+}, \qquad (25)$$

where $a = k_m/\gamma_p$ (see SM Sec. III A [77]). The above equation has the same form as Eq. (5) with the function $\mathscr{H} = (\frac{1-b(q-1)e^{-\gamma_p t_s}}{1-b(q-1)})^a$. Hence Eqs. (6) and (7) are satisfied by the generating functions $F_\pm(q) = \sum_{n_\pm=0}^\infty q^{n_\pm} P^{ss}_\pm(n_\pm)$, for the cyclo-stationary distributions $P^{ss}_\pm$ of proteins. We show in SM Sec. III B [77] that the series expansion of $F_+(q)$ about $q = 1$ [see Eq. (9)] yields the following exact recursion relation for

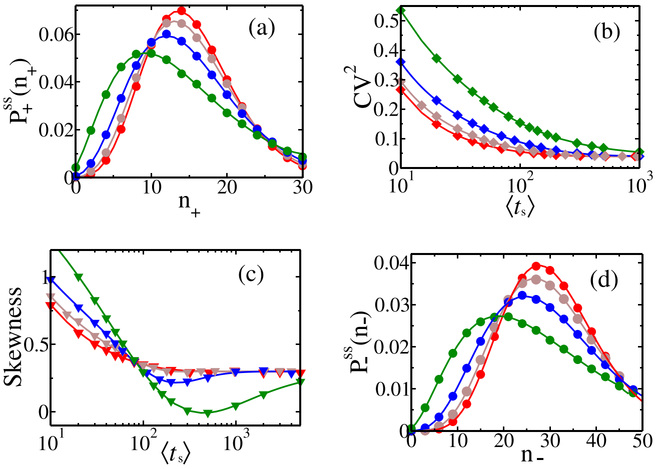

FIG. 5. (a) Cyclo-stationary distribution $P_+^{ss}(n_+)$ of protein count, for the four $g(t_s)$ shown in Fig. 3 (corresponding colors being the same). The corresponding variation of CV^2 (b) and skewness (c) with varying $\langle t_s \rangle$ are shown. (d) The distribution $P_-^{ss}(n_-)$ just before division. Here $k_m = 0.5$, $b = 2$, and $\gamma_p = 0.01$. The solid lines follow analytical formulas, while empty symbols represent KMC simulation data (using 5×10^7 histories).

$F_+^{(j)}(1)$:

$$F_+^{(k)}(1) = ak! \sum_{l=0}^{k} \sum_{j=0}^{k-l} (-1)^l \frac{b^{k-j}}{2^k} \times \frac{L_{l+j}(a+k-l-j-1)! F_+^{(j)}(1)}{(a-l)!l!j!(k-l-j)!}, \tag{26}$$

where $L_{l+j} = \int_0^\infty dt_s g(t_s) e^{-(l+j)\gamma_p t_s}$, a function of $(l+j)\gamma_p$, is the Laplace transform of the distribution $g(t_s)$.

Equation (26) is the key exact result—it is reducible to the form in Eq. (10) by combining terms. The coefficients $F_+^{(j)}(1)$ can be enumerated using Eq. (26) once the Laplace transform of the random cell cycle distribution $g(t_s)$ is known. But solving Eq. (26) in an analytical closed form is challenging. Hence, we used the numerical protocol (Sec. IV of SM [77]) to enumerate the $F_+^{(k)}(1)$ from Eq. (26). The coefficients were then used along with Borel summation method (see Sec. IV of SM [77]) for quicker convergence of the series in Eqs. (11) and (12) to obtain the cyclo-stationary distributions $P_\pm^{ss}(n_\pm)$. In Figs. 5(a) and 5(d), the protein distributions $P_+^{ss}(n_+)$ and $P_-^{ss}(n_-)$ (in solid lines) are well matched by simulation data (in symbols).

The exact $F_+^{(1)}(1)$, $F_+^{(2)}(1)$, and $F_+^{(3)}(1)$, are provided in Eqs. (63), (64), and (65) in SM Sec. III C [77] in terms of L_k. The exact CV^2 and sfor any $g(t_s)$ [using Eqs. (13)–(15)] for the protein count at cell birth, follow from those. For convenience of the user, we provide the following:

$$CV^2 = -1 + \frac{2-L_1}{ab(1-L_1)} + \frac{(2-L_1)}{a(4-L_2)(1-L_1)} \times \left[\frac{2-L_1}{1-L_1}[a(1-2L_1+L_2)+(1-L_2)] + 2a(L_1-L_2) \right]. \tag{27}$$

Like in the case of mRNA, the CV^2 of n_+ are hierarchical, with respect to the degree of fluctuations in $g(t_s)$ [Fig 5(b)], but they all approach a common asymptotic value. As all $L_j \to 0$ at large $\langle t_s \rangle$ for any $g(t_s)$, from Eq. (27) we conclude that $CV^2 \to (b+2)/ab$ (= 0.04 in the figure). Here too the additive terms $1/F_+^{(1)}(1)$, $F_+^{(2)}(1)/[F_+^{(1)}(1)]^2$ in CV^2 each monotonically decreases, implying CV^2 to be always monotonic with $\langle t_s \rangle$. In Sec. VI C, we analyze the contributions of intrinsic and extrinsic factors separately to CV^2 in Eq. (27).

Since the degradation rate γ_p of proteins are typically low, the terms in the skewness expression will take long time to reach their asymptotic value $2(1+b)/\sqrt{ab(2+b)}$ [= 0.3 in Fig. 5(c)] which follows by setting $L_j \to 0$ in Eq. (95) of SM [77]. Because of the relative difference of times of approach to the asymptotics of the groups of positive and negative terms [Eq. (95) of SM [77]), we have a nonmonotonic behavior in the skewness as a function of $\langle t_s \rangle$; see Fig. 5(c).

VI. VARIOUS EXTENSIONS OF THE ABOVE RESULTS

A. Cyclo-stationary distributions at any cell age τ

Although so far we have discussed the cyclo-stationary distribution of copy numbers in the new-born cells [$P_+^{ss}(y_+)$] and cells prior to division [$P_-^{ss}(y_-)$], the cyclo-stationary distribution $P^{ss}(y, \tau)$ of cells at any arbitrary "age" τ before the next cell division may be derived by using the $P_+^{ss}(y_+)$ as follows:

$$P^{ss}(y, \tau) = \sum_{y_+} P_+^{ss}(y_+) p(y, \tau | y_+). \tag{28}$$

Using its generating function $\tilde{G}(q, \tau) = \sum_y q^y P^{ss}(y, \tau) = \sum_k \frac{(q-1)^k}{k!} G_y^{(k)}(\tau)$, we derive (see SM Sec. V [77])

$$P^{ss}(y, \tau) = \sum_{k=y}^{\infty} \binom{k}{y} \frac{(-1)^{k-y}}{k!} G_y^{(k)}(\tau). \tag{29}$$

For the mRNAs (i.e., $y \equiv m$), with $F^{(k)}(1)$ from Eq. (21),

$$G_m^{(k)}(\tau) = \sum_{j=0}^{k} \binom{k}{j} \left(\frac{k_m}{\gamma_m} \right)^{k-j} F_+^{(j)}(1) e^{-j\gamma_m \tau} (1 - e^{-\gamma_m \tau})^{k-j}, \tag{30}$$

while for proteins (i.e., $y \equiv n$) with $F^{(k)}(1)$ from Eq. (26),

$$G_n^{(k)}(\tau) = ak! \sum_{l=0}^{k} \sum_{j=0}^{k-l} (-1)^l b^{k-j} e^{-(l+j)\gamma_p \tau} \times \frac{(a+k-l-j-1)! F_+^{(j)}(1)}{(a-l)!l!j!(k-l-j)!}. \tag{31}$$

See SM [77]) for details of the calculation. Thus, we have the analytical formulas for $P^{ss}(y, \tau)$ given by Eq. (29), along with Eqs. (30) and (31). Note that when $\tau \to 0$, $G_y^{(k)} \to F_+^{(k)}(1)$ and hence $P^{ss}(y, \tau) \to P_+^{ss}(y_+)$ (Eq. (11)) as expected. In Fig. 6 we show the plots of the theoretical $P^{ss}(m, \tau)$ and $P^{ss}(n, \tau)$ at three different ages, using the Erlang distribution for the $g(t_s)$. We performed independent single-lineage simulations and sampled 5×10^5 copy numbers at particular cell

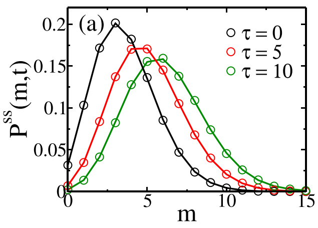

FIG. 6. Cyclo-stationary distributions $P^{ss}(m,\tau)$ of mRNA (m) and $P^{ss}(n,\tau)$ of protein (n), at three cell ages ($\tau = 0$, 5, and 10 min). The solid lines are from analytical theory [Eq. (29)] and circular symbols are obtained by averaging counts of multiple single lineage cells of specific ages. The $g(t_s)$ follows the Erlang distribution from Fig. 3. The parameters for mRNA are $k_m = 0.5$, $\gamma_m = 0.05$, and protein are $k_m = 0.5$, $b = 2$, and $\gamma_p = 0.01$.

ages τ in the cyclo-stationary state, and used that to obtain numerical $P^{ss}(m,\tau)$ and $P^{ss}(n,\tau)$ (see circular symbols), which match the theoretical curves in Fig. 6.

B. Age-averaged cyclo-stationary distributions

To obtain a distribution averaged over cell age, one would require a suitable cell-age distribution. For "constant" cell division times, the single-lineage age distribution is uniform over $\tau \in [0, T]$ [47,48,63], while for a population of exponentially growing cells it is $\phi(\tau) = \frac{\ln 2}{T}\, 2^{1-(\tau/T)}$ for $\tau \in [0, T]$ and has been used in various works [31,43,44,48,85].

However, we need the age distributions for random cell division times with a given $g(t_s)$. Recently, the general time-dependent cell age distribution for exponentially growing populations has been studied by Jafarpour *et al.* [86]. We are interested in the cyclo-stationary (or "steady") state limit and binary symmetric division, in which case the age distribution of the exponential population is $\phi(\tau) = 2\nu_m e^{-\nu_m\tau}\int_\tau^\infty g(t_s)dt_s$ with ν_m given implicitly by $2\int_0^\infty e^{-\nu_m t_s} g(t_s)dt_s = 1$, while the single-lineage age distribution is $\psi(\tau) = \int_\tau^\infty g(t_s)dt_s/\langle t_s\rangle$ [85–87]. For example, for the Erlang distribution [Eq. (16)], $\nu_m = \lambda(2^{\frac{1}{N}} - 1)$ and $\phi(\tau) = 2\nu_m e^{-\nu_m\tau}\Gamma(N, \tau\lambda)/\Gamma(N)$ and $\psi(\tau) = \Gamma(N, \tau\lambda)/(\Gamma(N)\langle t_s\rangle)$. These theoretical formulas are compared in Fig. 7(a) to the simulation data of age distributions (in symbols) of a population and of single lineage: the population distribution decays faster compared to the single lineage due to the additional exponential factor.

In Fig. 7(b) we first plot (see brown curve) the lineage "age-averaged" protein distribution $\overline{P_l^{ss}(n)} = \int_0^\infty d\tau\,\psi(\tau)P^{ss}(n,\tau)$, by averaging the lineage "age-specific" protein distribution $P^{ss}(n,\tau)$ [Eq. (29)] over the lineage age distribution $\psi(\tau)$; this is supported by lineage simulations (brown symbols). Randomly sampled protein counts from cells of any age of a simulated exponentially growing population (in the steady state) give the age-independent protein distribution $\overline{P_{\rm pop}^{ss}(n)}$ of the population [see green symbols in Fig. 7(b)]. As we do not have a theoretical age-specific protein distribution in a population, we give an approximate $\overline{P^{ss}(n)} = \int_0^\infty d\tau\,\phi(\tau)P^{ss}(n,\tau)$ (see solid black curve Fig. 7(b)), by averaging the lineage $P^{ss}(n,\tau)$ [Eq. (29)] over the population age distribution $\phi(\tau)$; the curve $\overline{P^{ss}(n)}$ is closer but different from $\overline{P_{\rm pop}^{ss}(n)}$ as expected.

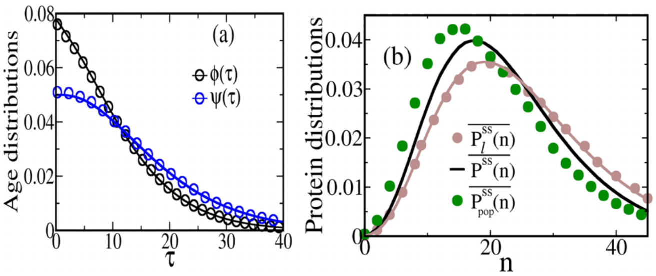

FIG. 7. (a) Steady-state age distributions $\phi(\tau)$ for a population (black) and $\psi(\tau)$ for single lineage (blue) against τ. Theoretical curves are solid lines, while simulation data are in symbols. (b) Age-averaged protein distributions, $\overline{P_l^{ss}(n)}$ (brown curve) for single lineage (theoretical, brown curve; simulation, brown symbols), theoretical $\overline{P^{ss}(n)}$ (black curve), and simulated $\overline{P_{\rm pop}^{ss}(n)}$ for an exponentially growing population (green circles). The $g(t_s)$ used is Erlang distributed with $N = 3$ [Eq. (16) and Fig. 3], and the model parameters for protein expression are $k_m = 0.5$, $b = 2$, and $\gamma_p = 0.01$.

C. Comparing the contributions to noise in protein copy number from intrinsic and extrinsic sources

In the expression of CV^2 for proteins at cell birth [Eq. (27)] the noise from gene expression, binomial partitioning, as well as random cell division times all contribute. Note that mean $\langle n_+\rangle = ab(1 - L_1)/(2 - L_1)$ is independent of whether the processes are deterministic or stochastic. If we wish to exclude the contribution of the partitioning noise, then the Binomial distribution may be replaced by a delta function [in Eq. (3)] such that $n_+ = n_-/2$ in every cycle. A separate calculation done for this case in SM Sec. VI [77] shows that $(q - 1)/2$ in Eq. (6) for the generating function is replaced by $(\sqrt{q} - 1)$, and subsequently yields

$$CV^2_{GE+DT} = CV^2 - \frac{2(2 - L_1)}{ab(4 - L_2)(1 - L_1)}, \tag{32}$$

where CV^2 is the total noise [from Eq. (27)]. The CV^2_{GE+DT} has a contribution from gene expression (GE) and division time (DT) randomness. Thus, the noise from binomial partitioning (BP) is

$$CV^2_{BP} = \frac{2(2 - L_1)}{ab(4 - L_2)(1 - L_1)}. \tag{33}$$

Note that CV^2_{BP} is dependent on $g(t_s)$, although mildly; see Table I for the different cases.

If in addition to removal of partitioning noise (i.e., setting $n_+ = n_-/2$), the gene expression is also deterministic, then the protein number at any time $n = n'_+ e^{-\gamma_p t} + ab(1 - e^{-\gamma_p t})$ (with initial count n'_+). In this case, the argument of $F_+(\cdot)$ on

TABLE I. Noise in n_+, for $\langle t_s\rangle = 20$ min, $a = 50$, $b = 2$.

$g(t_s)$	CV^2	CV^2_{BP}	CV^2_{DT}	CV^2_{GE}
Dirac delta	0.149	0.039	0	0.110
Beta exponential	0.174	0.039	0.025	0.110
Erlang ($N = 3$)	0.229	0.040	0.078	0.111
Exponential	0.372	0.042	0.217	0.113

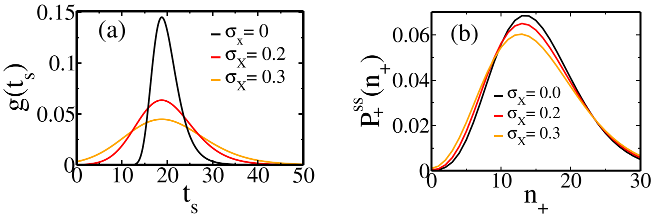

FIG. 8. (a) Plots of different $g(t_s)$ following Eq. (19), with various values of σ_X as shown. We have $\beta_0 = 0.035\,\text{min}^{-1}$, $\sigma_\beta = 0.0052\,\text{min}^{-1}$ for all of them. (b) The corresponding cyclo-stationary distributions $P_+^{ss}(n_+)$ [following the same colors as in (a)] are shown. The parameters for protein gene expression are $b = 2$, $k_m = 0.3$, and $\gamma_p = 0.01$.

the right-hand side of Eq. (6) is replaced by $(\sqrt{q})^{e^{-\gamma_p t_s}}$ (see SM Sec. VI [77]), and the calculation gives

$$CV_{DT}^2 = -1 + \frac{(2 - L_1)}{(4 - L_2)(1 - L_1)} \times \left[\frac{2 - L_1}{1 - L_1}(1 - 2L_1 + L_2) + 2(L_1 - L_2)\right]. \quad (34)$$

The CV_{DT}^2 is purely the extrinsic noise contribution coming from random variations in cell division times; hence if $g = \delta(t_s - T)$, then $L_2 = L_1^2$, and hence $CV_{DT}^2 = 0$ in Eq. (34) as expected. Also see Table I for other cases. The CV^2 in Eq. (27) is in excess of CV_{DT}^2 [Eq. (34)] by two terms. Comparing with Eq. (32), we have noise contribution from gene expression to total CV^2 as

$$CV_{GE}^2 = \frac{(2 - L_1)(2 - L_2)}{ab(1 - L_1)(4 - L_2)} + \frac{(2 - L_1)^2(1 - L_2)}{a(1 - L_1)^2(4 - L_2)}. \quad (35)$$

Note CV_{GE}^2 is dependent on $g(t_s)$, although mildly; see Table I for the different cases.

Finally percentage contributions of CV_{BP}^2, CV_{DT}^2, and CV_{GE}^2 to the total CV^2 depends on $g(t_s)$. From Table I, for the Beta exponential, they are, respectively, 23%, 14%, and 63%, for the Erlang, they are respectively, 18%, 34%, and 48%, and for the exponential they are respectively, 12%, 58%, and 30%. Thus, as division time noise rises, the relatively higher contribution of gene expression noise is offset by it. This qualitative trend is expected to remain the same, even if the percentage numbers presented here change with the values of a and b.

Here we have presented the results for the protein, but following the same type of calculations, one may derive corresponding results for the mRNA and compare with the full CV^2 from Eq. (24).

D. Division time distributions due to fluctuating threshold, and effect on cyclo-stationary protein statistics

In Sec. III we mentioned division time distributions arising from threshold crossing scenarios: Eq. (18) due to growth rate heterogeneity but fixed threshold, and Eq. (19) due to both growth rate heterogeneity and threshold fluctuations. Partly to demonstrate the applicability of our results to these cases too, and partly to see the effect of threshold fluctuations, we study them here.

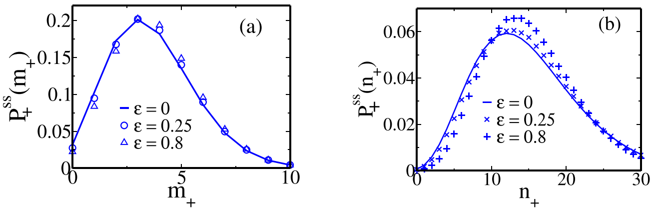

FIG. 9. Effect of correlation in cell cycle times on the cyclo-stationary distributions of (a) mRNA and (b) protein. The analytical predictions in the uncorrelated case are in solid lines. The symbols are from KMC simulations with various degrees of correlation indicated by the parameter ε.

In Fig. 8(a) for chosen $\beta_0 = 0.035\,\text{min}^{-1}$, $\sigma_\beta = 0.0052\,\text{min}^{-1}$ (with mean division time ≈20 min), three $g(t_s)$ distributions are shown with different widths σ_X of the size threshold in Eq. (19); parameters are in the ballpark of Ref. [72]. The corresponding cyclo-stationary protein distributions $P_+^{ss}(n_+)$ in new-born cells for the three cases are shown in Fig. 8(b). The copy number distributions broadens relatively slowly as threshold fluctuations (σ_X) rise and $g(t_s)$ quickly becomes broader.

E. Effect of correlations in cell cycle times

For deriving the basic recursive equations involving the cyclo-stationary distributions [Eqs. (3) and (4)] we had assumed that the random cell cycle times are uncorrelated, i.e., $g_2(t_{s_i}, t_{s_{i-1}}) = g(t_{s_i})g(t_{s_{i-1}})$. Given the complexities of cell-cycle control, this assumption may have to be relaxed. For example, recent work on specific human cell lines shows random cell-cycle times that are not correlated between the mother and daughter cells [88]. Interestingly, data showed modest correlations between the cell-cycle times of daughter cells, but this correlation is lost between cousin pair of cells [88]. Here we study through kinetic Monte Carlo simulations the effect of correlated cell cycle times. Correlation is achieved by choosing the time of the ith ($i > 1$) cycle t_{s_i} in terms of the $(i-1)$th cycle time $t_{s_{i-1}}$ in the following way:

$$t_{s_i} = (1 - \varepsilon) \times \text{Erl} + \varepsilon \times \text{t}_{\text{s}_{i-1}}, \quad (36)$$

where Erl is a random time chosen for the ith cycle, following the Erlang distribution in Fig. 3. For the first cycle $t_{s_1} = \text{Erl}$. By construction, the mean of the correlated t_{s_i} distribution remains the same as the Erlang distribution. Note that ε is the correlation coefficient between successive generation times, and by tuning it, we may vary the degree of correlation. In Figs. 9(a) and 9(b), we show the corresponding cyclo-stationary distributions of mRNA and protein. With a rise of ε, both distributions (in symbols) show some deviation (although not significant) from the uncorrelated analytical distributions (solid lines) that we have derived in this paper. The departure is more pronounced for protein than mRNA.

A natural way that correlations arise in generation times is through cell-size controlled division. For example, within the "adder" mechanism, mother-daughter generation times are known to have negative correlations [61,89]. Here we test how our theory of independent generation times compares with simulations of an adder model with correlated cell division

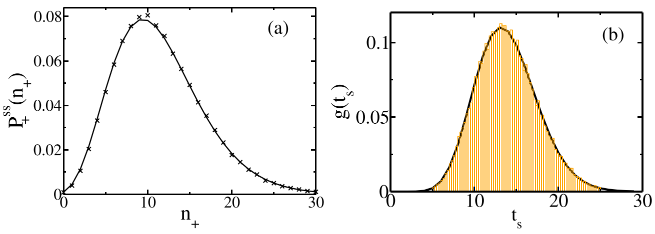

FIG. 10. (a) Cyclo-stationary distribution of protein count at birth, for correlated division times arising from a cell size-controlled "adder" model (× symbols), and for independent division times from our theory (solid line) with $g(t_s)$ in Eq. (37). The parameters for the adder model are $\alpha = 10$, $\bar{\Delta} = 10$, $\beta = 0.05$, while for the protein expression are $b = 2$, $k_m = 0.5$, and $\gamma_p = 0.01$. (b) Histogram of numerical $g(t_s)$ obtained from randomized generation times of the "adder" model, and the theoretical $g(t_s)$ from Eq. (37) (solid line).

times. In the simulated model, in each generation, the initial cell size s_b of a new-born cell grows exponentially (at rate β) by adding a random amount Δ drawn from a Gamma distribution $P_d(\Delta) = \frac{\alpha^\alpha}{\Gamma(\alpha)\bar{\Delta}}(\frac{\Delta}{\bar{\Delta}})^{\alpha-1}\exp(-\frac{\Delta\alpha}{\bar{\Delta}})$ with mean $\bar{\Delta}$, such that generation time t_s at division is given by $s_b + \Delta = s_b\exp(\beta t_s)$ [62]; the intial size of the next generation s_b' is obtained by equal partitioning of $s_b + \Delta$. We numerically found the mother-daughter generation time correlation coefficient to be -0.25 (as expected [89]). We do a parallel simulation of a bursty protein production (model in Sec. III) and binomial partitioning, following the sequence of correlated generation times provided by the adder cell-size model; the distribution of protein count n_+ at birth thus obtained in the cyclo-stationary state is shown with symbols in Fig. 10(a).

Next we take the correlated generation time sequence from the above adder model simulations, and randomize those and obtain a numerical $g(t_s)$ [see the histogram in Fig. 10(b)]. Corresponding to these data, a theoretical $g(t_s) = \int_0^\infty P_t(t_s|s_b)Q(s_b)ds_b$ with $P_t(t_s|s_b) = P_d(\Delta)\frac{d(\Delta)}{dt_s}|_{\Delta=s_b(e^{\beta t_s}-1)}$. The distribution $Q(s_b)$ of s_b is hard to derive exactly, so we approximate it to be a Gamma distribution like $P_d(\Delta)$, but satisfying the known constraints $\langle s_b\rangle = \bar{\Delta}$ and $CV_{s_b}^2 = CV_\Delta^2/3$, for adder models [89]. This gives after a few simple steps

$$g(t_s) = \frac{\Gamma(4\alpha)3^{3\alpha}\beta e^{\beta t_s}}{\Gamma(\alpha)\Gamma(3\alpha)}\frac{(e^{\beta t_s}-1)^{\alpha-1}}{(e^{\beta t_s}+2)^{4\alpha}}, \tag{37}$$

which [shown by the solid line in Fig. 10(b)] matches quite well the numerical $g(t_s)$. Using the above $g(t_s)$ [Eq. (37)] in Eqs. (26) and (11), we get our theoretical $P_+^{ss}(n_+)$ [solid line in Fig. 10(a)]. We observe that our theory for independent generation times gives quite close results to the simulated $P_+^{ss}(n_+)$ for the correlated generation times obtained through cell size control. Although surprising, this is consistent with the absence of significant deviation in Fig. 9(b) for correlation coefficient $\varepsilon = 0.25$ (same value) in the model of Eq. (36).

VII. CONCLUDING DISCUSSION

Analysis of single-cell transcriptomic and proteomic data requires an understanding of stochastic cellular processes that influence the variability of copy numbers of mRNA and proteins from cell to cell. Years of theoretical work on various models of gene expression, along with additional complexities of partitioning noise during division, dynamic cell growth, and gene duplication, have already enriched the means for analyzing the data. The key contribution of this paper is to present a method to incorporate the non-negligible aspect of noise in cell cycle times, within the evolving broader picture.

We have studied theoretically copy number statistics in cells obtained after many cycles of cell division, each involving binomial partitioning of copy numbers, when a cyclo-stationary condition has been attained. For any random (but uncorrelated) cell division times, we have presented a method to obtain exact series representations of the distributions of copy numbers. This is a significant theoretical advancement, as analytical solutions were known only for deterministic division times and a specific type of random time distribution (namely, the Erlang). Moreover treatments of the Erlang distribution relied on steady-state assumption of every cell cycle stage, and hence could only get cell age-averaged statistics. Our method here makes no such assumption and gives exact distributions at any cell age, and also leaves the scope of obtaining the age-averaged result using the suitable cell age distribution [85]. We have demonstrated that mild correlations in division times of successive cycles would not cause strong departure from the analytical results to have obtained for the uncorrelated case. Extension of the current work by developing efficient summation methods of the formal series we present would be a fruitful direction to pursue in future.

The following results follow as consequences of random cell cycle times. Along with the general series forms of cyclo-stationary distributions of mRNA and proteins for any random cell cycle distribution, we specifically showed that for fixed cycle times the mRNA distribution is Poisson. Thus, the Poisson form, well known for constitutive transcription, stays preserved after accounting for partitioning noise as long as the cell division time is a constant. Hence, the departure of the mRNA distribution from Poisson is a signature of the effect of cell cycle time variability (in cases where promoter regulation is absent). Next, the skewness has nonmonotonic variation with mean division time, if fluctuations in division time are strong, and degradation rate of copy numbers are weak. This behavior of skewness may serve as a signature of high variability of cell cycle times. In comparison, we found that CV^2 never exhibits nonmonotonicity, and it is high in consonance with higher fluctuations in cell division times. Splitting the net CV^2 into contributions from intrinsic and extrinsic sources, we showed in Table I that with rising fluctuations in division times, its contribution may be substantial (rising to around 50%) in comparison to noise due to gene expression and binomial partitioning.

Although the models of gene expression that we analyze are the basic ones for constitutive production, they demonstrate the new method with clarity. Further complexity in models of gene expression have to be introduced through the corresponding generating functions [analogous to Eq. (5)].

Although cell-size control was not elaborately treated in the paper, in Sec. VI E we also simulated an adder model that leads to mother-daughter generation times having negative correlations, and we obtained a cyclo-stationary protein distribution influenced by such correlated cell cycle times.

We found a very small deviation of this distribution from the theoretical distribution for independent generation times. To obtain analytically exact expressions however, we require further development of the framework presented here to theoretically incorporate intergenerational correlation in division times. Protein distributions in populations of exponentially growing cells and their relationship with lineage distributions we have derived can be studied in future work, like other models [48,90]. We have also not considered the bidirectional crosstalk of gene expression and generation times. Here we have considered proteins which do not participate in cell size or cell cycle regulation, and so their expression did not directly influence the generation times. On the other hand generation time is related to cell growth rate and cell growth may influence both volume-dependent gene expression rates [50] and the dilution rate of protein concentration [91]. Effect of cell volume may be considered in effective concentration-based models of gene products [50,75,91,92], where the average concentration remains the same after division although average volume and copy numbers become half. More details of different cell-cycle stages [93], DNA duplication, and dosage compensation [48,49] are realistic aspects which need to be considered in future within more complete modeling. Each of these extensions may follow the core idea presented here, but are expected to be lengthy and interesting calculations and hence left for future study.

We made an interesting connection between two entirely different biophysical problems, namely, the problem of cell division and that of synaptic vesicle release triggered by action potentials [80]. The underlying mathematical structure of these two problems is similar, as both involve repeated cycles of growth (or replenishment) and partitioning (or release). Yet the specific results are different as the details of the process of gene expression differs from that of the stochastic docking of synaptic vesicles.

Stochastic gene expression plays a critical role in a number of biologically and biomedically significant processes. Notable examples include stochastic cell fate determination in both bacteria and multicellular organisms [94,95], spontaneous prophage induction in bacteria [96,97], and the random expression of proteins that confer antibiotic resistance in bacteria [98] or chemotherapy resistance in cancer cells [99,100]. Analytical models of protein expression are thus potentially very useful for analyzing noise-driven processes in biology. Previous models for the distribution of mRNA and protein levels in cells did not fully account for the contributions from randomness in cell division times in conjunction with binomial partitioning. In this paper we incorporate these sources of extrinsic noise into an analytical exact theory of mRNA and protein gene expression. We anticipate that the results would have wide applicability in biomedical and biological research for systems where noise in cell cycle duration plays a functional role.

ACKNOWLEDGMENTS

D.D. acknowledges the visitor program of MPI-PKS Dresden, where a part of this work was done, and thanks K. Rijal and Madan Rao for discussions. We also thank one referee for very useful suggestions. S.Y.A. acknowledges IIT Bombay for financial support through the institute's postdoctoral fellowship. A.S. acknowledges support from NIH-NIGMS via Grant No. R35GM148351.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

---

[1] A. Raj, P. van den Bogaard, S. A. Rifkin, A. van Oudenaarden, and S. Tyagi, Imaging individual mRNA molecules using multiple singly labeled probes, Nat. Methods **5**, 877 (2008).

[2] A. Raj and A. Van Oudenaarden, Nature, nurture, or chance: Stochastic gene expression and its consequences, Cell **135**, 216 (2008).

[3] D. Lovatt, B. K. Ruble, J. Lee, H. Dueck, T. K. Kim, S. Fisher, C. Francis, J. M. Spaethling, J. A. Wolf, M. S. Grady *et al.*, Transcriptome *in vivo* analysis (TIVA) of spatially defined single cells in live tissue, Nat. Methods **11**, 190 (2014).

[4] J. Cao, J. S. Packer, V. Ramani *et al.*, Comprehensive single-cell transcriptional profiling of a multicellular organism, Science **357**, 661 (2017).

[5] D. Mahdessian, A. J. Cesnik, C. Gnann *et al.*, Spatiotemporal dissection of the cell cycle with single-cell proteogenomics, Nature (London) **590**, 649 (2021).

[6] T. K. Suen, B. Al, and K. Placek, Cell-to-cell proteome variability: Life in a cycle, Sig. Transduct. Target Ther. **6**, 229 (2021).

[7] A. Deloupy, V. Sauveplane, J. Robert, S. Aymerich, M. Jules, and L. Robert, Extrinsic noise prevents the independent tuning of gene expression noise and protein mean abundance in bacteria, Sci. Adv. **6**, eabc3478 (2020).

[8] B. Munsky, G. Neuert, and A. van Oudenaarden, Using gene expression noise to understand gene regulation, Science **336**, 183 (2012).

[9] O. Padovan-Merhar and A. Raj, Using variability in gene expression as a tool for studying gene regulation, WIREs Syst. Biol. Med. **5**, 751 (2013).

[10] M. M. Saint-Antoine and A. Singh, Network inference in systems biology: Recent developments, challenges, and applications, Curr. Opin. Biotechnol. **63**, 89 (2020).

[11] S. M. Shaffer, M. C. Dunagin, S. R. Torborg *et al.*, Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance, Nature (London) **546**, 431 (2017).

[12] K. S. Farquhar, D. A. Charlebois, M. Szenk, J. Cohen, D. Nevozhay, and G. Balázsi, Role of network-mediated stochasticity in mammalian drug resistance, Nat. Commun. **10**, 2766 (2019).

[13] I. El Meouche, P. Jain, M. K. Jolly, and J.-P. Capp, Drug tolerance and persistence in bacteria, fungi and cancer cells:

Role of non-genetic heterogeneity, Transl. Oncol. **49**, 102069 (2024).

[14] C. A. Chang, J. Jen, S. Jiang *et al.*, Ontogeny and vulnerabilities of drug-tolerant persisters in HER2+ breast cancer, Cancer Discov. **12**, 1022 (2022).

[15] I. El Meouche, Y. Siu, and M. J. Dunlop, Stochastic expression of a multiple antibiotic resistance activator confers transient resistance in single cells, Sci. Rep. **6**, 19538 (2016).

[16] A. Singh and L. S. Weinberger, Stochastic gene expression as a molecular switch for viral latency, Curr. Opin. Microbiol. **12**, 460 (2009).

[17] P. S. Swain, M. B. Elowitz, and E. D. Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression, Proc. Natl. Acad. Sci. USA **99**, 12795 (2002).

[18] J. M. Raser and E. K. O'Shea, Control of stochasticity in eukaryotic gene expression, Science **304**, 1811 (2004).

[19] A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas, and S. Tyagi, Stochastic mRNA synthesis in mammalian cells, PLoS Biol. **4**, e309 (2006).

[20] L. Cai, N. Friedman, and X. S. Xie, Stochastic protein expression in individual cells at the single molecule level, Nature (London) **440**, 358 (2006).

[21] Y. Taniguchi, P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie, Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells, Science **329**, 533 (2010).

[22] A. Sanchez and I. Golding, Genetic determinants and cellular constraints in noisy gene expression, Science **342**, 1188 (2013).

[23] G. Storz, J. A. Opdyke, and A. Zhang, Controlling mRNA stability and translation with small, noncoding rnas, Curr. Opin. Microbiol. **7**, 140 (2004).

[24] M. Thattai and A. Van Oudenaarden, Intrinsic noise in gene regulatory networks, Proc. Natl. Acad. Sci. USA **98**, 8614 (2001).

[25] J. Paulsson, Models of stochastic gene expression, Phys. Life Rev. **2**, 157 (2005).

[26] A. Sanchez, H. G. Garcia, D. Jones, R. Phillips, and J. Kondev, Effect of promoter architecture on the cell-to-cell variability in gene expression, PLoS Comp. Biol. **7**, e1001100 (2011).

[27] V. Shahrezaei and P. S. Swain, Analytical distributions for stochastic gene expression, Proc. Natl. Acad. Sci. USA **105**, 17256 (2008).

[28] P. Bokes, J. R. King, A. T. Wood, and M. Loose, Exact and approximate distributions of protein and mRNA levels in the low-copy regime of gene expression, J. Math. Biol. **64**, 829 (2012).

[29] A. Singh, C. A. Vargas, and R. Karmakar, Stochastic analysis and inference of a two-state genetic promoter model, in *2013 American Control Conference* (IEEE, New York, 2013), pp. 4563–4568.

[30] A. Singh, B. S. Razooky, R. D. Dar, and L. S. Weinberger, Dynamics of protein noise can distinguish between alternate sources of gene-expression variability, Mol. Syst. Biol. **8**, 607 (2012).

[31] Y. Wang, Z. Yu, R. Grima, and Z. Cao, Exact solution of a three-stage model of stochastic gene expression including cell-cycle dynamics, J. Chem. Phys. **159**, 224102 (2023).

[32] E. D. Hawkins, J. F. Markham, L. P. McGuinness, and P. D. Hodgkin, A single-cell pedigree analysis of alternative stochastic lymphocyte fates, Proc. Natl. Acad. Sci. USA **106**, 13457 (2009).

[33] R. Tsukanov, G. Reshes, G. Carmon, E. Fischer-Friedrich, N. S. Gov, I. Fishov, and M. Feingold, Timing of Z-ring localization in *Escherichia coli*, Phys. Biol. **8**, 066003 (2011).

[34] G. Reshes, S. Vanounou, I. Fishov, and M. Feingold, Timing the start of division in *E. coli*: A single-cell study, Phys. Biol. **5**, 046001 (2008).

[35] A. H. K. Roeder, V. Chickarmane, A. Cunha, B. Obara, B. S. Manjunath, and E. M. Meyerowitz, Variability in the control of cell division underlies sepal epidermal patterning in *Arabidopsis thaliana*, PLoS Biol. **8**, e1000367 (2010).

[36] E. B. Stukalin, I. Aifuwa, J. S. Kim, D. Wirtz, and S. X. Sun, Age-dependent stochastic models for understanding population fluctuations in continuously cultured cells, J. R. Soc. Interface **10**, 20130325 (2013).

[37] C. A. Yates, M. J. Ford, and R. L. Mort, A multi-stage representation of cell proliferation as a Markov process, Bull. Math. Biol. **79**, 2905 (2017).

[38] P. Thomas, Intrinsic and extrinsic noise of gene expression in lineage trees, Sci. Rep. **9**, 474 (2019).

[39] D. Huh and J. Paulsson, Non-genetic heterogeneity from stochastic partitioning at cell division, Nat. Genet. **43**, 95 (2011).

[40] D. Huh and J. Paulsson, Random partitioning of molecules at cell division, Proc. Natl. Acad. Sci. USA **108**, 15004 (2011).

[41] C. J. Zopf, K. Quinn, J. Zeidman, and N. Maheshri, Cell-cycle dependence of transcription dominates noise in gene expression, PLoS Comput. Biol. **9**, e1003161 (2013).

[42] N. A. Cookson, S. W. Cookson, L. S. Tsimring, and J. Hasty, Cell cycle-dependent variations in protein concentration, Nucleic Acids Res. **38**, 2676 (2010).

[43] O. G. Berg, A model for the statistical fluctuations of protein numbers in a microbial population, J. Theor. Biol. **71**, 587 (1978).

[44] D. R. Rigney, Stochastic model of constitutive protein levels in growing and dividing bacterial cells, J. Theor. Biol. **76**, 453 (1979).

[45] F. Bertaux, S. Marguerat, and V. Shahrezaei, Division rate, cell size and proteome allocation: Impact on gene expression noise and implications for the dynamics of genetic circuits, R. Soc. Open Sci. **5**, 172234 (2018).

[46] R. Dessalles, V. Fromion, and P. Robert, Models of protein production along the cell cycle: An investigation of possible sources of noise, PLoS ONE **15**, e0226016 (2020).

[47] M. Soltani and A. Singh, Effects of cell-cycle-dependent expression on random fluctuations in protein levels, R. Soc. Open Sci. **3**, 160578 (2016).

[48] C. H. L. Beentjes, R. Perez-Carrasco, and R. Grima, Exact solution of stochastic gene expression models with bursting, cell cycle and replication dynamics, Phys. Rev. E **101**, 032403 (2020).

[49] Z. Cao and R. Grima, Analytical distributions for detailed models of stochastic gene expression in eukaryotic cells, Proc. Natl. Acad. Sci. USA **117**, 4682 (2020).

[50] C. Jia, A. Singh, and R. Grima, Concentration fluctuations in growing and dividing cells: Insights into the emergence of concentration homeostasis, PLoS Comput. Biol. **18**, 1 (2022).

[51] P. Fantes and P. Nurse, Control of cell size at division in fission yeast by a growth-modulated size control over nuclear division, Exp. Cell Res. **107**, 377 (1977).

[52] I. G. Johnston and N. S. Jones, Closed-form stochastic solutions for non-equilibrium dynamics and inheritance of cellular components over many cell divisions, Proc. R. Soc. A: Math. Phys. Eng. Sci. **471**, 20150050 (2015).

[53] J. Männik, P. Kar, C. Amarasinghe, A. Amir, and J. Männik, Determining the rate-limiting processes for cell division in *Escherichia coli*, Nat. Commun. **15**, 9948 (2024).

[54] C. A. Vargas-Garcia, K. R. Ghusinga, and A. Singh, Cell size control and gene expression homeostasis in single-cells, Curr. Opin. Syst. Biol. **8**, 109 (2018).

[55] G. Reshes, S. Vanounou, I. Fishov, and M. Feingold, Cell shape dynamics in *Escherichia coli*, Biophys. J. **94**, 251 (2008).

[56] A. Zilman, V. V. Ganusov, and A. S. Perelson, Stochastic models of lymphocyte proliferation and death, PLoS ONE **5**, e12775 (2010).

[57] X. Liu, J. Yan, and M. W. Kirschner, Cell size homeostasis is tightly controlled throughout the cell cycle, PLoS Biol. **22**, e3002453 (2024).

[58] C. Nieto, C. A. Vargas-García, J. M. Pedraza, and A. Singh, Mechanisms of cell size regulation in slow-growing *Escherichia coli* cells: Discriminating models beyond the adder, npj Syst. Biol. Appl. **10**, 61 (2024).

[59] M. Campos, I. V. Surovtsev, S. Kato, A. Paintdakhi, B. Beltran, S. E. Ebmeier, and C. Jacobs-Wagner, A constant size extension drives bacterial cell size homeostasis, Cell **159**, 1433 (2014).

[60] C. Cadart, S. Monnier, J. Grilli, P. J. Sáez, N. Srivastava, R. Attia, E. Terriac, B. Baum, M. Cosentino-Lagomarsino, and M. Piel, Size control in mammalian cells involves modulation of both growth rate and cell cycle duration, Nat. Commun. **9**, 3275 (2018).

[61] S. Taheri-Araghi, S. Bradde, J. T. Sauls, N. S. Hill, P. A. Levin, J. Paulsson, M. Vergassola, and S. Jun, Cell-size control and homeostasis in bacteria, Curr. Biol. **25**, 385 (2015).

[62] C. Nieto, C. A. Vargas-García, and A. Singh, A generalized adder for cell size homeostasis: Effects on stochastic clonal proliferation, Biophys. J. **124**, 1376 (2025).

[63] M. Soltani, C. A. Vargas-Garcia, D. Antunes, and A. Singh, Intercellular variability in protein levels from stochastic expression and noisy cell cycle processes, PLoS Comput. Biol. **12**, e1004972 (2016).

[64] R. Perez-Carrasco, C. Beentjes, and R. Grima, Effects of cell cycle variability on lineage and population measurements of messenger RNA abundance, J. R. Soc. Interface. **17**, 20200360 (2020).

[65] C. Jia and R. Grima, Frequency domain analysis of fluctuations of mRNA and protein copy numbers within a cell lineage: Theory and experimental validation, Phys. Rev. X **11**, 021032 (2021).

[66] C. Nieto, C. Augusto Vargas-Garcia, and A. Singh, A moments-based analytical approach for cell size homeostasis, IEEE Control Syst. Lett. **8**, 2205 (2024).

[67] C. W. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*, Springer Series in Synergetics, Vol. 115 (Springer, Berlin, Heidelberg, 1985).

[68] K. R. Ghusinga, C. A. Vargas-Garcia, and A. Singh, A mechanistic stochastic framework for regulating bacterial cell division, Sci. Rep. **6**, 30229 (2016).

[69] S. Iyer-Biswas, G. E. Crooks, N. F. Scherer, and A. R. Dinner, Universality in stochastic exponential growth, Phys. Rev. Lett. **113**, 028101 (2014).

[70] S. Iyer-Biswas, C. S. Wright, J. T. Henry, K. Lo, S. Burov, Y. Lin, G. E. Crooks, S. Crosson, A. R. Dinner, and N. F. Scherer, Scaling laws governing stochastic growth and division of single bacterial cells, Proc. Natl. Acad. Sci. USA **111**, 15912 (2014).

[71] L. Luo, Y. Bai, and X. Fu, Stochastic threshold in cell size control, Phys. Rev. Res. **5**, 013173 (2023).

[72] K. Biswas and N. Brenner, Universality of phenotypic distributions in bacteria, Phys. Rev. Res. **6**, L022043 (2024).

[73] D. Antunes and A. Singh, Quantifying gene expression variability arising from randomness in cell division times, J. Math. Biol. **71**, 437 (2015).

[74] J. Jędrak, M. Kwiatkowski, and A. Ochab-Marcinek, Exactly solvable model of gene expression in a proliferating bacterial cell population with stochastic protein bursts and protein partitioning, Phys. Rev. E **99**, 042416 (2019).

[75] N. Friedman, L. Cai, and X. S. Xie, Linking stochastic dynamics to population distribution: An analytical framework of gene expression, Phys. Rev. Lett. **97**, 168302 (2006).

[76] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, Real-time kinetics of gene activity in individual bacteria, Cell **123**, 1025 (2005).

[77] See Supplemental Material at http://link.aps.org/supplemental/10.1103/ysj3-gkrr for derivation of the analytical cyclo-stationary distributions of mRNA and protein counts and the corresponding cumulants, a supporting simulation data, description of computational methods of series summation and kinetic Monte-Carlo simulations, and analytical results on age-specific distribution and deterministic limits of partitioning and gene expression.

[78] A. D. Bird, M. J. Wall, and M. J. Richardson, Bayesian inference of synaptic quantal parameters from correlated vesicle release, Front. Comput. Neurosci. **10**, 116 (2016).

[79] E. G. Krächan, A. U. Fischer, J. Franke, and E. Friauf, Synaptic reliability and temporal precision are achieved via high quantal content and effective replenishment: Auditory brainstem versus hippocampus, J. Physiol. **595**, 839 (2017).

[80] K. Rijal, N. I. C. Müller, E. Friauf, A. Singh, A. Prasad, and D. Das, Exact distribution of the quantal content in synaptic transmission, Phys. Rev. Lett. **132**, 228401 (2024).

[81] Z. Vahdat, O. Gambrell, J. Fisch, E. Friauf, and A. Singh, Inferring synaptic transmission from the stochastic dynamics of the quantal content: An analytical approach, PLoS Comput. Biol. **21**, e1013067 (2025).

[82] J. Yu, J. Xiao, X. Ren, K. Lao, and X. S. Xie, Probing gene expression in live cells, one protein molecule at a time, Science **311**, 1600 (2006).

[83] Interestingly, the first passage times of threshold crossing of bacterial population size undergoing autocatalytic growth follows the Beta exponential distribution, as was shown by Dellbruck years before [101].

[84] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. **81**, 2340 (1977).

[85] E. O. Powell, Growth rate and generation time of bacteria, with special reference to continuous culture, Microbiology **15**, 492 (1956).

[86] F. Jafarpour, C. S. Wright, H. Gudjonson, J. Riebling, E. Dawson, K. Lo, A. Fiebig, S. Crosson, A. R. Dinner, and S. Iyer-Biswas, Bridging the timescales of single-cell and population dynamics, Phys. Rev. X **8**, 021007 (2018).

[87] P. Thomas, Making sense of snapshot data: Ergodic principle for clonal cell populations, J. R. Soc. Interface **14**, 20170467 (2017).

[88] S. Chakrabarti, A. L. Paek, J. Reyes, K. A. Lasick, G. Lahav, and F. Michor, Hidden heterogeneity and circadian-controlled cell fate inferred from single cell lineages, Nat. Commun. **9**, 5372 (2018).

[89] S. Jun, F. Si, R. Pugatch, and M. Scott, Fundamental principles in bacterial physiology—History, recent progress, and the future with focus on cell size control: A review, Rep. Prog. Phys. **81**, 056601 (2018).

[90] A. Genthon, Analytical cell size distribution: Lineage-population bias and parameter inference, J. R. Soc. Interface **19**, 20220405 (2022).

[91] Z. Zhang, I. Zabaikina, C. Nieto, Z. Vahdat, P. Bokes, and A. Singh, Stochastic gene expression in proliferating cells: Differing noise intensity in single-cell and population perspectives, PLoS Comput. Biol. **21**, e1013014 (2025).

[92] K. Rijal, A. Prasad, A. Singh, and D. Das, Exact distribution of threshold crossing times for protein concentrations: Implication for biological timekeeping, Phys. Rev. Lett. **128**, 048101 (2022).

[93] O. Padovan-Merhar, G. P. Nair, A. G. Biaesch *et al.*, Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms, Mol. Cell. **58**, 339 (2015).

[94] R. Losick and C. Desplan, Stochasticity and cell fate, Science **320**, 65 (2008).

[95] K. Carniol, P. Eichenberger, and R. Losick, A threshold mechanism governing activation of the developmental regulatory protein σ^F in *Bacillus subtilis*, J. Biol. Chem. **279**, 14860 (2004).

[96] R. Miyazaki, M. Minoia, N. Pradervand, S. Sulser, F. Reinhard, and J. R. van der Meer, Cellular variability of RpoS expression underlies subpopulation activation of an integrative and conjugative element, PLoS Genet. **8**, e1002818 (2012).

[97] J. W. Little and C. B. Michalowski, Stability and instability in the lysogenic state of phage lambda, J. Bacteriol. **192**, 6064 (2010).

[98] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler, Bacterial persistence as a phenotypic switch, Science **305**, 1622 (2004).

[99] L. Schuh, M. Saint-Antoine, E. M. Sanford *et al.*, Gene networks with transcriptional bursting recapitulate rare transient coordinated high expression states in cancer, Cell Syst. **10**, 363 (2020).

[100] S. M. Shaffer, B. L. Emert, R. A. Reyes Hueros *et al.*, Memory sequencing reveals heritable single-cell gene expression programs associated with distinct cellular behaviors, Cell **182**, 947 (2020).

[101] M. Delbrück, Statistical fluctuations in autocatalytic reactions, J. Chem. Phys **8**, 120 (1940).

Supplementary Material

Cyclo-stationary distributions of mRNA and Protein counts for random cell division times

Syed Yunus Ali,[1,*] Aditya Saran,[1] Ashok Prasad,[2,†] Abhyudai Singh,[3,‡] and Dibyendu Das[1,4,§]
[1]*Department of Physics, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India*
[2]*School of Biomedical and Chemical Engineering,*
Colorado State University, Fort Collins, Colorado 80521, USA
[3]*Department of Electrical and Computer Engineering,*
University of Delaware, Newark, DE 19716, USA
[4]*Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany*

I. GENERAL FRAMEWORK TO STUDY THE CYCLO-STATIONARY DISTRIBUTIONS

A. Equation for the cyclo-stationary copy number distributions $P^{ss}_+(y_+)$ and $P^{ss}_-(y_-)$

In the main text, we present the equations for $P^i_+(y_{+,i},t_{s_i})$ and $P^i_-(y_{-,i},t_{s_i})$, the distributions of copy numbers just after and just before the i^{th} cell division respectively, as follows:

$$P^i_+(y_{+,i},t_{s_i}) = \sum_{y_{+,i-1}}\sum_{x_{+,i}} B(y_{+,i}+x_{+,i},\frac{1}{2},x_{+,i})p(y_{+,i}+x_{+,i},t_{s_i}|y_{+,i-1})\ P^{i-1}_+(y_{+,i-1},t_{s_{i-1}}), \tag{1}$$

$$P^i_-(y_{-,i},t_{s_i}) = \sum_{y_{+,i-1}} p(y_{-,i},t_{s_i}|y_{+,i-1})\ P^{i-1}_+(y_{+,i-1},t_{s_{i-1}}). \tag{2}$$

Integrating Eq. 1 over the joint probability distribution of $g_2(t_{s_i},t_{s_{i-1}})$ of successive division time intervals,

$$\int_0^\infty\int_0^\infty dt_{s_i}dt_{s_{i-1}}g_2(t_{s_i},t_{s_{i-1}})P^i_+(y_{+,i},t_{s_i}) = \int_0^\infty\int_0^\infty dt_{s_i}dt_{s_{i-1}}g_2(t_{s_i},t_{s_{i-1}})\sum_{y_{+,i-1}}\sum_{x_{+,i}}B(y_{+,i}+x_{+,i},\frac{1}{2},x_{+,i}) \\ p(y_{+,i}+x_{+,i}|y_{+,i-1},t_{s_i})P^i_+(y_{+,i-1},t_{s_{i-1}}). \tag{3}$$

If it is further assumed that successive division times are uncorrelated, i.e. $g_2(t_{s_i},t_{s_{i-1}}) = g(t_{s_i})g(t_{s_{i-1}})$, where $g(t_{s_i})$ is the normalized distributions of t_{s_i}, we have

$$\int_0^\infty\int_0^\infty dt_{s_i}dt_{s_{i-1}}g(t_{s_i})g(t_{s_{i-1}})P^i_+(y_{+,i},t_{s_i}) = \int_0^\infty\int_0^\infty dt_{s_i}dt_{s_{i-1}}g(t_{s_i})g(t_{s_{i-1}})\sum_{y_{+,i-1}}\sum_{x_{+,i}}B(y_{+,i}+x_{+,i},\frac{1}{2},x_{+,i}) \\ p(y_{+,i}+x_{+,i}|y_{+,i-1},t_{s_i})P^i_+(y_{+,i-1},t_{s_{i-1}}) \tag{4}$$

which simplifies to

$$\int_0^\infty dt_{s_i}g(t_{s_i})P^i_+(y_{+,i},t_{s_i}) = \sum_{y_{+,i-1},x_{+,i}}\int_0^\infty dt_{s_i}g(t_{s_i})p(y_{+,i}+x_{+,i}|y_{+,i-1},t_{s_i})\ B\left(y_{+,i}+x_{+,i},\frac{1}{2},x_{+,i}\right) \\ \times\int_0^\infty dt_{s_{i-1}}g(t_{s_{i-1}})P^i_+(y_{+,i-1},t_{s_{i-1}}) \tag{5}$$

---

[*] syed_24@iitb.ac.in
[†] ashok.prasad@colostate.edu
[‡] absingh@udel.edu
[§] dibyendudas@iitb.ac.in

For $i \gg 1$, as the cyclo-stationary regime is attained, we may define the distribution as cell birth, $P_+^{ss}(y_+) = \int_0^\infty dt_{s_i} g(t_{s_i}) P_+^i(y_{+,i}, t_{s_i})$. Dropping the subscripts i, and setting $t_{s_i} = t_s$ and $y_{+,i-1} = y'_+$, Eq. 5 gives

$$P_+^{ss}(y_+) = \sum_{y'_+}\sum_{x_+}\int_0^\infty dt_s g(t_s)\ B(y_+ + x_+, \frac{1}{2}, x_+) \times p(y_+ + x_+, t_s|y'_+)\ P_+^{ss}(y'_+). \tag{6}$$

In a similar way, one may derive from Eq. 2 above, the cyclo-stationary distribution just before division, defined as $P_-^{ss}(y_-) = \int_0^\infty dt_{s_i} g(t_{s_i}) P_-^i(y_{-,i}, t_{s_i})$, related to $P_+^{ss}(y_+)$:

$$P_-^{ss}(y_-) = \sum_{y'_+}\int_0^\infty dt_s g(t_s)\ p(y_-, t_s|y'_+, t_s)\ P_+^{ss}(y'_+). \tag{7}$$

Eqs. 6 and 7 are the two Eqs. 3 and 4 in the main text.

B. Self-consistent integral for the generating functions $F_+(q)$ and its relation to $F_-(q)$

We define generating function $F_\pm(q) = \sum_{y_\pm=0}^\infty q^{y_\pm} P_\pm^{ss}(y_\pm)$. Multiply $\sum_{y_+} q^{y_+}$ on both sides of Eq. 6 we get,

$$F_+(q) = \sum_{y'_+}\sum_{x_+}\int_0^\infty dt_s g(t_s) \sum_{y_+} q^{y_+} p(y_+ + x_+, t_s|y'_+)\ B(y_+ + x_+, \frac{1}{2}, x_+) P_+^{ss}(y'_+) \tag{8}$$

We defining a new variable $\tilde{y}_+ = y_+ + x_+$, where $x_+ \le \tilde{y}_+ \le \infty$. But since $B(\tilde{y}_+, \frac{1}{2}, x_+) = 0$ for $\tilde{y}_+ < x_+$, we put $\tilde{y}_+ \in [0, \infty)$. Using the explicit form of the binomial function, the recursion relation of $F_+(q)$ may then be written as:

$$\begin{aligned} F_+(q) &= \sum_{y'_+}\int_0^\infty dt_s g(t_s) \sum_{\tilde{y}_+} q^{\tilde{y}_+} p(\tilde{y}_+, t_s|y'_+) \sum_{x_+=0}^{\tilde{y}_+} \frac{1}{q^{x_+}} \binom{\tilde{y}_+}{x_+} \left(\frac{1}{2}\right)^{\tilde{y}_+ - x_+} \left(\frac{1}{2}\right)^{x_+} P_+^{ss}(y'_+) \\ &= \sum_{y'_+}\int_0^\infty dt_s g(t_s) \left(\sum_{\tilde{y}_+} q^{\tilde{y}_+} p(\tilde{y}_+, t_s|y'_+) \left(\frac{q+1}{2q}\right)^{\tilde{y}_+} \right) P_+^{ss}(y'_+) \\ &= \int_0^\infty dt_s g(t_s) \sum_{y'_+} F\left(\frac{q+1}{2}, t_s \middle| y'_+\right) P_+^{ss}(y'_+). \end{aligned} \tag{9}$$

Here $F(q, t|y') = \sum_y q^y p(y, t|y')$ is the generating function of the probability $p(y, t|y')$ tied to the process of gene expression. Let us assume that this generating function has a form $F(q, t|y') = \mathcal{H}(q-1, \gamma_y t) \times (1 + (q-1)e^{-\gamma_y t})^{y'}$. This leads to the following self-consistent integral for $F_+(q)$ (which is Eq. 6 in the main text):

$$F_+(q) = \int_0^\infty dt_s g(t_s) \mathcal{H}(\frac{q-1}{2}, \gamma_y t_s) F_+\left(1 + \frac{(q-1)}{2} e^{-\gamma_y t_s}\right). \tag{10}$$

In a similar way, by multiplying $\sum_{y_-} q^{y_-}$ on both sides of Eq. 7 we get,

$$\begin{aligned} F_-(q) &= \sum_{y_-} q^{y_-} \sum_{y'_+}\int_0^\infty dt_s g(t_s) \times p(y_-, t_s|y'_+) \times P_+^{ss}(y'_+) \\ &= \int_0^\infty dt_s g(t_s) \sum_{y'_+} F(q, t_s|y'_+) P_+^{ss}(y'_+) \\ &= \int_0^\infty dt_s g(t_s) \mathcal{H}(q-1, \gamma_y t_s) F_+\left(1 + (q-1)e^{-\gamma_y t_s}\right). \end{aligned} \tag{11}$$

Putting $q = 2q' - 1$ in Eq. 10 we have,

$$F_+(2q'-1) = \int_0^\infty dt_s g(t_s) \mathcal{H}(q'-1, \gamma_y t_s) F_+\left(1 + (q'-1)e^{-\gamma_y t_s}\right), \tag{12}$$

and comparing with Eq. 11, we obtain (the Eq. 7 in the main text):

$$F_-(q) = F_+(2q-1) \tag{13}$$

C. Closed form of $F_+(q)$ for Fixed cell cycle times, and the intractable nested integrals for Random cell cycle times

For fixed cell division times, i.e. $g(t_s)=\delta(t_s-T)$, Eq. 10 reduces to

$$F_+(q)=\mathcal{H}\left(\frac{q-1}{2},\gamma_y T\right)F_+\left(1+\frac{(q-1)}{2}e^{-\gamma_y T}\right). \tag{14}$$

If we set, $q-1=w$ then Eq. 14 becomes to

$$F_+(1+w)=\mathcal{H}\left(\frac{w}{2},\gamma_y T\right)F_+\left(1+\frac{w}{2}e^{-\gamma_y T}\right). \tag{15}$$

This recursion formula may be iterated to obtain

$$F_+(1+w)=\prod_{k=1}^{j}\mathcal{H}\left(\frac{w}{2}\frac{e^{-(k-1)\gamma_y T}}{2^{k-1}},\gamma_y T\right)F_+\left(1+w\left(\frac{e^{-\gamma_y T}}{2}\right)^j\right). \tag{16}$$

As $j\to\infty$, $F_+\left(1+w\left(\frac{e^{-\gamma_y T}}{2}\right)^j\right)\to F_+(1)=1$, and hence

$$F_+(1+w)=\prod_{k=1}^{\infty}\mathcal{H}\left(\frac{w}{2}\frac{e^{-(k-1)\gamma_y T}}{2^{k-1}},\gamma_y T\right). \tag{17}$$

Replacing back $w=q-1$ Eq. 17 gives the closed form in Eq. 8 in the main text.

For random division times t_s, with any general normalised function $g(t_s)$, when we substitute $w=q-1$, Eq. 10 becomes:

$$F_+(1+w)=\int_0^\infty dt_s\; g(t_s)\mathcal{H}\left(\frac{w}{2},\gamma_y t_s\right)F_+\left(1+\frac{w}{2}e^{-\gamma_y t_s}\right). \tag{18}$$

Iterating one step, and replacing the F_+ on the right side with an similar integral as Eq. 18, we obtain:

$$F_+(w+1)=\int_0^\infty dt_s^1\; g(t_s^1)\mathcal{H}\left(\frac{w}{2},\gamma_y t_s^1\right)\int_0^\infty dt_s^2\; g(t_s^2)\mathcal{H}\left(\frac{w_1(t_s^1)}{2},\gamma_y t_s^2\right)F_+\left(1+\frac{w_1(t_s^1)}{2}e^{-\gamma_y t_s^2}\right), \tag{19}$$

where $w_1(t_s^1)=\frac{w}{2}e^{-\gamma_y t_s^1}$. Continuing with the next iteration,

$$F_+(w+1)=\int_0^\infty dt_s^1\; g(t_s^1)\mathcal{H}\left(\frac{w}{2},\gamma_y t_s^1\right)\int_0^\infty dt_s^2\; g(t_s^2)\mathcal{H}\left(\frac{w_1}{2},\gamma_y t_s^2\right)\int_0^\infty dt_s^3\; g(t_s^3)\mathcal{H}\left(\frac{w_2}{2},\gamma_y t_s^3\right)F_+\left(1+\frac{w_2}{2}e^{-\gamma_y t_s^3}\right), \tag{20}$$

where $w_2=w_2(t_s^1,t_s^2)=\frac{w_1}{2}e^{-\gamma_y t_s^2}=\frac{w}{2^2}e^{-\gamma_y t_s^1}e^{-\gamma_y t_s^2}$. Repeating this indefinitely, as $j\to\infty$ we have $F_+(1+w_j)\to F_+(1)=1$, where $w_j=\frac{w_{j-1}}{2}e^{-\gamma_y t_s^j}$, and hence

$$F_+(1+w)=\prod_{k=1}^{\infty}\int_0^\infty dt_s^k\; g(t_s^k)\mathcal{H}\left(\frac{w_{k-1}}{2},\gamma_y t_s^k\right). \tag{21}$$

As $w_k=w_k(t_s^1,t_s^2,...,t_s^k)=\frac{w}{2^k}e^{-\gamma_y t_s^1}e^{-\gamma_y t_s^2}...e^{-\gamma_y t_s^k}$, the above nested integrals are in general intractable. This is why the problem has stayed challenging.

D. Deriving the cyclo-stationary distributions at birth and before division, from the series expansion of generating function about $q=1$

Although $P_+^{ss}(y_+)$ are themselves coefficients of the series expansion of $F_+(q)$ about $q=0$, we may start with an alternate expansion of $F_+(q)=\sum_{k=0}^{\infty}\frac{(q-1)^k}{k!}F_+^{(k)}(1)$ about $q=1$. In that case,

$$P_+^{ss}(y_+)=\frac{1}{y_+!}\left[\frac{\partial^{y_+}}{\partial q^{y_+}}F_+(q)\right]_{q=0}=\frac{1}{y_+!}\left[\frac{\partial^{y_+}}{\partial q^{y_+}}\sum_{k=0}^{\infty}\frac{(q-1)^k}{k!}F_+^{(k)}(1)\right]_{q=0}=\sum_{k=y_+}^{\infty}\binom{k}{y_+}\frac{(-1)^{k-y_+}}{k!}F_+^{(k)}(1) \tag{22}$$

Similarly, using Eq. 13, we have

$$P_-^{ss}(y_-) = \frac{1}{y_-!}\left[\frac{\partial^{y_-}}{\partial q^{y_-}}F_-(q)\right]_{q=0} = \frac{1}{y_-!}\left[\frac{\partial^{y_-}}{\partial q^{y_-}}F_+(2q-1)\right]_{q=0} = \frac{1}{y_-!}\left[\frac{\partial^{y_-}}{\partial q^{y_-}}\sum_{k=0}^{\infty}\frac{2^k(q-1)^k}{k!}F_+^{(k)}(1)\bigg|_{q=0}\right]$$
$$= \sum_{k=y_-}^{\infty}\binom{k}{y_-}\frac{(-1)^{k-y_-}}{k!}2^kF_+^{(k)}(1) \tag{23}$$

Thus above, we have the series expansions of $P_+^{ss}(y_+)$ and $P_-^{ss}(y_-)$ (Eq. 11 and 12 in the main text) involving the coefficients $F_+^{(k)}(1)$.

E. The first three cumulants of $P_+^{ss}(y_+)$ in terms of the coefficients $F_+^{(k)}(1)$

Using $F_+(q) = \sum_{y_+} P_+^{ss}(y_+)q^{y_+} = \sum_{k=0}^{\infty}\frac{(q-1)^k}{k!}F_+^{(k)}(1)$, we may obtain the cumulants as follows. The mean of y_+:

$$\langle y_+\rangle = \sum_{y_+} y_+P_+^{ss}(y_+) = q\frac{\partial}{\partial q}F_+(q)\bigg|_{q=1} = q\frac{\partial}{\partial q}\left(\sum_k\frac{(q-1)^k}{k!}F_+^{(k)}(1)\right)_{q=1} = q\sum_k\frac{k(q-1)^{k-1}}{k!}F_+^{(k)}(1)\bigg|_{q=1} = F_+^{(1)}(1) \tag{24}$$

The second moment

$$\langle y_+^2\rangle = q\frac{\partial}{\partial q}q\frac{\partial}{\partial q}F_+(q)\bigg|_{q=1} = q\sum_k\frac{k(q-1)^{k-1}}{k!}F_+^{(k)}(1) + q^2\sum_k\frac{k(k-1)(q-1)^{k-2}}{k!}F_+^{(k)}(1)\bigg|_{q=1} = F_+^{(1)}(1) + F_+^{(2)}(1) \tag{25}$$

Hence the Variance

$$\kappa_2 = \langle y_+^2\rangle - \langle y_+\rangle^2 = F_+^{(1)}(1) + F_+^{(2)}(1) - \left(F_+^{(1)}(1)\right)^2 \tag{26}$$

The third moment

$$\begin{aligned}\langle y_+^3\rangle &= q\frac{d}{dq}q\frac{d}{dq}q\frac{d}{dq}F_+(q)\bigg|_{q=1}\\ &= q\sum_k\frac{k(q-1)^{k-1}}{k!}F_+^{(k)}(1) + 3q^2\sum_k\frac{k(k-1)(q-1)^{k-2}}{k!}F_+^{(k)}(1) + q^2\sum_k\frac{k(k-1)(k-2)(q-1)^{k-3}}{k!}F_+^{(k)}(1)\bigg|_{q=1}\\ &= F_+^{(1)}(1) + 3F_+^{(2)}(1) + F_+^{(3)}(1)\end{aligned} \tag{27}$$

Hence the third cumulant is (see Eqs. 24, 26 and 27 above)

$$\kappa_3 = \langle(y_+ - \kappa_1)^3\rangle = [F_+^{(1)}(1) + 3F_+^{(2)}(1) + F_+^{(3)}(1)] - 3\kappa_1\kappa_2 - \kappa_1^3 \tag{28}$$

The above equations appear as Eq. 13, 14 and 15 in the main text. Using the above cumulants we obtain $CV^2 = \kappa_2/\kappa_1^2$ and Skewness $= \kappa_3/\kappa_2^{3/2}$ in our study.

II. STATISTICS OF THE mRNA NUMBER IN THE CYCLO-STATIONARY STATE

A. The generating function related to the model of transcription, and thereby obtaining function $\mathcal{H}$

The Master equation for the stochastic model of mRNA production and degradation is

$$\frac{dp(m,t|m'_+)}{dt} = k_m p(m-1,t|m'_+) + \gamma_m(m+1)p(m+1,t|m'_+) - (\gamma_m m + k_m)p(m,t|m'_+). \tag{29}$$

Here k_m is the transcription rate, and γ_m is the degradation rate of mRNAs. The generating function $F(q,t) = \sum_{j=0}^{\infty} q^m p(m,t|m'_+)$ of the distribution $p(m,t|m'_+)$ satisfies (using Eq. 29 above) the following:

$$\frac{\partial F(q,t)}{\partial t} + \gamma_m(q-1)\frac{\partial F}{\partial q} = k_m(q-1)F. \tag{30}$$

Eq. 30 can be solved by using the method of Lagrange characteristic, and one gets

$$F(q,t) = e^{\lambda(t)(q-1)}(1+(q-1)e^{-\gamma_m t})^{m'_+}, \tag{31}$$

where $\lambda(t) = (k_m/\gamma_m)[1-e^{-\gamma_m t}]$. For brevity we will use $\lambda(t) \equiv \lambda$ below. Thus comparing with Eq. 5 of the main text (also see below Eq. 9), we identify the function

$$\mathcal{H} = e^{\frac{k_m}{\gamma_m}[1-e^{-\gamma_m t}](q-1)}. \tag{32}$$

B. Obtaining the coefficients $F_+^{(k)}(1)$ and the series of the distributions $P_\pm^{ss}(m_\pm)$

Using $\mathcal{H}$ from Eq. 32 in Eq. 10, and $F_+(q) = \sum_{j=0}^{\infty} \frac{(q-1)^j}{j!} F_+^{(j)}(1)$ we have

$$\begin{aligned}
F_+(q) &= \int_0^\infty dt_s g(t_s) e^{\lambda((q-1)/2)} F_+(1+((q-1)/2)e^{-\gamma_m t_s}) \\
&= \int_0^\infty dt_s g(t_s) e^{\lambda((q-1)/2)} \sum_{j=0}^{\infty} \frac{F_+^{(j)}(1)}{j!}\left(\frac{q-1}{2}\right)^j e^{-j\gamma_m t_s} \\
&= \int_0^\infty dt_s g(t_s) \sum_{l=0}^{\infty}\sum_{j=0}^{\infty} \frac{\lambda^l \left(\frac{q-1}{2}\right)^l}{l!} \frac{F_+^{(j)}(1)}{j!}\left(\frac{q-1}{2}\right)^j e^{-j\gamma_m t_s} \\
&= \int_0^\infty dt_s g(t_s) \sum_{l=0}^{\infty}\sum_{j=0}^{\infty} \frac{F_+^{(j)}(1)}{l!j!}\left(\frac{q-1}{2}\right)^{l+j}\left(\frac{k_m}{\gamma_m}\right)^l e^{-j\gamma_m t_s}(1-e^{-\gamma_m t_s})^l
\end{aligned} \tag{33}$$

Changing summation indices to $k = l+j$ and defining $\Psi_{k,j} = \int_0^\infty dt_s g(t_s) e^{-j\gamma_m t_s}(1-e^{-\gamma_m t_s})^{k-j}$, Eq. 33 becomes

$$F_+(q) = \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{q-1}{2}\right)^k \sum_{j=0}^{k} \binom{k}{j}\left(\frac{k_m}{\gamma_m}\right)^{k-j} \Psi_{k,j} F_+^{(j)}(1) \tag{34}$$

Using the relation $F_+(q) = \sum_{k=0}^{\infty} \frac{(q-1)^k}{k!} F_+^k(1)$ on the left side of Eq. 34 above, and comparing coefficients we get the desired recursion relation (which appears in Eq. 21 of the main text):

$$F_+^{(k)}(1) = \frac{1}{2^k}\sum_{j=0}^{k}\left(\frac{k_m}{\gamma_m}\right)^{k-j}\binom{k}{j}\Psi_{k,j}F_+^{(j)}(1). \tag{35}$$

The first few coefficients are explicitly as follows. As $\sum_{m_+} P^{ss}(m_+) = 1$ we firstly have $F_+^{(0)}(1) = 1$. The next coefficient (from Eq. 35) is

$$F_+^{(1)}(1) = \frac{1}{2}\left(\Psi_{1,0}\frac{k_m}{\gamma_m} + \Psi_{1,1}F_+^{(1)}(1)\right) = \frac{\frac{k_m}{\gamma_m}\frac{1}{2}\Psi_{1,0}}{1-\frac{1}{2}\Psi_{1,1}} \tag{36}$$

Proceeding similarly we have $F_+^{(2)}(1)$ determined by $F_+^{(1)}(1)$ as follows:

$$F_+^{(2)}(1) = \frac{\left(\frac{k_m}{\gamma_m}\right)^2 \frac{1}{2^2}\Psi_{2,0}}{1-\frac{1}{2^2}\Psi_{2,2}} + \frac{\left(\frac{k_m}{\gamma_m}\right)^2 \frac{1}{2^3}\binom{2}{1}\Psi_{2,1}\Psi_{1,0}}{(1-\frac{1}{2}\Psi_{1,1})(1-\frac{1}{2^2}\Psi_{2,2})} \tag{37}$$

Next, the coefficient

$$F_{+}^{(3)}(1) = \left[\frac{\frac{1}{2^3}\left(\frac{k_m}{\gamma_m}\right)^3 \Psi_{3,0}}{1-\frac{1}{2^3}\Psi_{3,3}}\right] + \left[\frac{\frac{1}{2^4}\binom{3}{1}\Psi_{3,1}\Psi_{1,0}\left(\frac{k_m}{\gamma_m}\right)^3}{(1-\frac{1}{2^3}\Psi_{3,3})(1-\frac{1}{2}\Psi_{1,1})}\right] + \left[\frac{\frac{1}{2^5}\binom{3}{2}\Psi_{3,2}\Psi_{2,0}\left(\frac{k_m}{\gamma_m}\right)^3}{(1-\frac{1}{2^2}\Psi_{2,2})(1-\frac{1}{2^3}\Psi_{3,3})}\right]$$
$$+ \left[\frac{\frac{1}{2^6}\binom{3}{2}\Psi_{3,2}\Psi_{2,1}\Psi_{1,0}\left(\frac{k_m}{\gamma_m}\right)^3}{(1-\frac{1}{2}\Psi_{1,1})(1-(\frac{1}{2^2}\Psi_{2,2})(1-\frac{1}{2^3}\Psi_{3,3})}\right] \tag{38}$$

Observing the pattern of the successive coefficients, we obtain the general solution for $F_{+}^{(k)}(1)$ as follows:

$$F_{+}^{(k)}(1) = \left(\frac{k_m}{\gamma_m}\right)^k \frac{1}{2^k}\frac{1}{\left(1-\frac{1}{2^k}\Psi_{k,k}\right)}\left[\sum_{\{S_{k-1}\}}\frac{(\frac{1}{2})^{\sum_i j_i}\phi_{k,j_z}\ \phi_{j_z,j_{z-1}}\ ...\ \phi_{j_{1,0}}}{\prod_i\left(1-\frac{1}{2^{j_i}}\Psi_{j_i,j_i}\right)}+\phi_{k,0}\right] \tag{39}$$

where $\phi_{k,j} = \Psi_{k,j}\binom{k}{j}$. Here $\{S_{k-1}\}$ denotes the set of all the subsets $S_{k-1} = \{j_i\} = (j_z, j_{z-1}...j_1)$ of integers $j_i \in (1, 2, ...k-1)$ such that $j_z > j_{z-1} > ... > j_1$. For example for $k = 3$, the subsets are (1), (2), and (2,1) as is seen in Eq. 38.

With the coefficients given by Eq. 39, the cyclo-stationary distributions are formally given by the series:

$$P_{+}^{ss}(m_+) = \sum_{k=m_+}^{\infty}\binom{k}{m_+}\frac{(-1)^{k-m_+}}{k!}F_{+}^{(k)}(1) \tag{40}$$

$$P_{-}^{ss}(m_-) = \sum_{k=m_-}^{\infty}\binom{k}{m_-}\frac{(-1)^{k-m_-}2^k}{k!}F_{+}^{(k)}(1) \tag{41}$$

C. The cyclo-stationary mRNA distributions are Poisson for fixed cell-division times T

For fixed cell division times, $g(t_s) = \delta(t_s - T)$, we have $\Psi_{k,j} = e^{-j\gamma_m T}(1-e^{-\gamma_m T})^{k-j}$, and

$$\begin{aligned}\Psi_{k,j_z}\Psi_{j_z,j_{z-1}}\dots\Psi_{j_1,0} &= e^{-j_z\gamma_m T}(1-e^{-\gamma_m T})^{k-j_z}e^{-j_{z-1}\gamma_m T}(1-e^{-\gamma_m T})^{j_z-j_{z-1}}....e^{-0*\gamma_m T}(1-e^{-\gamma_m T})^{j_1}\\ &= e^{-\gamma_m T\sum_{i=1}^{z}j_i}(1-e^{-\gamma_m T})^k\end{aligned} \tag{42}$$

Consequently from Eqs. 40 and 39,

$$\begin{aligned}P_{+}^{ss}(m_+) &= \sum_{k=m_+}^{\infty}\frac{(-1)^{k-m_+}}{k!}\binom{k}{m_+}F_{+}^{(k)}(1)\\ &= \sum_{k=m_+}^{\infty}\frac{(-1)^{j-m_+}}{k!}\binom{k}{m_+}\left[\left(\frac{k_m}{\gamma_m}\right)^k\frac{1}{2^k}\frac{1}{\left(1-\frac{1}{2^k}\Psi_{k,k}\right)}\left[\sum_{\{S_{k-1}\}}\frac{(\frac{1}{2})^{\sum j_i}\phi_{k,j_z}\phi_{j_z,j_{z-1}}...\phi_{j_{1,0}}}{\prod_i\left(1-\frac{1}{2^{j_i}}\Psi_{j_i,j_i}\right)}+\phi_{k,0}\right]\right]\\ &= \sum_{k=m_+}^{\infty}\frac{(-1)^{k-m_+}}{j!}\binom{k}{m_+}\left(\frac{k_m}{\gamma_m}\right)^k\frac{\frac{1}{2^k}}{1-\frac{1}{2^k}\Psi_{k,k}}(1-e^{-\gamma_m T})^k\left[1+\sum_{\{S_{k-1}\}}\binom{k}{j_z}\binom{j_z}{j_{z-1}}...\binom{j_2}{j_1}\frac{\prod_i\left(\frac{1}{2}e^{-\gamma_m T}\right)^{j_i}}{\prod_i\left(1-(\frac{1}{2}e^{-\gamma_m T})^{j_i}\right)}\right]\end{aligned} \tag{43}$$

Using the the following identity [1], with $x = \frac{1}{2}e^{-\gamma_m T}$ in this case,

$$1+\sum_{\{S_{k-1}\}}\binom{k}{j_z}\binom{j_z}{j_{z-1}}...\binom{j_2}{j_1}\prod_i\frac{x^{j_i}}{1-x^{j_i}} = \frac{1-x^k}{(1-x)^k} \tag{44}$$

Thus the $(1-x^k)$ factors cancel from the numerator and denominator, and Eq. 43 simplifies to

$$\begin{aligned} P_{ss}(m_+) &= \sum_{k=m_+}^{\infty} \binom{k}{m_+} \frac{(-1)^{k-m_+}}{k!} \left(\frac{k_m}{\gamma_m}\right)^k \frac{\frac{1}{2^k}}{(1-e^{-\gamma_m T}\frac{1}{2})^k} (1-e^{-\gamma_m T})^k \\ &= \sum_{k=m_+}^{\infty} \frac{d^k}{m_+!(k-m_+)!} (-1)^{k-m_+} \end{aligned} \tag{45}$$

with $d = \left(\frac{k_m}{\gamma_m}\right) \frac{(1-e^{-\gamma_m T})}{(2-e^{-\gamma_m T})}$. Thus $F_+^{(k)}(1) = d^k$. Replacing $k - m_+ = r$, then Eq. 45 reduces to a Poisson distribution:

$$P_+^{ss}(m_+) = \frac{d^{m_+}}{m_+!} \sum_{r=0}^{\infty} \frac{(-1)^r d^r}{r!} = \frac{d^{m_+}}{m_+!} \exp(-d). \tag{46}$$

Since Eq. 41 has an extra factor of 2^k multiplying $F_+^{(k)}(1)$, we would have d replaced by $2d$ and the distribution:

$$P_-^{ss}(m_-) = \frac{(2d)^{m_-}}{m_-!} \exp(-2d). \tag{47}$$

D. Simplified form of the cyclo-stationary distributions for Exponentially distributed cell cycle times

For exponentially distributed division times, $g(t_s) = \lambda e^{-\lambda t_s}$ (with $\lambda = 1/T$), and

$$\Psi_{k,j} = \int_0^{\infty} dt_s g(t_s) e^{-\gamma_m t_s} (1-e^{-\gamma_m t_s})^{k-j} = \lambda/\gamma_m\, B\left(\frac{\lambda + j\gamma_m}{\gamma_m}, k-j+1\right) = \frac{\lambda}{\gamma_m} \frac{\Gamma(k-j+1)\Gamma\left(\frac{\lambda}{\gamma_m}+j\right)}{\Gamma\left(\frac{\lambda}{\gamma_m}+k+1\right)} \tag{48}$$

and $\Psi_{k,k} = \frac{\lambda}{\gamma_m} \frac{\Gamma(1)\Gamma\left(\frac{\lambda}{\gamma_m}+k\right)}{\Gamma\left(\frac{\lambda}{\gamma_m}+k+1\right)} = \frac{\lambda}{\lambda+\gamma_m k}$. Then

$$\phi_{k,j} = \binom{k}{j} \Psi_{k,j} = \left(\frac{\lambda}{\gamma_m}\right) \frac{\Gamma(k+1)\Gamma[\frac{\lambda}{\gamma_m}+j]}{\Gamma(j+1)\Gamma[\frac{\lambda}{\gamma_m}+k+1]} \tag{49}$$

Hence $\phi_{k,0} = k! \frac{\Gamma(\frac{\lambda}{\gamma_m})}{\Gamma(\frac{\lambda}{\gamma_m}+k+1)}$. Furthermore,

$$\phi_{n,j_z}\ \phi_{j_z,j_{z-1}}\ \ldots\ \phi_{j_1,0} = \frac{\lambda^{z+1}}{\gamma_m^{z+1}} \frac{k!\ \Gamma[\frac{\lambda}{\gamma_m}]}{\Gamma[\frac{\lambda}{\gamma_m}+k+1]} \prod_{i=1} \frac{\gamma_m}{\lambda + \gamma_m j_i} \tag{50}$$

Substituting the above, in Eq. 40 and 39, we have

$$\begin{aligned} P_+^{ss}(m_+) &= \sum_{k=m_+}^{\infty} \binom{k}{m_+} (-1)^{k-m_+} \frac{1}{k!} \left(\frac{k_m}{\gamma_m}\right)^k \frac{1/2^k}{(1-1/2^k \Psi_{k,k})} \left[\sum_{\{S_{k-1}\}} \frac{(\frac{1}{2})^{\sum j_i} \phi_{k,j_z}\ \phi_{j_z,j_{z-1}}\ \ldots\ \phi_{j_1,0}}{\prod_i (1-\frac{1}{2^{j_i}}\Psi_{j_i,j_i})} + \phi_{k,0} \right] \\ &= \sum_{k=m_+}^{\infty} \binom{k}{m_+} (-1)^{k-m_+} \frac{\frac{1}{2^k}\left(\frac{k_m}{\gamma_m}\right)^k}{(1-\frac{1}{2^k}\Psi_{k,k})} \left[\frac{\lambda}{\gamma_m}\ \frac{\Gamma(\frac{\lambda}{\gamma_m})}{\Gamma(\frac{\lambda}{\gamma_m}+k+1)} + \sum_{\{S_{k-1}\}} \frac{(\frac{1}{2})^{\sum j_i} \frac{\lambda^{z+1}}{\gamma_m^{z+1}} \frac{\Gamma(\lambda/\gamma_m)}{\Gamma(\lambda/\gamma_m+k+1)}}{\prod_i \left(1-\frac{1}{2^{j_i}}\Psi_{j_i,j_i}\right)} \prod_i \frac{\gamma_m}{\lambda+\gamma_m j_i} \right] \\ &= \sum_{k=m_+}^{\infty} \binom{k}{m_+} (-1)^{k-m_+} \frac{\frac{1}{2^k}\left(\frac{k_m}{\gamma_m}\right)^k}{(1-\frac{1}{2^k}\Psi_{k,k})}\ \frac{\Gamma(\frac{\lambda}{\gamma_m}+1)}{\Gamma(\frac{\lambda}{\gamma_m}+k+1)} \left[1 + \sum_{\{S_{k-1}\}} \prod_i \frac{\frac{1}{2^{j_i}}\Psi_{j_i,j_i}}{(1-\frac{1}{2^{j_i}}\Psi_{j_i,j_i})} \right] \\ &= \sum_{k=m_+}^{\infty} \binom{k}{m_+} (-1)^{k-m_+} \frac{\frac{1}{2^k}\left(\frac{k_m}{\gamma_m}\right)^k}{(1-\frac{1}{2^k}\Psi_{k,k})}\ \frac{\Gamma(\frac{\lambda}{\gamma_m}+1)}{\Gamma(\frac{\lambda}{\gamma_m}+k+1)} \frac{1}{\prod_i^{k-1}(1-\frac{1}{2^i}\Psi_{i,i})} \\ &= \sum_{k=m_+}^{\infty} \binom{k}{m_+} \frac{(-1)^{k-m_+}}{k!} \frac{k!\left(\frac{k_m}{2}\right)^k}{\prod_i^k(\lambda+\gamma_m i-\frac{1}{2^i}\lambda)}, \end{aligned} \tag{51}$$

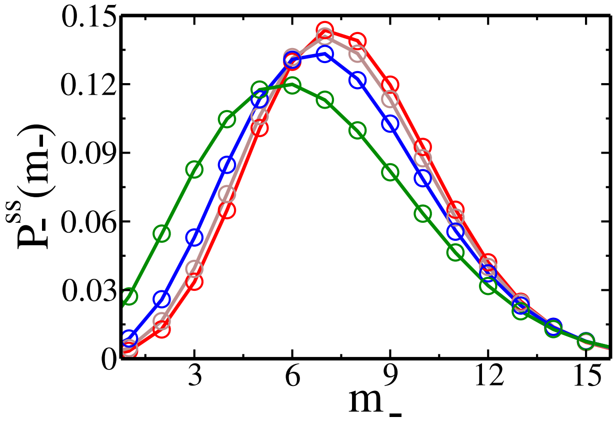

FIG. 1. Cyclo-stationary distribution $P_{-}^{ss}(m_{-})$ of mRNA for the four $g(t_s)$ shown in Fig. 3 in the main text (corresponding colours being the same)

where we have used an identity [1], with $f(j_i) = \frac{1}{2^{j_i}}\Psi_{j_i,j_i}$ as follows:

$$1 + \sum_{\{S_{k-1}\}} \prod_i \frac{f(j_i)}{(1-f(j_i))} = \frac{1}{\prod_i^{k-1}(1-f(i))} \tag{52}$$

Eq. 51 shows that $F_+^{(k)}(1) = k!\left(\frac{k_m}{2}\right)^k / \prod_i^k(\lambda + \gamma_m i - \frac{1}{2^i}\lambda)$. For obtaining the distribution before division, we note the extra factor of 2^k in Eq. 41, and that implies (comparing with Eq. 51)

$$P_-^{ss}(m_-) = \sum_{k=m_-}^{\infty} \binom{k}{m_-} \frac{(-1)^{k-m_-}}{k!} \frac{k!\,(k_m)^k}{\prod_i^k(\lambda + \gamma_m i - \frac{1}{2^i}\lambda)}. \tag{53}$$

E. Expressions of CV^2 and Skewness of the distribution $P_+^{ss}(m_+)$ of mRNA at cell birth.

Using the exact expressions of $F_+^{(1)}(1)$, $F_+^{(2)}(1)$, and $F_+^{(3)}(1)$ in Eqs. 36, 37 and 38, we have obtained the cumulants (from Eqs. 24, 26 and 28), and thus studied the mean, CV^2 and Skewness in the main text.

III. STATISTICS OF CYCLO-STATIONARY PROTEIN COUNT

A. The generating function for protein kinetics, and thereby obtaining function $\mathcal{H}$

The Master equation for the bursty protein translation is the following (with initial copy number being n'_+):

$$\frac{\partial P(n,t|n'_+)}{\partial t} = k_m\left[\sum_{r=1}^{n} \frac{b^r}{(b+1)^{r+1}} P(n-r,t|n'_+) - \frac{b}{b+1}P(n,t|n'_+)\right] + \gamma_p\left[(n+1)P(n+1,t|n'_+) - nP(n,t|n'_+)\right]. \tag{54}$$

The generating function $F(q,t|n'_+) = \sum_{j=0}^{\infty} q^n p(n,t|n'_+)$ is known to satisfy [2]

$$\frac{1}{v}\frac{\partial F}{\partial \tau} + \frac{\partial F}{\partial v} = \frac{ab}{1-bv}F \tag{55}$$

where $v = q-1$, $\tau = \gamma_p t$ and $a = \frac{k_m}{\gamma_p}$. The solution of Eq. 54 by the method of Lagrange characteristics yield [2]

$$F(q,t|n'_+) = \left(\frac{1-b(q-1)e^{-\gamma_p t_s}}{1-b(q-1)}\right)^a \times \left(1 + (q-1)\,e^{-\gamma_p t_s}\right)^{n'_+}. \tag{56}$$

Thus comparing with Eq. 5 of the main text (also see below Eq. 9), we identify the function

$$\mathcal{H} = \left(\frac{1-b(q-1)e^{-\gamma_p t_s}}{1-b(q-1)}\right)^a . \tag{57}$$

B. Obtaining the recursion relation for the coefficients $F_+^{(k)}(1)$ in the series of the distributions $P_\pm^{ss}(m_\pm)$

Using $\mathcal{H}$ from Eq. 57 in Eq. 10, and $F_+(q) = \sum_{j=0}^{\infty} \frac{(q-1)^j}{j!} F_+^{(j)}(1)$ we have

$$\begin{aligned}
F_+(q) &= \int_0^\infty dt_s g(t_s) \left(\frac{1-\frac{b}{2}(q-1)\,e^{-\gamma_p t_s}}{1-\frac{b}{2}(q-1)}\right)^a F_+\left(1+\frac{(q-1)}{2}e^{-\gamma_p t_s}\right) \\
&= \int_0^\infty dt_s g(t_s) \left(\frac{1-\frac{b}{2}(q-1)\,e^{-\gamma_p t_s}}{1-\frac{b}{2}(q-1)}\right)^a \sum_{j=0}^{\infty} \frac{F_+^j(1)\ (q-1)^j}{2^j j!} e^{-j\gamma_p t_s} \\
&= \int_0^\infty dt_s g(t_s) \sum_{l=0}^{\infty} \frac{\Gamma(a+1)}{\Gamma(a-l+1)} \left(\frac{b}{2}\right)^l \frac{(-1)^l (q-1)^l e^{-l\gamma_p t_s}}{l!} \sum_{s=0}^{\infty} \frac{\Gamma(a+s)}{\Gamma(a)} \left(\frac{b}{2}\right)^s \frac{(q-1)^s}{s!} \sum_{j=0}^{\infty} \frac{F_+^j(1) e^{-j\gamma_p t_s} (q-1)^j}{2^j j!} \\
&= \sum_{l=0}^{\infty}\sum_{j=0}^{\infty}\sum_{s=0}^{\infty} \int_0^\infty dt_s g(t_s) (-1)^l \frac{\Gamma(a+1)}{\Gamma(a-l+1)} \frac{\Gamma(a+s)}{\Gamma(a)} \left(\frac{b}{2}\right)^{l+s} (q-1)^{l+j+s} e^{-(l+j)\gamma_p t_s} \binom{a+k-1}{k} \frac{F_+^{(j)}(1)}{2^j\ l!\ s!\ j!}
\end{aligned} \tag{58}$$

We define $L_{l+j} = \int_0^\infty dt_s g(t_s) e^{-(l+j)\gamma_p t_s}$, and replace indices $l+j+s=k$, which leads to

$$F_+(q) = \sum_{k=0}^{\infty} \frac{(q-1)^k}{k!} \sum_{l=0}^{k}\sum_{j=0}^{k-l} a \frac{k!\Gamma(a+k-l-j)}{\Gamma(a-l+1)} (-1)^l \left(\frac{b}{2}\right)^{k-j} \frac{L_{l+j} F_+^{(j)}(1)}{l! j! (k-l-j)! 2^j} \tag{59}$$

A series expansion on the left side of Eq. 59 about $q=1$, and comparing with the right side, yields the desired recursion relation:

$$F_+^{(k)}(1) = ak! \sum_{l=0}^{k}\sum_{j=0}^{k-l} (-1)^l \frac{b^{k-j}}{2^k} L_{l+j} \frac{(a+k-l-j-1)!}{(a-l)!\ l!\ j!\ (k-l-j)!} F_+^{(j)}(1) \tag{60}$$

The above Eq. 60 appears as Eq. 26 in the main text. Once these coefficients $F_+^{(k)}(1)$ are solved for, they may be used to obtain the cyclo-stationary distributions

$$P_+^{(ss)}(n_+) = \sum_{k=n_+}^{\infty} \binom{k}{n_+} \frac{(-1)^{k-n_+}}{k!} F_+^{(k)}(1) \tag{61}$$

$$P_-^{(ss)}(n_-) = \sum_{k=n_-}^{\infty} \binom{k}{n_-} \frac{(-1)^{k-n_-} 2^k}{k!} F_+^{(k)}(1) \tag{62}$$

C. Expressions of first three $F_+^{(k)}(1)$ which determine exactly CV^2 and Skewness of the distribution $P_+^{ss}(n_+)$ of protein.

Firstly, $F_+^{(0)}(1) = 1$. Then using the above Eq. 60 recursively, we get

$$F_+^{(1)}(1) = ab\frac{1-L_1}{2-L_1} \tag{63}$$

then,

$$F_+^{(2)}(1) = \frac{ab^2}{4-L_2}\left[(1-L_2) + a\frac{(2-3L_1+L_1L_2)}{(2-L_1)}\right] \tag{64}$$

and then,

$$F_{+}^{(3)}(1)=\frac{1}{8-L_3}\Bigg[ab^3\Big((1+a)(2+a)-3a((1+a)L_1-(a-1)L_2)-(a-2)(a-1)L_3\Big)$$
$$+3ab^2\Big((1+a)L_1-2aL_2+(a-1)L_3\Big)F_{+}^{(1)}(1)+3ab\Big(L_2-L_3\Big)F_{+}^{(2)}(1)\Bigg] \tag{65}$$

The constants L_1, L_2 and L_3 may be evaluated given a $g(t_s)$. Then the above Eqs. 63, 64 and 65 are substituted in the Eqs. 24, 26 and 28, to obtain the cumulants and thus CV^2 and Skewness, which are studied in the main text.

IV. COMPUTATIONAL METHODS

A. Precautions to perform numerical sums of different series to obtain the coefficients $F_{+}^{(k)}(1)$ and the theoretical cyclo-stationary distributions $P_{\pm}^{ss}$

In this work we had to sum various series to obtain the desired coefficients and functions.

The equations for the coefficients $F_{+}^{(k)}(1)$ appear as Eq. 35 for mRNA, and Eq. 60 for protein, and are of the form

$$F_{+}^{(k)}(1)=\sum_{j=1}^{k-1}c_{k,j}F_{+}^{(j)}(1) \tag{66}$$

As the values of $F_{+}^{(j)}(1)$ grow very fast with j we loose precision soon in ordinary calculations. A better way to store large numbers is by taking logarithm, and we do so for terms in Eq. 66. Thus we store terms

$$u_{k,j}=\ln c_{k,j}+\ln F_{+}^{(j)}(1). \tag{67}$$

We specify very high precision for such calculation and storage in Mathematica (through the SetPrecision[d] command) up to $d=100$ decimal places in case of mRNA and $d=200$ decimal places for protein. We reconstruct back the coefficient

$$F_{+}^{(k)}(1)=\sum_{j=1}^{k-1}e^{u_{k,j}}. \tag{68}$$

Once the coefficients $F_{+}^{(k)}(1)$ are obtained by the above method, up to some desired k, we put them in the series in Eqs. 40 and 41 for mRNA, and Eqs. 61, 62 for protein, to obatin the cyclo-stationary distributions. For mRNA, convergence was attained for $\sim 30-50$ terms in the series of $P_{\pm}^{ss}(m_{\pm})$.

For protein, ordinary sum of the series of $P_{\pm}^{ss}(n_{\pm})$ were not enough with reasonable values of k. We used Borel sum formula as follows:

$$P_{\pm}^{ss}(n_{\pm})=\sum_{k=n_{\pm}}^{\infty}f(k,n_{\pm})=P_{\pm}^{ss}(n_{\pm})\Big|_{\text{Borel}}=\lim_{t\to\infty}e^{-t}\sum_{n=0}^{M\to\infty}\frac{t^n}{n!}\sum_{k'=0}^{n}f(k'+n_{\pm},n_{\pm}) \tag{69}$$

In calculations we choose $M\sim 200-250$ and $t=30$ to obtain convergence of the protein cyclo-stationary distributions.

B. Kinetic Monte Carlo Simulations

We perform Kinetic Monte Carlo (KMC) or Gillespie [3] simulations for the various models governing the transcription or translation models of mRNA and protein in this paper, undergoing Binomial partitioning after random time intervals t_s drawn from some distribution $g(t_s)$. Thus, at any instant, there are three possible events to either increase, decrease, or reset the copy number (due to cell partition). We typically use $\sim 10^7$ histories for getting the data for various distributions and cumulants, which were then compared with the theory.

For lineage protein distribution of any age, we simulate several copies ($\sim 10^5$) of single lineage samples and stop them at a fixed large time (roughly after ~ 20 generations) such that the cyclo-stationary state is reached. We get the protein counts from the samples and use them to obtain the desired distribution. We also maintain clocks with every cell and at the final time, get their current age to obtain the age distribution.

Similarly, we simulate an exponentially growing population by keeping track of the cell age, division times, and protein count for each cell. After a long time, we just stopped the simulation and acquired the data from the population to plot the age distribution and the age-independent protein number distribution.

V. AGE DEPENDENT CYCLO-STATIONARY DISTRIBUTIONS

The cyclo-stationary distribution $P^{ss}(y,\tau)$ of cells at an age τ before the next cell division, may written with respect to $P^{ss}_{+}(y_+)$ at birth, as follows:

$$P^{ss}(y,\tau) = \sum_{y_+} P^{ss}_{+}(y_+)p(y,\tau|y_+). \tag{70}$$

Using its generating function $\tilde{G}(q,\tau) = \sum_y q^y P^{ss}(y,\tau)$, from Eq. 70 (and using the same steps as in Eqs. 9 and 10)

$$\begin{aligned}\tilde{G}(q,\tau) = \sum_{y_+} P^{ss}_{+}(y_+)F(q,\tau|y_+) &= \mathcal{H}(q-1,\gamma_y\tau)\ F_+(1+(q-1)e^{-\gamma_y\tau}) \\ &= \sum_k \frac{(q-1)^k}{k!}G^{(k)}_y(\tau)\end{aligned} \tag{71}$$

The defined quantities $G^{(k)}_y(\tau)$ are obtained in Eq. 71 by expressing $F_+(q) = \sum_{j=0}^{\infty}\frac{(q-1)^j}{j!}F^{(j)}_+(1)$, and expanding $\mathcal{H}$ as a power series of $(q-1)$. The resulting expression of $G^{(k)}_y(\tau)$ are of the form of the integrands of Eqs. 33 or 58 (without the $\int_0^\infty dt_S g(t_s)/2^k$ factors), and are explicitly given for mRNA and protein in Eqs. 30 and 31 in the main text. Finally it is easy to obtain the desired age-dependent distributions in terms of $G^{(k)}_y(\tau)$ as

$$P^{ss}(y,\tau) = \frac{1}{y!}\frac{\partial^y\tilde{G}}{\partial q^y}\bigg|_{q=0} = \sum_{k=y}^{\infty}\binom{k}{y}\frac{(-1)^{k-y}}{k!}G^{(k)}_y(\tau). \tag{72}$$

VI. GENERATING FUNCTIONS OF PROTEIN DISTRIBUTION AT BIRTH, FOR DETERMINISTIC PARTITIONING, AND DETERMINISTIC GENE EXPRESSION

If we have a deterministic partitioning, we would replace the binomial distribution $B(\tilde{y}_+,\frac{1}{2},x_+)$ in Eqs. 8 and 9 by $\delta_{x_+,\tilde{y}_+/2}$ and as a result

$$\begin{aligned}F_+(q) &= \sum_{y'_+}\int_0^\infty dt_s g(t_s)\sum_{\tilde{y}_+} q^{\tilde{y}_+}p(\tilde{y}_+,t_s|y'_+)\sum_{x_+=0}^{\tilde{y}_+}\frac{1}{q^{x_+}}\delta_{x_+,\frac{\tilde{y}_+}{2}}P^{ss}_+(y'_+) \\ &= \sum_{y'_+}\int_0^\infty dt_s g(t_s)\left(\sum_{\tilde{y}_+} q^{\tilde{y}_+/2}p(\tilde{y}_+,t_s|y'_+)\right)P^{ss}_+(y'_+) \\ &= \int_0^\infty dt_s g(t_s)\sum_{y'_+}F\left(\sqrt{q},t_s\middle|y'_+\right)P^{ss}_+(y'_+).\end{aligned} \tag{73}$$

For proteins $y\equiv n$, and using the appropriate $F(q,t|n'_+)$ from Eq. 56, we have the counterpart of Eq. 10 as:

$$F_+(q) = \int_0^\infty dt_s g(t_s)\left(\frac{1-b\left(\sqrt{q}-1\right)e^{-\gamma_p t_s}}{1-b\left(\sqrt{q}-1\right)}\right)^a F_+\left(1+(\sqrt{q}-1)e^{-\gamma_p t_s}\right) \tag{74}$$

If in addition to deterministic partitioning, one also has deterministic protein kinetics

$$\frac{dn}{dt} = k_m b - \gamma_p n, \tag{75}$$

then $\tilde{n}_+ = n'_+ e^{-\gamma_p t_s} + ab(1 - e^{-\gamma_p t_s})$ and F in Eq. 73 gets replaced by $q^{\frac{1}{2}(\lambda_p + n'_+ e^{-\gamma_p t_s})}$ where $\lambda_p = ab(1 - e^{-\gamma_p t_s})$, i.e.

$$\begin{aligned}
F_+(q) &= \sum_{n'_+} \int_0^\infty dt_s g(t_s) \left(\sum_{\tilde{n}_+} q^{\tilde{n}_+/2} p(\tilde{n}_+, t_s | n'_+) \right) P_+^{ss}(n'_+) \\
&= \int_0^\infty dt_s g(t_s) \sum_{n'_+} P_+^{ss}(n'_+) q^{\frac{1}{2}(\lambda_p + n'_+ e^{-\gamma_p t_s})} \\
&= \int_0^\infty dt_s g(t_s) q^{\frac{1}{2}\lambda_p} \sum_{n'_+} P_+^{ss}(n'_+) \left(q^{\frac{1}{2} e^{-\gamma_p t_s}} \right)^{n'_+} \\
&= \int_0^\infty dt_s g(t_s) q^{\frac{1}{2}\lambda_p} F_+(q^{\frac{1}{2} e^{-\gamma_p t_s}}) \qquad (76)
\end{aligned}$$

The moment $\langle n_+ \rangle = q \frac{\partial}{\partial q} F_+(q) \Big|_{q=1}$ and $\langle n_+^2 \rangle = q \frac{\partial}{\partial q} q \frac{\partial}{\partial q} F_+(q) \Big|_{q=1}$, and hence taking derivatives of on two sides of Eq. 74 and 76 respectively, and setting $q = 1$, we may obtain the moments in the two cases above. The explicit forms of CV^2 thus obtained are shown in Eq. 32 in the main text (corresponding to Eq. 74) and in Eq. 34 in the main text (corresponding to Eq. 76).

[1] K. Rijal, N. I. C. Müller, E. Friauf, A. Singh, A. Prasad, and D. Das, Exact Distribution of the Quantal Content in Synaptic Transmission, Phys. Rev. Lett. **132**, 228401 (2024).

[2] V. Shahrezaei and P. S. Swain, Analytical distributions for stochastic gene expression, Proc. Natl. Acad. Sci. USA **105**, 17256 (2008).

[3] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. **81**, 2340 (1977).