The nonlinear dynamics and fluctuations of mRNA levels in cell cycle coupled transcription

Qiwen Sun[1,2], Feng Jiao[1,2], Genghong Lin[1,2], Jianshe Yu[1], Moxun Tang[2]*

1 Center for Applied Mathematics, Guangzhou University, Guangzhou, 510006, China, **2** Department of Mathematics, Michigan State University, East Lansing, Michigan, United States of America

* tang@math.msu.edu

Abstract

Gene transcription is a noisy process, and cell division cycle is an important source of gene transcription noise. In this work, we develop a mathematical approach by coupling transcription kinetics with cell division cycles to delineate how they are combined to regulate transcription output and noise. In view of gene dosage, a cell cycle is divided into an early stage $\mathbb{S}_1$ and a late stage $\mathbb{S}_2$. The analytical forms for the mean and the noise of mRNA numbers are given in each stage. The analysis based on these formulas predicts precisely the fold change r^* of mRNA numbers from $\mathbb{S}_1$ to $\mathbb{S}_2$ measured in a mouse embryonic stem cell line. When transcription follows similar kinetics in both stages, r^* buffers against DNA dosage variation and $r^* \in (1, 2)$. Numerical simulations suggest that increasing cell cycle durations up-regulates transcription with less noise, whereas rapid stage transitions induce highly noisy transcription. A minimization of the transcription noise is observed when transcription homeostasis is attained by varying a single kinetic rate. When the transcription level scales with cellular volume, either by reducing the transcription burst frequency or by increasing the burst size in $\mathbb{S}_2$, the noise shows only a minor variation over a wide range of cell cycle stage durations. The reduction level in the burst frequency is nearly a constant, whereas the increase in the burst size is conceivably sensitive, when responding to a large random variation of the cell cycle durations and the gene duplication time.

Author summary

Gene transcription in single cells is inherently a stochastic process, resulting in a large variability in the number of transcripts and constituting the phenotypic heterogeneity in cell population. Cell division cycle has global effects on transcriptional outputs, and is thought to be an additional source of transcription noise. In this work, we develop a hybrid model to delineate the combined contribution of transcription activities and cell divisions in the variability of transcript counts. By working with the analytical forms of the mean and the noise of mRNA numbers, we show that if the transcription kinetic rates do not change considerably, then the average mRNA level is increased about 1 to 2 folds from earlier to later cell cycle stages. When transcription homeostasis is attained by varying a single kinetic rate between the two cell cycle stages, we find no significant changes in the

transcription noise, and the homeostasis nearly minimizes the noise. In our continuous study on the transcript concentration homeostasis that the transcription level scales with the cellular volume, we find only minor variations of the noise if the homeostasis is maintained either by reducing the transcription burst frequency or by increasing the burst size in late cell cycle phase, in the face of a large cell cycle stage duration variation. The reduction in the burst frequency is relative robust, while the increase in the burst size is conceivably sensitive, to the large random variation of the cell cycle durations and the gene duplication time.

Introduction

Single cell studies over last decades have shown that gene transcription is inherently a stochastic process in a bursting fashion [1–5]. The transcriptional bursting, whereby a gene promoter transits randomly between short periods of mRNA production and long periods of no productions, has been widely studied and invoked to explain how the fluctuation of mRNA molecules arises among single cells of identical genes [6–8]. Early studies on the origin of variability in gene expression found that the noise is not solely due to the randomness in reactions intrinsic to gene expression [9].

Recent experiments have suggested that cell division cycle is an important source of gene expression noise [10–13]. In virtually all cells, from bacteria to mammalian cells, a conserved class of genes is involved in cell cycle stage-specific gene expression. For instance, *SWI5* and *CLB2* are responsible for mitotic progression, whose transcripts are stable during the interphase, but exhibit a 30-fold increase in degradation in the mitosis phase [12]. In budding yeast, acetylation of histone 3 suppresses transcription activity to buffer changes in DNA dose for expression homeostasis of other genes during DNA replication [13]. During cell division processes, genome duplication involves DNA dosage increase at discrete times in *S* phase, and introduces considerable variations in gene copies [13–15]. Moreover, the time spent between two successive cell-division events [11], the DNA replication catalyzed by DNA polymerases [16, 17], the variation in transcription kinetics between different cell cycle stages [9, 15, 18], and the partition of molecules between two daughter cells [19], are all observed to be stochastic and may contribute to cell-to-cell variability in transcript counts.

It remains largely unexplored how these random events govern mRNA outputs and their fluctuation among individual cells [1]. In this work, we initiate a mathematical approach by coupling the classical two-state model with cell division cycles to delineate the combined contribution of transcription activities and cell divisions in the variability of transcript counts [4, 6, 20]. In view of gene dosage, a cell cycle is divided into $\mathbb{S}_1$ and $\mathbb{S}_2$ stages. In each stage, the target gene transits randomly between active and inactive states with constant rates. As usual, we use the mean, the noise, and the noise strength to characterize stochastic gene transcription. For a given random variable N, we denote by $\mathrm{E}[N]$, $\mathrm{E}[N^2]$, and $\mathrm{Var}[N] = \mathrm{E}[N - \mathrm{E}[N]]^2$ its mean, the second moment, and variance, respectively. Its noise and the noise strength are defined by

$$\eta^2(N) = \frac{\mathrm{Var}[N]}{(\mathrm{E}[N])^2} = \frac{\mathrm{E}[N^2] - (\mathrm{E}[N])^2}{(\mathrm{E}[N])^2}, \quad \text{and} \quad \Phi(N) = \frac{\mathrm{Var}[N]}{\mathrm{E}[N]}. \tag{1}$$

We will formulate the master equations for the model and derive the differential equations of the mean and the second moment. The analytical forms of the mean, the noise, and the noise strength at steady-state will be given.

We measure the fold change of mRNA copy numbers from $\mathbb{S}_1$ to $\mathbb{S}_2$ by $r^* = m_2^*/m_1^*$, where m_1^* and m_2^* are the mean transcription levels at the two stages. Although r^* may take any pre-scribed value in theory, we find that when the transcription kinetic rates are similar in the two stages, the fold change buffers against the DNA dosage variation and stays within (1, 2), as observed in yeast [12] and mammalian cells [15]. Furthermore, if stage transitions are consid-erably slower than transcription state transitions and mRNA turnover, then $r^* \approx 2$. The accu-racy of our theoretical results is tested by numerical examples that generate nearly the same fold change measured in a mouse embryonic stem cell line [15]. Increasing either of the cell cycle durations up-regulates transcription with less noise, and rapid transitions between cell cycle stages are a major source of highly noisy transcription. Our numerical examples also demonstrate that transcription homeostasis does not bring significant changes in transcription noise. If transcription homeostasis is attained by varying a single kinetic rate in the two cell cycle stages, then the homeostasis nearly minimizes transcription noise. Motivated by increas-ing evidences that many cellular processes depend mainly on the concentration rather than the absolute number of enzymes [18, 21, 22], we continue to study the noise profile when the transcript concentration homeostasis is maintained. Our analysis reveals an interesting phe-nomenon that the transcription noise is relatively stable when the concentration homeostasis is maintained, either by reducing the transcription burst frequency or by increasing the burst size in late cell cycle phase, over a wide range of cell cycle stage durations. The reduction degree in the burst frequency is nearly a constant, while the increase in the burst size is con-ceivably sensitive, when responding to a large random variation of the cell cycle durations and the gene duplication time.
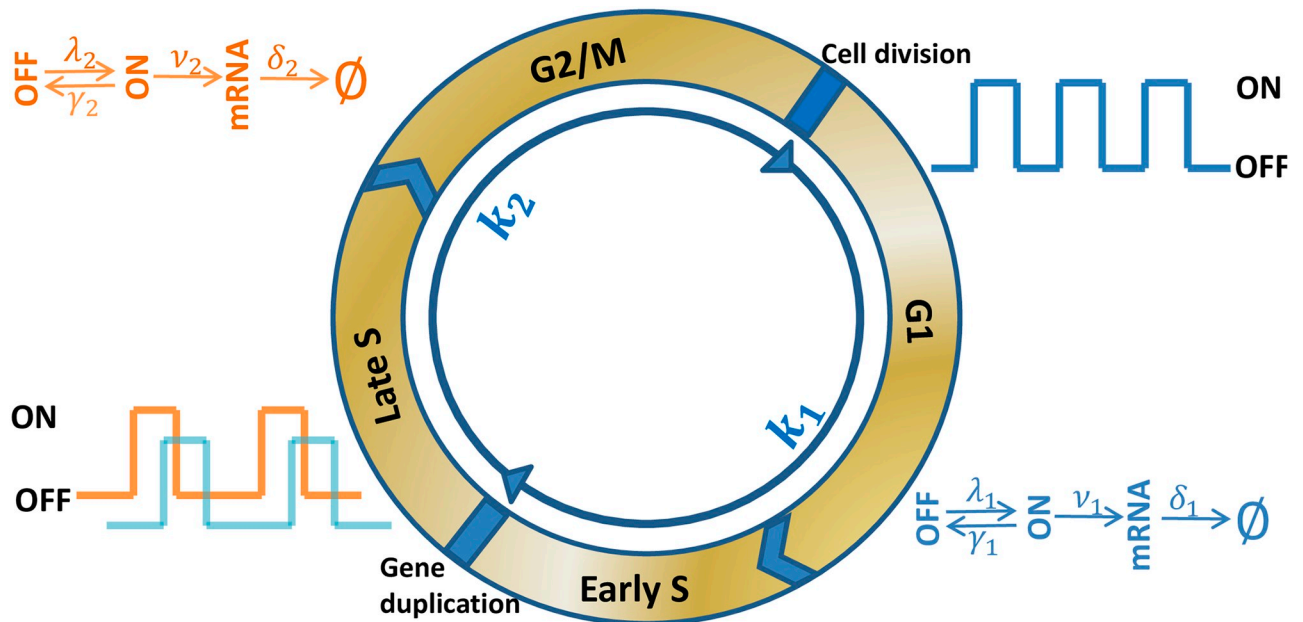
Models

Coupling transcription with cell division cycle

In past two decades, the two-state model has been a prevailing tool to characterize stochastic gene transcription in single cells, from bacteria, yeast, to mammalian cells [4–6, 8, 20, 23]. In the model, as depicted in the diagram

$$\text{gene OFF} \underset{\gamma}{\overset{\lambda}{\rightleftarrows}} \text{gene ON} \xrightarrow{v} \text{mRNA} \xrightarrow{\delta} \emptyset, \tag{2}$$
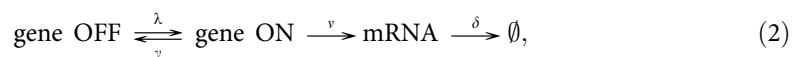
it is postulated that the gene promoter transits randomly between inactive (gene OFF) and active (gene ON) states with constant activation rate $\lambda > 0$ and inactivation rate $\gamma > 0$. The transcripts are produced only when the gene is active with a synthesis rate $v > 0$, and are turned over with a degradation rate $\delta > 0$. Apparently, as the four rates are all assumed to be constants, the transcription described by the model is independent of many important cellular processes such as cell growth and cell division.

Actively dividing eukaryote cells go through several stages known collectively as the cell division cycle, including Gap 1 phase (G_1) for cell growth, the synthesis phase (S) for DNA rep-lication, Gap 2 phase (G_2) for DNA repairing, and the mitotic phase (M) for cell division; see Fig 1. During S phase, each gene is duplicated into two copies that are transcribed indepen-dently in the same cell [15]. During M phase, a cell is divided into two daughter cells and residual mRNA molecules are randomly partitioned. Cell division cycle has global effects on mRNA and protein synthesis, and is also an important source of gene expression noise [10–13]. In recent years, many real-time monitoring methods, such as single molecule fluorescent in situ hybridization (smFISH), have been developed to estimate mRNA copy numbers in dif-ferent cell cycle stages. In mouse embryonic stem cells, nascent Oct4 and Nanog mRNAs were measured in different phases using smFISH method [15]. It was found that the ratio of the

Fig 1. Coupling gene transcription with cell cycle. Actively dividing eukaryote cells go through G_1, S, G_2 and M phases in one cell cycle. In our model, we divide one cycle into two stages: $\mathbb{S}_1$ (from last division to gene duplication) and $\mathbb{S}_2$ (from gene duplication to next division). Cells orderly rotate between $\mathbb{S}_1$ and $\mathbb{S}_2$ stages with constant rates κ_1 and κ_2. The promoter transits randomly between active and inactive states, and the transcription kinetics changes with the cell cycle stages. During $\mathbb{S}_1$ stage, the kinetics is parameterized by activation rate λ_1, inactivation rate γ_1, synthesis rate ν_1, and mRNA degradation rate δ_1. After DNA replication, the gene is duplicated into two identical copies that are transcribed independently with constant rates λ_2, γ_2, ν_2 and δ_2 in the same cell during $\mathbb{S}_2$ stage.

average number of mRNA copies in G_2 phase to the average in G_1 phase is 1.28 ± 0.09 for Oct4 mRNA, and 1.51 ± 0.15 for Nanog mRNA. In yeast cells, CLB2 mRNAs accumulate apace in late S phase and are degraded almost completely before cytokinesis [12]. From the measurements of [12], we estimated that the median of cytoplasmic CLB2 mRNA copy numbers is ~ 10 in G_2/M phase, and ~ 5 in S phase. It remains an essential and widely open question to quantify how the transition of cell cycle phases, the variation of DNA content and transcription kinetics in different phases, and the random partition of mRNAs in daughter cells affect the dynamics and noise of gene transcription.

In this work, we initiate a quantitative approach to this important question by developing a model that couples gene transcription with cell cycles. During DNA replication in S phase, the two complementary strands in each double helix are separated and serve as templates for the production of their counterparts. After the completion of the whole DNA replication process, which takes hours in some cells [24], each gene copy is doubled with two copies. Normally, the duplication of a single gene takes much shorter time and is completed within seconds to minutes [17, 25]. For instance, the genome of *Escherichia coli* K12 has ~ 4.64 million base pairs with ~ 4375 genes [26], and is replicated at $\sim 470 \pm 180$ bp/s [17]. The average duplication time of each gene takes $1.63 \sim 3.66$ seconds. In our model, we treat the short duplication process of our target gene as instantaneous, and accordingly, divide a cell cycle into two stages:

- $\mathbb{S}_1$ **stage**, consisting of the whole G_1 phase and the early S phase until the gene of our interest is duplicated.

- $\mathbb{S}_2$ **stage**, consisting of the late S phase after the gene is duplicated, and G_2/M phases.

We make the following assumptions to complete the description of the model:

i. The durations T_1 in $\mathbb{S}_1$ stage and T_2 in $\mathbb{S}_2$ stage are independently and exponentially distributed with respective rates $\kappa_1 > 0$ and $\kappa_2 > 0$.

ii. At the beginning of each stage, the transcription activity is turned off and the gene remains in the OFF state.

iii. In $\mathbb{S}_1$ stage, the transcription of the single gene copy is described by the two-state model with constant rates $\lambda_1, \gamma_1, \nu_1$ and δ_1.

iv. In $\mathbb{S}_2$ stage, the two identical copies of the gene are transcribed separately and independently, and the transcriptions are described by the two-state model with constant rates λ_2, γ_2, ν_2 and δ_2.

v. At the end of $\mathbb{S}_2$ stage, each mRNA molecule has an equal probability to be distributed to the two daughter cells; see Fig 1.

We do not assume constant durations in $\mathbb{S}_1$ and $\mathbb{S}_2$ stages in (i), because the time spent in each cell cycle phase is often not fixed, and the timing for the duplication of the target gene is random. The cell cycle duration in mouse embryonic stem cells measured by flow cytometry varies in $11 \sim 16$ hours [27, 28], that are roughly distributed in G_1 (26%), S (52%), and G_2/M (22%) estimated by the percentage of cells in these phases [28]. The times spent in cell cycle phases were also measured by time-lapse microscopy and single cell tracking in T and B lymphocytes from reporter mice, and the total division time data were found to be well approximated by the sum of consecutive independent exponential and Gaussian distributions [11]. We assume that the transcription is turned off at the beginning of each stage, as DNA synthesis is catalyzed by DNA polymerase in nucleosomes, and during late $\mathbb{S}_2$ stage, the chromatin shrinks into chromosome [29–31]. In either case, transcription factors and RNA polymerase II are usually prevented from reaching to gene promoters to initiate transcription [32]. Assumption (v) is equivalent to the binomial distribution of mRNA molecules in the two daughter cells, which has been assumed in most theoretical models, and supported by recent experiments. The partition in *Escherichia coli* measured by the MS2-GFP reporter strongly supports the assumption that each mRNA copy goes to one of the two daughter cells with equal probability [33].

The master equations

The transcription state of a gene of our interest in a single cell at a time $t \geq 0$ can be characterized by the number of active gene copies, the cell cycle stage, and its mRNA copy number. Without loss of generality, we assume that the gene has exactly one copy in $\mathbb{S}_1$ stage, and two copies in $\mathbb{S}_2$ stage, in any single cell of an isogenic cell population. We let $I(t)$ denote the number of active genes in a cell. In $\mathbb{S}_1$ stage, $I(t) = 0$ if the gene is OFF, and $I(t) = 1$ if it is ON. In $\mathbb{S}_2$ stage, $I(t) = 0$ if the two gene copies are OFF, $I(t) = 2$ if both are ON, and $I(t) = 1$ in the remaining cases. We let $U(t)$ specify the cell cycle stage, with $U(t) = 1$ in $\mathbb{S}_1$ stage, and $U(t) = 2$ in $\mathbb{S}_2$ stage. Let $M(t)$ denote the mRNA copy number for the gene in one cell. Then the transcription state can be fully quantified by the following joint probabilities

$$P_1(i, m, t) = \text{Prob}\{I(t) = i, M(t) = m, U(t) = 1\}, \qquad i = 0, 1; m = 0, 1, 2, \cdots, \qquad (3)$$

$$P_2(i, m, t) = \text{Prob}\{I(t) = i, M(t) = m, U(t) = 2\}, \qquad i = 0, 1, 2; m = 0, 1, 2, \cdots. \qquad (4)$$

For clarity and simplicity in the following calculations, we assume that all cells in the isogenic population are synchronized at the beginning of $\mathbb{S}_1$ stage, and count only newly

produced mRNA molecules from time zero. Accordingly, we have the initial condition

$$P_1(0,0,0) = 1, \quad P_1(0,m,0) = 0, \quad m > 0,$$
$$P_1(1,m,0) = P_2(0,m,0) = P_2(1,m,0) = P_2(2,m,0) = 0, \quad m \geq 0. \tag{5}$$

By using the standard procedure in stochastic process, we calculate the time evolutions of these probabilities based on the basic assumptions (i)-(v) in our model and derive the master equations:

$$P_1'(0,m,t) = \gamma_1 P_1(1,m,t) - (m\delta_1 + \lambda_1 + \kappa_1)P_1(0,m,t)$$
$$+ (m+1)\delta_1 P_1(0,m+1,t) + \kappa_2 \sum_{n=m}^{\infty} \left(\frac{1}{2}\right)^n \binom{n}{m} P_2(n,t), \tag{6}$$

$$P_1'(1,m,t) = \lambda_1 P_1(0,m,t) - (v_1 + m\delta_1 + \gamma_1 + \kappa_1)P_1(1,m,t)$$
$$+ v_1 P_1(1,m-1,t) + (m+1)\delta_1 P_1(1,m+1,t), \tag{7}$$

$$P_2'(0,m,t) = \kappa_1 P_1(m,t) - (m\delta_2 + 2\lambda_2 + \kappa_2)P_2(0,m,t)$$
$$+ (m+1)\delta_2 P_2(0,m+1,t) + \gamma_2 P_2(1,m,t), \tag{8}$$

$$P_2'(1,m,t) = 2\lambda_2 P_2(0,m,t) + 2\gamma_2 P_2(2,m,t) + (m+1)\delta_2 P_2(1,m+1,t)$$
$$+ v_2 P_2(1,m-1,t) - (v_2 + m\delta_2 + \lambda_2 + \gamma_2 + \kappa_2)P_2(1,m,t), \tag{9}$$

$$P_2'(2,m,t) = \lambda_2 P_2(1,m,t) - (2v_2 + m\delta_2 + 2\gamma_2 + \kappa_2)P_2(2,m,t)$$
$$+ 2v_2 P_2(2,m-1,t) + (m+1)\delta_2 P_2(2,m+1,t). \tag{10}$$

The last expression $P_2(n,t)$ in (6), defined by

$$P_2(n,t) = P_2(0,n,t) + P_2(1,n,t) + P_2(2,n,t),$$

gives the probability that the cell resides on $\mathbb{S}_2$ stage with n transcripts, and $P_1(m,t)$ in (8), defined by $P_1(m,t) = P_1(0,m,t) + P_1(1,m,t)$, represents the probability that the cell resides on $\mathbb{S}_1$ stage with m copies of mRNA molecules. The technical steps leading to (6)–(10) are given in S1 Text.

The determination of the moment functions

The transcription dynamics of a gene in a cell population is best characterized by the mean value $m(t) = \mathbf{E}[M(t)]$ of the random process $M(t)$ that counts the number of its mRNA copies. The second moment $\mu(t) = \mathbf{E}[M^2(t)]$ is essential in the calculation of its noise that quantifies the fluctuation of mRNA copy numbers among individual cells. More importantly, as the cell division cycle is integrated into our model, we can extend $m(t)$ and $\mu(t)$ to the two cell cycle stages $\mathbb{S}_1$ and $\mathbb{S}_2$. The comparison of these quantities in the two stages can help us understand how the gene duplication contributes to the variation of transcription levels and noises. To start with, we give the formal definitions of these concepts and present the differential equations that provide a framework from which they can be solved analytically.

For this purpose, we need various probabilities by adding the joint probabilities $P_j(i,m,t)$ introduced in (3)–(4) when i, j, or m runs through all possible values. We use a conventional simplification of notations: If any of i, j and m is removed from $P_j(i,m,t)$, then the new

probability is defined by summing $P_j(i, m, t)$ over the range of the removed index. For instance,

$$P_1(m, t) = P_1(0, m, t) + P_1(1, m, t),$$
$$P_2(m, t) = P_2(0, m, t) + P_2(1, m, t) + P_2(2, m, t) \tag{11}$$

are the respective probabilities that the cell resides on $\mathbb{S}_1$ and $\mathbb{S}_2$ stages with m copies of mRNA molecules, without specifying the promoter state. A further summation of the two probabilities in (11) defines

$$P(m, t) = P_1(m, t) + P_2(m, t) \tag{12}$$

as the probability that there are m copies of mRNA molecules in the cell. Similarly, we can define $P_1(i, t)$ and $P_2(i, t)$. To avoid the confusion with these probabilities defined in (11), we change them to $P_{1i}(t)$ and $P_{2i}(t)$ with

$$P_{1i}(t) = \sum_{m=0}^{\infty} P_1(i, m, t), \quad P_{2i}(t) = \sum_{m=0}^{\infty} P_2(i, m, t). \tag{13}$$

By adding the probabilities in (13) we have

$$P_1(t) = P_{10}(t) + P_{11}(t), \quad P_2(t) = P_{20}(t) + P_{21}(t) + P_{22}(t) \tag{14}$$

as the respective probabilities that the cell resides on $\mathbb{S}_1$ and $\mathbb{S}_2$ stages. By adding the master Eqs (6)–(10) in m, we obtain a closed system of $P_{1i}(t)$ and $P_{2i}(t)$,

$$\begin{cases} P'_{10}(t) = & \kappa_2 P_2(t) + \gamma_1 P_{11}(t) - (\lambda_1 + \kappa_1) P_{10}(t), \\ P'_{11}(t) = & \lambda_1 P_{10}(t) - (\gamma_1 + \kappa_1) P_{11}(t), \\ P'_{20}(t) = & \kappa_1 P_1(t) + \gamma_2 P_{21}(t) - (2\lambda_2 + \kappa_2) P_{20}(t), \\ P'_{21}(t) = & 2\lambda_2 P_{20}(t) - (\lambda_2 + \gamma_2 + \kappa_2) P_{21}(t) + 2\gamma_2 P_{22}(t), \\ P'_{22}(t) = & \lambda_2 P_{21}(t) - (2\gamma_2 + \kappa_2) P_{22}(t). \end{cases} \tag{15}$$

The initial condition for this system can be derived by a summation of the initial data given in (5). This linear system of ordinary differential equations with constant coefficients can be solved analytically, and its solution subject to the corresponding initial condition determines uniquely $P_{1i}(t)$ and $P_{2i}(t)$.

Due to the technical complexity, we break down the process of determining $m(t)$, $\mu(t)$, and their extensions in $\mathbb{S}_1$ and $\mathbb{S}_2$ in several steps, and move most involving calculations to S1 Text.

Step 1: The determination of the mean level $m(t)$: With $P_1(m, t)$, $P_2(m, t)$, and $P(m, t)$ defined in (11) and (12), we have

$$m(t) = \mathbf{E}[M(t)] = \sum_{m=0}^{\infty} m P(m, t) = n_1(t) + n_2(t), \tag{16}$$

where

$$n_1(t) = \sum_{k=0}^{\infty} k P_1(k, t), \quad \text{and} \quad n_2(t) = \sum_{k=0}^{\infty} k P_2(k, t). \tag{17}$$

As we show in S1 Text, $n_1(t)$ and $n_2(t)$ satisfy the following system of inhomogeneous linear ordinary differential equations with constant coefficients:

$$
\begin{cases}
n_1'(t) = & -(\delta_1 + \kappa_1)n_1(t) + \dfrac{\kappa_2}{2}n_2(t) + v_1 P_{11}(t), \\[2mm]
n_2'(t) = & \kappa_1 n_1(t) - (\delta_2 + \kappa_2)n_2(t) + v_2[P_{21}(t) + 2P_{22}(t)].
\end{cases}
\tag{18}
$$

As $P_{11}(t)$, $P_{21}(t)$, and $P_{22}(t)$ can be solved uniquely from (15), we can find $n_1(t)$ and $n_2(t)$ by solving (18) subject to the initial condition $n_1(0) = n_2(0) = 0$, and find $m(t)$ by (16).

Step 2: The determination of the second moment $\mu(t)$: Similar to the definition of $m(t)$ in (16), we have

$$
\mu(t) = \mathbf{E}[M^2(t)] = \sum_{m=0}^{\infty} m^2 P(m,t) = \omega_1(t) + \omega_2(t),
\tag{19}
$$

where

$$
\omega_1(t) = \sum_{k=0}^{\infty} k^2 P_1(k,t), \quad \text{and} \quad \omega_2(t) = \sum_{k=0}^{\infty} k^2 P_2(k,t).
\tag{20}
$$

As we show in S1 Text, the time evolutions of $\omega_1(t)$ and $\omega_2(t)$ are given by the system

$$
\begin{cases}
\omega_1'(t) = & -(2\delta_1 + \kappa_1)\omega_1(t) + \dfrac{\kappa_2}{4}\omega_2(t) + \delta_1 n_1(t) \\[2mm]
& + \dfrac{\kappa_2}{4}n_2(t) + v_1[2n_{11}(t) + P_{11}(t)], \\[2mm]
\omega_2'(t) = & \kappa_1 \omega_1(t) - (2\delta_2 + \kappa_2)\omega_2(t) + \delta_2 n_2(t) \\[2mm]
& + v_2[P_{21}(t) + 2P_{22}(t) + 2n_{21}(t) + 4n_{22}(t)],
\end{cases}
\tag{21}
$$

where

$$
n_{1i}(t) = \sum_{m=0}^{\infty} m P_1(i,m,t), \quad i = 0, 1, \quad n_{2i}(t) = \sum_{m=0}^{\infty} m P_2(i,m,t), \quad i = 0, 1, 2,
\tag{22}
$$

and

$$
n_1(t) = n_{10}(t) + n_{11}(t), \qquad n_2(t) = n_{20}(t) + n_{21}(t) + n_{22}(t).
$$

Apparently, (21) is not a closed system, and finding $\omega_1(t)$ and $\omega_2(t)$ requires the following system of $n_{1i}(t)$ and $n_{2i}(t)$:

$$
\begin{cases}
n_{10}'(t) = & \dfrac{\kappa_2}{2}n_2(t) - (\delta_1 + \lambda_1 + \kappa_1)n_{10}(t) + \gamma_1 n_{11}(t), \\[2mm]
n_{11}'(t) = & \lambda_1 n_{10}(t) + v_1 P_{11}(t) - (\delta_1 + \gamma_1 + \kappa_1)n_{11}(t), \\[2mm]
n_{20}'(t) = & \kappa_1 n_1(t) + \gamma_2 n_{21}(t) - (\delta_2 + 2\lambda_2 + \kappa_2)n_{20}(t), \\[2mm]
n_{21}'(t) = & 2\lambda_2 n_{20}(t) + 2\gamma_2 n_{22}(t) + v_2 P_{21}(t) - (\delta_2 + \lambda_2 + \gamma_2 + \kappa_2)n_{21}(t), \\[2mm]
n_{22}'(t) = & \lambda_2 n_{21}(t) + 2v_2 P_{22}(t) - (\delta_2 + 2\gamma_2 + \kappa_2)n_{22}(t),
\end{cases}
\tag{23}
$$

This system is obtained by multiplying (6)–(10) with m and then taking sums. As $P_{11}(t)$, $P_{21}(t)$, and $P_{22}(t)$ can be solved from (15), it is a closed system of $n_{1i}(t)$ and $n_{2i}(t)$. By

substituting its unique solution subject to the zero initial condition into (21), we can determine $\omega_1(t)$ and $\omega_2(t)$, and therefore the second moment $\mu(t)$.

Step 3: The moment functions on $\mathbb{S}_1$ and $\mathbb{S}_2$ stages: To extend the definitions of $m(t)$ and $\mu(t)$ to the two cell cycle stages $\mathbb{S}_1$ and $\mathbb{S}_2$, we define the conditional probabilities

$$p_1(i, m, t) = \text{Prob}\{I(t) = i, M(t) = m \mid U(t) = 1\} = \frac{P_1(i, m, t)}{P_1(t)}, \tag{24}$$

$$p_2(i, m, t) = \text{Prob}\{I(t) = i, M(t) = m \mid U(t) = 2\} = \frac{P_2(i, m, t)}{P_2(t)}, \tag{25}$$

for the probabilities $P_1(t)$ and $P_2(t)$ defined in (14). Then

$$p_1(m, t) = p_1(0, m, t) + p_1(1, m, t), \quad p_2(m, t) = p_2(0, m, t) + p_2(1, m, t) + p_2(2, m, t)$$

are the probabilities that there are m copies of mRNA molecules when the cell resides on $\mathbb{S}_1$ or $\mathbb{S}_2$ stage. The average transcription levels in $\mathbb{S}_1$ and $\mathbb{S}_2$ stages are defined by

$$m_1(t) = \sum_{k=0}^{\infty} k p_1(k, t), \quad m_2(t) = \sum_{k=0}^{\infty} k p_2(k, t), \tag{26}$$

and the second moments are defined by

$$\mu_1(t) = \sum_{k=0}^{\infty} k^2 p_1(k, t), \quad \mu_2(t) = \sum_{k=0}^{\infty} k^2 p_2(k, t). \tag{27}$$

By comparing (26) with the definition of $n_1(t)$ and $n_2(t)$ in (17), and (27) with the definition of $\omega_1(t)$ and $\omega_2(t)$ in (20), we find the simple relation

$$m_1(t) = \frac{n_1(t)}{P_1(t)}, \quad m_2(t) = \frac{n_2(t)}{P_2(t)}, \quad \mu_1(t) = \frac{\omega_1(t)}{P_1(t)}, \quad \mu_2(t) = \frac{\omega_2(t)}{P_2(t)}. \tag{28}$$

As a cell is either on $\mathbb{S}_1$ or on $\mathbb{S}_2$ stage, we have $P_1(t) + P_2(t) \equiv 1$. From the basic assumption (i), the two stages $\mathbb{S}_1$ and $\mathbb{S}_2$ transit each other by constant rates κ_1 and κ_2. It implies that $P_1(t)$ and $P_2(t)$ are simply related by

$$P_1'(t) = \kappa_2 P_2(t) - \kappa_1 P_1(t) = \kappa_2 - (\kappa_1 + \kappa_2) P_1(t).$$

This simple equation can also be derived by adding equations in (15). By the assumption that all cells are synchronized on $\mathbb{S}_1$ initially, we have $P_1(0) = 1$. Hence

$$P_1(t) = \frac{\kappa_2}{\kappa_1 + \kappa_2} + \frac{\kappa_1}{\kappa_1 + \kappa_2} e^{-(\kappa_1 + \kappa_2)t}, \quad P_2(t) = \frac{\kappa_1}{\kappa_1 + \kappa_2} - \frac{\kappa_1}{\kappa_1 + \kappa_2} e^{-(\kappa_1 + \kappa_2)t}. \tag{29}$$

Our methods for finding $n_1(t)$ and $n_2(t)$ in Step 1, and $\omega_1(t)$ and $\omega_2(t)$ in Step 2, combined with (28) and (29), constitute a complete analytical approach for computing the mean values $m_1(t)$ and $m_2(t)$, and the second moments $\mu_1(t)$ and $\mu_2(t)$, in the two cell cycle stages.

Results

Analytic expressions of the first and the second moments

Our discussion in the previous section offers a clear analytical approach for finding the mean value $m(t)$ and the second moment $\mu(t)$ of mRNA number $M(t)$, along with their extensions to the two cell cycle stages $\mathbb{S}_1$ and $\mathbb{S}_2$. However, neither of these functions has a simple analytical

expression. For simplicity, we will only present their steady-state values in exact forms, and use their temporal forms in numerical simulations. Although the steady-state values are much simpler than the temporal forms, they are still rather complex and capture the delicate involvement of the system parameters as shown by the next two theorems. For a function $f(t)$ that has a finite limit as $t \to \infty$, we let f^* denote its limit.

Theorem 1 *If the transcription of a gene obeys the model described in* Fig 1, *then the mean transcription level of the gene in a population of isogenic cells at steady-state is*

$$m^* = m_1^* \cdot \frac{\kappa_2}{\kappa_1 + \kappa_2} + m_2^* \cdot \frac{\kappa_1}{\kappa_1 + \kappa_2}, \tag{30}$$

a linear combination of the mean levels m_1^* in $\mathbb{S}_1$ stage and m_2^* in $\mathbb{S}_2$ stage, and

$$m_1^* = \frac{2v_1\lambda_1(\delta_2 + \kappa_2)(\lambda_2 + \gamma_2 + \kappa_2) + 2v_2\lambda_2\kappa_1(\lambda_1 + \gamma_1 + \kappa_1)}{[2(\delta_1 + \kappa_1)(\delta_2 + \kappa_2) - \kappa_1\kappa_2](\lambda_1 + \gamma_1 + \kappa_1)(\lambda_2 + \gamma_2 + \kappa_2)}, \tag{31}$$

$$m_2^* = \frac{2v_1\lambda_1\kappa_2(\lambda_2 + \gamma_2 + \kappa_2) + 4v_2\lambda_2(\delta_1 + \kappa_1)(\lambda_1 + \gamma_1 + \kappa_1)}{[2(\delta_1 + \kappa_1)(\delta_2 + \kappa_2) - \kappa_1\kappa_2](\lambda_1 + \gamma_1 + \kappa_1)(\lambda_2 + \gamma_2 + \kappa_2)}. \tag{32}$$

Theorem 2 *If the transcription of a gene obeys the model described in* Fig 1, *then the second moment of its mRNA copy number $M(t)$ at steady-state is*

$$\mu^* = \mu_1^* \cdot \frac{\kappa_2}{\kappa_1 + \kappa_2} + \mu_2^* \cdot \frac{\kappa_1}{\kappa_1 + \kappa_2}, \tag{33}$$

where μ_1^* and μ_2^* are the second moments in $\mathbb{S}_1$ and $\mathbb{S}_2$ stages given by

$$\mu_1^* = m_1^* + \frac{8v_1(\kappa_2 + 2\delta_2) \cdot m_{s1}^* + 2v_2\kappa_1 \cdot m_{s2}^*}{4(\kappa_1 + 2\delta_1)(\kappa_2 + 2\delta_2) - \kappa_1\kappa_2}, \tag{34}$$

$$\mu_2^* = m_2^* + \frac{8v_1\kappa_2 \cdot m_{s1}^* + 8v_2(\kappa_1 + 2\delta_1) \cdot m_{s2}^*}{4(\kappa_1 + 2\delta_1)(\kappa_2 + 2\delta_2) - \kappa_1\kappa_2}, \tag{35}$$

with

$$m_{s1}^* = \frac{(\delta_1 + \lambda_1 + \kappa_1)m_1^* - \kappa_1 m_2^*/2}{\delta_1 + \lambda_1 + \gamma_1 + \kappa_1}, \tag{36}$$

$$m_{s2}^* = \frac{(\delta_2 + \kappa_2 + 2\lambda_2)m_2^* - \kappa_2 m_1^* + 2v_2 p_{22}^*}{\delta_2 + \lambda_2 + \gamma_2 + \kappa_2}, \tag{37}$$

and $p_{22}^* = 2\lambda_2^2/[(\kappa_2 + \lambda_2 + \gamma_2)(\kappa_2 + 2\lambda_2 + 2\gamma_2)]$.

The proofs of Theorems 1 and 2 are given in S1 Text. By using definition (1), combined with the analytical expressions (31) and (32) of the stationary mean transcription levels, and (34) and (35) for the second moments, we derive the noise strengths of mRNA copy numbers

in $\mathbb{S}_1$ and $\mathbb{S}_2$ as

$$\Phi_1^* = 1 - m_1^* + \frac{1}{m_1^*} \cdot \frac{8v_1(\kappa_2 + 2\delta_2) \cdot m_{s1}^* + 2v_2\kappa_1 \cdot m_{s2}^*}{4(\kappa_1 + 2\delta_1)(\kappa_2 + 2\delta_2) - \kappa_1\kappa_2}, \tag{38}$$

$$\Phi_2^* = 1 - m_2^* + \frac{1}{m_2^*} \cdot \frac{8v_1\kappa_2 \cdot m_{s1}^* + 8v_2(\kappa_1 + 2\delta_1) \cdot m_{s2}^*}{4(\kappa_1 + 2\delta_1)(\kappa_2 + 2\delta_2) - \kappa_1\kappa_2}. \tag{39}$$

The noises η_1^{2*} and η_2^{2*} are given by

$$\eta_1^{2*} = \frac{1}{m_1^*} - 1 + \frac{1}{(m_1^*)^2} \cdot \frac{8v_1(\kappa_2 + 2\delta_2) \cdot m_{s1}^* + 2v_2\kappa_1 \cdot m_{s2}^*}{4(\kappa_1 + 2\delta_1)(\kappa_2 + 2\delta_2) - \kappa_1\kappa_2}, \tag{40}$$

$$\eta_2^{2*} = \frac{1}{m_2^*} - 1 + \frac{1}{(m_2^*)^2} \cdot \frac{8v_1\kappa_2 \cdot m_{s1}^* + 8v_2(\kappa_1 + 2\delta_1) \cdot m_{s2}^*}{4(\kappa_1 + 2\delta_1)(\kappa_2 + 2\delta_2) - \kappa_1\kappa_2}. \tag{41}$$

The fold change of transcripts

The broad range of the fold change. We use $r^* = m_2^*/m_1^*$, the ratio of the mean transcription levels in the two stages, as a measure to quantify the fold change of gene transcription from $\mathbb{S}_1$ stage to $\mathbb{S}_2$ stage [12, 15, 34]. Our formulas (31) and (32) imply

$$r^* = \frac{m_2^*}{m_1^*} = \frac{v_1\lambda_1\kappa_2(\lambda_2 + \gamma_2 + \kappa_2) + 2v_2\lambda_2(\delta_1 + \kappa_1)(\lambda_1 + \gamma_1 + \kappa_1)}{v_1\lambda_1(\delta_2 + \kappa_2)(\lambda_2 + \gamma_2 + \kappa_2) + v_2\lambda_2\kappa_1(\lambda_1 + \gamma_1 + \kappa_1)}. \tag{42}$$

All ten system parameters are involved in this formula. If the transcripts are turned over extremely fast in $\mathbb{S}_1$ stage such that δ_1 dominates all other parameter values, then r^* can be made arbitrarily large. On the contrary, if the transcripts are turned over extremely fast in $\mathbb{S}_2$ stage, then r^* can be made sufficiently small. Thus r^* could take any arbitrary positive number as parameter values vary, and the range of the fold change predicted by (42) is the whole set (0, ∞) of positive numbers.

In many bacterial cells, it has been observed that the transcription kinetic rates of some genes remain the same in the whole cell cycle [35], that is,

$$v_i = v, \ \ \delta_i = \delta, \ \ \lambda_i = \lambda, \ \ \gamma_i = \gamma, \ \ i = 1, 2. \tag{43}$$

If (43) holds, then (42) can be simplified to

$$r^* = \frac{2(\delta + \kappa_1)(\lambda + \gamma + \kappa_1) + \kappa_2(\lambda + \gamma + \kappa_2)}{\kappa_1(\lambda + \gamma + \kappa_1) + (\delta + \kappa_2)(\lambda + \gamma + \kappa_2)}. \tag{44}$$

What is the range of r^* in this case? During $S/G_2/M$ phases, a cell contains twice as many copies of each gene as that in G_1 phase. Intuitively, one may envisage that the number of mRNA copies in $\mathbb{S}_2$ stage doubles the number in $\mathbb{S}_1$ stage and so $r^* \approx 2$. Our next theorem shows that r^* can deviate from 2 largely, and surprisingly, the constraint (43) does not reduce the range of r^*.

Theorem 3 *For any constant $C > 0$, there exist system parameters under the constraint* (43) *to make $r^* = C$.*

This counter-intuitive result predicts that, even when the transcription kinetic rates do not change in different cell cycle phases, the fold change of transcription levels from $\mathbb{S}_1$ to $\mathbb{S}_2$ can be made, in theory, arbitrarily large or small. A complete proof of this result is given in

[S1 Text](). In the proof, we specify the stage transition rates κ_1 and κ_2 in terms of the kinetic rates to make $r^* = C$ for $C \leq 1$, $C \in (1, 3/2]$, $C \in (3/2, 2)$, and $C \geq 2$ separately. For instance, when $C \leq 1$, we take

$$\lambda + \gamma = \kappa_1, \quad \kappa_2 = \frac{4\kappa_1}{C}, \quad \delta = \frac{(16 - 12C - 2C^3)\kappa_1}{C^3} > 0. \tag{45}$$

Substituting these parameters into (44) gives

$$
\begin{aligned}
r^* &= \frac{2[(16 - 12C - 2C^3)\kappa_1/C^3 + \kappa_1](\kappa_1 + \kappa_1) + 4\kappa_1 C \cdot (\kappa_1 + 4\kappa_1/C)}{\kappa_1(\kappa_1 + \kappa_1) + [(16 - 12C - 2C^3)\kappa_1/C^3 + 4\kappa_2/C](\kappa_1 + 4\kappa_1/C)} \\
&= \frac{4(16 - 12C - C^3)\kappa_1^2/C^3 + 4(C + 4)\kappa_1^2/C^2}{2\kappa_1^2 + (C + 4)(16 - 12C + 4C^2 - 2C^3)\kappa_1^2/C^4} = C.
\end{aligned}
$$

We note that (45) requires $\kappa_1 = \lambda + \gamma$ and $\kappa_2 \geq 4\kappa_1 = 4(\lambda + \gamma)$, which corresponds to faster cell cycle stage transitions comparing to gene promoter transitions, and a much shorter $\mathbb{S}_2$ stage than $\mathbb{S}_1$ stage. These conditions may not hold regularly in real cells, and a fold change $r^* < 1$ has been rarely observed under the constraint (43). However, there have been measurements reporting a slightly larger than 1-fold change. In 2016, Skinner et al. [15] quantified mature and nascent mRNA levels of *Oct*4 in individual mouse ES cells, and found that the increase from earlier cell cycle stage to later stage was only about 1.3-fold.

Cell cycle stage transitions are often slower than gene promoter state transitions, implying $\kappa_1 < \lambda + \gamma$ and $\kappa_2 < \lambda + \gamma$. If we divide the numerator and denominator in (44) by $\delta(\lambda + \gamma)$, then we may change (44) to

$$r^* = \frac{2 + 2\kappa_1/(\lambda + \gamma) + (2\kappa_1 + \kappa_2)/\delta + (2\kappa_1^2 + \kappa_2^2)/[\delta(\lambda + \gamma)]}{1 + \kappa_2/(\lambda + \gamma) + (\kappa_1 + \kappa_2)/\delta + (\kappa_1^2 + \kappa_2^2)/[\delta(\lambda + \gamma)]}.$$

It is evident that $r^* \approx 2$ if κ_1, κ_2 are considerably smaller than δ and $\lambda + \gamma$. It indicates that when the transcription kinetics are unchanged in the two cell cycle stages, and the stage transition is considerably slower than the transcription state transition and mRNA turnover, the mRNA number in $\mathbb{S}_2$ stage doubles the number in $\mathbb{S}_1$ stage at steady-state. This doubling property has been observed in several experimental measurements. In 2004, Vintersten et al. [34] reported a strong expression of a developed RFP variant, DsRed.T3, in mouse ES cells, and found that the nascent mRNA level exhibited a 2-fold increase from $\mathbb{S}_1$ stage to $\mathbb{S}_2$ stage. In 2011, Trcek et al. [12] measured the cytoplasmic mRNA level of *CLB2* in yeast and also found an approximate 2-fold increase from late G_2/M phases to S phase.

Next, we characterize the dependance of r^* on the stage transition rates κ_1 and κ_2. To emphasize its dependance on these parameters, we write $r^* = r^*(\kappa_1, \kappa_2)$.

Theorem 4 *Let* (43) *hold. Then we have*

a. *When κ_1 increases from 0 to ∞, r^* increases from $r^*(0, \kappa_2) < 2$ until it peaks uniquely and then decreases to approach 2 at ∞. In particular, $r^* > 2$ if and only if*

$$\kappa_1 > \kappa_2 + \frac{\kappa_2(\kappa_2 + \lambda + \gamma)}{2\delta}. \tag{46}$$

b. *When κ_2 increases from 0 to ∞, r^* decreases from $r^*(\kappa_1, 0) > 2$ until it bottoms out uniquely and then increases to approach 1 at ∞. In particular, $r^* < 1$ if and only if*

$$\kappa_2 > 2\kappa_1 + \lambda + \gamma + \frac{\kappa_1(\lambda + \gamma + \kappa_1)}{\delta}. \tag{47}$$

c. *When $\kappa_1 \le \kappa_2$, r^* has an upper bound strictly less than 2.*

Theorem 4 gives a precise description on the nonlinear dependance of the fold change r^* on the stage transition rates κ_1 and κ_2. Conditions (46) and (47) give the respective sufficient and necessary conditions for $r^* > 2$ and $r^* < 1$. The complete proof of this theorem is given in S1 Text.

Discussion

The fold change r^* and its dependance on the stage durations

The accuracy of estimating the fold change. We use a set of experimental data to test how accurately our formula (42) may help estimate the fold change r^* of mRNA copy numbers from the early cell cycle stage $\mathbb{S}_1$ to the later stage $\mathbb{S}_2$. In mouse embryonic stem cells, it was observed that the mean OFF duration in the transcription of *Oct*4 increased from 108 min in $\mathbb{S}_1$ stage to 173 min in $\mathbb{S}_2$ stage, and the mean ON duration lasted about 56 min in the two stages [15], suggesting

$$\lambda_1 = 0.5556 \text{ hr}^{-1}, \ \lambda_2 = 0.3468 \text{ hr}^{-1}, \ \gamma_1 = \gamma_2 = 1.0714 \text{ hr}^{-1}. \tag{48}$$

In [15], Skinner found that the average duration of $\mathbb{S}_1$ stage lasted about ~ 560 min by using smFISH. Cartwright et al. [36] measured the cell division time in the same cell line under similar growth conditions, and estimated the average time for one cell cycle at ~ 13 hr. These data suggest an average $\mathbb{S}_2$ stage duration at ~ 220 min. As the mean life of newly transcribed *Oct*4 mRNA before being converted to mature RNA was found to be close to 3.5 min, we estimate the other parameters as follows:

$$\delta_1 = \delta_2 = 17.14 \text{ hr}^{-1}, \ \kappa_1 = 0.1071 \text{ hr}^{-1}, \ \kappa_2 = 0.2727 \text{ hr}^{-1}. \tag{49}$$

As the synthesis rate remains about the same in the two cell cycle stages, we can simplify (42) to

$$r^* = \frac{\lambda_1\kappa_2(\lambda_2 + \gamma_2 + \kappa_2) + 2\lambda_2(\delta_1 + \kappa_1)(\lambda_1 + \gamma_1 + \kappa_1)}{\lambda_1(\delta_2 + \kappa_2)(\lambda_2 + \gamma_2 + \kappa_2) + \lambda_2\kappa_1(\lambda_1 + \gamma_1 + \kappa_1)}. \tag{50}$$

Substituting (48) and (49) into (50) gives $r^* = 1.2791$, which matches precisely the experimental measurement $r^* = 1.28 \pm 0.09$ observed in [15].

The condensation of r^* within [1, 2]. Although Theorem 1 predicts that r^* could take any positive value even if the transcription kinetic rates remain the same in all cell cycle phases, we use numerical simulation to demonstrate that it is more likely to observe $r^* \in [1, 2]$ in some cells. As suggested by [15] in the transcription of *Oct*4 gene in mouse embryonic stem cells, we fix

$$\lambda = 0.5556 \text{ hr}^{-1}, \quad \gamma = 1.0714 \text{ hr}^{-1}, \quad \nu = 1.89 \text{ min}^{-1}, \quad \delta = 0.14 \text{ hr}^{-1}. \tag{51}$$

The mean OFF and ON durations are ~ 108 min and ~ 56 min, respectively, and the mean mRNA lifetime is ~ 7.14 hr. To characterize the dependance of the ratio r^* on the stage
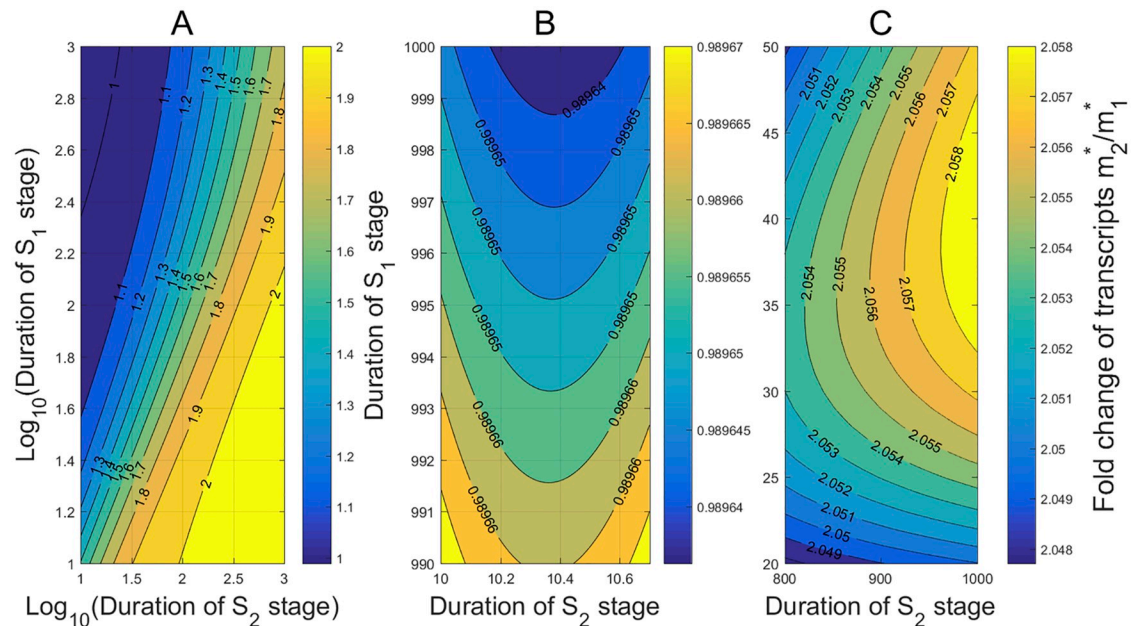
Fig 2. The condensation of r^* within [1, 2] and its dependance on the stage durations. All kinetic rates are estimated from [15] in the transcription of *Oct*4 gene in mouse embryonic stem cells. (A) For almost all the $\mathbb{S}_1$ and $\mathbb{S}_2$ stage durations ranging from 10 to 1,000 min, r^* is within a narrow range [1, 2]. When $1/\kappa_1 \gg 1/\kappa_2$ (or $1/\kappa_1 \ll 1/\kappa_2$), r^* varies slowly over a narrow region around 1 (or 2). When the two durations are close, r^* changes rapidly over (1.1, 2). (B) When the $\mathbb{S}_1$ stage duration is far greater than the $\mathbb{S}_2$ stage duration, $r^* < 1$ and is very close to 1. (C) When the $\mathbb{S}_2$ stage duration is far greater than the $\mathbb{S}_1$ stage duration, $r^* > 2$ and is very close to 2.

durations, we let the mean $\mathbb{S}_1$ stage duration $1/\kappa_1$ and the $\mathbb{S}_2$ stage duration $1/\kappa_2$ vary from 10 to 1,000 min. The contours of r^* on stage durations are shown in Fig 2, where the durations are rescaled in logarithm in Fig 2A. Very interestingly, the contours show that the ratio r^* is within the narrow range [1, 2] over almost all the durations, in a good agreement with many experimental measurements [9, 13, 15, 34]. It is also seen that for fixed $\mathbb{S}_2$ stage duration, r^* decreases and tends to a stable value (≈ 1) when the $\mathbb{S}_1$ stage duration increases. On the other hand, for fixed $\mathbb{S}_1$ stage duration, r^* increases and tends to a larger stable value (≈ 2) when the $\mathbb{S}_2$ stage duration increases.

To see more clearly how r^* changes in the limit cases, we enlarged the left upper corner of Fig 2A in Fig 2B, and the right lower corner of Fig 2A in Fig 2C. The contours in Fig 2B and 2C exhibit some minor deviations from our observation above for Fig 2A. In Fig 2B, where the $\mathbb{S}_1$ stage duration is far greater than the $\mathbb{S}_2$ stage duration, $r^* < 1$ and is very close to 1. Also, for fixed large $\mathbb{S}_1$ duration, r^* displays a non-monotone growth as claimed by Theorem 4(b) when the $\mathbb{S}_2$ duration increases from 10 min to 10.6 min. In Fig 2C, where the $\mathbb{S}_2$ duration is far greater than the $\mathbb{S}_1$ duration, $r^* > 2$ and is very close to 2. For fixed large $\mathbb{S}_2$ duration, r^* displays a non-monotone growth as claimed by Theorem 4(c) when the $\mathbb{S}_1$ duration increases from 20 min to 45 min.

The dependance of the fold change on the timing of gene duplication. In the discussion above, the cell cycle duration changes synchronously when the durations of $\mathbb{S}_1$ or $\mathbb{S}_2$ change. In some cases, the cell division time may not change significantly, while the variation of $\mathbb{S}_1$ and $\mathbb{S}_2$ durations is mainly caused by the random timing of gene duplication that may occur in the early *S* phase or at the end of *S* phase. During *S* phase, DNA synthesis is catalyzed by DNA polymerases [17, 37], and the genome is replicated at a stable rate on leading and lagging

strands [25]. In cell cycle-synchronized budding yeast, the total cell division time is ~ 70 min, comprising by ~ 35 min G_1 phase, ~ 18 min S phase and ~ 17 min G_2/M phase [13]. We examine how the mean levels m_1^*, m_2^* and the ratio r^* change when the gene duplication time increases from the start of S phase at 35 min to the end of S phase at 53 min, and transcripts are produced and turned over with the rates:

$$v_1 = 2.52 \text{ min}^{-1}, \ v_2 = 2.14 \text{ min}^{-1}, \ \delta_1 = \delta_2 = 1.18 \text{ hr}^{-1}.$$

We simplify our approach by taking $\lambda_1, \lambda_2 \to \infty$ and obtain from (31) and (32)

$$m_1^* = \frac{2v_1(\delta_2 + \kappa_2) + 2v_2\kappa_1}{2(\delta + \kappa_1)(\delta + \kappa_2) - \kappa_1\kappa_2}, \quad m_2^* = \frac{2v_1\kappa_2 + 4v_2(\delta + \kappa_1)}{2(\delta + \kappa_1)(\delta + \kappa_2) - \kappa_1\kappa_2}.$$

It follows that

$$r^* = \frac{v_1\kappa_2 + 2v_2(\delta + \kappa_1)}{v_1(\delta + \kappa_2) + v_2\kappa_1}.$$

As shown in Fig 3A, although the mean transcription level m_2^* in $\mathbb{S}_2$ decreases in the gene duplication time as we expect, it is surprisingly to see that the mean level m_1^* in $\mathbb{S}_1$ remains almost a constant. Very interestingly, as shown in Fig 3B, the ratio r^*, which changes moderately in a narrow range from 1.31 to 1.53, decreases almost linearly in the duplication time.

The effect of cell cycles on transcription noise

Our gene transcription model coupling with cell division cycles offers six quantities to characterize the fluctuations of mRNA numbers in single cells: The noise η^{2*} and the noise strength Φ^* in cells without referring to cell cycle stages, along with η_1^{2*} and Φ_1^* in $\mathbb{S}_1$ stage, and η_2^{2*} and Φ_2^* in $\mathbb{S}_2$ stage. Theorems 1 and 2 provide the basic formulas by which these quantities can be computed from the system parameters.
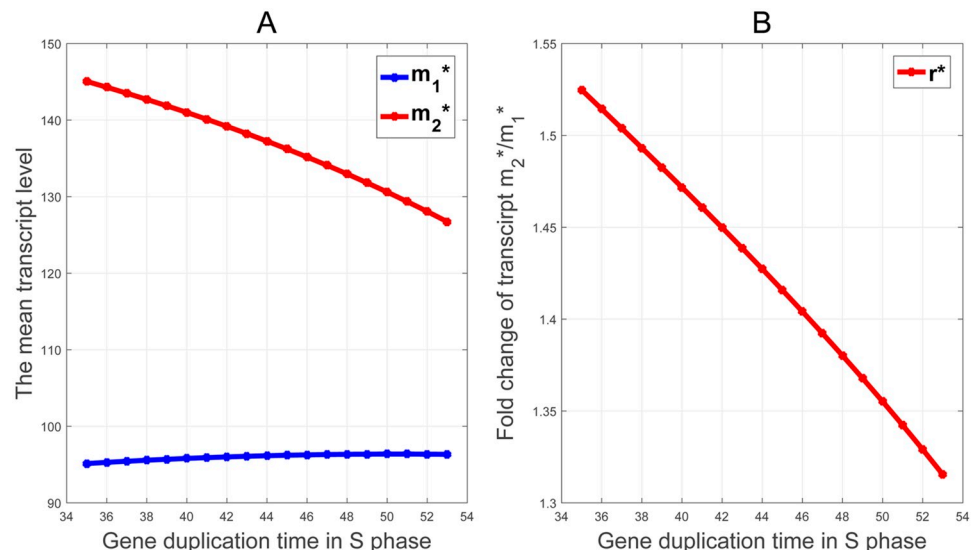
Fig 3. The dependance of the mean mRNA copy numbers m_1^*, m_2^* and the ratio r^* on the timing of gene duplication during S phase. (A) The mean level m_1^* stays nearly at a constant level, while m_2^* decreases linear as the gene duplication time increases from 35 min to 53 min in S phase. (B) The ratio r^* decreases almost linearly from 1.53 to 1.31 as the gene duplication time increases.

The relations between these quantities are far more complicated than our intuition may envisage. We use simple example to demonstrate the delicacy of their relations: Assume (43) and fix the kinetic rates as in (51), and take the stage transition rates $\kappa_1 = \kappa_2 = 1/1250 \text{ hr}^{-1}$. Then applying Theorem 2, (38) and (39) gives $\Phi_1^* = 43.2008$ and $\Phi_2^* = 43.8410$. Since Φ_1^* and Φ_2^* are nearly equal, one might expect by intuition that Φ^* is about equal to each of Φ_1^* and Φ_2^*, which is in conflict with

$$\Phi^* = 89.3329 > \Phi_1^* + \Phi_2^*$$

obtained by using Theorems 1 and 2. Moreover $\Phi^* = 89.3329$ given here is significantly higher than $\Phi^* \approx 1$ reported in various single cell measurements, including the classical studies by Taniguchi et al. [38] and Yu et al. [39]. In these studies, the genes were active most of time, so that m^* and η^{2*} exhibited a strict reciprocal relation, implying $\Phi^* \approx 1$. The stochastic switching between gene promoter ON and OFF states, combined with transitions between cell cycle stages, may induce much noisier transcriptions.

Due to the wide range of the six noise measures and their complex relations, we will discuss their profiles in three special cases below:

1. The transcription kinetic rates remain constants in all cell cycle phases.

2. The transcription homeostasis is maintained that $m_1^* = m_2^*$.

3. The concentration homeostasis is maintained that the mean transcription levels scale with the average volumes in the two stages.

The noise profile with constant kinetic rates. DNA replication during cell division cycles induces a doubling of gene copy numbers from $\mathbb{S}_1$ to $\mathbb{S}_2$. In bacteria, the transcription kinetic rates of many genes do not change considerably in different cell cycle phases [13]. In this case, mRNA production follows gene dosage with a rapid increase after DNA replication. In a sharp contrast, the response of the transcription noises to the temporal variation of gene dosage is rather complicated and few results have been given to elucidate the complexity. Let the transcription kinetic rates remain constants in all cell cycle phases and fix the rate constants as in (51). We use numerical simulations based on our analytical formulas to depict the profiles of the transcription noise measures and their relations by varying the cell cycle durations over the large time interval (10 min, 10^4 min).

In Fig 4A, the contours of the mean transcription level m^* in cells display a simple and monotonic variation over the stage durations. It increases in both stage durations and the increase is more sensitive on the second stage duration. Similarly, the contours of the noise η^{2*} also display a simple and monotonic variation over the stage durations. However, it decreases in both stage durations and the dependance on the two durations are nearly symmetric. Such an inverse relation between m^* and η^{2*} has been observed in many living cell measurements and theoretical studies [23, 35], and a strict reciprocal relation between them has been observed in Taniguchi et al. [38] and Yu et al. [39] with $\Phi^* \approx 1$. Nevertheless, as the transcription specified by (51) is frequently interrupted by long OFF periods, the contours of the noise strength Φ^* in Fig 4C indicate that Φ^* is significantly higher than 1 with $\Phi^* \in [20, 100]$. Moreover, Φ^* displays a highly nonlinear, non-monotonic variation over the cell cycle stages. Intriguingly, among the durations shown in Fig 4C, Φ^* peaks in a small region when $\mathbb{S}_1$ stage lasts between 10 and 100 minutes on average, and $\mathbb{S}_2$ stage lasts between 125 and 250 minutes.

In Fig 4D–4F, the fold changes m_2^*/m_1^*, η_2^{2*}/η_1^{2*}, and Φ_2^*/Φ_1^* from $\mathbb{S}_1$ to $\mathbb{S}_1$ all vary within a narrow range with m_2^*/m_1^*, $\Phi_2^*/\Phi_1^* \in [1, 2]$, and $\eta_2^{2*}/\eta_1^{2*} \in [0.5, 1]$. The right upper corners in
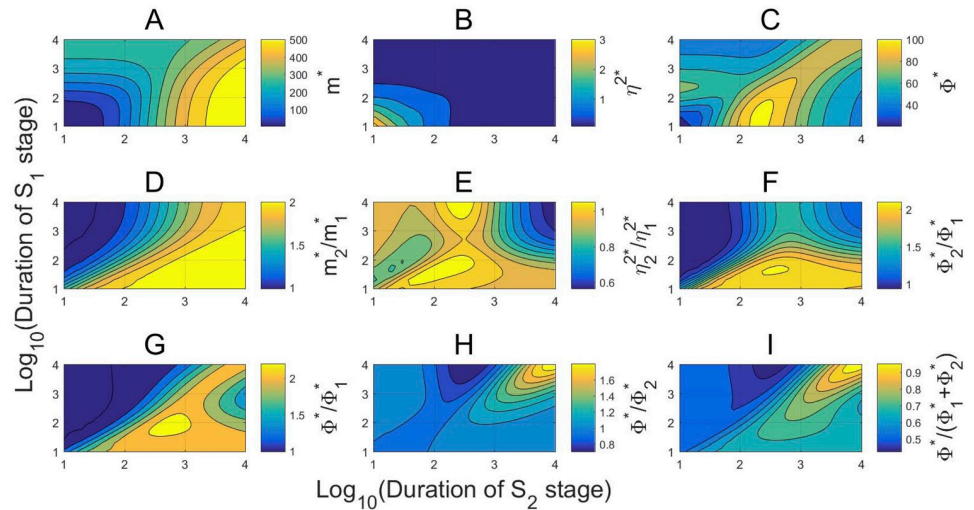
Fig 4. The nonlinear behavior of the transcription noises. The transcription kinetics are specified by (43) and (51), and the stage durations vary from 10 to 10^4 min. (A) The mean m^* increases in the cell cycle durations. (B) The noise η^{2*} is about equally sensitive to the variation of the stage durations and decreases from 3 to 0.5 as the stage durations increase. (C) The noise strength Φ^* varies widely from 20 to 100 with a highly nonlinear dependance on the stage durations. (D)-(F) The ratios m_2^*/m_1^*, $\Phi_2^*/\Phi_1^* \in [1, 2]$, and $\eta_2^{2*}/\eta_1^{2*} \in [0.5, 1]$ over most durations. (G)-(I) Φ^* is larger than both Φ_1^* and Φ_2^* but less than $\Phi_1^* + \Phi_2^*$ over most durations.

https://doi.org/10.1371/journal.pcbi.1007017.g004

Fig 4D–4F seem to suggest that, as both stage durations become large,

$$m_2^*/m_1^* \approx 2, \qquad \eta_2^{2*}/\eta_1^{2*} \approx 1/2, \qquad \Phi_2^*/\Phi_1^* \approx 1. \tag{52}$$

Indeed, (52) can be supported by simple mathematical calculations. By taking limits in (31) and (32), we obtain

$$\lim_{\kappa_1 \to 0} m_1^*(\kappa_1, \kappa_2) = \frac{v\lambda}{\delta(\lambda + \gamma)} = \frac{1}{2} \lim_{\kappa_2 \to 0} m_2^*(\kappa_1, \kappa_2).$$

and verify the first part of (52). Taking limits in (34) and (35) gives

$$\lim_{\kappa_1 \to 0} \mu_1^*(\kappa_1, \kappa_2) = \frac{v\lambda}{\delta(\lambda + \gamma)} \left[1 + \frac{v(\delta + \lambda)}{\delta(\delta + \lambda + \gamma)} \right],$$

and

$$\lim_{\kappa_2 \to 0} \mu_2^*(\kappa_1, \kappa_2) = \frac{2v\lambda}{\delta(\lambda + \gamma)} \left[1 + \frac{v}{\delta} \cdot \left(\frac{\lambda}{\lambda + \gamma} + \frac{\delta + \lambda}{\delta + \lambda + \gamma} \right) \right].$$

By the definition of the noise strength in (1), we find

$$\lim_{\kappa_1 \to 0} \Phi_1^*(\kappa_1, \kappa_2) = \lim_{\kappa_2 \to 0} \Phi_2^*(\kappa_1, \kappa_2) = 1 + \frac{v\gamma}{(\lambda + \gamma)(\delta + \lambda + \gamma)} \tag{53}$$

and verify the third part in (52). Consequently, the second part in (52) is also verified since the noise equals the noise strength divided by the mean.

The approximate identities in (52) indicate that when the cell cycle stage transitions are sufficiently slow, the mean transcription levels are doubled and the noises are halved from $\mathbb{S}_1$ to $\mathbb{S}_2$, while the noise strengths remain about the same. However, this simple statement is invalid

in general as shown by the contours in Fig 4D–4F, where η_2^{2*}/η_1^{2*} and Φ_2^*/Φ_1^* display highly nonlinear and wild variations.

In Fig 4G–4I, we compare Φ^* with Φ_1^*, Φ_2^*, and $\Phi_1^* + \Phi_2^*$. In most of time,

$$\Phi_1^* < \Phi^* < 2\Phi_1^*, \quad \Phi_2^* < \Phi^* < 2\Phi_2^*, \quad (\Phi_1^* + \Phi_2^*)/2 < \Phi^* < \Phi_1^* + \Phi_2^*.$$

However, these estimates do not hold universally. For instance, Fig 4H indicates that Φ^*/Φ_2^* can be as low as 0.8 and so $\Phi^* < \Phi_2^*$ in some cases. Also, we have shown one example early that $\Phi^* > \Phi_1^* + \Phi_2^*$ may hold. Overall, Φ^* is rarely identical to either Φ_1^* or Φ_2^*, and stays between $\max\{\Phi_1^*, \Phi_2^*\}$ and $\Phi_1^* + \Phi_2^*$ most of time.

The noise profile in gene transcription homeostasis. It has been observed repeatedly that gene transcription in eukaryotic cells, ranging from yeast to mammals, has a limited dependency on DNA dosage [13]. Cells have a DNA dosage-compensating mechanism to precisely reduce mRNA production in late cell cycle stage, resulting in a gene transcription homeostasis that overall transcription remains constant across $\mathbb{S}_1$ and $\mathbb{S}_2$ stages [18]. Various compensating mechanisms for gene transcription homeostasis have been found [12, 13, 15, 18], including reducing transcription activation or mRNA synthesis rates, and increasing inactivation or mRNA degradation rates in $\mathbb{S}_2$ stage. For all of the genes measured in foreskin fibroblast cells, Padovan-Merhar et al. [18] found that the number of active sites per gene copy in $\mathbb{S}_2$ stage was approximately half of that in $\mathbb{S}_1$. In mouse embryonic stem cells, Skinner et al. [15] found that the activation rates of *Oct4* and *Nanog* genes were reduced from 0.5556 hr^{-1} and 0.1124 hr^{-1} in $\mathbb{S}_1$ stage to 0.3468 hr^{-1} and 0.08 hr^{-1} in $\mathbb{S}_2$ stage, respectively. In cell cycle-synchronized budding yeast, Voichek et al. [13] found the reduction of mRNA synthesis rate during *S* phase. In yeasts, Trcek et al. [12] found a 30-fold increase in the mRNA degradation rates of *SWI5* and *CLB2* during prometaphase/metaphase.

Mathematically, we define the gene transcription homeostasis by $m_1^* = m_2^*$ that the mean transcription levels in $\mathbb{S}_1$ and $\mathbb{S}_2$ stages remain the same. If the homeostasis is brought by varying a single pair of corresponding kinetic rates, then substituting (31) and (52) into $m_1^* = m_2^*$ yields the relations of the two varied rates as follows:

$$v_2 = v_1 \cdot \frac{\delta(\lambda + \gamma + \kappa_2)}{(2\delta + \kappa_1)(\lambda + \gamma + \kappa_1)}, \tag{54}$$

$$\delta_2 = \delta_1 \cdot \frac{2(\lambda + \gamma + \kappa_1)}{\lambda + \gamma + \kappa_2} + \frac{\kappa_1(\lambda + \gamma + \kappa_1)}{\lambda + \gamma + \kappa_2}, \tag{55}$$

$$\frac{1}{\lambda_2} = \frac{1}{\lambda_1} \cdot \frac{(2\delta + \kappa_1)(\gamma + \kappa_1)}{\delta(\gamma + \kappa_2)} + \frac{\delta + \kappa_1}{\delta(\gamma + \kappa_2)}. \tag{56}$$

$$\gamma_2 = \gamma_1 \cdot \frac{2\delta + \kappa_1}{\delta} + \frac{\delta\lambda + 2\delta\kappa_1 + \lambda\kappa_1 + \kappa_1^2 - \delta\kappa_2}{\delta}, \tag{57}$$

In (54)–(57), the parameters with no subscripts are assumed to be independent of cell cycle stages. It is seen that v_2 and v_1 change proportionally, whereas δ_2 and δ_1, γ_2 and γ_1, and the two OFF durations depend linearly.

We test how the noise η^{2*} and the noise strength Φ^* respond when the parameter pair defined in each of (54)–(57) change linearly to maintain the homeostasis. We again fix the parameters with no subscripts in (54)–(57) as in (51), and let

$$\kappa_1^{-1} = 560 \text{ min}, \qquad \kappa_2^{-1} = 220 \text{ min}, \tag{58}$$
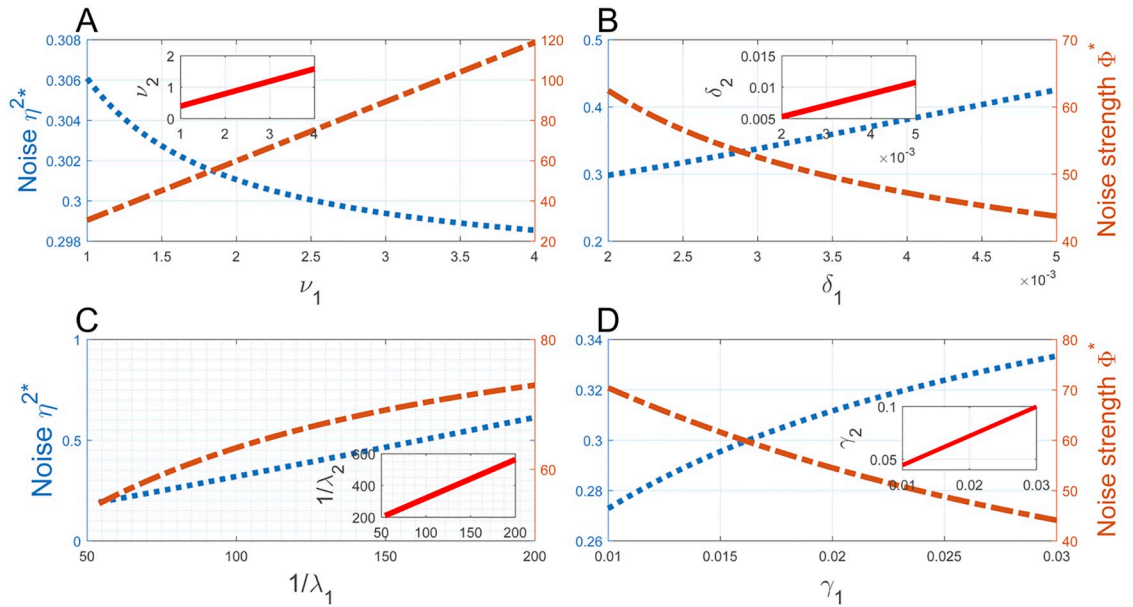
Fig 5. The noise profile in gene transcription homeostasis. The inserts in (A)-(D) depict the linear relations (54)–(57), where ν_1, δ_1, λ_1, and γ_1 have the same unit min^{-1} and vary near the corresponding values specified in (51). Both η^{2*} and Φ^* change monotonically on these intervals in opposite directions, except in (C) where they both decrease in λ_1. The noise η^{2*} takes small values near 0.30 in all panels and has only insignificant variations. The noise strength Φ^* takes large values but also shows insignificant variations except in Fig 5A.

https://doi.org/10.1371/journal.pcbi.1007017.g005

that are approximately the average duration of $\mathbb{S}_1$ and $\mathbb{S}_2$ stages in a mouse embryonic stem cell line [15, 36]. In Fig 5A–5D, we vary ν_1, δ_1, λ_1, and γ_1 near the corresponding parameter values specified in (51), with $\nu_1 \in [1, 4]$ in (A), $\delta_1 \in [2 \times 10^{-3}, 5 \times 10^{-3}]$ in (B), $\lambda_1 \in [0.005, 0.02]$ in (C), and $\gamma_1 \in [0.01, 0.03]$ in (D), all sharing the same unit min^{-1}. The linear relations of ν_1 and ν_2, δ_1 and δ_2, $1/\lambda_1$ and $1/\lambda_2$, and γ_1 and γ_2 defined by (54)–(57) are depicted in the inserts of these panels. It is easily seen that both η^{2*} and Φ^* vary monotonically on these intervals. Interestingly, η^{2*} and Φ^* change in opposite directions in Fig 5A, 5B and 5D, and they both decrease in the activation rate λ_1 in Fig 5C. The noise η^{2*} takes small values near 0.30 in all panels and shows only insignificant variations. The noise strength Φ^* takes considerably larger values but the variation is also insignificant, except in Fig 5A where it increases from about 30 to 120 when ν_1 increases from 1 to 4. Overall, these data suggest that if the transcription homeostasis is induced by varying a single pair of corresponding kinetic rates, then the variation of the rates does not bring significant changes in transcription noise.

We examine the relation between the transcription noise η^{2*} and transcription homeostasis further by varying ν_2/ν_1, δ_2/δ_1, λ_2/λ_1, and γ_2/γ_1 from 0.1 to 10 in Fig 6A–6D. All other parameters are kept as in (51) and (58). We label the points at which transcription homeostasis is reached as H, and the noise minimizing points as Z. In all panels from (A)-(D), it is seen that the two points stay very close. It suggests that if transcription homeostasis is attained by varying a single kinetic rate in the two cell cycle stages, then the homeostasis nearly minimizes transcription noise.

The noise profile in transcript concentration homeostasis. Cells reproduce by cell division. During cell mitosis, newly divided daughter cells are naturally smaller than their parent cells. Thus, a cell in $\mathbb{S}_2$ stage is normally larger than itself in $\mathbb{S}_1$ stage, although it has also been reported that cellular volumes are weakly related to cell cycle phases [18]. If the transcription
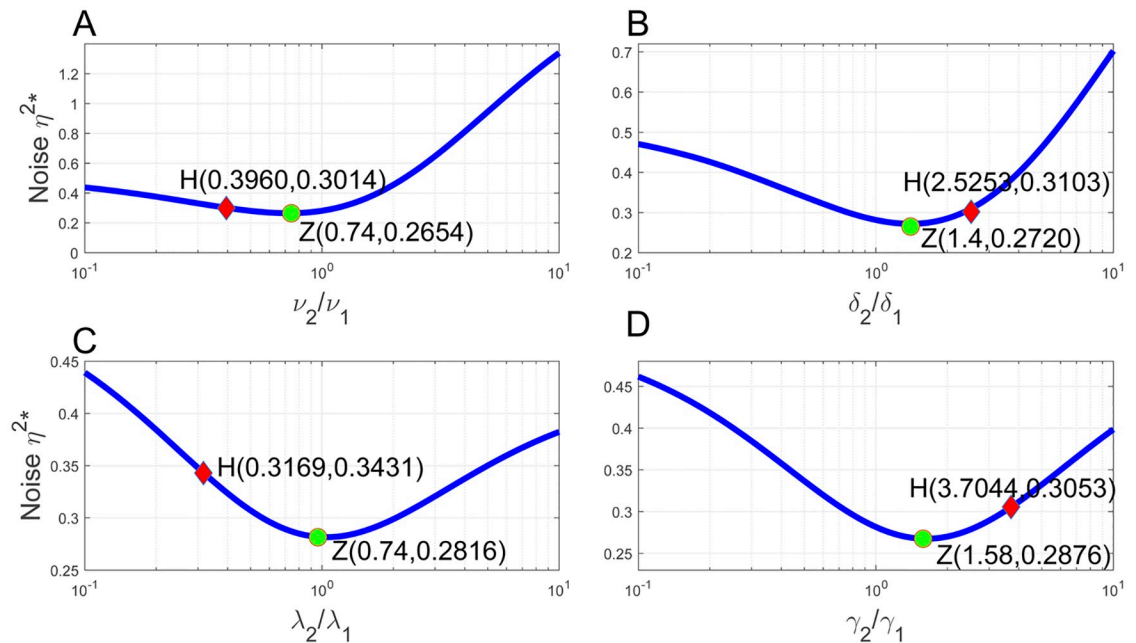
Fig 6. Transcription homeostasis nearly minimizes transcription noise. The ratios ν_2/ν_1, δ_2/δ_1, λ_2/λ_1, and γ_2/γ_1 increase from 0.1 to 10 in (A)-(D), whereas other parameters are kept as in (51) and (58). In (A)-(D), the point H at which transcription homeostasis is reached and the noise minimizing point Z stay very close.

https://doi.org/10.1371/journal.pcbi.1007017.g006

homeostasis is maintained as we discussed above, then the transcript concentration in a cell decreases as it grows from $\mathbb{S}_1$ to $\mathbb{S}_2$ stage. However, in agree with the paradigm that biochemical reaction rates are determined mostly by the concentration of reactants and enzymes, increasing evidences have shown that many cellular processes depend on the concentration of gene expression products rather than their absolute numbers [18, 21, 22]. In cell populations with a large variability in cellular volumes, the numbers of most molecules need to scale with volumes to maintain a stable concentration for proper cellular functions [18, 21].

Let V_1 and V_2 denote the average volumes of individual cells on $\mathbb{S}_1$ and $\mathbb{S}_2$ stages, respectively, and let

$$r_v = V_2/V_1$$

denote the fold change of the volumes on the two stages. Recall that $r^* = m_2^*/m_1^*$ is the fold change of the mean transcription levels at steady-state. We define the transcript concentration homeostasis by

$$r^* = r_v^*, \tag{59}$$

that is, the mean transcription levels scale with the volumes on both cell cycle stages.

We need to find an analytical formula of r_v in terms of the cell cycle stage transition rates κ_1 and κ_2. By combining it with the expression of r^* in (42) and the identity (59), we obtain an explicit relation of the system parameters for the transcript concentration homeostasis. The expression of r_v depends on how cells grow and how the two cell cycle stage durations distribute. Various cell growth models have been proposed [22, 40–42], including the prevailing exponential growth model that cell volumes grow exponentially. Assume that a newly divided cell has a volume $V_0 = 1$ at time $t = 0$. Then the cellular volume $V(t)$ during the first cell cycle is

given by

$$V(t) = e^{at}, \tag{60}$$

for a constant growth rate $a > 0$. Let $(0, T_1)$ be the duration of $\mathbb{S}_1$, and $(T_1, T_1 + T_2)$ be the duration of $\mathbb{S}_2$. We assume that the cell volume is doubled at the end of the cell cycle. Then $V(T_1 + T_2) = 2$. Recall that T_1 and T_2 are independently and exponentially distributed with rates κ_1 and κ_2. With these specifications, we have

$$a = \frac{\kappa_1 + \kappa_2 - \sqrt{\kappa_1^2 + \kappa_2^2}}{2}, \qquad r_v = \frac{\kappa_2}{\kappa_2 - a} = \frac{2\kappa_2}{\sqrt{\kappa_1^2 + \kappa_2^2} - \kappa_1 + \kappa_2}. \tag{61}$$

We note that a increases in both κ_1 and κ_2, and $r_v \in (1, 2)$. When $\kappa_1 \to 0$, the cell stays sufficiently long in $\mathbb{S}_1$ and has little growth in $\mathbb{S}_2$. In this case, (61) gives $r_v \to 1$ as we expect. On the other hand, $r_v \to 2$ as $\kappa_2 \to 0$.

To prove (61), we begin with the observation that

$$r_v = \frac{\mathrm{E}[V(T_1 + T_2)]}{\mathrm{E}[V(T_1)]},$$

where E denotes expectation. We omit the detail of deriving this identity, which follows from the independent exponential distributions of T_1 and T_2, the renewal theory in stochastic processes, and Proposition 3.4.5 in [43]. From (60), we obtain

$$\mathrm{E}[V(T_1)] = \int_0^\infty V(t)\kappa_1 e^{-\kappa_1 t}dt = \int_0^\infty \kappa_1 e^{(a-\kappa_1)t}dt = \frac{\kappa_1}{\kappa_1 - a}.$$

The distribution function of $T_1 + T_2$ is given by

$$
\begin{aligned}
F(t) &= \mathrm{Prob}\{T_1 + T_2 \le t\} \\
&= \int_0^t \mathrm{Prob}\{T_1 \le t - x\}d\mathrm{Prob}\{T_2 \le x\} \\
&= \int_0^t (1 - e^{-\kappa_1(t-x)})\kappa_2 e^{-\kappa_2 x}dx \\
&= 1 - \frac{\kappa_2}{\kappa_2 - \kappa_1}e^{-\kappa_1 t} - \frac{\kappa_1}{\kappa_1 - \kappa_2}e^{-\kappa_2 t},
\end{aligned}
$$

when $\kappa_1 \ne \kappa_2$ [5, 8]. By taking limits, it applies to the case $\kappa_1 = \kappa_2$ as well. Thus

$$\mathrm{E}[V(T_1 + T_2)] = \int_0^\infty V(t)dF(t) = \frac{\kappa_1 \kappa_2}{(\kappa_1 - a)(\kappa_2 - a)}.$$

Since it also holds that $\mathrm{E}[V(T_1 + T_2)] = 2$, we derive

$$\frac{\kappa_1 \kappa_2}{(\kappa_1 - a)(\kappa_2 - a)} = 2$$

from which the expression of a in (61) is obtained. The second part in (61) can be verified by dividing the expressions of $\mathrm{E}[V(T_1 + T_2)]$ and $\mathrm{E}[V(T_1)]$ directly.

We apply our mathematical formulas (42) and (61) to examine how the transcript concentration homeostasis is attained by varying the transcription frequency λ_i, or the burst size v_i/γ_i, $i = 1, 2$, in different cell cycle phases, and how the noise responds to the variation. In a recent interesting study elucidating the mechanism underlying the concentration homeostasis, Padovan-Merhar et al. [18] found a DNA-linked *cis*-acting factor in mammalian cells, that can lead
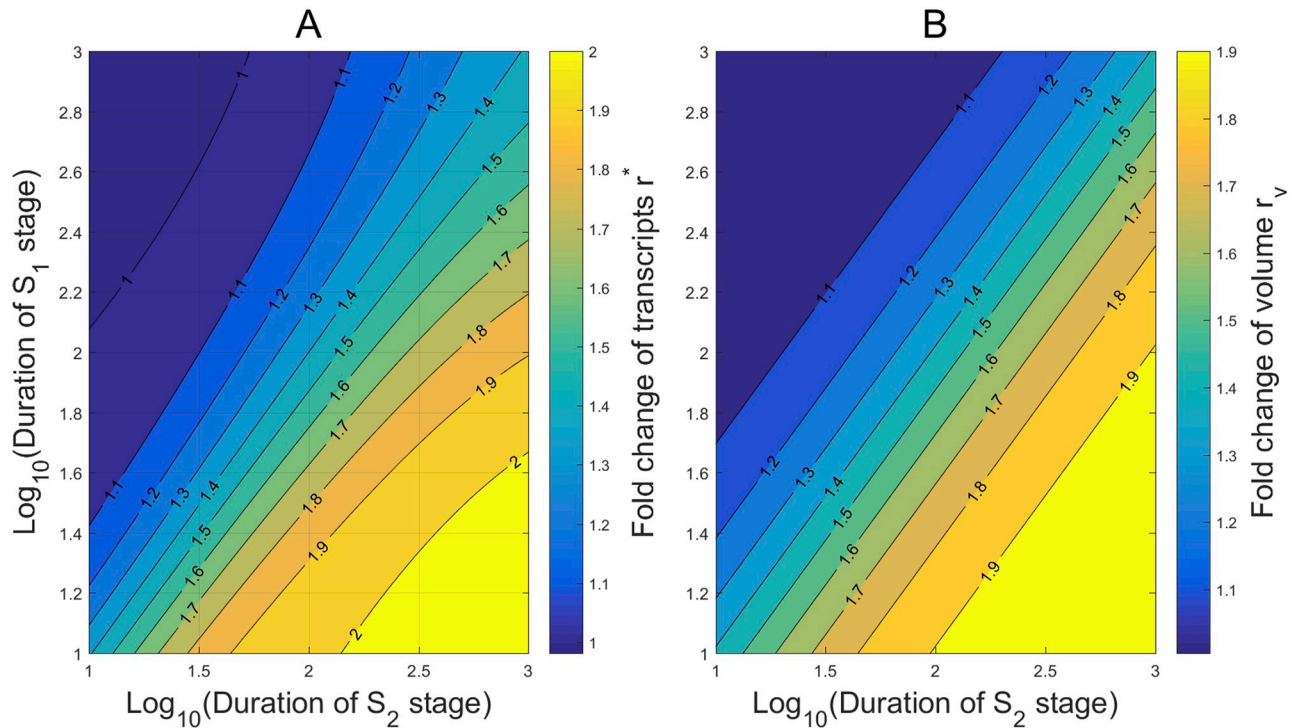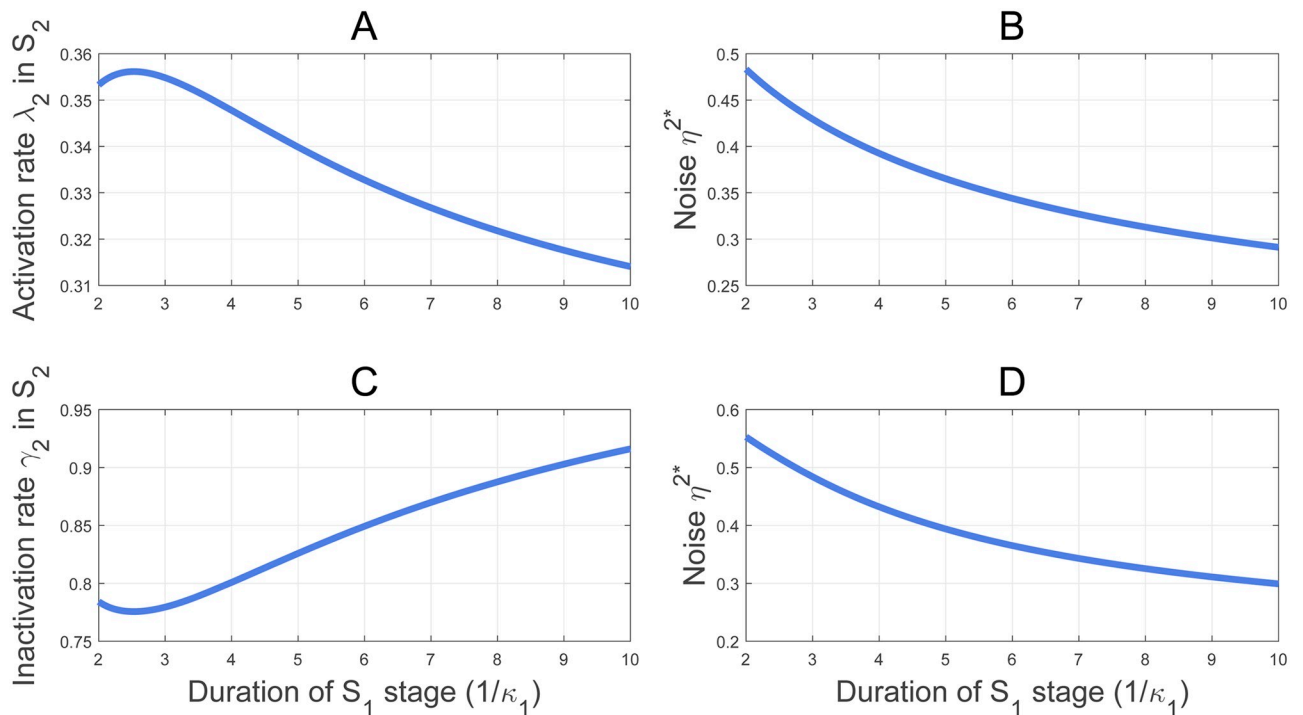
Fig 7. The transcript concentration homeostasis by reducing the transcription burst frequency in G_2 phase. The burst frequency λ_2 is reduced from 0.5556 in (51) to 0.3445, and other kinetic rates are taken from (51). (A) The contours of the fold change r^* for the mean transcription levels from $\mathbb{S}_1$ to $\mathbb{S}_2$. (B) The contours of the fold change r_v for the volumes from $\mathbb{S}_1$ to $\mathbb{S}_2$. The two sets of contours in (A) and (B) display a large similarity, indicating that the transcript concentration homeostasis is almost reached by reducing the frequency uniformly over most area of the stage durations.

to the reduction of transcription frequency of each gene copy in G_2 phase, and therefore an overall balanced burst frequency in all cell cycle stages. Motivated by this observation, we reduce λ_2 in (51) from 0.5556 to 0.3445, which is approximately a 38% reduction, and keep other kinetic rates as in (51). As the stage durations vary from 10 min to 1, 000 min, the contours of the fold change r^* for the mean transcription levels and the fold change r_v for the volumes are shown in Fig 7. Very interestingly, the contour of r^* in Fig 7A and the contour of r_v in Fig 7B are very similar over most area of the stage durations. This surprising similarity suggests that the concentration homeostasis could be reached by merely reducing the transcription burst frequency in G_2 phase, in a uniform reduction degree that is robust against a large variation of cell stage durations and the randomness of gene duplication time.

To test further the robustness in the reduction of the transcription burst frequency in $\mathbb{S}_2$ against the variation of cell cycle stage durations, we fix the kinetic rates except λ_2 as in (51), the duration of $\mathbb{S}_2$ at $\kappa_2^{-1} = 220$ min as in (58), but let the duration of $\mathbb{S}_1$ vary from 2 to 10 hours. By applying (42) and (61) to $r^* = r_v$, we determine a unique λ_2 corresponding to each $\mathbb{S}_1$ duration for the transcript concentration homeostasis. The curve of λ_2 versus the $\mathbb{S}_1$ duration $1/\kappa_1$ is shown in Fig 8A. It is interesting to see that λ_2 exhibits only a minor variation within a narrow interval (0.315, 0.355) over a large span of the $\mathbb{S}_1$ duration from 2 to 10 hours. As $\lambda_1 = 0.5556$, this corresponds to a stable reduction, between 36% and 42%, of the transcription frequency from $\mathbb{S}_1$ to $\mathbb{S}_2$. We note also that $\lambda_2 = 0.3445$ fixed in Fig 7 is within (0.315, 0.355). More interestingly, the transcription noise sketched in Fig 8B, corresponding to the λ_2 values of Fig 8A, also exhibits a minor variation within the narrow range (0.28, 0.48).

Fig 8. Minor variations of λ_2, γ_2 and η^{2*} versus the large variation of the cell cycle durations in the concentration homeostasis. (A) The transcription frequency for the concentration homeostasis stays within (0.315, 0.355) over a large span of the $\mathbb{S}_1$ duration from 2 to 10 hours; $\kappa_2^{-1} = 220$ min, and other kinetic rates are taken from (58). (B) The noise corresponding to (A) remains within (0.28, 0.48). (C) The inactivation rate γ_2 for the concentration homeostasis with the parameters in (A) except $\lambda_2 = \lambda_1/2$ to equalize the burst frequency in $\mathbb{S}_1$ and $\mathbb{S}_2$. (D) The noise corresponding to (C) remains within (0.29, 0.56).

In Fig 8A and 8B, the transcription burst size v_i/γ_i does not change in $\mathbb{S}_1$ and $\mathbb{S}_2$ stages. However, Caveney et al. [21] and Padovan-Merhar et al. [18] found that the transcription burst size increases in larger cells, while the burst frequency in $\mathbb{S}_2$ shows no dramatic changes. These findings suggest that the increase in the burst size could be responsible for the concentration homeostasis. We are motivated to test how the burst size increase induces the concentration homeostasis by reducing the gene OFF rate γ_2. We set λ_1, γ_1, and the mRNA synthesis and degradation rates as in (51), and define $\lambda_2 = \lambda_1/2$ to equalize the burst frequency on the two stages. As in Fig 8A and 8B, we fix $\kappa_2^{-1} = 220$ min, and let the average duration of $\mathbb{S}_1$ vary from 2 to 10 hours. For each $\mathbb{S}_1$ duration $1/\kappa_1$, there is a unique γ_2 determined by the concentration homeostasis identity $r^* = r_v$ as shown in Fig 8C. Again, γ_2 displays only a minor variation within the narrow interval (0.78, 0.92) in Fig 8C where the $\mathbb{S}_1$ duration increases from 2 to 10 hours. However, in contrast to the minor variation of γ_2, the corresponding increase in the bust size is noticeable. As $\gamma_1 = 1.0714$, the increase in the burst size in $\mathbb{S}_2$, given by $(\gamma_1 - \gamma_2)/\gamma_2$ ranges from 16.45% to 37.35%. Very interestingly, the corresponding transcription noise sketched in Fig 8D does not show a noticeable variation but stays within the narrow range (0.29, 0.56).

The minor variation of η^{2*} in Fig 8B and 8D suggests that the transcription noise is relatively stable when the transcript concentration homeostasis is maintained, either by reducing the transcription burst frequency or by increasing the burst size in late cell cycle phase, in the face of a large cell cycle stage duration variation. By comparing the profiles of η^{2*} in Fig 8B and 8D it is also seen that the burst size variation creates a slightly larger noise than reducing the burst frequency, similar to the observation by Caveney et al. [21] in cell-free synthetic

reaction chambers. Furthermore, our simulations in Figs 7 and 8A and 8C show that the reduction degree in the burst frequency for the transcript concentration homeostasis is relative robust, while the increase in the burst size is conceivably sensitive, to the random variation of the cell cycle durations.

Conclusion

Gene transcription involves inherently various probabilistic steps that create fluctuations in mRNA and protein counts [1–3]. The random transitions between the active and inactive promoter states have been widely invoked to explain the fluctuation of mRNA numbers among individual cells of identical genes [6, 7]. Recent experimental studies have revealed that the cell division cycle has global effects on transcriptional outputs, and is thought to be an additional source of transcription noise [10–13].

In this work, we integrated cell division cycles into an extended two-state model to delineate the combined contribution of transcription activities and cell divisions in the variability of transcript counts [4, 6]. In the model, a cell division cycle is divided into $\mathbb{S}_1$ stage before the duplication of a target gene and the late stage $\mathbb{S}_2$, on which the durations are independently and exponentially distributed with rates κ_1 and κ_2. When a cell divides, each mRNA molecule has an even chance of being partitioned to one of the two daughter cells. We defined two joint probabilities to quantify the system state, and derived the master equations of their time evolutions. From the master equations we obtained the differential equations of the mean and the second moment of mRNA numbers in single cells. By solving these equations we presented in Theorem 1 the steady-state mean transcription level m^* in cells, together with the means m_1^* and m_2^* on the two stages. The analytical forms of the second moments are presented in Theorem 2, which in turn help us determine six noise measures: the noise η^{2*} and the noise strength Φ^* without referring to cell cycle stages, along with η_1^{2*} and Φ_1^* in $\mathbb{S}_1$, and η_2^{2*} and Φ_2^* in $\mathbb{S}_2$. The fold change of mRNA counts from $\mathbb{S}_1$ to $\mathbb{S}_2$ is quantified by $r^* = m_2^*/m_1^*$. As a cell contains twice as many copies of each gene in $\mathbb{S}_2$ as that in $\mathbb{S}_1$, one may envisage by intuition that $r^* \approx 2$. However, our Theorem 3 shows that r^* can take any prescribed value in theory, although we also found that $r^* \approx 2$ when the transcription kinetics are unchanged in the two stages, and stage transitions are considerably slower than mRNA turnover and transcription state transitions. The dependance of r^* on κ_1 and κ_2 is examined deeply in Theorem 4, where the necessary and sufficient conditions for $r^* < 1$ or $r^* > 2$ are identified. In particular, it is proved that r^* has an upper bound strictly less than 2 when $\kappa_1 \leq \kappa_2$.

We tested the accuracy of our analytical results against various experimental data. For a gene in a mouse embryonic stem cell line, our result predicts $r^* = 1.2791$, which offers a good match with $r^* = 1.28 \pm 0.09$ measured in [15]. The analysis also indicates that if the transcription kinetics do not change considerably in the two cell cycle stages, then the average mRNA counts increase about 1 to 2 folds from $\mathbb{S}_1$ stage to $\mathbb{S}_2$ stage as observed in mouse embryonic cells [15] and yeast [12]. The mean m^* increases while the noise η^{2*} decreases in each of the cell cycle durations. Rapid transitions between cell cycle stages were identified as a major source of highly noisy transcription. Eukaryotic cells have a DNA dosage-compensating mechanism to reduce mRNA production in late cell cycle stage, resulting in gene transcription homeostasis that overall transcription remains constant across $\mathbb{S}_1$ and $\mathbb{S}_2$ stages [13, 18]. Our analysis reveals that transcription homeostasis does not bring significant changes in transcription noise. If transcription homeostasis is attained by varying a single kinetic rate in the two cell cycle stages, then the homeostasis nearly minimizes transcription noise. As many cellular processes depend on the concentration of enzymes rather than their absolute numbers for proper cellular function [18, 21, 22], we also studied the noise profile when the transcript

concentration homeostasis is maintained that the mean transcription level scales with the cellular volume in $\mathbb{S}_1$ and $\mathbb{S}_2$. We found that the transcription noise is relatively stable when the transcript concentration homeostasis is maintained, either by reducing the transcription burst frequency or by increasing the burst size in late cell cycle phase, in the face of a large cell cycle stage duration variation. The reduction degree in the burst frequency is relative robust, while the increase in the burst size is conceivably sensitive, to the large random variation of the cell cycle durations and the gene duplication time.

This work provides one of the first theoretical explorations on how the coupling of stochastic promoter state transitions and cell cycle progressions regulates transcription abundance and noise. It presents a core model for further inclusion of more complex transcription kinetics and cell cycle progressions. The kinetic rates may display large variations in different cell cycle phases or within the same phase, or oscillate periodically in the cell cycle progression [44]. With the expansion of the model, motivated and tested by more upcoming experimental data, the approach initiated here is expected to be developed further to help understand the role played by the cell cycle dependent gene expression in cell functions and cell fate decision [45, 46].

Supporting information

S1 Text. The supporting information consists of two parts: The derivation of differential equations and the proof of theorems. The detailed derivation of master equations and differential equations is given in the former part. Theorems 1-4 are restated in the second part, and the detailed mathematical proofs are given behind these theorems.
(PDF)

Author Contributions

Conceptualization: Jianshe Yu, Moxun Tang.

Formal analysis: Qiwen Sun, Feng Jiao, Genghong Lin.

Software: Qiwen Sun, Feng Jiao, Genghong Lin.

Supervision: Jianshe Yu, Moxun Tang.

Writing – original draft: Qiwen Sun, Moxun Tang.

Writing – review & editing: Qiwen Sun, Moxun Tang.

References

1. Kaufmann BB, van Oudenaarden A. Stochastic gene expression: from single molecules to the proteome. Current Opinion in Genetics & Development. 2007; 17:107–112. https://doi.org/10.1016/j.gde.2007.02.007

2. Maheshri N, O'Shea EK. Living with noisy genes: how cells function reliably with inherent variability in gene expression. Annual Review of Biophysics and Biomolecular Structure. 2007; 36:413–434. https://doi.org/10.1146/annurev.biophys.36.040306.132705 PMID: 17477840

3. Neuert G, Munsky B, Tan RZ, Teytelman L, Khammash M, van Oudenaarden A. Systematic identification of signal-activated stochastic gene regulation. Science. 2013; 339:584–587. https://doi.org/10.1126/science.1231456 PMID: 23372015

4. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. Stochastic mRNA synthesis in mammalian cells. PLoS Biology. 2006; 4:1707–1719. https://doi.org/10.1371/journal.pbio.0040309

5. Sun Q, Tang M, Yu J. Modulation of gene transcription noise by competing transcription factors. Journal of Mathematical Biology. 2012; 62:469–494.

6. Dar RD, Razooky BS, Singh A, Trimeloni TV, McCollum JM, Cox CD, et al. Transcriptional burst frequency and burst size are equally modulated across the human genome. Proc Natl Acad Sci U S A. 2012; 109:17454–17459. https://doi.org/10.1073/pnas.1213530109 PMID: 23064634

7. Munsky B, Neuert G, van Oudenaarden A. Using gene expression noise to understand gene regulation. Science. 2012; 336:183–187. https://doi.org/10.1126/science.1216379 PMID: 22499939

8. Sun Q, Tang M, Yu J. Temporal profile of gene transcription noise modulated by cross-talking signal transduction pathways. Bulletin of Mathematical Biology. 2012; 74:375–398. https://doi.org/10.1007/s11538-011-9683-z PMID: 21870200

9. Zopf CJ, Quinn K, Zeidman J, Maheshri N. Cell-cycle dependence of transcription dominates noise in gene expression. PLoS Computational Biology. 2013; 9(7):e1003161. https://doi.org/10.1371/journal.pcbi.1003161 PMID: 23935476

10. Bertoli C, Skotheim JM, de Bruin RAM. Control of cell cycle transcription during G1 and S phases. Nature Reviews Molecular Cell Biology. 2013; 14:518–528. https://doi.org/10.1038/nrm3629 PMID: 23877564

11. Dowling MR, Kan A, Heinzel S, Zhou JHS, Marchingo JM, Wellard CJ, et al. Stretched cell cycle model for proliferating lymphocytes. Proc Natl Acad Sci U S A. 2014; 111:6377–6382. https://doi.org/10.1073/pnas.1322420111 PMID: 24733943

12. Trcek T, Larson DR, Moldón A, Query CC, Singer RH. Single-molecule mRNA decay measurements reveal promoter-regulated mRNA stability in yeast. Cell. 2011; 147:1484–1497. https://doi.org/10.1016/j.cell.2011.11.051 PMID: 22196726

13. Voichek Y, Bar-Ziv R, Barkai N. Expression homeostasis during DNA replication. Science. 2016; 351:1087–1090. https://doi.org/10.1126/science.aad1162 PMID: 26941319

14. Marguerat S, Bälher J. Coordinating genome expression with cell size. Trends in Genetics. 2012; 28:560–565. https://doi.org/10.1016/j.tig.2012.07.003 PMID: 22863032

15. Skinner SO, Xu H, Nagarkar-Jaiswal S, Freire PR, Zwaka TP, Golding I. Single-cell analysis of transcription kinetics across the cell cycle. eLife. 2016; 5:e12175. https://doi.org/10.7554/eLife.12175 PMID: 26824388

16. Fraser HB. Cell-cycle regulated transcription assocates with DNA replication timing in yeast and human. Genome Biology. 2013; 14:R111. https://doi.org/10.1186/gb-2013-14-10-r111 PMID: 24098959

17. Graham JE, Marians KJ, Kowalczykowski SC. Independent and stochastic action of DNA polymerases in the replisome. Cell. 2017; 169:1201–1213. https://doi.org/10.1016/j.cell.2017.05.041 PMID: 28622507

18. Padovan-Merhar O, Nair GP, Biaesch AG, Mayer A, Scarfone S, Foley SW, et al. Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. Molecular Cell. 2015; 58(2): 339–352. https://doi.org/10.1016/j.molcel.2015.03.005 PMID: 25866248

19. Huh D, Paulsson J. Random partitioning of molecules at cel division. Proc Natl Acad Sci U S A. 2011; 108:15004–15009.

20. Yu J, Sun Q, Tang M. The nonlinear dynamics and fluctuations of mRNA levels in cross-talking pathway activated transcription. Journal of Theoretical Biology. 2014; 363:223–234. https://doi.org/10.1016/j.jtbi.2014.08.024 PMID: 25152214

21. Caveney PM, Norred SE, Chin CW, Boreyko JB, Razooky BS, Retterer ST, et al. Resource sharing controls gene expression bursting. ACS Synthetic Biology. 2017; 6:334–343. https://doi.org/10.1021/acssynbio.6b00189 PMID: 27690390

22. Kempe H, Schwabe A, Crémazy F, Verschure PJ, Bruggeman FJ. The volumes and transcript counts of single cells reveal concentration homeostasis and capture biological noise. Molecular Biology of the Cell. 2015; 26:797–804. https://doi.org/10.1091/mbc.E14-08-1296 PMID: 25518937

23. Kuang J, Tang M, Yu J. The mean and noise of protein numbers in stochastic gene expression. Journal of Mathematical Biology. 2013; 67:261–291. https://doi.org/10.1007/s00285-012-0551-8 PMID: 22638878

24. Cameron IL, Greulich RC. Evidence for an essentially constant duration of DNA synthesis in renewing epithelia of the adult mouse. The Journal of Cell Biology. 1963; 18:31–40. https://doi.org/10.1083/jcb.18.1.31 PMID: 14018040

25. Pham TM, Tan KW, Sakumura Y, Okumura K, Maki H, Akiyama MT. A single-molecule approach to DNA replication in *Escherichia coli* cells demonstrated that DNA polymerase III is a major determinant of fork speed. Molecular Microbiology. 2013; 90:584–596. https://doi.org/10.1111/mmi.12386 PMID: 23998701

26. Rogozin IB, Makarova KS, Natale DA, Spiridonov AN, Tatusov RL, Wolf YI, et al. Congruent evolution of different classes of non-coding DNA in prokaryotic genomes. Nucleic Acids Research. 2002; 30:4264–4271. https://doi.org/10.1093/nar/gkf549 PMID: 12364605

27. Jovic D, Sakaue-Sawano A, Abe T, Cho CS, Nagaoka M, Miyawaki A, et al. Direct observation of cell cycle progression in living mouse embryonic stem cells on an extracellular matrix of E-cadherin. SpringerPlus. 2013; 2:585. https://doi.org/10.1186/2193-1801-2-585 PMID: 25674414

28. White J, Dalton S. Cell cycle control of embryonic stem cells. Stem Cell Reviews. 2005; 1:131–138. https://doi.org/10.1385/SCR:1:2:131 PMID: 17142847

29. Duderstadt KE, Geertsema HJ, Stratmann SA, Punter CM, Kulczyk AW, Richardson CC, et al. Simultaneous real-time imaging of leading and lagging strand synthesis reveals the coordination dynamics of single replisomes. Molecular Cell. 2016; 64:1035–1047. https://doi.org/10.1016/j.molcel.2016.10.028 PMID: 27889453

30. Knezetic JA, Luse DS. The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. Cell. 1986; 45:95–104. https://doi.org/10.1016/0092-8674(86)90541-6 PMID: 3955658

31. Lorch Y, LaPointe JW, Korgberg RD. Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell. 1987; 49:203–210. https://doi.org/10.1016/0092-8674(87)90561-7 PMID: 3568125

32. Shemoen AW, O'Farrell PH. Progression of the cell cycle through mitosis leads to abortion of nascent transcripts. Cell. 1991; 67(2):303–310. https://doi.org/10.1016/0092-8674(91)90182-X

33. Golding I, Paulsson J, Zawilski SM, Cox EC. Real-time kinetics of gene activity in individual bacteria. Cell. 2005; 123:1025–1036. https://doi.org/10.1016/j.cell.2005.09.031 PMID: 16360033

34. Vintersten K, Monetti C, Gertsenstein M, Zhang P, Laszlo L, Biechele S, et al. Mouse in red: red fluorescent protein expression in mouse ES cells, embryos, and adult animals. Genesis. 2004; 40:241–246. https://doi.org/10.1002/gene.20095 PMID: 15593332

35. Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea E, Pilpel Y, et al. Noise in protein expression scales with natural protein abundance. Nature Genetics. 2006; 38:636–643. https://doi.org/10.1038/ng1807 PMID: 16715097

36. Cartwright P, McLean C, Sheppard A, Rivett D, Jones K, Dalton S. LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. Development. 2005; 132:885–896. https://doi.org/10.1242/dev.01670 PMID: 15673569

37. Dohrmann PR, Correa R, Frisch RL, Rosenberg SM, McHenry CS. The DNA polymerase III holoenzyme contains γ and is not a trimeric polymerase. Nucleic Acids Research. 2016; 44:1285–1297. https://doi.org/10.1093/nar/gkv1510 PMID: 26786318

38. Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M, Hearn J, et al. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. Science. 2010; 329:533–538. https://doi.org/10.1126/science.1188308 PMID: 20671182

39. Yu J, Xiao J, Ren X, Lao K, Xie XS. Probing gene expression in live cells, one protein molecule at a time. Science. 2006; 311:1600–1603. https://doi.org/10.1126/science.1119623 PMID: 16543458

40. Campos M, Surovtsev IV, Kato S, Paintdakhi A, Beltran B, Ebmeier SE, Jacobs-Wagner C. A constant size extension drives bacterial cell size homeostasis. Cell. 2014; 159:1433–1446. https://doi.org/10.1016/j.cell.2014.11.022 PMID: 25480302

41. Schmoller KM. The phenomenology of cell size control. Current Opinion in Cell Biology. 2017; 49:53–58. https://doi.org/10.1016/j.ceb.2017.11.011 PMID: 29232627

42. Si F, Li D, Cox SE, Sauls JT, Azizi O, Sou C, et al. Invariance of initiation mass and predictability of cell size in Escherichia coli. Current Biology. 2017; 27:1–10. https://doi.org/10.1016/j.cub.2017.03.022

43. Ross SM. Stochastic Processes. 2nd ed. John Wiley & Sons, Inc.; 1996.

44. Müller GA, Stangner K, Schmitt T, Wintsche A, Engeland K. Timing of transcription during the cell cycle: Protein complexes binding to E2F, E2F/CLE, CDE/CHR, or CHR promoter elements define early and late cell cycle gene expression. Oncotarget. 2017; 8:97736–97748. https://doi.org/10.18632/oncotarget.10888 PMID: 29228647

45. Gonzales KAU, Liang H, Lim YS, Chan YS, Yeo JC, Tan CP, et al. Deterministic restriction on pluripotent state dissolution by cell-cycle pathways. Cell. 2015; 162:564–579. https://doi.org/10.1016/j.cell.2015.07.001 PMID: 26232226

46. Pauklin S, Vallier L. The cell-cycle state of stem cells determines cell fate propensity. Cell. 2012; 155(1):135–147. https://doi.org/10.1016/j.cell.2013.08.031