



Research article

Analysis of a non-Markov transcription model with nuclear RNA export and RNA nuclear retention

Qiqi Deng¹, Aimin Chen², Huahai Qiu³ and Tianshou Zhou^{1,*}

¹ Guangdong Province Key Laboratory of Computational Science, School of Mathematics, Sun Yat-sen University, Guangzhou 510275, China

² School of Mathematics and Statistics, Henan University, Kaifeng 475004, China

³ School of Mathematical and Physical Sciences, Wuhan Textile University, Wuhan 430200, China

* **Correspondence:** Email: mcszhtsh@mail.sysu.edu.cn.

Abstract: Transcription involves gene activation, nuclear RNA export (NRE) and RNA nuclear retention (RNR). All these processes are multistep and biochemical. A multistep reaction process can create memories between reaction events, leading to non-Markovian kinetics. This raises an unsolved issue: how does molecular memory affect stochastic transcription in the case that NRE and RNR are simultaneously considered? To address this issue, we analyze a non-Markov model, which considers multistep activation, multistep NRE and multistep RNR can interpret many experimental phenomena. In order to solve this model, we introduce an effective transition rate for each reaction. These effective transition rates, which explicitly decode the effect of molecular memory, can transform the original non-Markov issue into an equivalent Markov one. Based on this technique, we derive analytical results, showing that molecular memory can significantly affect the nuclear and cytoplasmic mRNA mean and noise. In addition to the results providing insights into the role of molecular memory in gene expression, our modeling and analysis provide a paradigm for studying more complex stochastic transcription processes.

Keywords: transcription; RNA nuclear retention; molecular memory; non-Markov model; transcription noise

1. Introduction

Transcription is not only the most important but also the most complex step in gene expression.

This double characteristic makes gene transcription receive lasting and extensive attention. With the development of measurement technologies (e.g., single-cell and single-molecule technologies), more molecular details of transcription have been experimentally uncovered. Nevertheless, some intermediate processes would not have been specified due to the complexity of gene transcription. Thus, traditional Markov models of gene transcription such as extensively studied ON-OFF models [1–4] would not interpret experimental phenomena nor reveal the stochastic mechanisms of transcription. More biologically reasonable mathematical models need to be developed.

It is well known that gene transcription involves RNA nuclear retention (RNR) and nuclear RNA export (NRE). However, these two important processes were often ignored in previous studies [1–8]. Main reasons include that 1) NRE was before considered to be a transient process, compared to other processes occurring in transcription. It was reported that the NRE phase lasted about 20min on average and were gene-specific [9,10]; 2) For RNR, less than 30% poly (A+) RNA is nuclear-retained and undetectable in the cytoplasm [11]. Currently, more and more experimental evidences have indicated that RNR play a key role in biological processes, e.g., in *S. cerevisiae* cells, RNR may play a precautionary role during stress situations [12]; in plants, the RNR process of NLP7 can orchestrate the early response to nitrate [13]; and in the signaling pathway of antiviral innate immunity, the RNA helicase DDX46 acts like a negative regulator to induce nuclear retention of antiviral innate transcripts [14]. These experimental facts suggest that RNR or NRE cannot be neglected when one makes theoretical predictions of gene expression (including expression levels and noise).

Several works have revealed the respective roles of NRE and RNR in modulating stochastic gene expression [15,16–19]. One report exhibited that transcriptional burst attributed to promoter state switching could be substantially attenuated by the transport of mRNA from nucleus to cytoplasm [17]. Another report showed that slow pre-mRNA export from the nucleus could be an effective mechanism to attenuate protein variability arising from transcriptional burst [15]. In addition, RNR was also identified to buffer transcriptional burst in tissues and mammalian cells [16,18]. However, it has been experimentally confirmed that NRE and RNR can occur simultaneously in eukaryotes [20]. How these two dynamic processes cooperatively affect gene expression remains elusive and even unexplored.

As a matter of fact, gene activation, NRE and RNR are multistep processes. In general, transcription begins only when the chromatin template accumulates over time until the promoter becomes active [21,22], where the accumulating process is a multistep biochemical process in which some intermediate steps would not have been specified due to experimental technologies. A representative example is that inactive phases of the promoter involving the prolactin gene in a mammalian cell are differently distributed, showing strong memory [23]. Similarly, both the export of mRNAs generated in the nucleus to the cytoplasm through nuclear pores and the retention of mRNAs among nuclear speckles or paraspeckles are in general also multistep reaction processes [24]. All these multistep processes can create memories between reaction events, leading to non-Markov dynamics. Traditional Markov models are no longer suitable to the modeling of gene transcription with molecular memory, and non-Markov models can well model multistep processes involved in gene transcription [7].

In this paper, we introduce a non-Markov model of stochastic gene transcription. It considers not only RNR and NRE but also molecular memories created due to the multistep NRE process, due to RNR process, or due to the multistep activation process, thus including previous transcription models [1–4] as its special case. In order to solve this non-Markov model, we introduce effective transition rates, which explicitly decode the effect of molecular memory and by which we can transform the original non-Markov issue into an equivalent yet mathematical tractable Markov one. Based on this useful

technique, we derive analytical results, which extend previous results [3,8,24,25] and provide insights into the role of molecular memory in affecting the nuclear and cytoplasmic mRNA mean and noise. The overall modeling and analysis provide a paradigm for studying more complex stochastic transcription processes.

2. Method

2.1. Model hypotheses

Most previous studies [15,20,26] of gene transcription focused on the dynamics of NRE processes, where mature mRNAs were released to the cytoplasm with the help of nuclear pore complex (NPC) (Figure 1) [16,27]. The number of NPCs or the count of those assistant proteins that controlled NPC determined the speed of exporting mRNA. Measuring the exporting rate was often replaced by measuring the retention time in the nucleus, which however may vary with environmental changes [16–19]. Other previous studies of gene transcription [9–11,28,29] focused on the dynamics of transcription initiation and elongation, where elongation time (T) was measured by the length of a gene using the length of bases (L) and the average rate of elongation (v), i.e., $T = L/v$. These studies assumed that all mature mRNAs were first exported to the cytoplasm and then translated into proteins. However, biological experiments indicated that there were a considerable part of mature mRNAs that stayed in the nucleus in a probabilistic manner and lasted a long period (Figure 1) [24].

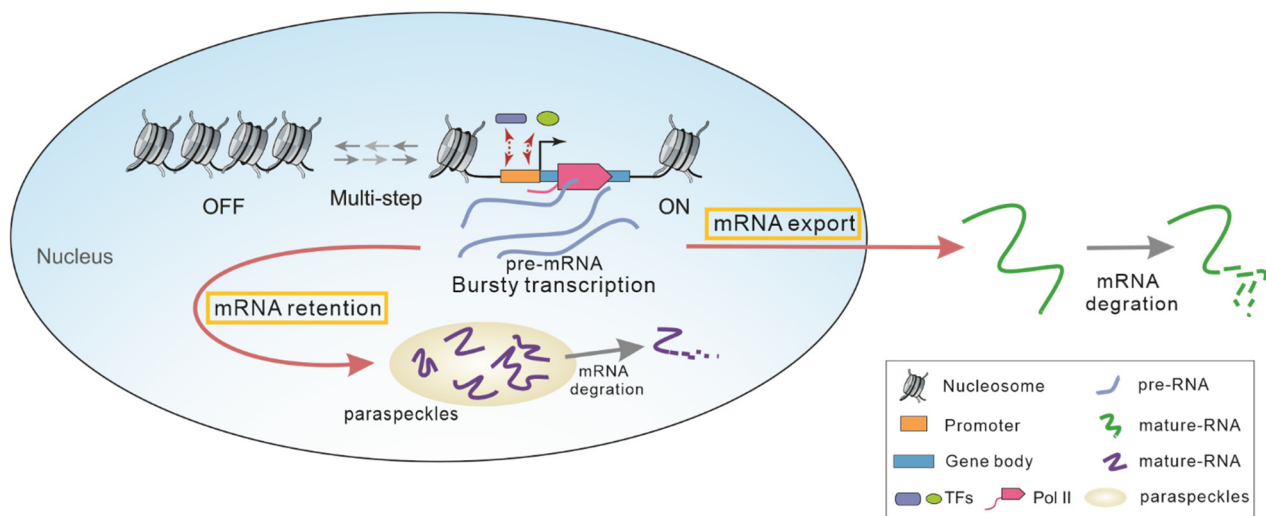


Figure 1. Schematic diagram for a model of stochastic gene transcription. First, chromatin (consisting of nucleosomes) opens in a multistep manner, and then, DNA is transcribed into mRNAs also in a multistep manner. Some of these mRNAs are remained in the nucleus (forming so-called paraspeckles) in a multistep manner, and the others are exported into the cytoplasm through the nuclear pores also in a multistep manner.

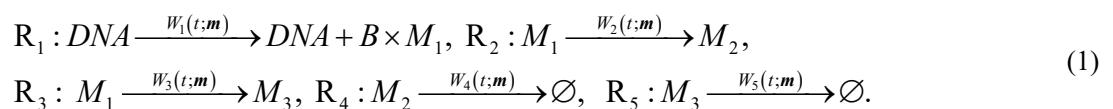
Here, we consider two cases: one case where NRE dominates over RNR and the other case where RNR dominates over NRE. For both cases, the gene is assumed to have one “off” state (corresponding to the inactive form of the promoter) and one “on” state (corresponding to the active form), and the

promoter is assumed to switch randomly between these two states. Only in “on” state, can the gene generate pre-mRNA. After an alternative splicing (AS) process or an alternative polyadenylation (APA) process, which occurs frequently at the 3' UTR, a portion of mature mRNAs (one type of transcripts) may be transported to the cytoplasm through NPC wherein they execute translation tasks. The rest mature mRNAs (another type of transcripts) may be remained in the nucleus for a long time, possibly assembling on the sub-cellular region (wherein they form nuclear speckles or paraspeckles [30–32]) with the assistance of proteins, some of which would have been unspecified. When the intracellular environment is changed, most of mature mRNAs will release to the cytoplasm in response to this change. In addition, most genes (especially in eukaryotic cells) are expressed in a bursty manner [1–4].

As pointed out in the introduction, gene transcription, NRE and RNR are all multistep reaction processes. In order to model these processes, we introduce a non-exponential waiting-time distribution for each intermediate reaction as done in ref. [7,33]. Since non-exponential waiting times lead to non-Markovian dynamics, the existing Markov theory cannot be used directly.

2.2. Mathematical models

Assume that burst size in gene transcription follows a distribution described by $\text{prob}\{B = i\} = \alpha_i$, where each α_i is a nonnegative constant and $i = 0, 1, 2, \dots$. Let M_1 , M_2 and M_3 represent pre-mRNA, mature mRNA (one type of transcripts) transported to the cytoplasm and mature mRNA (another type of transcripts) retained in the nucleus respectively, and denote by m_1 , m_2 and m_3 their molecular numbers respectively. Thus, $\mathbf{m} = (m_1, m_2, m_3)^T$ represents the micro-state of the underlying system. Let $W_1(t; \mathbf{m})$, $W_2(t; \mathbf{m})$ and $W_3(t; \mathbf{m})$ be waiting-time distributions for the syntheses of pre-mRNAs, mature mRNA transported to the cytoplasm, and mature mRNA retained in the nucleus, respectively. Let $W_4(t; \mathbf{m})$ and $W_5(t; \mathbf{m})$ be waiting-time distributions for degradations of M_2 and M_3 , respectively. To that end, the gene-expression model to be studied are described by the following five biochemical reactions labelled by R_i ($1 \leq i \leq 5$)



Let $\langle B \rangle$ represent the mean burst size. Note that if $\alpha_1 = 1$ and $\alpha_k = 0$ for all $k \neq 1$, this implies $\langle B \rangle \equiv 1$. In this case, the promoter is in the ON state all the time, and Eq (1) describes the constitutive gene expression. The other cases correspond to bursty gene expression. This is because $B = 0$ implies that the promoter is in the OFF state (meaning that pre-mRNAs are not generated), whereas $B > 0$ implies that the promoter is at the ON state (meaning that pre-mRNAs are generated).

For each reaction, there is a memory function [7,33]. Denote by $M_i(t; \mathbf{m})$ memory function for reaction R_i ($1 \leq i \leq 5$). These memory functions can be expressed by waiting-time distributions in Eq (1).

In fact, if we let $\tilde{M}_i(s; \mathbf{m})$ be the Laplace transform of memory function $M_i(t; \mathbf{m})$, then $\tilde{M}_i(s; \mathbf{m})$ can be expressed as $\tilde{M}_i(s; \mathbf{m}) = s\tilde{\varphi}_i(s; \mathbf{m}) / \left[1 - \sum_{i=1}^5 \tilde{\varphi}_i(s; \mathbf{m}) \right]$, where $\tilde{\varphi}_i(s; \mathbf{m})$ is the Laplace

transform of function $\varphi_i(t; \mathbf{m})$ that can be expressed as $\varphi_i(t; \mathbf{m}) = W_i(t; \mathbf{m}) \prod_{k \neq i} \left[1 - \int_0^t W_k(t'; \mathbf{m}) dt' \right]$

($1 \leq i \leq 5$) [7]. Let $P(\mathbf{m}; t)$ be the probability that the system is in state \mathbf{m} at time t and $\tilde{P}(\mathbf{m}; s)$ be the Laplace transform of $P(\mathbf{m}; t)$. With $\tilde{M}_i(s; \mathbf{m})$, we can show that the chemical master equation in the sense of Laplace transform takes the form

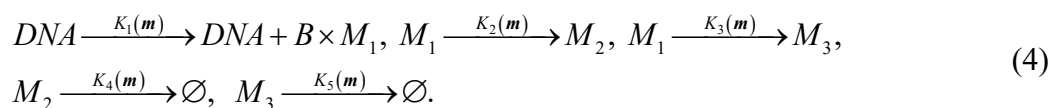
$$s\tilde{P}(\mathbf{m}; s) - P(\mathbf{0}; s) = \left(\sum_{i=0}^{m_1} \alpha_i \mathbb{E}_1^{-i} - \mathbb{I} \right) \left[\tilde{M}_1(s; \mathbf{m}) \tilde{P}(\mathbf{m}; s) \right] + \left(\mathbb{E}_1 \mathbb{E}_2^{-1} - \mathbb{I} \right) \left[\tilde{M}_2(s; \mathbf{m}) \tilde{P}(\mathbf{m}; s) \right] \\ + \left(\mathbb{E}_1 \mathbb{E}_3^{-1} - \mathbb{I} \right) \left[\tilde{M}_3(s; \mathbf{m}) \tilde{P}(\mathbf{m}; s) \right] + \sum_{j=4}^5 \left(\mathbb{E}_j - \mathbb{I} \right) \left[\tilde{M}_j(s; \mathbf{m}) \tilde{P}(\mathbf{m}; s) \right], \quad (2)$$

where \mathbb{E} is the step operator and \mathbb{E}^{-1} is the inverse of \mathbb{E} , and \mathbb{I} is the unit operator.

Interestingly, we find that limit $\lim_{s \rightarrow 0} \tilde{M}_i(s; \mathbf{m})$ always exists, and if the limit function is denoted by $K_i(\mathbf{m})$, then $K_i(\mathbf{m})$ can be explicitly expressed by given waiting-time distributions $W_k(t; \mathbf{m})$ ($1 \leq k \leq 5$), that is,

$$K_i(\mathbf{m}) = \frac{\int_0^{+\infty} W_i(t; \mathbf{m}) \left[\prod_{j \neq i} \int_t^{+\infty} W_j(t'; \mathbf{m}) dt' \right] dt}{\int_0^{+\infty} \left[\prod_{j=1}^5 \int_t^{+\infty} W_j(t'; \mathbf{m}) dt' \right] dt}, \quad 1 \leq i \leq 5. \quad (3)$$

Note that $s \rightarrow 0$ corresponds to $t \rightarrow +\infty$ according to the definition of Laplace transform. However, $t \rightarrow +\infty$ corresponds to the steady-state case, which is our interest. We point out that function $K_i(\mathbf{m})$, which will be called **effective transition rate** for reaction R_i , explicitly decodes the effect of molecular memory, where $1 \leq i \leq 5$. More importantly, using these effective transition rates, we can construct a Markov reaction network with the same topology as the original non-Markov reaction network:



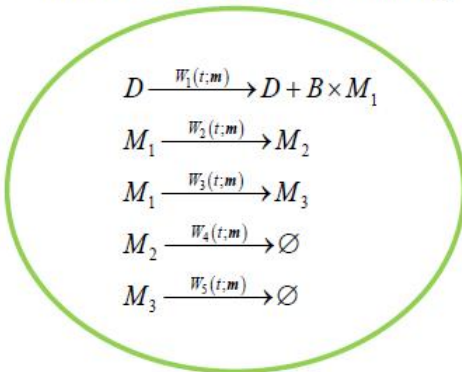
Moreover, two reaction networks have exactly the same chemical master equation at steady state:

$$\left(\sum_{i=0}^{m_1} \alpha_i \mathbb{E}_1^{-i} - \mathbb{I} \right) \left[K_1(s; \mathbf{m}) P(\mathbf{m}) \right] + \left(\mathbb{E}_1 \mathbb{E}_2^{-1} - \mathbb{I} \right) \left[K_2(\mathbf{m}) P(\mathbf{m}) \right] \\ + \left(\mathbb{E}_1 \mathbb{E}_3^{-1} - \mathbb{I} \right) \left[K_3(\mathbf{m}) P(\mathbf{m}) \right] + \sum_{j \in \{4,5\}} \left(\mathbb{E}_j - \mathbb{I} \right) \left[K_j(\mathbf{m}) P(\mathbf{m}) \right] = 0, \quad (5)$$

implying that both stationary behaviors are exactly the same (referring to Figure 2), where $P(\mathbf{m})$ is the stationary probability density function corresponding to dynamic density probability function $P(\mathbf{m}; t)$.

In summary, by introducing an effective transition rate ($K_i(\mathbf{m})$) for each reaction R_i , given by Eq (3), a mathematically difficult non-Markov issue is transformed into a mathematical tractable Markov issue. This brings us convenience for theoretical analysis. In the following, we will focus on analysis of Eq (5).

Non-Markovian reaction system



stationary behavior
is exactly the same

Markovian reaction system

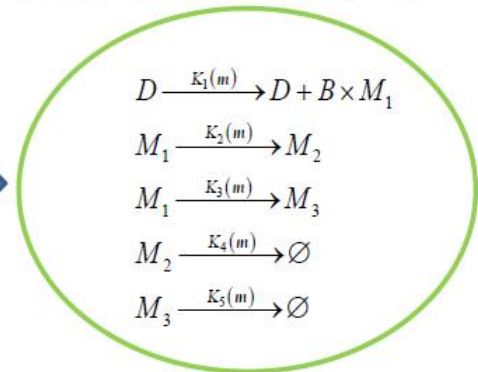


Figure 2. Schematic diagram for two reaction networks with the same topology and the same reactive species, where W -type functions represent reaction-event waiting-time distributions on the left-hand side whereas K -type functions are effective transition rates on the right-hand side (see the main text for details), D represents DNA, and the other symbols are explained in the context. The results in reference [7] imply that the stationary behaviors of two reaction networks are exactly the same although reaction events in the two cases take place in different manners.

Note that Gamma functions can well model multistep processes [34,35]. This is because the convolution of several exponential distributions is an Erlang distribution (a special case of Gamma distribution). Therefore, in order to model the effect of molecular memory on the mRNA expression, we assume that waiting-time distributions for gene activation, NRE and RNR processes are Gamma distributions: $W_1(t;\mathbf{m}) = [\Gamma(L_0)]^{-1} t^{L_0-1} (k_0)^{L_0} e^{-k_0 t}$, $W_2(t;\mathbf{m}) = [\Gamma(L_c)]^{-1} t^{L_c-1} (m_1 k_c)^{L_c} e^{-m_1 k_c t}$ and $W_3(t;\mathbf{m}) = [\Gamma(L_r)]^{-1} t^{L_r-1} (m_1 k_r)^{L_r} e^{-m_1 k_r t}$, and that waiting-time distributions for the other processes are exponential ones, $W_4(t;\mathbf{m}) = m_2 d_c e^{-m_2 d_c t}$, and $W_5(t;\mathbf{m}) = m_3 d_r e^{-m_3 d_r t}$. Here $\Gamma(\cdot)$ is the common Gamma function, k_0 is the mean transcription rate, k_c and k_r denote the mean synthesis rates for mRNAs in the nucleus and mRNAs in the cytoplasm respectively, d_c and d_r are the mean degradation rates of mRNA in the nucleus and mRNA in the cytoplasm respectively. Throughout this paper, L_0 , L_c and L_r are called memory indices since, e.g., $L_0 = 1$ corresponds to the memoryless case whereas $L_0 \neq 1$ to the memory case.

2.3. Parameter settings

Let k_1 be the total synthesis rate of mRNAs in the cytoplasm, which is composed of two parts: one part is the rate at which pre-mRNA generates a transcript through the AS process or the APA process, and the other is the rate at which the mature mRNA is finally exported to the cytoplasm with the help of NPC. The rate of generating mature mRNAs, determined by the gene length, is generally fixed. In contrast, the rate of exporting mRNAs to the cytoplasm may change in a large range, depending on cellular environments. This is because some types of mRNAs are exported in a fast manner due to RNA binding proteins or linked splicing factors, and other types of mRNAs are exported in a slow manner and the corresponding genes are most intron-containing ones [19]. Thus, we can use k_1 to characterize the export rate indirectly. Similarly, if we let k_2 be the synthesis rate of mRNAs in the nucleus, then it also includes two parts: one part is the rate of pre-mRNA produced through an AS or APA process, and the other is the rate at which transcripts are transported to some sub-cellular regions (e.g., nuclear speckles or paraspeckles) under the assistance of some proteins. Here, we assume that k_2 changes a little so that the involved processes are simplified. Owing to AS or APA processes, the lengths of mature mRNAs of the two kinds can be significantly different. Usually, the rate k_1 is faster than the rate k_2 . The retention and export of transcripts are random, we introduce another parameter p_r , called remaining probability throughout this paper, to characterize this randomness. Then, the practical export rate and the practical retention rate should be $k_c = k_1(1 - p_r)$ and $k_r = k_2 p_r$ respectively, where $p_r \in (0, 1)$.

Based on the experimental data from Halpern's group [26] that measured the whole genome-wide catalogue of nuclear and cytoplasmic mRNA from MIN6 cells, we know that most genes (~70%) has more cytoplasmic transcripts than nuclear transcripts. Thus, we can get an approximate formula for remaining probability p_r : $p_r = N_n / (N_n + N_c)$ where N_n is the number of transcripts in the nucleus and N_c the number of transcripts in the cytoplasm. By considering gene INS1 for which the value of ratio N_c / N_n is the maximal (13.2 ± 4.6), we can know that the value of p_r is about 5%. In that paper, the authors also mentioned that about 30% of the genes in MIN6 cells have more transcripts in the nucleus than in the cytoplasm. By considering gene ChREBP for which the value of ratio cytoplasm/nucleus is about 0.05, we can know that the value of p_r is about 95%. Therefore, the range of remaining probability (p_r) in the whole genome is about 5~95%. It is reasonable that the 50% value of p_r is set as a threshold. For convenience, we categorize models of eukaryotic gene expression into two classes: one class where the RNE process is dominant and the other class where the RNR process is dominant. For the former, $p_r < 0.5$ holds, implying that most mRNAs are exported to the cytoplasm through nuclear pores, whereas for the latter, $p_r > 0.5$ holds, implying that most mRNAs retain in the nucleus.

In the following analysis, memory indices L_i ($i = 0, c, r$), and remaining probability p_r will be taken as key parameters while the other parameter values will be kept fixed. Without loss of generality,

we assume that two degradation rates for mRNAs in the nucleus and cytoplasm are equal and denote by d the common degradation rate, i.e., $d_c = d_r = d$.

3. Analytical results

First, if we let x_i represent the concentration of reactive species M_i , i.e., $x_i = \lim_{\substack{m_i \rightarrow \infty \\ \Omega \rightarrow \infty}} m_i / \Omega$ ($i=1,2,3$) where Ω represents the volume of the system, then the rate equations corresponding to the constructed-above Markov reaction network can be expressed as

$$\frac{d\mathbf{x}}{dt} = \mathbf{S}\mathbf{K}(\mathbf{x}), \quad (6)$$

where $\mathbf{x} = (x_1, x_2, x_3)^T$ is a column vector, $\mathbf{S} = (S_{ij})_{3 \times 5} = \begin{pmatrix} \langle B \rangle & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix}$ is a matrix, and

$\mathbf{K}(\mathbf{x}) = (K_1(\mathbf{x}), K_2(\mathbf{x}), K_3(\mathbf{x}), K_4(\mathbf{x}), K_5(\mathbf{x}))^T$ is a column vector of effective transition rates. The steady states or equilibria of the system described by Eq (6), denote by \mathbf{x}^s , are determined by solving the algebraic equation group, i.e., $\mathbf{S}\mathbf{K}(\mathbf{x}^s) = \mathbf{0}$.

Second, if denoting by $\langle X \rangle$ the mean of random variable X and taking approximation $\langle M_i \rangle \approx x_i^s$ ($i=1,2,3$), then we can derive the following matrix equation (see Appendix A):

$$\mathbf{A}_s \boldsymbol{\Sigma}_s + \boldsymbol{\Sigma}_s \mathbf{A}_s^T + \Omega \mathbf{D}_s = \mathbf{0}, \quad (7)$$

where two square matrices $\mathbf{A}_s = (A_{ij})_{3 \times 3}$ and $\mathbf{D}_s = (D_{ij})_{3 \times 3}$ evaluated at steady state are known, and covariance matrix $\boldsymbol{\Sigma}_s = (\sigma_{ij})$ with $\sigma_{ij} = \langle (M_i - \langle M_i \rangle)(M_j - \langle M_j \rangle) \rangle$ is unknown. Note that diagonal elements σ_{22} and σ_{33} represent variances for random variables M_2 (corresponding to mRNA in the cytoplasm) and M_3 (corresponding to mRNA in the nucleus), which are our interest. In addition, we can also derive formulae similar to Eq (3) in the case of continuous variables.

In order to show the explicit effect of molecular memory on the mRNA expression in different biological processes, we consider two special cases: 1) The process of gene activation with memory and other processes are memoryless. 2) The process of nuclear RNA export with memory and other processes are memoryless. In other cases, there are in general no analytical results.

Case 1 $L_0 \neq 1$, $L_c = 1$ and $L_r = 1$. In this case, the five effect transition rates become:

$$K_1(\mathbf{x}) = \frac{(k_c x_1 + k_r x_1 + dx_2 + dx_3)(k_0)^{L_0}}{(k_c x_1 + k_r x_1 + dx_2 + dx_3 + k_0)^{L_0} - (k_0)^{L_0}}, \quad K_2(\mathbf{x}) = k_c x_1, \quad K_3(\mathbf{x}) = k_r x_1, \quad K_4(\mathbf{x}) = dx_2, \quad \text{and}$$

$K_5(\mathbf{x}) = dx_3$. Note that in the case of continuous variables, the corresponding effect transition rates

$K_i(\mathbf{x})$ ($1 \leq i \leq 5$) have the same expressions except for variable notations.

We can show that the means of mRNAs in the nucleus and in the cytoplasm are given respectively by (see Appendix B)

$$\langle M_2 \rangle = \tilde{k}_0 \tilde{k}_c, \quad \langle M_3 \rangle = \tilde{k}_0 \tilde{k}_r, \quad (8)$$

where $\tilde{k}_0 = \frac{1}{2} \left[(1 + 2\langle B \rangle)^{1/L_0} - 1 \right] \frac{k_0}{k_c + k_r} > 0$ with $\langle B \rangle$ being the expectation of burst size B

(random variable), $\tilde{k}_c = \frac{k_c}{d}$ and $\tilde{k}_r = \frac{k_r}{d}$. Apparently, $\langle M_i \rangle$ ($i = 2, 3$) is a monotonically decreasing

function of memory index L_0 , implying that molecular memory always reduces the mRNA expression

level in the nucleus and cytoplasm. In addition, by noting $k_c = k_1(1 - p_r)$ and $k_r = k_2 p_r$, we can know

that if k_2/k_1 is fixed, then $\langle M_3 \rangle$ (i.e., the mean of mRNAs in the nucleus) is a monotonically

increasing function of remaining probability p_r whereas $\langle M_2 \rangle$ (i.e., the mean of mRNAs in the

cytoplasm) is a monotonically decreasing function of p_r . In addition, $\langle M_3 \rangle$ is a monotonically

decreasing function of $\rho = k_c/k_r$ whereas $\langle M_2 \rangle$ is a monotonically increasing function of ρ .

These results are in agreement with intuition.

Interestingly, we find that σ_{22} and σ_{33} , the variances for mRNAs in the nucleus and in the cytoplasm respectively, have the following relationship (see Appendix B)

$$\sigma_{22} = \left(\frac{\tilde{k}_r}{\tilde{k}_c} \right)^2 \sigma_{33} + \tilde{k}_0 (\tilde{k}_c + \tilde{k}_r) \frac{\tilde{k}_c}{\tilde{k}_r}, \quad (9)$$

indicating that the mRNA variance in the cytoplasm, σ_{22} , is larger than that in the nucleus, σ_{33} .

From the viewpoint of experiments, the cytoplasmic mRNAs are easy to measure whereas the nuclear mRNAs are difficult to measure. Therefore, we are interested in the cytoplasmic mRNA expression (including the level and noise). By complex calculations, we can further show that the cytoplasmic mRNA variance is given by

$$\sigma_{22} = \frac{\tilde{k}_0 \tilde{k}_c \tilde{k}_r^3}{\tilde{k}_c^3 + \tilde{k}_r^3} \left\{ 2 + \frac{\tilde{k}_c}{\tilde{k}_r} + \frac{1}{2} \frac{2\tilde{b} + [\gamma - (1 + \tilde{b})^2] (\tilde{k}_c + \tilde{k}_r)}{1 + (1 + \tilde{b})(1 + 2\tilde{b})(\tilde{k}_c + \tilde{k}_r)} \right\}. \quad (10)$$

where $\tilde{b} = \frac{1}{4\langle B \rangle} \left[L_0 + 2(L_0 - 1)\langle B \rangle - L_0(1 + 2\langle B \rangle)^{(L_0-1)/L_0} \right] > 0$ and $\gamma = \frac{\langle B^2 \rangle + \langle B \rangle}{\langle B \rangle}$ with $\langle B^2 \rangle$ being

the second-order raw moment of burst size B . Furthermore, if we define the noise intensity as the ratio of the variance over the squared mean, then the noise intensity for the cytoplasmic mRNA, denoted by η_c , can be analytically expressed as

$$\eta_c = \frac{1}{\tilde{k}_0 \tilde{k}_c} \frac{\tilde{k}_r^3}{\tilde{k}_c^3 + \tilde{k}_r^3} \left\{ 2 + \frac{\tilde{k}_c}{\tilde{k}_r} + \frac{1}{2} \frac{2\tilde{b} + \left[\gamma - (1 + \tilde{b})^2 \right] (\tilde{k}_c + \tilde{k}_r)}{1 + (1 + \tilde{b})(1 + 2\tilde{b})(\tilde{k}_c + \tilde{k}_r)} \right\}. \quad (11)$$

Note that if $L_0 = 1$, which correspond to the Markov case, then $\tilde{k}_0 = \frac{k_0 \langle B \rangle}{k_c + k_r}$ and $\tilde{b} = 0$. Thus, the

cytoplasmic mRNA noise in the Markov case, denoted by $\eta_c|_{L_0=1}$, is given by

$$\eta_c|_{L_0=1} = \frac{1}{\tilde{k}_0 \tilde{k}_c} \frac{\tilde{k}_r^3}{\tilde{k}_c^3 + \tilde{k}_r^3} \left(\frac{\tilde{k}_c + 2\tilde{k}_r}{\tilde{k}_r} + \frac{\gamma}{2} \frac{\tilde{k}_c + \tilde{k}_r}{1 + \tilde{k}_c + \tilde{k}_r} \right). \quad (12)$$

Therefore, the ratio of the noise in non-Markov ($L_0 \neq 1$) case over that in the Markov ($L_0 = 1$) case

$$\frac{\eta_c}{\eta_c|_{L_0=1}} = \left\{ \frac{\tilde{k}_c + 2\tilde{k}_r}{\tilde{k}_r} + \frac{1}{2} \frac{2\tilde{b} + \left[\gamma - (1 + \tilde{b})^2 \right] (\tilde{k}_c + \tilde{k}_r)}{1 + (1 + \tilde{b})(1 + 2\tilde{b})(\tilde{k}_c + \tilde{k}_r)} \right\} \left(\frac{\tilde{k}_c + 2\tilde{k}_r}{\tilde{k}_r} + \frac{\gamma}{2} \frac{\tilde{k}_c + \tilde{k}_r}{1 + \tilde{k}_c + \tilde{k}_r} \right)^{-1}, \quad (13)$$

which may be more than the unit but may also be less than the unit, depending on the size of remaining probability. However, if L_0 is large enough (e.g., $L_0 > 2$), then the ratio in Eq (13) is always larger than the unit, implying that molecular memory amplifies the cytoplasmic mRNA noise.

Case 2 $L_0 = 1$, $L_c \neq 1$ and $L_r = 1$. In this case, five effect transition rates reduce to $K_1(\mathbf{x}) = k_0$,

$$K_2(\mathbf{x}) = \frac{(k_0 + k_r x_1 + dx_2 + dx_3)(k_c x_1)^{L_c}}{(k_0 + k_c x_1 + k_r x_1 + dx_2 + dx_3)^{L_c} - (k_c x_1)^{L_c}}, \quad K_3(\mathbf{x}) = k_c x_1, \quad K_4(\mathbf{x}) = dx_2, \quad \text{and} \quad K_5(\mathbf{x}) = dx_3.$$

It seems to us there are no analytical results as in Case 1. However, if $p_r = 0$ (i.e., if we do not consider

nuclear retention), then we can show that the steady state is given by $x_1 = \frac{k_0}{k_1} \omega$, $x_2 = \frac{k_0 \langle B \rangle}{d}$, $x_3 = 0$,

where $\omega = \frac{(1 + \langle B \rangle) \langle B \rangle^{1/L_c}}{(1 + 2\langle B \rangle)^{1/L_c} - \langle B \rangle^{1/L_c}}$ is a factor depending on both transcriptional burst and molecular

memory. Moreover, the mRNA noise in the cytoplasm is given by (see Appendix C for derivation)

$$\eta_c = \frac{d(1+\langle B \rangle)}{2k_0 \langle B \rangle} \frac{2d\omega(1+\omega+\langle B \rangle) + \gamma L_c k_1 \langle B \rangle k_1 (1+2\langle B \rangle)}{d\omega(1+\omega+\langle B \rangle) + L_c \langle B \rangle (1+2\langle B \rangle) [d\omega + k_1 (1+\langle B \rangle)]}. \quad (15)$$

In order to see the contribution of molecular memory to the cytoplasmic mRNA noise, we calculate the ratio of the noise in non-Markov ($L_c \neq 1$) case over that in the Markov ($L_c = 1$) case

$$\frac{\eta_c}{\eta_c|_{L_c=1}} = \frac{d+k_1}{2d+\gamma k_1} \frac{(1+\langle B \rangle) [2d\omega(1+\omega+\langle B \rangle) + \gamma L_c k_1 \langle B \rangle k_1 (1+2\langle B \rangle)]}{d\omega(1+\omega+\langle B \rangle) + L_c \langle B \rangle (1+2\langle B \rangle) [d\omega + k_1 (1+\langle B \rangle)]}, \quad (16)$$

which is in general larger than the unit for large $L_c > 1$ (corresponding to strong memory), indicating that molecular memory in general enlarges the cytoplasmic mRNA noise.

4. Numerical results

4.1. Impact of memory index L_c on cytoplasmic mRNA expression

Here we numerically investigate the effect of molecular memory (L_c) from nuclear RNA export on the cytoplasm mRNA (M_2) in the cases that the other model parameter values are fixed. Numerical results are demonstrated in Figure 3. From Figure 3(a), we observe that the mean level of the cytoplasm mRNA is a monotonically decreasing function of L_c , independent of the value choice of remaining probability (p_r) (even though we only show the two cases of p_r values). This is in agreement with our intuition since a more reaction step for mRNA synthesis inevitably leads to less mRNA molecules. On the other hand, we observe from Figure 3(b) that molecular memory reduces the cytoplasm mRNA noise (η_c) for smaller values of L_c but enlarges η_c for larger values of L_c , implying that there is an optimal L_c such that η_c arrives at the minimum. We emphasize that the dependences shown in Figure 3(a),(b) are qualitative since they are independent of the value choice of remaining probability.

Importantly, Figure 3 indicates that the results obtained by the linear noise approximation (solid lines) are well in agreement with the results obtained by the Gillespie algorithm [36]. Therefore, the linear noise approximation can be used in fast evaluation of the expression noise, and in the following, we will focus on results obtained by the linear noise approximation. In addition, we point out that most results obtained here and thereafter are qualitative since they are independent of the choice of parameter values. However, to demonstrate interesting phenomena clearly, we will choose special values of some model parameters.

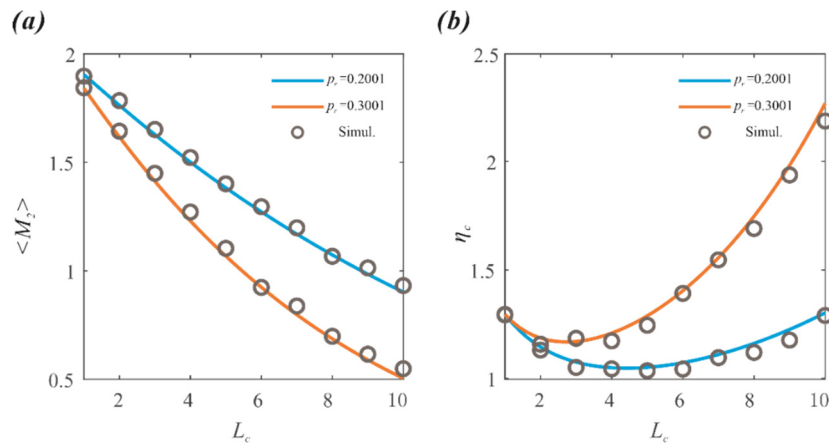


Figure 3. Influence of molecular memory (L_c) from multistep RNA export on the cytoplasmic mRNA (M_2), where solid lines represent theoretical results obtained by linear noise approximation (Appendix A), and empty circles represent numerical results obtained by a Gillespie algorithm [36]. Parameter values are set as $\langle B \rangle = 2$, $k_1 = 4$, $k_2 = 0.8$, $k_0 = 1$, $d_c = d_r = 1$. (a) Impact of molecular memory from the RNA export on the mean cytoplasmic mRNA (M_2) for two values of remaining probability (p_r), where the blue solid line corresponds to $p_r = 0.2001$ and the orange solid line to $p_r = 0.3001$. (b) Effect of molecular memory from the RNA export on the cytoplasmic mRNA noise (η_c) for two values of p_r .

4.2. Effects of molecular memory and remaining probability on cytoplasmic mRNA expression: The case of $L_c \neq 1$ and $L_0 = L_r = 1$

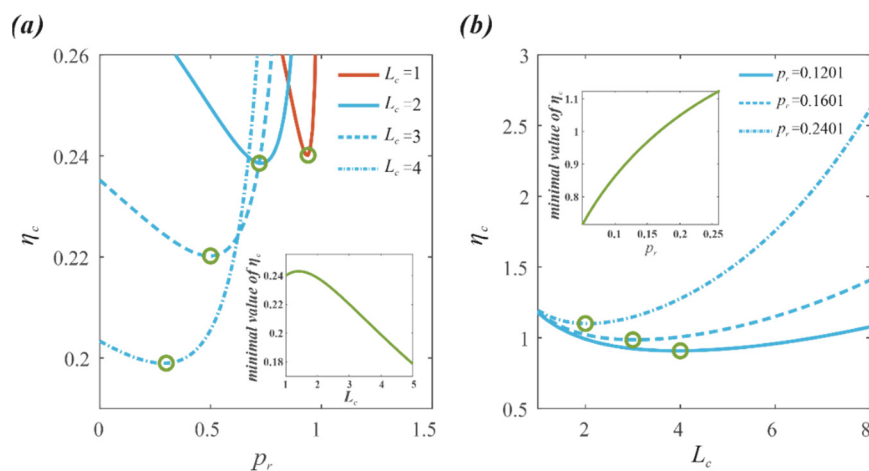


Figure 4. Influence of remaining probability and molecular memory on mature mRNA transported to the cytoplasm (M_2). Solid lines represent theoretical results obtained by our linear noise approximation (Appendix A). Empty circles represent the minimum of the noise of the cytoplasmic mRNA. (a) Parameter values are set as $\langle B \rangle = 40$, $k_1 = 5$, $k_2 = 0.8$, $k_0 = 2.5$, $d_c = d_r = 1$. (b) Parameter values are set as $\langle B \rangle = 2$, $k_1 = 2$, $k_2 = 0.8$, $k_0 = 1$, $d_c = d_r = 1$.

Here we focus on numerically investigating joint effects of molecular memory (L_c) and remaining probability (p_r) on the cytoplasm mRNA (M_2). Figure 4(a) demonstrates effects of p_r on the M_2 noise for three values of L_c . We observe that with the increase of remaining probability, the cytoplasmic mRNA noise (η_c) first decreases and then increases, implying that there is a critical p_r such that η_c arrives at the minimum (referring to empty circles in Figure 4(a)) or that remaining probability can minimize the cytoplasmic mRNA noise. Moreover, this minimum is independent of the values of memory index L_c . In addition, we find that the minimal η_c first increases and then decreases with increasing L_c (the inset of Figure 4(a)). In other words, noise of cRNA can derive a optimal value with the decrease of remaining probability and the increase of memory index. Figure 4(b) shows the dependences of η_c on L_c for three different values of remaining probability. We find that molecular memory can also make cytoplasmic mRNA noise reach the minimum (referring to empty circles in Figure 4(b)), and this minimal noise is monotonically increasing function of p_r .

4.3. Effects of molecular memory and remaining probability on cytoplasmic mRNA expression: The case of $L_0 \neq 1$, $L_c = 1$ and $L_r = 1$

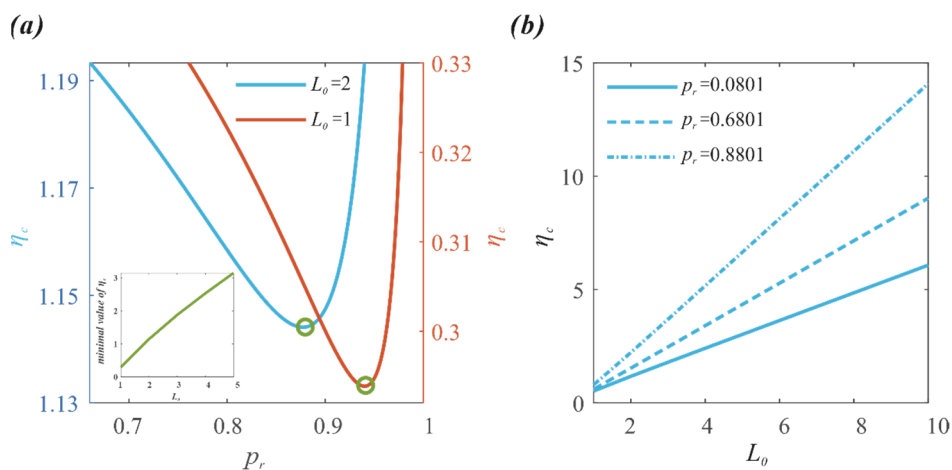


Figure 5. Influence of remaining probability for nuclear RNA retention and molecular memory from multistep gene activation on the cytoplasmic mRNA (M_2), where lines represent the results obtained by linear noise approximation (Appendix A). Empty circles represent the minimum of the noise of the cytoplasm mRNA. (a) The dependence of cytoplasmic mRNA noise η_c on remaining probability p_r for two values of memory index L_0 , where the inset is an enlarged diagram showing the dependence of the minimal η_c on L_0 . Parameter values are set as $\langle B \rangle = 20$, $k_1 = 10$, $k_2 = 1$, $k_0 = 2.5$, $d_c = d_r = 1$. (b) The dependence of cytoplasmic mRNA noise η_c on remaining probability p_r for three values of remaining probability p_r . Parameter values are set as $\langle B \rangle = 2$, $k_1 = 4$, $k_2 = 1$, $k_0 = 2.5$, $d_c = d_r = 1$.

Here we focus on numerically analyzing joint effects of memory index L_0 and remaining probability p_r on the cytoplasm mRNA (M_2). Figure 5(a) demonstrates effects of p_r on the M_2

noise for two representative values of L_0 (note: $L_0 = 1$ corresponds to the memoryless case whereas $L_0 = 2$ corresponds to the memory case). We observe that with the increase of remaining probability, the cytoplasmic mRNA noise (η_c) first decreases and then increases, implying that there is a critical p_r such that η_c reaches the minimum (referring to empty circles in Figure 5(a)) or that remaining probability can minimize the cytoplasmic mRNA noise. Moreover, this minimum (referring to empty circles) is a monotonically increasing function of memory index L_0 .

Figure 5(b) demonstrates that the cytoplasmic mRNA noise (η_c) is always a monotonically increasing function of memory index L_0 , independent of remaining probability. In addition, we observe that η_c is a monotonically increasing function of remaining probability (this can be seen by comparing three lines).

4.4. Effects of molecular memory and remaining probability on cytoplasmic mRNA expression: The case of $L_r \neq 1$, $L_c = 1$ and $L_0 = 1$

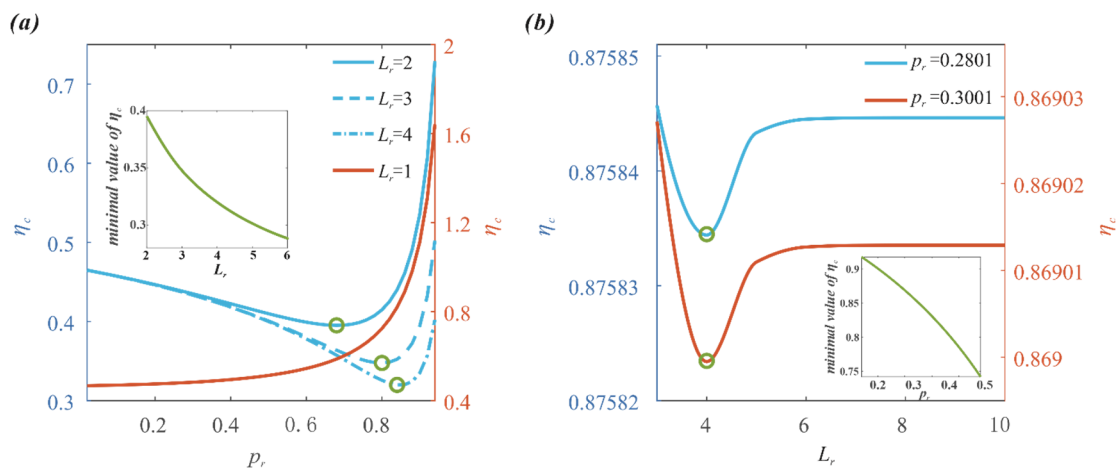


Figure 6. Influence of remaining probability and molecular memory on mature mRNA transported to the cytoplasm (M_2), solid lines represent theoretical results obtained by linear noise approximation (Appendix A). Empty circles represent the minimum of the noise of the cytoplasm mRNA. (a) Parameter values are set as $\langle B \rangle = 2$, $k_1 = 2$, $k_2 = 0.8$, $k_0 = 2.5$, $d_c = d_r = 1$. (b) Parameter values are set as $\langle B \rangle = 2$, $k_1 = 2$, $k_2 = 0.8$, $k_0 = 1$, $d_c = d_r = 1$.

Here we consider the case that RNR is a multistep process, i.e., $L_r \neq 1$. Numerical results are demonstrated in Figure 6. We observe from Figure 6(a) that except for the case of $L_r = 1$, which corresponds to the Markov process and for which the cytoplasmic mRNA noise (η_c) is a monotonically increasing function of remaining probability (p_r), the dependences of η_c on p_r are not monotonic in the cases of $L_r \neq 1$ (corresponding to non-Markov processes) but there is a threshold of p_r such that η_c reaches the minimum (referring to empty circles), similarly to the case of Figure 5(a). Moreover, this minimal noise is a monotonically decreasing function of memory index L_r (referring to the inset of Figure 6(a)) but the monotonicity is opposite to that in the case of Figure 5(a).

Figure 6(b) shows how the cytoplasmic mRNA noise (η_c) depends on memory index L_r for two

different values of remaining probability. Interestingly, we observe that there is an optimal value of L_r such that the cytoplasmic mRNA noise reaches the minimum. Moreover, the minimal η_c is a monotonically decreasing function of remaining probability (p_r), referring to the inset in the bottom right-hand corner.

5. Conclusions and discussion

Gene transcription in eukaryotes involve many molecular processes, some of which are well known and others are little known and even unknown [37,38]. In this paper, we have introduced a non-Markov model of stochastic transcription, which simultaneously considers RNA nuclear retention and nuclear RNA export processes and in which we have used non-exponential waiting-time distributions (e.g., Gamma distributions) to model some unknown or unspecified molecular processes involved in, e.g., the synthesis of pre-mRNA and the export of mRNAs generated in the nucleus to the cytoplasm and the retention of mRNA in the nucleus. Since non-exponential waiting times can lead to non-Markov kinetics, we have introduced effective transition rates for the reactions underlying transcription to transform a mathematically difficult issue to a mathematically tractable one. As a result, we have derived the analytical expressions of mRNA means and noise in the nucleus and cytoplasm, which revealed the importance of molecular memory in controlling or fine-tuning the expressions of two kinds of mRNA. Our modeling and analysis provided a heuristic framework for studying more complex gene transcription processes.

Our model considered main events occurring in gene transcription such as bursty expression (burst size follows a general distribution), alternative splicing (by which two kinds of transcripts are generated), RNR (a part of RNA molecules that are kept in the nucleus) and RNE (another part of RNA molecules that are exported to the cytoplasm). Some popular experimental technologies such as single-cell sequence data [39], single-molecule fluorescence in-situ hybridization (FISH) [40] and electron micrographs (EM) of fixed cells [41] have indicated that RNR and NRE are two complex biochemical processes, each involving regulation by a large number of proteins or complexes [42]. In particular, the mRNAs exported to the cytoplasm involve the structure of nuclear pore complex (NPC) [43]. A number of challenging questions still remain unsolved, e.g., how do RNR and NRE cooperatively regulate the expressions of nuclear and cytoplasmic mRNAs? Why are these two dynamical processes necessary for the whole gene-expression processes when the cells survive in complex environments? And what advantages do they have in contrast to a single NRE process?

Despite simple, our model can not only reproduce results for pre-mRNA (nascent mRNA) means at steady state in previous studies but also give results in agreement with experimental data on the mRNA Fano factors (define as the ratio of variance over mean) of some genes. However, we point out that some results of Fano factors obtained using our model is not always in agreement with the experimental data, e.g., for five genes, RBP3, TAF5, TAF6, TAF12 and KAP104, results obtained by our model seem not in agreement with experimental data but results obtained by a previous theoretical model [44] seems better (data are not shown). In addition, for the PRB8 gene, results on Fano factor, obtained by our model and the previous model, are poorly in agreement with experimental data (data are not shown). This indicates that constructing a theoretical model for the whole transcription process still needs more work.

In spite of differences, our results are wholly in agreement with some experimental data or observations. First, the qualitative result that RNR always reduces the nuclear pre-mRNA noise

and always amplifies the cytoplasmic mRNA noise is in agreement with some experimental observations [28,42,45] and also with intuition since the retention naturally increases the mean number of the nuclear pre-mRNAs but decreases the mean number of the cytoplasmic mRNAs. Second, we compare our theoretical predictions with experimental results [28,45]. Specifically, we use previously published experimental data for two yeast genes, RBP2 and MDN1 [28,45] to calculate the cytoplasmic mRNA Fano factors. Parameter k_1 is set as $k_1 \approx 0.29 \pm 0.013 / \text{min}$, which is based on experimental data [28] and the degradation rates of the cytoplasmic mRNAs for RBP2 and MDN1 are set according to $d_c = \ln 2 / t_{1/2}$, where $t_{1/2}$ is an experimental mRNA half-life. Then, we can find that the results on the Fano factors of genes RBP2 and MDN1 are well in agreement with the experimental data [45].

At the whole genome scale, about 70% mRNAs in the nucleus are transported to the cytoplasm whereas about 30% mRNAs are retained in the nucleus [26]. This fact implies that the changing range of remaining probability is moderate or small. In addition, the nuclear export rate of a different gene is in general different. If this rate is not too large, then following the increase of remaining probability, the increase in the cytoplasmic mRNA noise is inevitable. This result indirectly interprets the reason why the noise at the protein level is quite large as shown in previous studies of gene expression [46].

Finally, for some genes, the relative changing ranges of remaining probability and nuclear export rate may be large at the transcription level. In this case, it is in theory sufficient that adjusting one of nuclear export rate and remaining probability can fine-tune the cytoplasmic mRNA noise if the mean burst size is fixed, but differences would exist between theoretical and experimental results since NRE and RNR occur simultaneously in gene expression and are functionally cooperative. In addition, since biological regulation may be different from the theoretical assumption made here, the nuclear or cytoplasmic mRNA noise predicted in theory may be overestimated.

Acknowledgments

This work was partially supported by National Nature Science Foundation of China (11931019), and Key-Area Research and Development Program of Guangzhou (202007030004).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. J. Peccoud, B. Ycart, Markovian modelling of gene product synthesis, *Theor. Popul. Biol.*, **48** (1995), 222–234. <https://doi.org/10.1006/tpbi.1995.1027>
2. V. Shahrezaei, P. S. Swain, Analytical distributions for stochastic gene expression, *Proc. Natl. Acad. Sci. U.S.A.*, **105** (2008), 17256–17261. <https://doi.org/10.1073/pnas.0803850105>
3. J. Paulsson, Models of stochastic gene expression, *Phys. Life Rev.*, **2** (2005), 157–175. <https://doi.org/10.1016/j.plrev.2005.03.003>

4. T. B. Kepler, T. C. Elston, Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations, *Biophys. J.*, **81** (2001), 3116–3036. [https://doi.org/10.1016/S0006-3495\(01\)75949-8](https://doi.org/10.1016/S0006-3495(01)75949-8)
5. N. Friedman, L. Cai, X. S. Xie, Linking stochastic dynamics to population-distribution: An analytical framework of gene expression, *Phys. Rev. Lett.*, **97** (2006), 168302. <https://doi.org/10.1103/PhysRevLett.97.168302>
6. J. J. Zhang, T. S. Zhou, Promoter architecture-mediated transcriptional dynamics, *Biophys. J.*, **106** (2014), 479–488. <https://doi.org/10.1016/j.bpj.2013.12.011>
7. J. J. Zhang, T. S. Zhou, Markovian approaches to modeling intracellular reaction processes with molecular memory, *Proc. Natl. Acad. Sci. U.S.A.*, **116** (2019), 23542–23550. <https://doi.org/10.1073/pnas.1913926116>
8. T. Jia, R. V. Kulkarni, Intrinsic noise in stochastic models of gene expression with molecular memory and bursting, *Phys. Rev. Lett.*, **106** (2011), 058102. <https://doi.org/10.1103/PhysRevLett.106.058102>
9. S. Choubey, Nascent RNA kinetics: Transient and steady state behavior of models of transcription, *Phys. Rev. E.*, **97** (2018), 022402. <https://doi.org/10.1103/PhysRevE.97.022402>
10. B. Zoller, S. C. Little, T. Gregor, Diverse spatial expression patterns emerge from unified kinetics of transcriptional bursting, *Cell*, **175** (2018), 835–847. <https://doi.org/10.1016/j.cell.2018.09.056>
11. M. Z. Ali1, S. Choubey, Decoding the grammar of transcriptional regulation from RNA polymerase measurements: models and their applications, *Phys. Biol.*, **16** (2019), 061001. <https://doi.org/10.1088/1478-3975/ab45bf>
12. T. B. Kallehauge, M. C. Robert, E. Bertrand, T. H. Jensen, Nuclear retention prevents premature cytoplasmic appearance of mRNA, *Mol. Cell.*, **48** (2012), 145–152. <https://doi.org/10.1016/j.molcel.2012.07.022>
13. C. Marchive, F. Roudier, L. Castaings, V. Brehaut, E. Blondet, V. Colot, et al., Nuclear retention of the transcription factor NLP7 orchestrates the early response to nitrate in plants, *Nat. Commun.*, **4** (2013), 1713. <https://doi.org/10.1038/ncomms2650>
14. Q. L. Zheng, J. Hou, Y. Zhou, Z. Y. Li, X. T. Cao, The RNA helicase DDX46 inhibits innate immunity by entrapping m6A-demethylated antiviral transcripts in the nucleus, *Nat. Immunol.*, **18** (2017), 1094–1103. <https://doi.org/10.1038/ni.3830>
15. A. Singh, P. Bokes, Consequences of mRNA transport on stochastic variability in protein levels, *Biophys. J.*, **103** (2012), 1087–1096. <https://doi.org/10.1016/j.bpj.2012.07.015>
16. A. Mor, S. Suliman, R. Ben-Yishay, S. Yunger, Y. Brody, Y. Shav-Tal, Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells, *Nat. Cell. Biol.*, **12** (2010), 543–552. <https://doi.org/10.1038/ncb2056>
17. M. Oeffinger, D. Zenklusen, To the pore and through the pore: a story of mRNA export kinetics, *Biochim. Biophys. Acta.*, **1819** (2012), 494–506. <https://doi.org/10.1016/j.bbagr.2012.02.011>
18. V. O. Wickramasinghe, R. A. Laskey, Control of mammalian gene expression by selective mRNA export, *Nat. Rev. Mol. Cell. Biol.*, **16** (2015), 431–442. <https://doi.org/10.1038/nrm4010>
19. T. Chen, B. van Steensel, Comprehensive analysis of nucleocytoplasmic dynamics of mRNA in *Drosophila* cells, *PLoS Genet.*, **13** (2017), e1006929. <https://doi.org/10.1371/journal.pgen.1006929>
20. N. Battich, T. Stoeger, L. Pelkmans, Control of transcript variability in single mammalian cells, *Cell*, **163** (2015), 1596–1610. <https://doi.org/10.1016/j.cell.2015.11.018>

21. G. L. Hager, C. Elbi, T. A. Johnson, T. Voss, A. K. Nagaich, R. L. Schiltz et al., Chromatin dynamics and the evolution of alternate promoter states. *Chromosome Res.*, **14** (2006), 107–116. <https://doi.org/10.1007/s10577-006-1030-0>
22. R. Metivier, G. Reid, F. Gannon, Transcription in four dimensions: nuclear receptor-directed Transcription in four dimensions, *EMBO Rep.*, **7** (2006), 161–167. <https://doi.org/10.1038/sj.embor.7400626>
23. C. V. Harper, B. Finkenstädt, M. R. White, Dynamic analysis of stochastic transcription cycles, *PLoS Biol.*, **9** (2011), e1000607. <https://doi.org/10.1371/journal.pbio.1000607>
24. Q. L. Wang, T. S. Zhou, Dynamical analysis of mCAT2 gene models with CTN-RNA nuclear retention, *Phys. Biol.*, **12** (2015), 016010. <https://doi.org/10.1088/1478-3975/12/1/016010>
25. H. H. Qiu, B. G. Zhang, T. S. Zhou, Analytical results for a generalized model of bursty gene expression with molecular memory, *Phys. Rev. E*, **100** (2019), 012128. <https://doi.org/10.1103/PhysRevE.100.012128>
26. K. B. Halpern, I. Caspi, D. Lemze, M. Levy, S. Landen, E. Elinav, et al., Nuclear retention of mRNA in mammalian tissues, *Cell. Rep.*, **13** (2015), 2653–2662. <https://doi.org/10.1016/j.celrep.2015.11.036>
27. F. Alber, S. Dokudovskaya, L. M. Veenhoff, W. H. Zhang, J. Kipper, D. Devos, et al., The molecular architecture of the nuclear pore complex, *Nature*, **450** (2007), 695–701. <https://doi.org/10.1038/nature06405>
28. S. Choubey, J. Kondev, A. Sanchez, Deciphering transcriptional dynamics in vivo by counting nascent RNA molecules, *PLoS Comput. Biol.*, **11** (2015), e1004345. <https://doi.org/10.1371/journal.pcbi.1004345>
29. H. Xu, S. O. Skinner, A. M. Sokac, I. Golding, Stochastic kinetics of Nascent RNA, *Phys. Rev. Lett.*, **117** (2016), 128101. <https://doi.org/10.1103/PhysRevLett.117.128101>
30. D. L. Spector, Nuclear domains, *J. Cell. Sci.*, **114** (2001), 2891–2893. <https://doi.org/10.1242/jcs.114.16.2891>
31. A. H. Fox, Y. W. Lam, A. K. L. Leung, C. E. Lyon, J. Andersen, M. Mann, et al., Paraspeckles: A novel nuclear domain, *Curr. Biol.*, **12** (2002), 13–25. [https://doi.org/10.1016/S0960-9822\(01\)00632-7](https://doi.org/10.1016/S0960-9822(01)00632-7)
32. A. I. Lamond, D. L. Spector, Nuclear speckles: A model for nuclear organelles, *Nat. Rev. Mol. Cell. Biol.*, **4** (2003), 605–612. <https://doi.org/10.1038/nrm1172>
33. T. Aquino, M. Dentz, Chemical continuous time random walks, *Phys. Rev. Lett.*, **119** (2017), 230601. <https://doi.org/10.1103/PhysRevLett.119.230601>
34. B. C. Arnold, Majorization: Here, there and everywhere, *Statist. Sci.*, **22** (2007), 407–413. <https://doi.org/10.1214/0883423060000000097>
35. A. David, S. Larry, The least variable phase type distribution is Erlang, *Stoch. Models*, **3** (1987), 467–473. <https://doi.org/10.1080/15326348708807067>
36. D. T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, *J. Comput. Phys.*, **22** (1976), 403–434. [https://doi.org/10.1016/0021-9991\(76\)90041-3](https://doi.org/10.1016/0021-9991(76)90041-3)
37. I. Brouwer, T. L. Lenstra, Visualizing transcription: Key to understanding gene expression dynamics, *Curr. Opin. Chem. Biol.*, **51** (2019), 122–129. <https://doi.org/10.1016/j.cbpa.2019.05.031>

38. E. M. Wissink, A. Vihervaara, N. D. Tippens, J. T. Lis, Nascent RNA analyses: Tracking transcription and its regulation, *Nat. Rev. Genet.*, **20** (2019), 705–723. <https://doi.org/10.1038/s41576-019-0159-6>
39. D. Grün, L. Kester, A. van Oudenaarden, Validation of noise models for single-cell transcriptomics, *Nat. Methods.*, **11** (2014), 637. <https://doi.org/10.1038/nmeth.2930>
40. D. R. Larson, R. H. Singer, D. Zenklusen, A single molecule view of gene expression, *Cell. Biol.*, **19** (2009), 630–637. <https://doi.org/10.1016/j.tcb.2009.08.008>
41. A. E. Hage, S. L. French, A. L. Beyer, D. Tollervey, Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis, *Genes. Dev.*, **24** (2010), 1546–1558. <https://doi.org/10.1101/gad.573310>
42. D. Grunwald, R. H. Singer, M. Rout, Nuclear export dynamics of RNA-protein complexes, *Nature*, **475** (2011), 333–341. <https://doi.org/10.1038/nature10318>
43. R. Benyishay, A. J. Ashkenazy, Y. Shav-Tal, Dynamic encounters of genes and transcripts with the nuclear pore, *Trends. Genet.*, **32** (2016), 419–431. <https://doi.org/10.1016/j.tig.2016.04.003>
44. Q. L. Wang, T. S. Zhou, Alternative-splicing-mediated gene expression, *Phys. Rev. E.*, **89** (2014), 012713. <https://doi.org/10.1103/PhysRevE.89.012713>
45. S. J. Gandhi, D. Zenklusen, T. Lionnet, R. H. Singer, Transcription of functionally related constitutive genes is not coordinated, *Nat. Struct. Mol. Biol.*, **18** (2011), 27–34. <https://doi.org/10.1038/nsmb.1934>
46. T. Stoeger, N. Battich, L. Pelkmans, Passive noise filtering by cellular compartmentalization, *Cell*, **164** (2016), 1151–1161. <https://doi.org/10.1016/j.cell.2016.02.005>
47. N. G. van Kampen, *Stochastic processes in physics and chemistry*, North-Holland, Amsterdam, 2007. <https://doi.org/10.1016/B978-044452965-7/50006-4>
48. J. Elf, M. Ehrenberg, Fast evaluation of fluctuations in biochemical networks with the linear noise approximation, *Gen. Res.*, **13** (2003), 2475–2484. <https://doi.org/10.1101/gr.1196503>

Appendix

A. Linear noise approximation for the gene model in the main text

First, the chemical master equation for the constructed Markov reaction network reads

$$\begin{aligned} \frac{\partial P(\mathbf{m};t)}{\partial t} = & \left(\sum_{i=0}^m \alpha_i \mathbb{E}_1^{-i} - \mathbb{I} \right) [K_1(\mathbf{m})P(\mathbf{m};t)] + (\mathbb{E}_1 \mathbb{E}_2^{-1} - \mathbb{I}) [K_2(\mathbf{m})P(\mathbf{m};t)] \\ & + (\mathbb{E}_1 \mathbb{E}_3^{-1} - \mathbb{I}) [K_3(\mathbf{m})P(\mathbf{m};t)] + \sum_{j \in \{4,5\}} \{ (\mathbb{E}_j - \mathbb{I}) [K_j(\mathbf{m})P(\mathbf{m};t)] \}, \end{aligned} \quad (\text{A1})$$

Second, the steady state or equilibrium of the system described by Eq (6) in the main text, denoted by $\mathbf{x}^S = (x_1^S, x_2^S, x_3^S)^\top$, can be obtained by solving the algebraic equation group

$$\begin{aligned} \langle B \rangle K_1(\mathbf{x}^S) - K_2(\mathbf{x}^S) - K_3(\mathbf{x}^S) &= 0 \\ K_2(\mathbf{x}^S) - K_4(\mathbf{x}^S) &= 0 \\ K_3(\mathbf{x}^S) - K_5(\mathbf{x}^S) &= 0 \end{aligned} \quad (\text{A2})$$

Then, we perform the Ω -expansions [47] to derive a Lyapunov matrix equation for covariance matrix between M_i and M_j with $i, j = 1, 2, 3$, i.e., for matrix $\Sigma = \left(\left\langle (M_i - \langle M_i \rangle)^T (M_j - \langle M_j \rangle) \right\rangle \right) \equiv (\sigma_{ij})$. Note that

$$K_i(\mathbf{n}) = K_i(\mathbf{x} + \Omega^{-1/2} \mathbf{z}) = K_i(\mathbf{x}) + \Omega^{-1/2} \sum_{j=p,c,r} z_j \frac{\partial K_i(\mathbf{x})}{\partial x_j} + o(\Omega^{-1}), \quad i = 1, 2, 3 \quad (\text{A3a})$$

$$\sum_{i=0}^{m_1} \alpha_i \mathbb{E}_1^{-i} - \mathbb{I} = -\Omega^{-1/2} \sum_{i=0}^{m_1} i \alpha_i \frac{\partial}{\partial z_1} + \frac{1}{2} \Omega^{-1} \sum_{i=0}^{m_1} i^2 \alpha_i \frac{\partial^2}{\partial z_1^2} + o(\Omega^{-3/2}), \quad (\text{A3b})$$

$$\mathbb{E}_1 \mathbb{E}_2^{-1} - \mathbb{I} = -\Omega^{-1/2} \left(\frac{\partial}{\partial z_2} - \frac{\partial}{\partial z_1} \right) + \frac{1}{2} \Omega^{-1} \left(\frac{\partial^2}{\partial z_1^2} + \frac{\partial^2}{\partial z_2^2} - 2 \frac{\partial^2}{\partial z_1 \partial z_2} \right) + o(\Omega^{-3/2}), \quad (\text{A3c})$$

$$\mathbb{E}_1 \mathbb{E}_3^{-1} - \mathbb{I} = -\Omega^{-1/2} \left(\frac{\partial}{\partial z_3} - \frac{\partial}{\partial z_1} \right) + \frac{1}{2} \Omega^{-1} \left(\frac{\partial^2}{\partial z_1^2} + \frac{\partial^2}{\partial z_3^2} - 2 \frac{\partial^2}{\partial z_1 \partial z_3} \right) + o(\Omega^{-3/2}), \quad (\text{A3d})$$

$$\mathbb{E}_j - \mathbb{I} = -\Omega^{-1/2} \frac{\partial}{\partial z_j} + \frac{1}{2} \Omega^{-1} \frac{\partial^2}{\partial z_j^2} + o(\Omega^{-3/2}), \quad j = 2, 3. \quad (\text{A3e})$$

Hereafter $o(y)$ represents the infinitesimal quantity of the same order as $y \rightarrow 0$. We denote by $\Pi(\mathbf{z}; t)$ the probability density function for new random variable \mathbf{z} . Then, the relationship between variables $P(\mathbf{m}; t)$ and $\Pi(\mathbf{z}; t)$ is

$$\frac{\partial P(\mathbf{m}; t)}{\partial t} = \frac{\partial \Pi(\mathbf{z}; t)}{\partial t} - \Omega^{1/2} \sum_{i=1,2,3} \frac{dx_i}{dt} \frac{\partial \Pi(\mathbf{z}; t)}{\partial z_i}. \quad (\text{A4})$$

By substituting Eqs (A3) and (A4) into Eq (A1) and comparing the coefficients of $\Omega^{1/2}$, we have

$$\begin{aligned} \sum_{i=1,2,3} \frac{dx_i}{dt} \frac{\partial \Pi(\mathbf{z}; t)}{\partial z_i} &= \langle B \rangle K_1(\mathbf{x}) \frac{\partial \Pi(\mathbf{z}; t)}{\partial z_1} + K_2(\mathbf{x}) \left(\frac{\partial \Pi(\mathbf{z}; t)}{\partial z_2} - \frac{\partial \Pi(\mathbf{z}; t)}{\partial z_1} \right) \\ &+ K_3(\mathbf{x}) \left(\frac{\partial \Pi(\mathbf{z}; t)}{\partial z_3} - \frac{\partial \Pi(\mathbf{z}; t)}{\partial z_1} \right) + \sum_{j \in \{2,3\}} K_j \frac{\partial \Pi(\mathbf{z}; t)}{\partial z_j}, \end{aligned} \quad (\text{A5})$$

which naturally holds due to Eq (6) in the main text, where $\langle B \rangle = \sum_{i=0}^{m_1} i \alpha_i$ is the mean burst size.

Comparing the coefficients of Ω^0 , we have

$$\begin{aligned}
\frac{\partial \Pi(\mathbf{z}; t)}{\partial t} = & -\langle B \rangle \sum_{j=1,2,3} \frac{\partial K_1(\mathbf{x})}{\partial x_j} \frac{\partial [z_j \Pi(\mathbf{z}; t)]}{\partial z_1} - \sum_{j=1,2,3} \frac{\partial K_2(\mathbf{x})}{\partial x_j} \left(\frac{\partial [z_j \Pi(\mathbf{z}; t)]}{\partial z_2} - \frac{\partial [z_j \Pi(\mathbf{z}; t)]}{\partial z_1} \right) \\
& - \sum_{j=1,2,3} \frac{\partial K_3(\mathbf{x})}{\partial x_j} \left(\frac{\partial [z_j \Pi(\mathbf{z}; t)]}{\partial z_3} - \frac{\partial [z_j \Pi(\mathbf{z}; t)]}{\partial z_1} \right) - \sum_{j \in \{2,3\}} \sum_{k=1,2,3} \frac{\partial K_j(\mathbf{x})}{\partial x_k} \frac{\partial [z_k \Pi(\mathbf{z}; t)]}{\partial z_j} \\
& + \frac{1}{2} \langle B^2 \rangle K_1(\mathbf{x}) \frac{\partial^2 \Pi(\mathbf{z}; t)}{\partial z_1^2} + \frac{1}{2} K_2(\mathbf{x}) \left(\frac{\partial^2 \Pi(\mathbf{z}; t)}{\partial z_1^2} + \frac{\partial^2 \Pi(\mathbf{z}; t)}{\partial z_2^2} - 2 \frac{\partial^2 \Pi(\mathbf{z}; t)}{\partial z_1 \partial z_2} \right) \\
& + \frac{1}{2} K_3(\mathbf{x}) \left(\frac{\partial^2 \Pi(\mathbf{z}; t)}{\partial z_1^2} + \frac{\partial^2 \Pi(\mathbf{z}; t)}{\partial z_3^2} - 2 \frac{\partial^2 \Pi(\mathbf{z}; t)}{\partial z_1 \partial z_3} \right) + \frac{1}{2} \sum_{j \in \{2,3\}} K_j \frac{\partial^2 \Pi(\mathbf{z}; t)}{\partial z_j^2}
\end{aligned} \tag{A6}$$

where $\langle B^2 \rangle = \sum_{i=0}^{m_1} i^2 \alpha_i$ is the second moment of burst size. Since $K_i(\mathbf{x})$ is independent of \mathbf{z} , Eq (A6)

can be rewritten as

$$\frac{\partial \Pi(\mathbf{z}; t)}{\partial t} = - \sum_{i,j \in \{1,2,3\}} A_{ij} \frac{\partial [z_j \Pi(\mathbf{z}; t)]}{\partial z_i} + \frac{1}{2} \sum_{i,j \in \{1,2,3\}} D_{ij} \frac{\partial^2 \Pi(\mathbf{z}; t)}{\partial z_i \partial z_j}, \tag{A7}$$

where the elements of matrix $\mathbf{A} = (A_{ij})$ take the form

$$\mathbf{A} = (A_{ij}) = \begin{pmatrix} \langle B \rangle \frac{\partial K_1}{\partial x_1} - \frac{\partial K_2}{\partial x_1} - \frac{\partial K_3}{\partial x_1} & \langle B \rangle \frac{\partial K_1}{\partial x_2} - \frac{\partial K_2}{\partial x_2} - \frac{\partial K_3}{\partial x_2} & \langle B \rangle \frac{\partial K_1}{\partial x_3} - \frac{\partial K_2}{\partial x_3} - \frac{\partial K_3}{\partial x_3} \\ \frac{\partial K_2}{\partial x_1} - \frac{\partial K_4}{\partial x_1} & \frac{\partial K_2}{\partial x_2} - \frac{\partial K_4}{\partial x_2} & \frac{\partial K_2}{\partial x_3} - \frac{\partial K_4}{\partial x_3} \\ \frac{\partial K_3}{\partial x_1} - \frac{\partial K_5}{\partial x_1} & \frac{\partial K_3}{\partial x_2} - \frac{\partial K_5}{\partial x_2} & \frac{\partial K_3}{\partial x_3} - \frac{\partial K_5}{\partial x_3} \end{pmatrix} \tag{A8a}$$

and matrix $\mathbf{D} = (D_{ij})$ takes the form

$$\mathbf{D} = \begin{pmatrix} \langle B^2 \rangle K_1 + K_2 + K_3 & -K_2 & -K_3 \\ -K_2 & K_2 + K_4 & 0 \\ -K_3 & 0 & K_3 + K_5 \end{pmatrix}. \tag{A8b}$$

If we consider the stationary equation of Eq (A7), denote by \mathbf{A}_s and \mathbf{D}_s the corresponding matrices.

Third, the steady-state Fokker Planck equation allows a solution of the following form

$$\Pi(\mathbf{z}) = \frac{1}{\sqrt{(2\pi)^3 \det(\Sigma_s)}} \exp\left(-\frac{1}{2} \mathbf{z}^T \Sigma_s^{-1} \mathbf{z}\right). \tag{A9}$$

Here, matrix $\Sigma_s = \left(\left\langle (\mathbf{M} - \mathbf{x}^s)^T (\mathbf{M} - \mathbf{x}^s) \right\rangle \right) = (\sigma_{ij})$ (covariance matrix) is determined by solving the following Lyapunov matrix equation

$$\mathbf{A}_s \Sigma_s + \Sigma_s \mathbf{A}_s^T + \mathbf{D}_s = \mathbf{0}. \tag{A10}$$

Note that the diagonal elements of matrix Σ_S are just the variances of the state variables, and the vector of the mean concentrations of the reactive species is given approximately by $\langle \mathbf{M} \rangle \approx \mathbf{x}^S$. Eq (A10) is an extension of the linear noise approximation in the Markov case [48].

B. Analytical results for statistical quantities in the case of $L_0 \neq 1$ and $L_c = L_r = 1$

In this case, we can show that effective transition rates are given by

$$K_1(\mathbf{x}) = \frac{(x_1 k_c + x_1 k_r + x_2 d + x_3 d)(k_0)^{L_0}}{(x_1 k_c + x_1 k_r + x_2 d + x_3 d + k_0)^{L_0} - (k_0)^{L_0}}, \quad K_2(\mathbf{x}) = x_1 k_c, \quad K_3(\mathbf{x}) = x_1 k_r, \quad K_4(\mathbf{x}) = x_2 d, \quad \text{and}$$

$K_5(\mathbf{x}) = x_3 d$, where $\mathbf{x} = (x_1, x_2, x_3)^T$. Thus, according to Eq (A2), we know that the steady state is given by

$$x_1^S = \frac{ak_0}{k_c + k_r}, \quad x_2^S = \frac{k_c}{d} x_1^S, \quad x_3^S = \frac{k_r}{d} x_1^S, \quad (\text{B1})$$

where $a = \frac{(1 + 2\langle B \rangle)^{1/L_0} - 1}{2}$. Note that

$$\frac{\partial K_1(\mathbf{x})}{\partial x_1} = \frac{(k_c + k_r)(k_0)^{L_0}}{[k_0(1 + 2a)]^{L_0} - (k_0)^{L_0}} - \frac{2ak_0 L_0 (k_c + k_r)(k_0)^{L_0} [k_0(1 + 2a)]^{L_0 - 1}}{\{[k_0(1 + 2a)]^{L_0} - (k_0)^{L_0}\}^2}$$

Therefore,

$$\left. \frac{\partial K_1(\mathbf{x})}{\partial x_1} \right|_{\mathbf{x}=\mathbf{x}^S} = \frac{\partial K_1(\mathbf{x}^S)}{\partial x_1} = (k_c + k_r) \frac{2\langle B \rangle - L_0 \left[(1 + 2\langle B \rangle) - (1 + 2\langle B \rangle)^{(L_0 - 1)/L_0} \right]}{4\langle B \rangle^2}. \quad (\text{B2a})$$

Completely similarly, we have

$$\frac{\partial K_1(\mathbf{x})}{\partial x_2} = \frac{\partial K_1(\mathbf{x})}{\partial x_3} = d \frac{2\langle B \rangle - L_0 \left[(1 + 2\langle B \rangle) - (1 + 2\langle B \rangle)^{(L_0 - 1)/L_0} \right]}{4\langle B \rangle^2}. \quad (\text{B2b})$$

Thus, matrix \mathbf{A}_S reduces to

$$\mathbf{A} = (A_{ij}) = \begin{pmatrix} (b\langle B \rangle - 1)(k_c + k_r) & bd\langle B \rangle & bd\langle B \rangle \\ k_c & -d & 0 \\ k_r & 0 & -d \end{pmatrix}, \quad (\text{B3})$$

where $b = \frac{-1}{4\langle B \rangle^2} \left[L_0 + 2(L_0 - 1)\langle B \rangle - L_0(1 + 2\langle B \rangle)^{(L_0 - 1)/L_0} \right]$. Meanwhile, the matrix \mathbf{D}_S in Eq (A10)

becomes

$$\mathbf{D}_S = \tilde{k}_0 \begin{pmatrix} \gamma(k_c + k_r) & -k_c & -k_r \\ -k_c & 2k_c & 0 \\ -k_r & 0 & 2k_r \end{pmatrix}, \quad (\text{B4})$$

where $\tilde{k}_0 = \frac{ak_0}{k_c + k_r}$ and $\gamma = \frac{\langle B^2 \rangle + \langle B \rangle}{\langle B \rangle}$. We can directly derive the following relationships from Eq (A10):

$$(b\langle B \rangle - 1)(k_c + k_r)\sigma_{11} + bd\langle B \rangle(\sigma_{12} + \sigma_{13}) = -\frac{\tilde{k}_0}{2}\gamma(k_c + k_r) \quad (\text{B5a})$$

$$(b\langle B \rangle - 1)(k_c + k_r)\sigma_{12} + bd\langle B \rangle\sigma_{22} + bd\langle B \rangle\sigma_{23} + k_c\sigma_{11} - d\sigma_{12} = \tilde{k}_0 k_c \quad (\text{B5b})$$

$$(b\langle B \rangle - 1)(k_c + k_r)\sigma_{13} + bd\langle B \rangle\sigma_{23} + bd\langle B \rangle\sigma_{33} + k_r\sigma_{11} - d\sigma_{13} = \tilde{k}_0 k_r \quad (\text{B5c})$$

and obtain the following relationships

$$\sigma_{12} = \frac{d}{k_c}\sigma_{22} - \tilde{k}_0, \quad \sigma_{13} = \frac{d}{k_r}\sigma_{33} - \tilde{k}_0 \quad \text{and} \quad \sigma_{23} = \frac{1}{2}\left(\frac{k_r}{k_c}\sigma_{22} + \frac{k_c}{k_r}\sigma_{33}\right) - \frac{k_c + k_r}{2d}\tilde{k}_0.$$

Substituting these relationships into Eq (B4a)–Eq (B4c) yields

$$(b\langle B \rangle - 1)(k_c + k_r)\sigma_{11} + bd\langle B \rangle\left(\frac{d}{k_c}\sigma_{22} + \frac{d}{k_r}\sigma_{33}\right) = 2bd\langle B \rangle\tilde{k}_0 - \frac{\tilde{k}_0}{2}\gamma(k_c + k_r), \quad (\text{B6a})$$

$$\frac{k_c}{d}\sigma_{11} + \left(2b\langle B \rangle - 1 + \frac{3b\langle B \rangle - 2}{2}\frac{k_r}{k_c} - \frac{d}{k_c}\right)\sigma_{22} + \frac{b\langle B \rangle}{2}\frac{k_c}{k_r}\sigma_{33} = \frac{k_c}{d}\tilde{k}_0 + \left[\frac{3b\langle B \rangle - 2}{2}\frac{k_c + k_r}{d} - 1\right]\tilde{k}_0, \quad (\text{B6b})$$

$$\frac{k_r}{d}\sigma_{11} + \frac{b\langle B \rangle}{2}\frac{k_r}{k_c}\sigma_{22} + \left(2b\langle B \rangle - 1 + \frac{3b\langle B \rangle - 2}{2}\frac{k_c}{k_r} - \frac{d}{k_r}\right)\sigma_{33} = \frac{k_r}{d}\tilde{k}_0 + \left[\frac{3b\langle B \rangle - 2}{2}\frac{k_c + k_r}{d} - 1\right]\tilde{k}_0. \quad (\text{B6c})$$

The combination of Eq (B6a),(B6b) gives

$$\sigma_{22} = \left(\frac{k_r}{k_c}\right)^2 \sigma_{33} + \frac{k_c}{k_r}\frac{k_c + k_r}{d}\tilde{k}_0, \quad \text{or} \quad \sigma_{33} = \left(\frac{k_c}{k_r}\right)^2 \left(\sigma_{22} - \frac{k_c + k_r}{d}\frac{k_c}{k_r}\tilde{k}_0\right). \quad (\text{B7a})$$

The sum of Eq (B6b),(B6c) gives

$$\frac{k_c + k_r}{d}\sigma_{11} + \left[(2b\langle B \rangle - 1)\frac{k_c + k_r}{d} - 1\right]\left(\frac{d}{k_c}\sigma_{22} + \frac{d}{k_r}\sigma_{33}\right) = \left[\frac{3b\langle B \rangle - 1}{2}\frac{k_c + k_r}{d} - 1\right]\tilde{k}_0. \quad (\text{B7b})$$

The combination of Eq (B7b) and (B6a) yields

$$\frac{d}{k_c}\sigma_{22} + \frac{d}{k_r}\sigma_{33} = 1 - \frac{b\langle B \rangle + \frac{1}{2}\left[(b\langle B \rangle - 1)^2 - \gamma\right]\frac{k_c + k_r}{d}}{1 + (b\langle B \rangle - 1)(2b\langle B \rangle - 1)\frac{k_c + k_r}{d}} \quad (\text{B7c})$$

By substituting this equation into Eq (B7a), we finally obtain

$$\sigma_{22} = \frac{\tilde{k}_0 \tilde{k}_c \tilde{k}_r^3}{\tilde{k}_c^3 + \tilde{k}_r^3} \left\{ 2 + \frac{\tilde{k}_c}{\tilde{k}_r} + \frac{1}{2} \frac{2\tilde{b} + [\gamma - (1 + \tilde{b})^2] (\tilde{k}_c + \tilde{k}_r)}{1 + (1 + \tilde{b})(1 + 2\tilde{b})(\tilde{k}_c + \tilde{k}_r)} \right\}, \quad (\text{B8a})$$

and further

$$\sigma_{33} = \frac{\tilde{k}_0 \tilde{k}_r \tilde{k}_c^3}{\tilde{k}_c^3 + \tilde{k}_r^3} \left\{ 1 - \frac{\tilde{k}_c^3}{\tilde{k}_r^3} - \frac{\tilde{k}_c^4}{\tilde{k}_r^4} + \frac{1}{2} \frac{2\tilde{b} + [\gamma - (1 + \tilde{b})^2] (\tilde{k}_c + \tilde{k}_r)}{1 + (1 + \tilde{b})(1 + 2\tilde{b})(\tilde{k}_c + \tilde{k}_r)} \right\}. \quad (\text{B8b})$$

where $\tilde{b} = -b \langle B \rangle = \frac{1}{4 \langle B \rangle} \left[L_0 + 2(L_0 - 1) \langle B \rangle - L_0 (1 + 2 \langle B \rangle)^{(L_0 - 1)/L_0} \right] > 0$ with $b < 0$, $\tilde{k}_c = \frac{k_c}{d}$,

$$\tilde{k}_r = \frac{k_r}{d}, \quad \gamma = \frac{\langle B^2 \rangle + \langle B \rangle}{\langle B \rangle}.$$

Thus, the cytoplasmic mRNA noise is given by

$$\eta_c = \frac{\sigma_{22}}{\langle M_2 \rangle^2} = \frac{1}{\tilde{k}_0 \tilde{k}_c} \frac{\tilde{k}_r^3}{\tilde{k}_c^3 + \tilde{k}_r^3} \left\{ 1 + \frac{\tilde{k}_c + \tilde{k}_r}{\tilde{k}_r} + \frac{1}{2} \frac{2\tilde{b} + [\gamma - (1 + \tilde{b})^2] (\tilde{k}_c + \tilde{k}_r)}{1 + (1 + \tilde{b})(1 + 2\tilde{b})(\tilde{k}_c + \tilde{k}_r)} \right\}, \quad (\text{B9a})$$

and the nuclear mRNA noise in the nucleus by

$$\eta_r = \frac{\sigma_{33}}{\langle M_3 \rangle^2} = \frac{1}{\tilde{k}_0 \tilde{k}_r} \frac{\tilde{k}_c^3}{\tilde{k}_c^3 + \tilde{k}_r^3} \left\{ 1 - \frac{\tilde{k}_c^3}{\tilde{k}_r^3} - \frac{\tilde{k}_c^4}{\tilde{k}_r^4} + \frac{1}{2} \frac{2\tilde{b} + [\gamma - (1 + \tilde{b})^2] (\tilde{k}_c + \tilde{k}_r)}{1 + (1 + \tilde{b})(1 + 2\tilde{b})(\tilde{k}_c + \tilde{k}_r)} \right\}. \quad (\text{B9b})$$

C. Analytical results for statistical quantities in case of $L_c \neq 1$ and $L_0 = L_r = 1$

In this case, we can show that five effect transition rates take the forms: $K_1(\mathbf{x}) = k_0$,

$$K_2(\mathbf{x}) = \frac{(k_0 + x_1 k_r + dx_2 + dx_3)(x_1 k_c)^{L_c}}{(k_0 + x_1 k_c + x_1 k_r + dx_2 + dx_3)^{L_c} - (x_1 k_c)^{L_c}}, \quad K_3(\mathbf{x}) = x_1 k_r, \quad K_4(\mathbf{x}) = dx_2, \quad \text{and} \quad K_5(\mathbf{x}) = dx_3. \quad \text{In}$$

order to derive analytical results, we assume that remaining probability is so small that $p_r \approx 0$,

implying $k_r = 0$, $k_c = k_1$ and $K_3(\mathbf{x}) = 0$. By solving the steady-state deterministic equation

$$\begin{cases} \langle B \rangle K_1 - K_2 - K_3 = 0 \\ K_2 - K_4 = 0 \\ K_3 - K_5 = 0, \end{cases} \quad (\text{C1})$$

we obtain the analytical expression of steady state (\mathbf{x}^S) given by

$$x_1^s = \frac{k_0(1+\langle B \rangle)\langle B \rangle^{1/L_c}}{k_1[(1+2\langle B \rangle)^{1/L_c} - \langle B \rangle^{1/L_c}]}, \quad x_2^s = \frac{k_0\langle B \rangle}{d} \quad \text{and} \quad x_3^s = 0. \quad (\text{C2})$$

Note that the elements of Jacob matrix in the linear noise approximation reduce to

$$a_{11} = -\frac{\partial K_2(\mathbf{x}^s)}{\partial x_1}, \quad a_{12} = -\frac{\partial K_2(\mathbf{x}^s)}{\partial x_2}, \quad a_{13} = -\frac{\partial K_2(\mathbf{x}^s)}{\partial x_3}, \quad a_{21} = \frac{\partial K_2(\mathbf{x}^s)}{\partial x_1}, \quad a_{22} = \frac{\partial K_2(\mathbf{x}^s)}{\partial x_2} - d,$$

$$a_{23} = \frac{\partial K_2(\mathbf{x}^s)}{\partial x_3}, \quad a_{31} = 0, \quad a_{32} = 0, \quad \text{and} \quad a_{33} = -d. \quad \text{Differentiating function } K_2(\mathbf{x}) \text{ with regard to } x_1$$

yields

$$\frac{\partial K_2(\mathbf{x})}{\partial x_1} = \frac{L_c(k_0 + x_2d)(k_1)^{L_c}(x_1)^{L_c-1} - L_c(k_0 + x_2d)(x_1k_1)^{L_c} \left[k_1(k_0 + x_1k_1 + x_2d)^{L_c-1} - (k_1)^{L_c}(x_1)^{L_c-1} \right]}{(k_0 + x_1k_1 + x_2d)^{L_c} - (x_1k_1)^{L_c} \left[(k_0 + x_1k_1 + x_2d)^{L_c} - (x_1k_1)^{L_c} \right]^2}.$$

Therefore,

$$\left. \frac{\partial K_2(\mathbf{x})}{\partial x_1} \right|_{\mathbf{x}=\mathbf{x}^s} = -\frac{\partial K_2(\mathbf{x}^s)}{\partial x_1} = \frac{L_c k_0 \langle B \rangle}{x_1} \left[\frac{1+2\langle B \rangle}{1+\langle B \rangle} - \frac{\langle B \rangle}{1+\langle B \rangle} \left(\frac{1+2\langle B \rangle}{\langle B \rangle} \right)^{(L_c-1)/L_c} \right]. \quad (\text{C3})$$

Completely similarly, we have

$$\frac{\partial K_2(\mathbf{x})}{\partial x_2} = \frac{\partial K_2(\mathbf{x})}{\partial x_3} = \frac{d\langle B \rangle}{1+\langle B \rangle} - \frac{dL_c\langle B \rangle^2}{1+\langle B \rangle} \frac{k_0}{k_1 x_1} \left(\frac{1+2\langle B \rangle}{\langle B \rangle} \right)^{(L_c-1)/L_c}. \quad (\text{C4})$$

Furthermore, the Jacob matrix becomes

$$\mathbf{A}_s = \begin{pmatrix} a_{11} & a_{12} & a_{12} \\ -a_{11} & -a_{12} - d & -a_{12} \\ 0 & 0 & -d \end{pmatrix}, \quad (\text{C5})$$

where $a_{11} = -\frac{\partial K_2(\mathbf{x}^s)}{\partial x_1}$ and $a_{12} = -\frac{\partial K_2(\mathbf{x}^s)}{\partial x_2}$ are given by Eqs (C3) and (C4).

Meanwhile, the matrix \mathbf{D}_s in the linear noise approximation is given by

$$\mathbf{D}_s = \begin{pmatrix} k_0\langle B^2 \rangle + k_0\langle B \rangle & -K_2 & 0 \\ -K_2 & 2K_2 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (\text{C6})$$

It follows from matrix equation $\mathbf{A}_s \boldsymbol{\Sigma}_s + \boldsymbol{\Sigma}_s \mathbf{A}_s^\top + \mathbf{D}_s = \mathbf{0}$ that

$$\sigma_{11} = -\frac{k_0\langle B^2 \rangle + k_0\langle B \rangle}{2a_{11}} - \frac{a_{12}}{a_{11}^2} K_2 + \frac{a_{12}(a_{12} + d)}{a_{11}^2} \sigma_{22}, \quad (\text{C7a})$$

$$\sigma_{22} = k_0 \frac{-a_{11} (\langle B^2 \rangle + \langle B \rangle) + 2d \langle B \rangle}{2d(a_{12} + d - a_{11})}. \quad (\text{C7b})$$

Substituting the expressions of a_{11} and a_{12} into Eq (C7b) yields

$$\sigma_{22} = \frac{k_0 \langle B \rangle}{2d} \frac{\frac{L_c k_1}{\omega} \frac{1+2\langle B \rangle}{1+\omega+\langle B \rangle} (\langle B^2 \rangle + \langle B \rangle) + 2d}{\frac{d}{1+\langle B \rangle} + \frac{L_c \langle B \rangle (1+2\langle B \rangle)}{1+\omega+\langle B \rangle} \left(\frac{d}{1+\langle B \rangle} + \frac{k_1}{\omega} \right)}, \quad (\text{C8})$$

where $\omega = \frac{(1+\langle B \rangle) \langle B \rangle^{1/L_c}}{(1+2\langle B \rangle)^{1/L_c} - \langle B \rangle^{1/L_c}}$.



AIMS Press

©2022 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)