Stochasticity in single gene expression with both intrinsic noise and fluctuation in kinetic parameters

Jinzhí Lei *

Zhou Pei-Yuan Center for Applied Mathematics, Tsinghua University, 100084, PR China

ARTICLE INFO

Article history:
Received 20 September 2007
Received in revised form
31 October 2008
Accepted 31 October 2008
Available online 12 November 2008

Keywords:
Gene expression
Stochasticity
Intrinsic noise
Extrinsic noise
Chemical Langevin equation

ABSTRACT

Stochasticity is one of the most important properties in gene expression. Noise originates from two sources: thermal fluctuation inherent in the system (intrinsic noise) and variabilities in factors external to the system that usually result to the fluctuation in the kinetic parameters (extrinsic noise). This paper studies analytically the stationary fluctuation of the number of protein molecules through a mathematical model involving both sources of noises. The results in this paper show that the two sources of noises interlock to each other to generate total fluctuation in protein numbers. In particular, the extrinsic noises effect the total fluctuation in multiple ways, including the extrinsic fluctuation, the correlation with intrinsic noise, the alternation of the time averaging of transcription and translation, and the amplification of the total fluctuation by an impact factor. The impact factor is pronounced when the fluctuations in the degradation rates of mRNA or protein are large. Moreover, the extrinsic noise to the translational rate generates large fluctuation when the translational efficiency is too low, which is contrast to the translational bursting in high translational efficiency because of intrinsic noise. These results suggest that it is important to control the mRNA and protein degradation rate as well as the translational efficiency in order to attenuate the fluctuation in gene expression in the present of both intrinsic and extrinsic noises.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

The importance of stochasticity in gene expression is increasingly recognized (Kærn et al., 2005; Paulsson, 2004, 2005). For a particular gene of interest, the amount of proteins it produces will vary from cell to cell in a population and over time in a single cell. Because cellular components interact with one another in complex regulatory networks, the fluctuation in the amount of even a single component may affect the performance of the entire system. The mechanisms through which a natural genetic network can operate reliably despite noisy environments and stochasticity in gene expression are not known and remain a difficult challenge in genetic network engineering (Kærn et al., 2005). In this paper, we will study the stochasticity of the simples building block in genetic network expression of a single gene.

There are two sources of stochasticity in gene expression: the intrinsic noise and extrinsic noise. First, gene expression is essentially a sequence of biochemical reactions that are inherently stochastic due to the random births, deaths, and collisions of the molecules. The inherent stochasticity is often referred to as intrinsic noise. Second, the cell environment is complicate and the expression of a particular gene may be regulated by other molecular species that are, themselves, gene products with populations that vary over time and from cell to cell. The variation in the environmental conditions may produce complicate effects on the fluctuations in gene expression and will be referred to as extrinsic noise. The extrinsic sources of noise arise independently of the gene and usually relate to the variability in the rate constants that are associated with the biochemical reactions in the gene's expression. In general, the total variation in gene expression is the joint effect of both intrinsic and extrinsic noise.

Although the stochastic nature of gene expression has long been postulated (Spudich and Koshland, 1976), most present studies have concentrated on intrinsic noise (Bar-Even et al., 2006; Blake et al., 2003; Elowitz et al., 2002; Kærn et al., 2005; Kepler and Elston, 2001; Ozbudak et al., 2002; Pedraza and Paulsson, 2008; Paulsson, 2004, 2005; Raser and O'Shea, 2004; Spudich and Koshland, 1976; Thattai and van Oudenaarden, 2001, 2002; Tian and Burrage, 2006; Wang et al., 2005). One can refer Kærn et al. (2005) and Paulsson (2005) for reviews on this subject.

In single gene expression, the intrinsic noise comes from fluctuations generated by stochastic promoter activation, promoter inactivation, mRNA and protein production and decay. The magnitude of the intrinsic noise is proportional to the inverse of the system size. Translational bursting and transcriptional bursting are two consequences of intrinsic noise that are associated
with high translational efficiency and slow promoter kinetics, respectively (Kærn et al., 2005).

Despite extensive studies of intrinsic noise, studies of extrinsic noise in gene expression have only begun in recent years (Elowitz et al., 2002; Hasty et al., 2000; Pedraza and van Oudenaarden, 2005; Shahrezaei et al., 2008; Swain et al., 2002; Tao, 2004). How extrinsic noise affects the stochasticity in gene expression remains unclear. Experiments have shown that both intrinsic and extrinsic noise contribute substantially to the overall variation (Elowitz et al., 2002). Therefore, the effect of extrinsic noise is important and should not be neglected.

Experimentally, it is difficult to distinguish intrinsic noise from extrinsic noise in vivo. Swain et al. (2002) have suggested a way to decompose the experimentally measured noise into a direct sum of intrinsic and extrinsic contributions. Accordingly, the two components of fluctuations in a single cell were experimentally differentiated (Elowitz et al., 2002). However, how these two components depend on system parameters and extrinsic noise is not clear. Therefore, it is difficult to predict the experimental results from the molecule properties.

In this paper, we will study the total fluctuation in gene expression through a mathematical model that involves both intrinsic and extrinsic noises. The extrinsic noise is usually originated from the total noise of the protein concentration from upstream gene expression and transcription factors and can be very complicated. In the current study, as we focus on the stochasticity in single gene expression, we will model the overall effect of extrinsic noise as the fluctuations in kinetic parameters. In doing this, we will start from a set of chemical Langevin equations for the chemical reactions in gene expression. From these equations, we are able to obtain the mathematical formulation for the stationary fluctuation in terms of reactions rates and extrinsic noise strengths. This result enables us to study the stochasticity in gene expression in noisy environments, and could be useful towards genetic network engineering.

In this paper, we will focus on the analytical studied and always assume the extrinsic noise perturbation to be white noise. But experiments have shown that the noises are usually colored and have long autocorrelation time (Kaufmann et al., 2007; Rosenfeld et al., 2005; Sigal et al., 2006). We will discuss the effect of colored noise numerically.

2. Model

Essentially, gene expression is a sequence of biochemical reactions, which are associated with mRNA and protein production, transitions between promoter states, and the decay of mRNA and protein. The main steps in gene expression are illustrated in Fig. 1. The stochasticity in gene expression comes from fluctuations generated by random promoter activation and inactivation, mRNA and protein production and decay, and the fluctuation in the reaction rates. Let $X_1, X_2, X_3$ be the amounts of active genes, mRNAs and proteins, respectively. Above processes can be modeled by chemical Langevin equations as follows (refer Appendix A for detail):

$$\frac{dx_1}{dt} = \lambda_1^n (n - x_1) - \lambda_1 x_1 + \sqrt{\lambda_1^n (n - x_1) \xi_1(t)} - \sqrt{\lambda_1 x_1 \xi_2(t)} + f_1 \gamma_3 (n - x_1) \eta_1(t) - \lambda_1 x_1 \eta_3(t),$$  

$$\frac{dx_2}{dt} = \lambda_2 x_1 - \delta_2 x_2 + \sqrt{\lambda_2 x_1 \xi_3(t)} - \sqrt{\delta_2 x_2 \xi_4(t)} + f_2 \gamma_3 x_1 \eta_3(t),$$  

$$\frac{dx_3}{dt} = \lambda_3 x_2 - \delta_3 x_3 + \sqrt{\lambda_3 x_2 \xi_5(t)} - \sqrt{\delta_3 x_3 \xi_6(t)} + f_3 \gamma_3 x_2 \eta_3(t),$$

The parameters are $n$, the gene copy number that is assumed to be fixed for the gene of interest; $\lambda_1, \lambda_1^n$, the rates of gene activation and inactivation, respectively; $\lambda_2, \lambda_3$, the production rates of mRNA and protein, respectively; $\delta_2, \delta_3$, the degradation rates of mRNA and protein, respectively. The $\gamma_i$'s are constants representing the standard deviations of noise perturbation. The $\xi_i$'s and $\eta_i$'s are independent white noises standing for intrinsic and extrinsic noises, respectively. Moreover, we apply Itô interpretation for the intrinsic noises, and Stratonovich interpretation for the extrinsic noises (refer Appendix A for discussion).

Eqs. (1)–(3) constitute a highly simplified representation of the problem. In biological systems, the time course of reaction rates depends on the environment condition and the amount of other proteins produced from upstream gene expressions, which are usually random and time-correlated in complex ways. Experimental observation has shown that external noise can have long correlation time (colored noise), which is at a time scale of about one cell cycle (from 40 min in bacteria to 200 min in human) (Kaufmann et al., 2007; Rosenfeld et al., 2005; Sigal et al., 2006). In the present study, we neglect the time correlation and adopt the simplest assumption that the perturbations are independent white noises. The effects of colored extrinsic noise will be discussed numerically in Section 3.4.

From the above mathematical model (1)–(3), we are able to obtain the average amounts $\langle X_i \rangle$ and stationary fluctuations $\eta_i^2$ using linear noise approximation (refer Appendix B for detail). The stochasticity in gene expression is measured by the stationary fluctuations and will be discussed below.

3. Results

3.1. Extrinsic noises alter the mean protein number

According to Eqs. (1)–(3), we obtain the average number of the molecules at stationary state as

$$\langle X_1 \rangle = \frac{g_1}{1 + g_1} n, \quad \langle X_2 \rangle = g_2 \langle X_1 \rangle, \quad \langle X_3 \rangle = g_3 \langle X_2 \rangle,$$

where

$$g_1 = \frac{\lambda_1^n}{\delta_1^n}, \quad g_2 = \frac{\lambda_2}{\delta_2 - \delta_1^n}, \quad g_3 = \frac{\lambda_3}{\delta_3 - \frac{\delta_1^n}{\delta_2}},$$

Here $g_2$ and $g_3$ are referred to as the transcriptional efficiency and the translational efficiency, respectively.

From (4), the averages of the molecule numbers depend on extrinsic noise strengths. Fluctuations in $\lambda_1$, $\delta_2$ or $\delta_3$ may increase the mean protein number, while fluctuation in $\lambda_1^n$ may decrease the mean protein number (Fig. 2). Mathematically, this effect is originated from the Stratonovich interpretation for the extrinsic noises. The same effect was also found recently by stochastic
simulation (Shahrezaei et al., 2008). The above consistence reveals the justification of applying Stratonovich interpretation for the extrinsic noises.

It is easy to obtain from (4) that the stationary solution exists only when
\[ f_{z_1}^2 + f_{z_1}^3 < 2(\lambda_{z_1} + \zeta_{z_1}), \quad f_{z_1}^3 < 2\delta_z, \quad f_{z_1}^2 < 2\delta_z, \]
which means that the noise perturbation to the degradation rates are weak.

3.2. Extrinsic noises effect the total fluctuation through multiple ways

In the present of both intrinsic and extrinsic noises, the total fluctuations can be obtained from (1) to (3) through the Itô formula (refer Appendix B for detail). The results are
\[
\begin{align*}
\eta_{i_1}^2 &= k_i^2 \frac{\lambda_{i_1} \text{obs}}{\sigma_{i_1}} \left[ \frac{1}{(X_1)} + \frac{1}{2} \left( \frac{\zeta_{i_1}}{\bar{g}_1} + \zeta_{i_1} \right) \right] + \frac{1}{4} \left( \frac{\zeta_{i_1}}{\bar{g}_1} + \zeta_{i_1} \right)^2, \\
\eta_{i_2}^2 &= k_i^2 \left[ \frac{1}{(X_2)} + \frac{1}{2} \left( \frac{\zeta_{i_2}}{\bar{g}_2} + \frac{\zeta_{i_2}}{\bar{g}_2} (1 + \eta_{i_2}^2) + \zeta_{i_2} \right) + \frac{1}{4} \left( \frac{\zeta_{i_2}}{\bar{g}_2} + \eta_{i_2}^2 \right)^2, \\
\eta_{i_3}^2 &= k_i^2 \left[ \frac{1}{(X_3)} + \frac{1}{2} \left( \frac{\zeta_{i_3}}{\bar{g}_3} + \frac{\zeta_{i_3}}{\bar{g}_3} (1 + \eta_{i_3}^2) + \zeta_{i_3} \right) + \frac{1}{4} \left( \frac{\zeta_{i_3}}{\bar{g}_3} + \eta_{i_3}^2 \right)^2 \right],
\end{align*}
\]
where
\[
\begin{align*}
\eta_{i_1} &= \frac{\tau_1 + \tau_2 + \tau_3}{\tau_1 + \tau_2 + \tau_3} \eta_{i_1}^2, \\
\eta_{i_2} &= \frac{\tau_1 \tau_3}{\tau_1 + \tau_2 + \tau_3} \frac{\tau_1}{\tau_1 + \tau_2 + \tau_3} \eta_{i_2}^2 + \frac{\tau_2}{\tau_2 + \tau_3} \eta_{i_2}^2, \\
\eta_{i_3} &= \frac{\tau_1 \tau_2}{\tau_1 + \tau_2 + \tau_3} \frac{\tau_1}{\tau_1 + \tau_2 + \tau_3} \eta_{i_3}^2 + \frac{\tau_2}{\tau_2 + \tau_3} \eta_{i_3}^2.
\end{align*}
\]
Here \(\eta_{i_1}^2\), \(\eta_{i_2}^2\) and \(\eta_{i_3}^2\) are the stationary fluctuations in the numbers of active genes, mRNAs and proteins, respectively. Other notations are defined in Appendix B. In particular, the \(\zeta\)'s are the noise strengths, \(\tau_i\) (\(i = 1, 2, 3\)) are average lifetimes of active gene, mRNA and protein, respectively, and \(k_i\) are impact factors defined by
\[
k_i = \frac{1}{1 - \frac{\zeta_{i}}{\bar{g}_{i}}} \quad (i = 1, 2, 3).
\]

The impact factors (9) suggest that results (6)–(8) are valid only when
\[
\zeta_i < 2 \quad (i = 1, 2, 3).
\]
In fact, it is easy to prove that the stationary solution is stable only if condition (10) is satisfied. We note that (10) automatically implies (3). The discussions below will limit to restriction (10). Comparisons between (8) and numerical simulation are given in Fig. 2.

From (6) to (8), we are able to decompose the total fluctuation in protein numbers into different components as following. In (8), the stationary fluctuation consists the contributions from the intrinsic fluctuation \(k_i^2/(X_3)\), the extrinsic fluctuation \(k_i^2/2\eta_i^2\), the correlation between intrinsic and extrinsic noises \(k_i^2/4\zeta_i\), and those transmitted from the fluctuations in the numbers of upstream molecules \(k_i \eta_i\). Here the transmission fluctuation depends on the time averaging of both transcription and translation, which are altered by the extrinsic noises. Mathematically, this effect is again originated from the Stratonovich interpretation for the extrinsic noise. Both contributions from the extrinsic fluctuation and from the correlation are proportional to the extrinsic noise strengths. All components are proportional to an impact factor \(k_i\), which increases rapidly when \(\zeta_i\) approaches 2. Similar results are also held for the stationary fluctuations in the number of active genes and mRNA molecules.

From the above discussion, the extrinsic noise effect the total fluctuation in protein numbers through multiple ways, including the extrinsic fluctuation, the correlation with intrinsic noises, the modification of the time averaging of transcription and translation, and the amplification of the overall fluctuation by the impact factors. In particular, the effect of impact factors is pronounced when the fluctuations in the degradation rates of mRNA or protein are large. This result suggests that it is important to reduce the fluctuation in mRNA and protein degradation rates in order to attenuate the fluctuation in gene expression.

Fig. 2. Extrinsic noise changes the mean and the probability distribution for protein numbers in gene expression. (A) A histogram of protein numbers generated by stochastic simulation. Only intrinsic fluctuations are included. (B) A histogram of protein numbers generated from intrinsic fluctuation and extrinsic noise perturbation to the translational rate \(\lambda_3\). (C) A histogram of protein numbers generated from intrinsic fluctuation and extrinsic noise perturbation to the protein degradation rate \(\zeta_3\). In all panels, the results obtained from simulation and theoretical prediction are shown. In both (B) and (C), the Fano factors of the extrinsic noises are 0.25. The reaction rates are given in Fig. 1, with and the gene copy number \(n = 1\). Simulation method is given in Appendix C.
3.3. Extrinsic fluctuation of the translational rate can induce large fluctuation when the translational efficiency is low

From previous discussion, the contribution of the total fluctuation that is originated from the variation of the translational rate is proportional to the inverse of translational efficiency. Consequently, the fluctuation in protein number is large when $\zeta_{i,1} > 0$ and the translational efficiency is low. This is because for the genes expressed at the same average number of protein molecules, the gene with lower translational efficiency needs greater number of mRNA molecules. The noise perturbs the translation of each mRNA molecule with the same strength, and therefore generates larger fluctuation in the protein number when there are more mRNA molecules.

The above result is relevant when we try to regulate the noise in gene expression by controlling the translational efficiency (Ozbudak et al., 2002). We note that when the intrinsic noise is the main source of fluctuation, the genes with high translational efficiency and low number of mRNA are known to have large fluctuation in protein number. This is known as the mechanism of translational bursting (Blake et al., 2003; Ozbudak et al., 2002; Thattai and van Oudenaarden, 2001, 2002). On the other hand, previous discussion shows that decreasing translational efficiency can result in large fluctuation as well when there is extrinsic noise. Therefore, the total fluctuation could be minimized if the translational efficiency takes intermediate value.

Substituting (7) into (8) and minimizing $\eta_{i,2}^2$, we find that the fluctuation $\eta_{i,2}^2$ achieves its minimum when the translational efficiency satisfies

$$ g_3 = C_0 \sqrt{\zeta_{i,3} E_{X,i}}. \quad (11) $$

where $C$ is a constant independent to the noise strength $\zeta_{i,2}$ and the average protein number $E_{X,i}$. Relation (11) shows that in order to minimize the fluctuation in protein number, the translational efficiency should be low when the noise for the translational rate is small, and high when the noise is large (Fig. 3). This result suggests that well control of translational efficiency according to the extrinsic noise strength is important to attenuate the fluctuation in gene expression.

![Fig. 3. Stationary fluctuations in the protein numbers with different translational efficiency. The results for weak extrinsic noise (with Fano factor 0.04) and strong extrinsic noise (with Fano factor 0.25) are shown. Arrows show the values of translational efficiency such that the fluctuation is minimum. The parameters are given at Fig. 1, $\lambda_3$ and $\lambda_4$ are adjusted such that the translational efficiency varies over wide range, with average protein number unchanged. All extrinsic noise strengths except the fluctuation in $\lambda_4$ are set to zero.]

3.4. Colored noise can be important for the fluctuation in gene expression

One of the main assumption in the previous studies is that the extrinsic noise is white. However, many experimental observations show that the noises are colored, having an autocorrelation time that is not negligible but comparable to the cell cycle (Rosenfeld et al., 2005). The effect of colored noise is difficult to predict analytically (Shahrezaei et al., 2008). Here, we examine the effect of colored noise numerically. To this end, we assume the extrinsic noises in (1)–(3) to be specific processes, the Ornstein–Uhlenbeck processes (Kauffman et al., 2007; Rosenfeld et al., 2005; van Kampen, 1992). The effects of colored noise to the translational rate (Fig. 4A and B) and to the protein degradation rate (Fig. 4C and D) are shown in Fig. 4.

When the colored noise is added to the translational rate, the total fluctuation increase obviously when the autocorrelation time increases from 0 to 2 h. Extending the autocorrelation time ($\tau > 2$ h) shows minor effect (Fig. 4A). The colored noise to the translational rate can induce large fluctuation when the translational efficiency is low (Fig. 4B). This consists with our previous result for white extrinsic noise. Note that the color noise perturbation to the protein production rate has been postulated in the previous model. This comes from the fluctuations in the number of active genes and mRNA molecules and is usually colored because of the nonzero average lifetime $\tau_i$. Here, we specify the effect of color noise perturbation to the translational rate that is originated from the fluctuations in the upstream control factors.

When the colored noise is added to the protein degradation rate, the effect of nonzero autocorrelation time can be dramatic even when $\tau$ is as small as 10 s. Such effect includes increasing the mean protein number and amplifying the total fluctuation (Fig. 4C and D). Moreover, long autocorrelation time can destabilize the stationary solution and the current model is invalid (data not shown).

Analytic explanations for the above simulation results are not known yet. These results suggest that colored noise can be important for the fluctuation in gene expression. Nonzero autocorrelation time creates some profound differences with the Langevin equation and needs for further studies, both mathematically and experimentally.

4. Relation to other works

4.1. Intrinsic noise fluctuation

Fluctuation in gene expression due to the intrinsic noise has extensively studied by many authors (Elowitz et al., 2002; Kærn et al., 2005; Kepler and Elston, 2001; Ozbudak et al., 2002; Pedraza and Paulsson, 2008; Paulsson, 2004, 2005; Raser and O’Shea, 2004; Thattai and van Oudenaarden, 2001, 2002). In our previous discussions, setting the extrinsic noise to zero, we obtain again the stationary fluctuation in protein number that was already known (for example, refer Paulsson, 2004, 2005):

$$ \eta_{i,3}^2 = \frac{1}{(X_3)} + \frac{\tau_2}{\tau_3 + \tau_2} + \frac{1}{(X_2)} + \frac{1}{\tau_1 + \tau_3} \left( 1 - P_{on} \right) \frac{1}{(X_1)} $$

$$ + \frac{1 - P_{on}}{\tau_1 + \tau_3} + \frac{\tau_2}{\tau_2} + \frac{\tau_3}{\tau_3 + \tau_3} \right) + \tau_3 + \tau_3 + \tau_3 + t_3 + t_3 + \tau_3 $$

(12)

where $\tau_i$ ($i = 1, 2, 3$) are average lifetimes of active gene, mRNA, and protein, respectively, and $P_{on} = \lambda_1 / (\lambda_1 + \lambda_2)$. On can refer Paulsson (2004, 2005) for more discussions on (12), and refer Kærn et al. (2005) for a review on intrinsic noise in gene expression.
4.2. Extrinsic noise fluctuation

The effects of extrinsic noise have been discussed in recent years, both numerically (Shahrezaei et al., 2008) and analytically (Pedraza and van Oudenaarden, 2005; Scott et al., 2006; Swain et al., 2002; Tao, 2004). In most studies, the sub-system of the present study is considered. The complete result similar to that of the case of intrinsic noise is not known. In our model, if we set $\xi(t)$ to zero, we have a system with only extrinsic noise, and the stationary fluctuations can be obtained as

$$\eta_{\text{ext}, 1}^2 = k_1 \frac{\lambda_{1, \text{obs}}}{\sigma_{1, \text{obs}}} \left( \frac{\tau_1 + \tau_2}{\tau_1 + \tau_2} \right), \quad (13)$$

$$\eta_{\text{ext}, 2}^2 = k_2 \left[ \frac{1}{2} \left( \frac{\tau_1}{\tau_2} \left( 1 + \eta_{\text{ext}, 2}^2 \right) + \eta_{\text{ext}, 1} \right) + \frac{\tau_1}{\tau_1 + \tau_2} \eta_{\text{ext}, 1}^2 \right], \quad (14)$$

$$\eta_{\text{ext}, 3}^2 = k_3 \left[ \frac{1}{2} \left( \frac{\tau_1}{\tau_3} \left( 1 + \eta_{\text{ext}, 2}^2 \right) + \eta_{\text{ext}, 1} \right) + \frac{\tau_1}{\tau_1 + \tau_2 + \tau_3} \eta_{\text{ext}, 1}^2 + \frac{\tau_2}{\tau_2 + \tau_3} \eta_{\text{ext}, 2}^2 \right]. \quad (15)$$

Eq. (15) gives the fluctuation in protein number due to the variation in the reaction rates.

5. Conclusion

In this paper, we have studied the stochasticity in single gene expression with both intrinsic and extrinsic noise. The stochastic process of gene expression is modeled by chemical Langevin equation, through which the stationary fluctuation in protein number is derived from the system parameters and noise strengths.

Our mathematical model shows that extrinsic noise in gene expression can alter the mean protein numbers, which consists with recent simulation results (Shahrezaei et al., 2008). This effect is mathematically originated from the Stratonovich interpretation of the extrinsic noises.

The stationary fluctuation of protein number depends on intrinsic and extrinsic noise in a complicated way, and no simple method is available to distinguish these two sources of noises. The fluctuation generated from intrinsic noise is usually proportional to the inverse of the molecular number. The extrinsic noise effects the total fluctuation in different manner, including the extrinsic fluctuation that is proportional to the extrinsic noise strengths, the correlation with intrinsic noises that is proportional to the extrinsic noise strengths and the inverse of molecular number, the modification of the time averaging of transcription and translation that is originated from the Stratonovich interpretation of the extrinsic noises, and the amplification of all components by the impact factors that can be pronounced when the fluctuations in mRNA or protein degradation rates are large. Moreover, the extrinsic fluctuation of the translational rate can induce large fluctuation when the translational efficiency is low. This effect is in contrast to translational bursting that usually happen when the translational efficiency is high because of the intrinsic noise. These results suggest that it is important to reduce the fluctuation in degradation rates of mRNA and protein and to control the translation efficiency while we try to attenuate the fluctuation in genetic network.

We conclude this paper with a final remark. We note that the mathematical model proposed in the present study constitutes a simplified representation of the problem. Further work remains to be done for more realistic situations. For example, when the extrinsic noise is strong, the system can be destabilized and more complicated dynamical behaviors may occur. In this case, the formula for the stationary fluctuation obtained in this paper is not...
valid. The random perturbations in real situations are not exactly white noises, and time correlations are important. New mathematical tools need to be developed for further study of these situations.

Acknowledgements

The author would like to thank the anonymous referees for their valuable comments and suggestions. The author is supported by the National Science Foundation of China (NSFC 10601029).

Appendix A. Theory

Gene expression is essentially a sequences of biochemical reactions. Here, we introduce some basic mathematical formulations of chemical reactions. Most of the contexts in this discussion are standard and one can refer (Gillespie, 2000; van Kampen, 1992) for detail.

A.1. Chemical Langevin equation

In usual, the stochastic dynamical behavior of a well-stirred mixture of N molecular species that chemically interact through M reaction channels can be modeled as a Markov process through chemical master equation (van Kampen, 1992). Alternatively, Gillespie proposed following chemical Langevin equation whenever two explicit dynamical conditions are satisfied (Gillespie, 2000)

\[
\frac{dX_i(t)}{dt} = \sum_{j=1}^{M} v_{ji}a_j(X(t)) + \sum_{j=1}^{M} v_{ji}^{1/2}/2 \chi_j(t), \quad (i = 1, 2, \ldots, N),
\]  

(16)

where \(X_i(t)\) is the number of the \(i\)th molecules, \(a_j(X)\) is the propensity function of the \(j\)th reaction channel, \(v_{ji}\) is the change in the \(i\)th molecule caused by the \(j\)th reaction channel, the \(\chi_j(t)\) are temporally uncorrelated, statistically independent Gaussian white noises satisfying

\[
\langle \chi_j(t) \rangle = 0, \quad \langle \chi_j(t) \chi_j(t') \rangle = \delta_{jj} \delta(t - t').
\]  

(17)

The validity of (16) has been discussed in Gillespie (2000, 2002). In general, when the molecular population is large, chemical Langevin equation (16) give reasonable approximation of the time course of the state of the system. If we only concern ourselves with the second moment of the system (as we did in this paper), chemical Langevin equation provides consistent results with the classical chemical master equation.

We can extend the chemical Langevin equation (16) to include extrinsic fluctuations. Assume that the propensity functions depend on the reaction rates \([c_1, \ldots, c_K]\), and there are noise perturbations to the reaction rates, i.e.,

\[
c_i = c_i^0 + f_{ci} \eta_i(t) \quad (i = 1, \ldots, K),
\]

where \(c_i^0 = \langle c_i \rangle\) and the noise perturbations are represented by \(f_{ci} \eta_i(t)\). Expanding the propensity function \(a_j(X)\) as a Taylor series of perturbations, and taking the first order approximation, we obtain following equations:

\[
\frac{dX_i(t)}{dt} = \sum_{j=1}^{M} v_{ji}a_j(X) + \sum_{j=1}^{M} v_{ji}^{1/2}/2 \chi_j(t) + \sum_{j=1}^{K} f_{ci}b_j(X)\eta_i(t),
\]

(18)

where \(a_j(X) = a_j(X)|_{c^0}\) and

\[
b_j = \sum_{k=1}^{M} v_{ik} \frac{\partial a_k(X)}{\partial c_j} |_{c^0}.
\]

(19)

In this paper, we always assume that \(\eta_i\) are independent white noises:

\[
(\eta_i(t)|_{t = t'}) = \delta_{ii} \delta(t - t').
\]

(20)

Therefore the coefficients \(f_{ci}\) is the standard deviation of the perturbations. Further, we will assume that \(f_{ci}\) is small compare with \(c_i^0 (i = 1, \ldots, K)\).

Eq. (18) are desired formulations of chemical reactions with both intrinsic and extrinsic noise. Mathematically, (18) should be interpreted as stochastic integral equations. The last two terms in each of the equations can be interpreted under the meaning of either Itô or Stratonovich, and are discriminated based on their sources. In the present study, we adopt the suggestion from van Kampen (1992), and apply Stratonovich interpretation for the extrinsic noise, and Itô interpretation for the intrinsic noise. Justifications of these interpretations are shown in the text.

A.2. Variance in stationary solution

This paper will focus on the fluctuation of the level of single gene expression under stochastic perturbation, which is measured by the variance of the molecular populations in stationary state.

Let \(\sigma\) be the matrix of variance. Consider the chemical reaction in the previous section in the absence of extrinsic noise. The classical result shows that \(\sigma\) satisfies (Paulsson, 2005; Keizer, 1987)

\[
\frac{d\sigma}{dt} = A\sigma + \sigma A^T + B,
\]

(21)

by linear approximation. Here \(A\) is the Jacobian matrix for the dynamics of the averages and \(B\) is a diffusion matrix that depends on the size of the random events.

Eq. (21) was obtained from chemical master equation (Keizer, 1987; van Kampen, 1992). As has been emphasized in Paulsson (2005), the key approximation to obtain (21) is that matrices \(A\) and \(B\) are interpreted assuming that fluctuations can be ignored. The average rates are then approximated by the rates at the average concentration. This approximation does not affect analysis of linear systems, and it may lead to incorrect answer for nonlinear systems with large fluctuations. When there is extrinsic noise, however, linear noise approximation is not enough and (21) has to be revised even for linear systems.

In the present of extrinsic noises, the variance in the stationary solution can be obtained from (18) by using Itô formula (Gardiner, 1983), and noticing that extrinsic noises are interpreted as Stratonovich interpretation. The calculations are standard and tedious and will be omit here. The final result is as follows.

Let \(\langle X \rangle\) be the average number of molecules and satisfy

\[
\frac{d\langle X \rangle}{dt} = \sum_{j=1}^{M} v_{ji}(a_j(X)) + \frac{1}{2} \sum_{k=1}^{K} f_{ci}^2 d_{ii}(X)\]

(22)

Define the matrices \(A, B\) and \(F^T\) as

\[
A_{ij} = \frac{\partial}{\partial X_i} \left[ \sum_{k=1}^{M} v_{ik} a_k(X) + \frac{1}{2} \sum_{k=1}^{K} f_{ci}^2 d_{ii}(X) \right]_{X = \langle X \rangle}
\]

and

\[
B_{ij} = G_{ij}(X), \quad F^T_{pj} = \frac{\partial^2 G_{p}(X)}{\partial X_i \partial X_q}.
\]

(23)

(24)
where
\[ G_0(X) = \sum_{k=1}^{N} v_k b_k \partial_k + \sum_{k=1}^{N} f_k^2 b_k(X)b_k(X). \]

Then the variance matrix \( \sigma \) satisfies
\[ \frac{d\sigma}{dt} = (A\sigma + \sigma A^T + B) + \frac{1}{2} \text{trace}(F^T\sigma). \] (25)

The stationary variance is obtained by letting time derivative in (25) equal to 0 and solving for \( \sigma \), i.e., the solution of
\[ (A\sigma + \sigma A^T + B) + \frac{1}{2} \text{trace}(F^T\sigma) = 0 \quad (i, j = 1, \ldots, N). \] (26)

From above results, when the system is linear and no extrinsic noise is presented (\( f_i = 0 \)), we have \( F^T = 0 \) and obtain again (21), which can be obtained from the chemical master equation (Keizer, 1987). This consistency indicates that the chemical Langevin equation (16) and the Itô interpretation for the intrinsic noise are justified up to the second moment of the system. For linear systems as we consider in this paper, Eq. (25) is exact for the variance matrix in the present of extrinsic noise.

**Appendix B. Stationary fluctuation in single gene expression**

To obtain the stationary fluctuation in single gene expression, we introduce the constants below. The observed rate constants
\[ \lambda_1^+_{\text{obs}} = \lambda_1^+ - \frac{f_1^2}{\lambda_1^+}, \quad \lambda_1^-_{\text{obs}} = \lambda_1^- - \frac{f_1^2}{\lambda_1^-}, \]
\[ \delta_1^+_{\text{obs}} = \lambda_1^+_{\text{obs}} + \lambda_1^-_{\text{obs}}, \quad \delta_1^-_{\text{obs}} = \delta_1 - \frac{f_1^2}{\delta_1}, \]

and net production rates
\[ b_1 = \lambda_1^+_{\text{obs}} - \frac{f_1^2}{\lambda_1^+}, \quad b_2 = \lambda_2 - \frac{f_2^2}{\lambda_2}, \quad b_3 = \lambda_3. \] (28)

The noise strengths are defined as
\[ \zeta_{11} = \frac{f_1^2}{\lambda_1^+}, \quad \zeta_{12} = \frac{f_1^2}{\lambda_2}, \quad \zeta_{13} = \frac{f_1^2}{\lambda_3}. \] (29)

The impact factors
\[ k_i = \frac{1}{1 - \frac{f_i^2}{\lambda_i}}, \quad (i = 1, 2, 3), \]

where
\[ \zeta_{ih} = \frac{f_i^2}{\lambda_i^+} = \frac{f_i^2}{\lambda_i^-} = \frac{1}{1 + \frac{f_i^2}{\lambda_i}} \]
is the average noise strength of the transition process.

In what follows, we will see that the noise strengths are relevant for the total fluctuation in protein number. It is more straightforward to associate them with experiment observations by the Fano factors, which are defined as
\[ F_c = \frac{f_i^2}{c} \quad (c = \lambda_1^+, \lambda_1^-, \lambda_2, \lambda_3, d_3). \] (30)

The noise strengths and the Fano factors for extrinsic noises are connected by
\[ \zeta_{ij} = \frac{1}{1 - \frac{f_i^2}{2f_i}}, \quad \zeta_{ji} = \frac{1}{1 - \frac{f_i^2}{2f_i}}, \]
\[ \zeta_{ik} = \frac{1}{1 - \frac{f_i^2}{2f_i}}, \quad \zeta_{ki} = \frac{1}{1 - \frac{f_i^2}{2f_i}}. \] (31)

From (1) to (3) and applying Stratonovich interpretation to the extrinsic noise, we obtain the equations for the averages \( \langle X_i \rangle \)
\[ \frac{dx_i}{dt} = \dot{\lambda}_1^+ - \dot{\lambda}_1^- - \lambda_2 - \lambda_3, \]
\[ \frac{dx_1}{dt} = \dot{\lambda}_2, \quad \frac{dx_2}{dt} = \dot{\lambda}_3. \] (32)

Thus, the average numbers of molecules at the stationary state are
\[ \langle X_1 \rangle = \frac{g_1}{1 + g_1}, \quad \langle X_2 \rangle = g_2, \quad \langle X_3 \rangle = g_3. \] (33)

The stationary variances can be obtained by calculating \( A, B \) and \( F \) from (23) and (24), and solving the equation
\[ (A\sigma + \sigma A^T + B) + \frac{1}{2} \text{trace}(F^T\sigma) = 0. \] (34)

It is easy to have
\[ A = \begin{bmatrix} -\dot{\lambda}_{\text{obs}} & 0 & 0 \\ \dot{\lambda}_2 & -\delta_{\text{obs}} & 0 \\ 0 & \dot{\lambda}_3 & -\delta_{\text{obs}} \end{bmatrix}, \]

and \( B \) is a diagonalization matrix, with
\[ B_{11} = 2\lambda_{\text{obs}}^+ \left(1 + \frac{\zeta_{11} + \zeta_{12}}{4}\right)\lambda_1 + \left(1 - g_1^2/\lambda_1^+ + g_1^2/\lambda_1^-\right)\lambda_1^-. \]
\[ B_{12} = 2\delta_{\text{obs}} \left(1 + \frac{\zeta_{12}}{4}\right)\lambda_2 + \left(1 - g_1^2/\lambda_2 + g_1^2/\lambda_3\right)\lambda_2, \]
\[ B_{13} = 2\delta_{\text{obs}} \left(1 + \frac{\zeta_{13}}{4}\right)\lambda_3 + \left(1 - g_1^2/\lambda_2 + g_1^2/\lambda_3\right)\lambda_3. \]

The matrix \( F \) is 0 when \( i \neq j \), and
\[ F_{11} = \text{diag}(2\frac{f_1^2}{\lambda_1^+} + 2\frac{f_1^2}{\lambda_1^-}, 0, 0), \]
\[ F_{12} = \text{diag}(2\frac{f_1^2}{\lambda_2}, 2\frac{f_1^2}{\lambda_3}, 0), \]
\[ F_{13} = \text{diag}(0, 2\frac{f_1^2}{\lambda_2}, 2\frac{f_1^2}{\lambda_3}). \]

Substitute \( A, B, F \) into (36) and solve for \( \sigma \), we have
\[ \sigma_{11} = \frac{B_{11}}{2\lambda_{\text{obs}}^+ - \left(f_1^2 + f_1^2\right)} \]
\[ \sigma_{12} = \frac{\dot{\lambda}_2}{\delta_{\text{obs}}^+ - \delta_{\text{obs}}^-} \sigma_{11}, \]
\[ \sigma_{13} = \frac{1}{2\lambda_{\text{obs}} - f_1^2} \left(\sigma_{11} + \sigma_{12}\right) \]
\[ \sigma_{23} = \frac{\dot{\lambda}_2}{\delta_{\text{obs}}^-} \sigma_{12}, \]
\[ \sigma_{33} = \frac{1}{2\lambda_{\text{obs}} - f_1^2} \left(\sigma_{11} + \sigma_{12}\right) \]

Experimentally, the noise is usually measured by the variance over squared average at stationary state, \( \eta^2 = \sigma^2/\langle X \rangle^2 \), and is often referred to as stationary fluctuation. Here \( \eta \) is referred to as coefficient of variance.
Define the matrix of stationary fluctuation $\eta$ as

$$
\eta_{ij} = \frac{\sigma_{ij}}{(X_i/X_j)} \quad (i,j = 1, 2, 3).
$$

Then the stationary fluctuations are

$$
\eta_{11} = k_1 \left( \frac{1}{2} \zeta_{21} + \zeta_{11} \right),
$$

$$
\eta_{22} = k_2 \left( \frac{1}{2} \zeta_{21} \right),
$$

$$
\eta_{33} = k_3 \left( \frac{1}{2} \zeta_{21} \right),
$$

and

$$
\eta_{12} = \frac{\zeta_{11}}{\tau_1},
$$

$$
\eta_{23} = \frac{\zeta_{22}}{\tau_2}.
$$

where $\tau_i = 1/\delta_{obs} \quad (i = 1, 2, 3)$ are average lifetimes.

Eqs. (42)–(46) give complete formulas for the stationary fluctuation in gene expression with both intrinsic and extrinsic noises. In the text, we use $\eta_{ij}^2 = \eta_{ij} \quad (i = 1, 2, 3)$ for the stationary fluctuation in the number of active genes, mRNA and proteins molecules, respectively.

From the above discussion, the most significant effect of the Stratonovich interpretation for the extrinsic noise is to slow down both mRNA and protein degradation and the activation–inactivation transition of the genes (refer (27)). We see that Stratonovich interpretation results in significant fluctuation when the extrinsic noise is strong. In the weak noise situation, however, the difference between Stratonovich and Itô interpretation is minor.

### Appendix C. Method

It is sometimes misleading if we simulate the gene expression process by solving Eqs. (1)–(3) directly. This is because the conditions for chemical Langevin equation are not satisfied when the gene copy number $n$ is small.

In this paper, we apply a hybrid method to perform the simulation. We assume no fluctuation in the reaction rates $\dot{x}_i$ and $x_i$ in the simulations in the present study. The process of promoter activation–inactivation is simulated by Gillespie’s (1977) algorithm, while the change of the numbers of mRNA and protein during the gap between the changes of promoter activity are simulated by solving the chemical Langevin equation (2)–(3) with constant $X_1$. In particular, we apply the strong order 1.0 Runge–Kutta scheme to solve the chemical Langevin equation (Kloeden and Platen, 1992). Moreover, since the number of molecules is always non-negative, we set $X_i = 0$ whenever $X_i < 0 \quad (i = 2, 3)$ in the simulation.

Thus, our algorithm is

1. Initialize the number of all molecules. Set time $t = 0$.
2. Calculate the next reaction time $\tau$ of the promoter activation–inactivation according to Gillespie’s algorithm.
3. Let $X_i \equiv X_i(t)$, and solve the chemical Langevin equation (2)–(3) from $t$ to $t + \tau$ by the strong order 1.0 Runge–Kutta method.
4. Change the number of $X_i$ according to the occurrence of promoter activation or inactivation. Change $t$ to $t + \mu$.
5. Go to step 2.

To study the effect of colored noise, we replace the extrinsic noise in Eqs. (2)–(3) by the Ornstein–Uhlenbeck processes that is defined by the stochastic differential equation (van Kampen, 1992; Kauffmann et al., 2007; Rosenfeld et al., 2005)

$$
\frac{d\eta}{dt} = -\eta/t + \sqrt{2/\tau} \xi(t),
$$

where $\tau$ is the autocorrelation time and $\xi(t)$ is the white noise. It is easy to see that the processes defined by (47) satisfies $\langle \eta(t) \rangle = 0$ and $\langle \eta(t) \eta(t') \rangle = \exp[-|t_1 - t_2|/\tau]$.

### References


