Switching and Diffusion Models for Gene Regulation Networks

Somkid Intep* Desmond J. Higham † Xuerong Mao ‡

June 3, 2009

Abstract

We analyze a hierarchy of three regimes for modeling gene regulation. The most complete model is a continuous time, discrete state space, Markov jump process. An intermediate 'switch plus diffusion' model takes the form of a stochastic differential equation driven by an independent continuous time Markov switch. In the third 'switch plus ODE' model the switch remains but the diffusion is removed. The latter two models allow for multi-scale simulation where, for the sake of computational efficiency, system components are treated differently according to their abundance. The 'switch plus ODE' regime was proposed by Paszek (Modeling stochasticity in gene regulation: characterization in the terms of the underlying distribution function, Bulletin of Mathematical Biology, 2007), who analyzed the steady state behavior, showing that the mean was preserved but the variance only approximated that of the full model. Here, we show that the tools of stochastic calculus can be used to analyze first and second moments for all time. A technical issue to be addressed is that the state space for the discrete-valued switch is infinite. We show that the new 'switch plus diffusion' regime preserves the biologically relevant measures of mean and variance, whereas the 'switch plus ODE' model uniformly underestimates the variance in the protein level. We also show that, for biologically relevant parameters, the transient behaviour can differ significantly from the steady state, justifying our time-dependent analysis. Extra computational results are also given for a protein dimerization model that is beyond the scope of the current analysis.

*Department of Mathematics, University of Strathclyde, Glasgow, G1 1XH, Scotland, UK (rs.sint@maths.strath.ac.uk). SI was sponsored by Thailand's Commission on Higher Education.

†Department of Mathematics, University of Strathclyde, Glasgow, G1 1XH, Scotland, UK (djh@maths.strath.ac.uk). DJH was supported by Engineering and Physical Sciences Research Council grants GR/S62383/01 and EP/E049370/1.

‡Department of Statistics and Modelling Science, University of Strathclyde, Glasgow, G1 1XH, Scotland, UK (xuerong@stams.strath.ac.uk).

Keywords: diffusion, hybrid model, Gillespie's algorithm, Ito lemma, Markov chain, slow scale simulation, stochastic simulation algorithm, transcription, transition rate, translation.

1 Introduction

Gene regulation is typically modeled using the language of chemical kinetics. At one extreme, discrete-valued stochastic models can be adopted, giving rise to a *Chemical Master Equation* (CME), from which sample paths can be simulated via Gillespie's algorithm [11, 12, 36]. At the other extreme, continuous-valued deterministic modeling leads to a set of ordinary differential equations (ODEs) that are sometimes said to arise through the *law of mass action* [6].

The ODE framework is typically (a) more amenable to analysis [1, 20], (b) cheaper to simulate with [32, 33] and (c) better suited to the important inverse problem of estimating rate constants and comparing models based on sparsely observed data [34]. However, in the case where small numbers of molecules are present, the modeling assumptions that give rise to the mass action ODE are not valid [11, 12, 22] and the discrete/stochastic effects captured by the CME should not be ignored. For example, the stochastic version of a bi-stable ODE model can account for switching between "almost stable" states [31, 35].

Although progress is being made on solving the CME [23] and on optimizing Gillespie's direct simulation method [9, 15], the fully discrete CME setting remains computationally infeasible for most realistic systems. Tau-leaping [5, 14] was introduced in an attempt to speed up stochastic simulation without resorting to a fully deterministic model. This tau-leaping approach can also be used as a means to derive an intermediate stochastic differential equation (SDE) model, known as the *Chemical Langevin Equation* (CLE) [13]. In the more general context of population dynamics this type of *diffusion limit* has also been defined as an approximation to a Markov jump process [22, 28].

It is intuitively appealing, and potentially extremely beneficial, to mix together these modeling regimes so that different species, different reactions or different time periods are treated by simulation methods that are as cheap as possible while preserving the overall accuracy [4, 7]. An interesting example that applies specifically to a simple gene regulation setting was proposed by Paszek [26]. Here, a hybrid model was put forward that uses the CME regime for low copy number species and the ODE framework for relatively abundant species. In this work, which follows on from a simpler context in [18], we exploit the fact that the hybrid model may be regarded as a system of ODEs driven by an independent Markovian switch. The switch has an infinite state space, but we show that existence and uniqueness, and numerical simulation theories carry through. This viewpoint makes it possible to analyze the first and second moments of the model using the tools of stochastic calculus, and to consider an alternative where

the ODE is replaced by a diffusion approximation. Our main findings are that

- the 'switch plus ODE' model uniformly underestimates the variance,
- the steady-state error in the variance for the 'switch plus ODE' model may significantly underestimate the error in the transient,
- replacing the 'switch plus ODE' model with a 'switch plus diffusion' model recovers the correct means and variances, for all time.

Overall, by studying a minimal, but biologically relevant, model that is tractable to analysis, we provide support for the use of the 'switch plus diffusion' regime as a means to incorporate stochasticity in a computationally viable manner.

This work is organized as follows. In the next section we describe the chemical system that models transcription and translation and state the ODEs for the evolution of first and second moments of the CME. We then analyze the hybrid 'switch plus diffusion' and 'switch plus ODE' models. Section 3 considers an alternative model where genes can alternate between active and inactive states. In section 4 we give numerical results for two models involving protein dimerization that lie outside the first order framework analyzed in sections 2 and 3. We give some conclusions in section 5, and in an appendix we collect some of the technical results that are needed to establish the validity of modeling and simulation in the presence of a switch with infinite state space.

We remark that in order to keep the analysis compact, we implicitly assume that initial conditions are deterministic and equal across all modeling regimes.

2 Gene Regulation Model

In order to understand the effect of intrinsic noise in gene regulation, recent authors [27, 30, 36] have modeled the processes of transcription and translation as first order reaction networks involving three species:

M denotes the amount of mRNA, we will also call this X_1,

P denotes the amount of protein, we will also call this X_2,

D denotes the amount of gene, we will also call this X_3.

In particular, Thattai and van Oudenaarden [30] proposed a simple descriptive set of reactions that takes the form

$$D \xrightarrow{k_1} D + M \tag{1}$$

$$M \xrightarrow{k_2} M + P \tag{2}$$

$$M \xrightarrow{k_3} \emptyset \tag{3}$$

$$P \xrightarrow{k_4} \emptyset \tag{4}$$

In words, (1) says that a gene can create a molecule of mRNA with rate constant k_1, without destroying itself. Reaction (2) says that a molecule of mRNA can create a protein with rate constant k_2, without destroying itself. In (3) and (4) a molecule of mRNA, or protein, can degrade with rate constant k_3 or k_4, respectively.

In this model, the amount of gene stays fixed, so X_3 remains constant. We may therefore take the state vector to be

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}.$$

The stoichiometric vectors [17, 36] for the four reactions are

$$\boldsymbol{\nu}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{\nu}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{\nu}_3 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \boldsymbol{\nu}_4 = \begin{bmatrix} 0 \\ -1 \end{bmatrix},$$

with corresponding propensity functions

$$a_1(X) = k_1 X_3, \quad a_2(X) = k_2 X_1, \quad a_3(X) = k_3 X_1, \quad a_4(X) = k_4 X_2.$$

Because X_3 is fixed, we will re-name $k_1 X_3$ as k_1.

Although they are clearly gross simplifications of the underlying biological processes, models such as this have proved useful for characterizing the level of intrinsic noise in a gene regulation network in various parameter regimes, and we note that "noise strength" in this context is typically summarized in terms of the ratio of variance to mean [27, 30].

2.1 Master Equation Moments

In this subsection, we interpret the system (1)–(4) as a Markov jump process defined by the CME, letting $M(t)$ and $P(t)$ denote the stochastic processes that specify the levels of mRNA and protein, respectively. The system fits into the framework of a *first-order reaction network*. More precisely, (1) and (2) involve *catalytic production from a source*, and (3) and (4) are of *degradation* type. Therefore we may use the general result of [8] to obtain a closed system of ODEs that describe the evolution of the first and second moments and correlations. This gives

$$\frac{d}{dt}\mathbb{E}[M(t)] = -k_3\mathbb{E}[M(t)] + k_1, \tag{5}$$

$$\frac{d}{dt}\mathbb{E}[P(t)] = k_2\mathbb{E}[M(t)] - k_4\mathbb{E}[P(t)], \tag{6}$$

$$\frac{d}{dt}\mathbb{E}[P(t)^2] = k_2\mathbb{E}[M(t)] + k_4\mathbb{E}[P(t)] + 2k_2\mathbb{E}[M(t)P(t)] - 2k_4\mathbb{E}[P(t)^2], \tag{7}$$

$$\frac{d}{dt}\mathbb{E}[M(t)^2] = k_1 + (2k_1 + k_3)\mathbb{E}[M(t)] - 2k_3\mathbb{E}[M(t)^2], \tag{8}$$

$$\frac{d}{dt}\mathbb{E}[M(t)P(t)] = k_2\mathbb{E}[M(t)^2] + k_1\mathbb{E}[P(t)] - (k_3 + k_4)\mathbb{E}[M(t)P(t)]. \tag{9}$$

2.2 Hybrid Diffusion Moments

Now we look at a hybrid model based on (1)–(4) where the number of mRNA molecules is modeled as a Markov jump process, as in subsection 2.1, but the evolution of the protein level in (2) and (4) is modeled with the CLE regime. We are motivated by the assumption that the protein is typically more abundant than the mRNA—Paszek [26] adopted this approach, but used an ODE in the protein regime, as discussed in subsection 2.3. This gives rise to an Ito SDE driven by an independent switch, of the form

$$dP^\star(t) = (k_2 r(t) - k_4 P^\star(t))dt + \sqrt{k_2 r(t)}dW_1(t) - \sqrt{k_4 P^\star(t)}dW_2(t). \qquad (10)$$

Here, $r(t)$ denotes the number of mRNA molecules present at time t, when reactions (1) and (3) are interpreted through the CME, and $P^\star(t)$ denotes the number of protein molecules present at time t, when reactions (2) and (4) are interpreted through the CLE. We use $P^\star(t)$ to distinguish this process from the protein level $P(t)$ arising from the full CME regime; this emphasizes that $P(t)$ and $P^\star(t)$ are different stochastic processes; in particular $P(t)$ is discrete-valued and $P^\star(t)$ is continuous-valued. In (10), $W_1(t)$ and $W_2(t)$ are mutually independent Brownian motions that are also independent of $r(t)$.

The switch $r(t)$ can take values in the set of non-negative integers $\{0, 1, 2, 3, \ldots\}$, with no upper limit. We let γ_{ij} denote the transition rate for the switch from state i to j so that, for $i \neq j$,

$$\mathbb{P}\left(r(t+h) = j \mid r(t) = i\right) := \gamma_{ij}h + o(h), \qquad (11)$$

and $\gamma_{ii} := -\sum_{j\neq i}\gamma_{ij}$ is such that

$$\mathbb{P}\left(r(t+h) = i \mid r(t) = i\right) := 1 + \gamma_{ii}h + o(h). \qquad (12)$$

For this switch, the only possible changes of state are increase or decrease by one. The chance of decay is proportional to the current number of molecules, and new molecules are being produced at a rate that is independent of the state. We therefore find that

$$\gamma_{i,i-1} = ik_3, \qquad \gamma_{i,i+1} = k_1, \qquad \gamma_{i,i} = -ik_3 - k_1, \qquad (13)$$

and all other transition rates are zero.

Now, let $\mathcal{L}$ denote the infinitesimal generator of a Markov process, [10, 37]. Then

$$\begin{aligned}\mathcal{L}r(t) &= \lim_{h\to 0}\frac{1}{h}\mathbb{E}\left[r(t+h) - r(t) \mid r(t) = r\right] \\ &= \lim_{h\to 0}\frac{1}{h}\left[\sum_{j\neq r} j(\gamma_{rj}h + o(h)) + r(1 + \gamma_{rr}h + o(h)) - r\right]\end{aligned}$$

$$
\begin{aligned}
&= \lim_{h\to 0}\frac{1}{h}\left[\sum_{j=0}^{\infty} j\gamma_{rj}h + o(h)\right] \\
&= \sum_{j=0}^{\infty} j\gamma_{rj}. \qquad (14)
\end{aligned}
$$

Therefore, by Dynkin's formula [37, Theorem 2.7], using (13) and (14),

$$
\begin{aligned}
dr(t) &= (\mathcal{L}r(t))dt + d(\text{mart.}) \\
&= \left(\sum_{j=0}^{\infty} j\gamma_{rj}\right)dt + d(\text{mart.}) \\
&= ((r-1)\gamma_{r,r-1} + r\gamma_{rr} + (r+1)\gamma_{r,r+1})dt + d(\text{mart.}) \\
&= (k_1 - k_3 r)dt + d(\text{mart.}), \qquad (15)
\end{aligned}
$$

where mart. denotes a martingale whose precise form is not relevant to our analysis.

Applying the generalised Ito lemma [37, Section 2.5], we find we have

$$d(P^\star r) = (k_2 r^2 + k_1 P^\star - (k_3 + k_4)P^\star r)dt + d(\text{mart.}) \qquad (16)$$

and

$$d(P^{\star 2}) = (2k_2 P^\star r - 2k_4 P^{\star 2} + k_2 r + k_4 P^\star)dt + d(\text{mart.}). \qquad (17)$$

So,

$$
\begin{aligned}
\frac{d}{dt}\mathbb{E}[P^\star(t)] &= k_2\mathbb{E}[r(t)] - k_4\mathbb{E}[P^\star(t)], \\
\frac{d}{dt}\mathbb{E}[P^{\star 2}(t)] &= 2k_2\mathbb{E}[P^\star(t)r(t)] - 2k_4\mathbb{E}[P^{\star 2}(t)] + k_2\mathbb{E}[r(t)] + k_4\mathbb{E}[P^\star(t)], \\
\frac{d}{dt}\mathbb{E}[P^\star(t)r(t)] &= k_2\mathbb{E}[r^2(t)] + k_1\mathbb{E}[P^\star(t)] - (k_3+k_4)\mathbb{E}[P^\star(t)r(t)].
\end{aligned}
$$

Since the switch $r(t)$ is identical to $M(t)$ from the full CME, we see from (6), (7) and (9) that this hybrid regime exactly reproduces the first two moments.

2.3 Hybrid ODE Moments

Here we consider the case where, as in subsection 2.2, the number of mRNA molecules is modeled a Markov jump process, but now the evolution of the protein level is modeled with the law of mass action. This regime was introduced and studied by Paszek [26]. We have an ODE driven by an independent switch, of the form

$$d\widehat{P}(t) = (k_2 r(t) - k_4\widehat{P}(t))dt, \qquad (18)$$

where, as in subsection 2.2, $r(t)$ denotes the number of mRNA molecules when (1) and (3) are modeled through the CME. We use $\widehat{P}(t)$ to denote the continuous-valued stochastic process that represents the protein level.

Instead of (16) and (17), we now have

$$d(\widehat{P}r) = (k_2 r^2 + k_1\widehat{P} - (k_3+k_4)\widehat{P}r)dt + d(\text{mart})$$

and

$$d(\widehat{P}^2) = (2k_2\widehat{P}r - 2k_4\widehat{P}^2)dt + d(\text{mart}).$$

So,

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}[\widehat{P}(t)] &= k_2\mathbb{E}[r(t)] - k_4\mathbb{E}[\widehat{P}(t)], && (19)\\
\frac{d}{dt}\mathbb{E}[\widehat{P}^2(t)] &= 2k_2\mathbb{E}[\widehat{P}(t)r(t)] - 2k_4\mathbb{E}[\widehat{P}^2(t)], && (20)\\
\frac{d}{dt}\mathbb{E}[\widehat{P}(t)r(t)] &= k_2\mathbb{E}[r(t)^2] + k_1\mathbb{E}[\widehat{P}(t)] - (k_3+k_4)\mathbb{E}[\widehat{P}(t)r(t)]. && (21)
\end{aligned}$$

Comparing these ODEs to (6), (7) and (9), and recalling that $r(t)$ is identical to $M(t)$, we see that this hybrid model matches the means and correlation of the full CME, but does not reproduce the correct second moment.

In the remainder of this section we analyze the discrepancy between the second moments in the CME and hybrid 'switch plus ODE' modes. First, we show that the error is always one-sided.

Theorem 2.1 *For the system (1)–(4), the variances for the protein arising from the CME and the hybrid model (18),* $\mathsf{var}[P(t)]$ *and* $\mathsf{var}[\widehat{P}(t)]$*, satisfy* $\mathsf{var}[\widehat{P}(t)] \leq \mathsf{var}[P(t)]$ *for all time, independently of the rate constants and initial conditions.*

Proof Letting $y(t) := \mathsf{var}[P(t)] - \mathsf{var}[\widehat{P}(t)]$, because the means match we have $y(t) = \mathbb{E}[P^2(t)] - \mathbb{E}[\widehat{P}^2(t)]$. We then see from (7) and (20) that

$$\frac{dy(t)}{dt} = k_2\mathbb{E}[M(t)] + k_4\mathbb{E}[P(t)] - 2k_4 y(t). \qquad (22)$$

Now, by construction, the CME does not allow molecules to become negative, so $h(t) := k_2\mathbb{E}[M(t)] + k_4\mathbb{E}[P(t)] \geq 0$. Using an integrating factor in (22) we find that

$$y(t) = e^{-2k_4 t}\int_0^t e^{2k_4 s}h(s)ds,$$

and the result follows.

To obtain a precise expression for the error in the variance, we may first solve for $\mathbb{E}[M(t)]$ in (5) and then for $\mathbb{E}[P(t)]$ in (6). Substituting in (22) then gives

$$\begin{aligned}\mathsf{var}[P(t)] - \mathsf{var}[\widehat{P}(t)] &= \frac{k_1 k_2}{k_3 k_4}\left(1 - e^{-k_4 t}\right) \\ &\quad + \left(\mathbb{E}[M(0)] - \frac{k_1}{k_3}\right)\frac{k_2}{k_4 - k_3}\left(e^{-k_3 t} - e^{-k_4 t}\right) \\ &\quad + \mathbb{E}[P(0)]\left(e^{-k_4 t} - e^{-2k_4 t}\right), \end{aligned} \tag{23}$$

when $k_3 \neq k_4$, and

$$\begin{aligned}\mathsf{var}[P(t)] - \mathsf{var}[\widehat{P}(t)] &= \frac{k_1 k_2}{k_3 k_4}\left(1 - e^{-k_4 t}\right) \\ &\quad + \left(\mathbb{E}[M(0)] - \frac{k_1}{k_3}\right) k_2 t e^{-k_4 t} \\ &\quad + \mathbb{E}[P(0)]\left(e^{-k_4 t} - e^{-2k_4 t}\right), \end{aligned} \tag{24}$$

when $k_3 = k_4$.

We note from (23) and (24) that $\lim_{t\to\infty} \mathsf{var}[P(t)] - \mathsf{var}[\widehat{P}(t)] = k_1 k_2/(k_3 k_4)$, in agreement with the steady state analysis in [26].

To interpret the expressions (23) and (24) further, we focus on the case where the initial conditions satisfy $\mathbb{E}[M(0)] = k_1/k_3$ and $\mathbb{E}[P(0)] > k_1 k_2/(k_3 k_4)$. The error in the variance then simplifies to

$$\mathsf{var}[P(t)] - \mathsf{var}[\widehat{P}(t)] = \frac{k_1 k_2}{k_3 k_4}\left(1 - e^{-k_4 t}\right) + \mathbb{E}[P(0)]\left(e^{-k_4 t} - e^{-2k_4 t}\right).$$

This expression has a unique maximum at time

$$t^\star := \frac{1}{k_4}\log\left(\frac{2k_3 k_4 \mathbb{E}[P(0)]}{k_3 k_4 \mathbb{E}[P(0)] - k_1 k_2}\right)$$

and the ratio of the maximum transient error to the steady state error is given by

$$\frac{\mathsf{var}[P(t^\star)] - \mathsf{var}[\widehat{P}(t^\star)]}{\lim_{t\to\infty} \mathsf{var}[P(t)] - \mathsf{var}[\widehat{P}(t)]} = \frac{1}{2} + \frac{k_3 k_4 \mathbb{E}[P(0)]}{4k_1 k_2} + \frac{k_1 k_2}{4k_3 k_4 \mathbb{E}[P(0)]}. \tag{25}$$

We see from (25) that the transient error in the variance can exceed the steady state error when $\mathbb{E}[P(0)]$ is large. In Figure 1, using biologically valid rate constants from [29], which are $k_1 = 0.3, k_2 = 0.1734, k_3 = 0.0115$ and $k_4 = 6.42 \times 10^{-5}$, we show how the error in the variance evolves when $\mathbb{E}[M(0)] = k_1/k_3$ and $\mathbb{E}[P(0)] = 4k_1 k_2/(k_3 k_4)$. Here the right hand side of (25) is $25/16 \approx 3/2$, and we see that the maximum temporal error is about 50% above the steady state value. We also show the case where $\mathbb{E}[M(0)] = 2$ and $\mathbb{E}[P(0)] = 4$, for

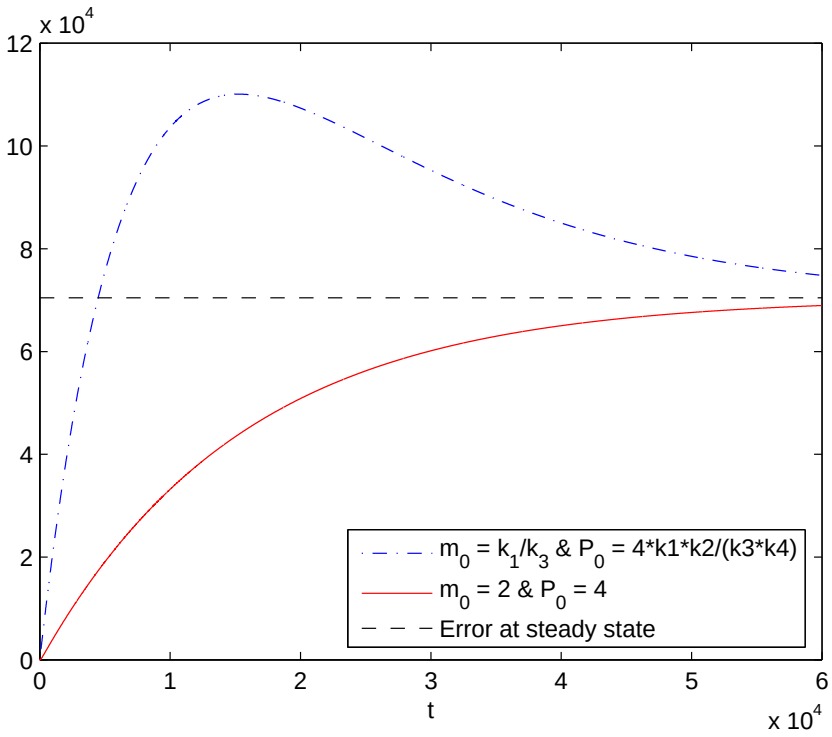

Figure 1: Modeling error in the protein variance for the 'switch plus ODE' hybrid (18), using rate constants from [29].

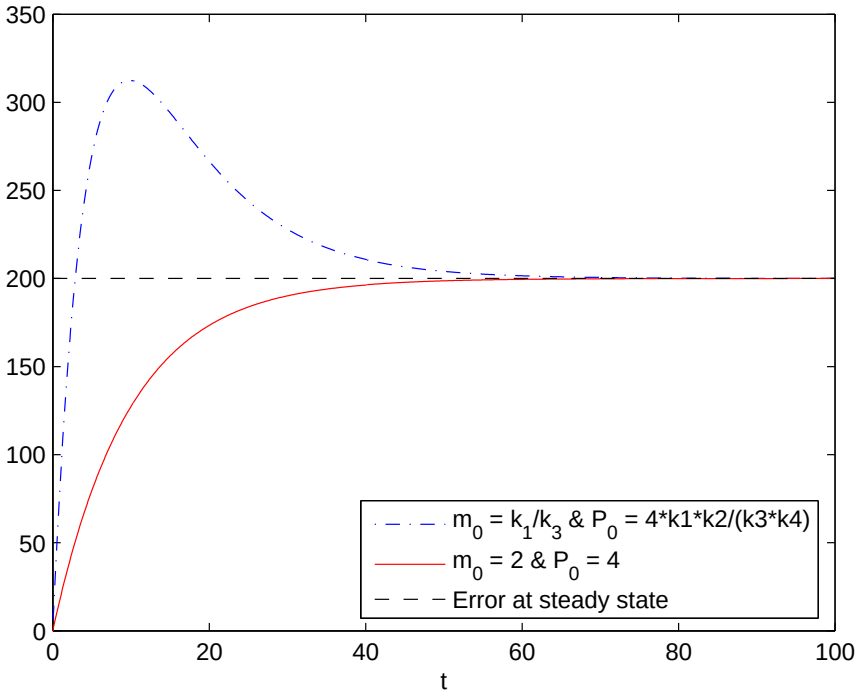

Figure 2: Modeling error in the protein variance for the 'switch plus ODE' hybrid (18), using rate constants from [27].

which it can be shown that the steady sate value is an upper bound for the error. Figure 2 shows similar behaviour for rate constants appearing in [27], which are $k_1 = 10, k_2 = 10, k_3 = 5$ and $k_4 = 0.1$.

We conclude this section with the results of some numerical experiments to demonstrate numerically that an Euler–Maruyama based method can successfully integrate the 'switch plus CLE' model. For the same parameters as Figure 1, we show in Figure 3 the absolute error in the sample mean for $P^\star(T)^2$, at $T = 5$, arising from the numerical method outlined in the Appendix, for $\Delta t = 2^{-4}$, 2^{-5}, 2^{-6} and 2^{-7}. The time interval $[0, 5]$ is different from that in Figure 1 because we are now interested in finite-time convergence of a numerical method and wish to observe asymptotic, small-stepsize, behavior. We used 10^7 sample paths and all 95% confidence intervals, shown as vertical lines, were less than 0.055. The errors are plotted on a log-log scale, and we see that the results are consistent with a weak order of 1. A least squares fit gave an error behaviour of $\propto \Delta t^{1.1}$ with residual of 0.08. Similarly, we show in Figure 4 the second moment of the error in $P^\star(T)^2$ for $\Delta t = 32 \times 2^{-10}$, 16×2^{-10}, 8×2^{-10} and 4×2^{-10}. Here, we used 10^4 sample paths and all 95% confidence intervals, shown as vertical lines, were less than 0.04. The errors are plotted on a log-log scale, and we see that the results are consistent with a strong order of $\frac{1}{2}$; that is, mean-square of order 1. A least squares fit gave a mean-square error behaviour of $\propto \Delta t^{1.3}$ with residual of 0.03.

3 A Related Active/Inactive Gene Model

Raser and O'Shea [27] extended the system (1)–(4) to the case where genes may alternate between an inactive state, where no mRNA is produced, and an active state. If there are m genes in total, and we let $D_i^\star$ denote the active state of the ith gene, this system may be written

$$\left.\begin{array}{l} D_i \xrightarrow{k_a} D_i^\star \\ D_i \xleftarrow{k_b} D_i^\star \\ D_i^\star \xrightarrow{k_1} D_i^\star + M \end{array}\right\} \quad 1 \leq i \leq m \tag{26}$$

and

$$M \xrightarrow{k_2} M + P \tag{27}$$

$$M \xrightarrow{k_3} \emptyset \tag{28}$$

$$P \xrightarrow{k_4} \emptyset \tag{29}$$

Here, the initial condition for the ith gene must be either $D_i(0) = 0$ and $D_i^\star(0) = 1$ (active) or $D_i(0) = 1$ and $D_i^\star(0) = 0$ (inactive), and $D_i(t) + D_i^\star(t) \equiv 1$ for all time.

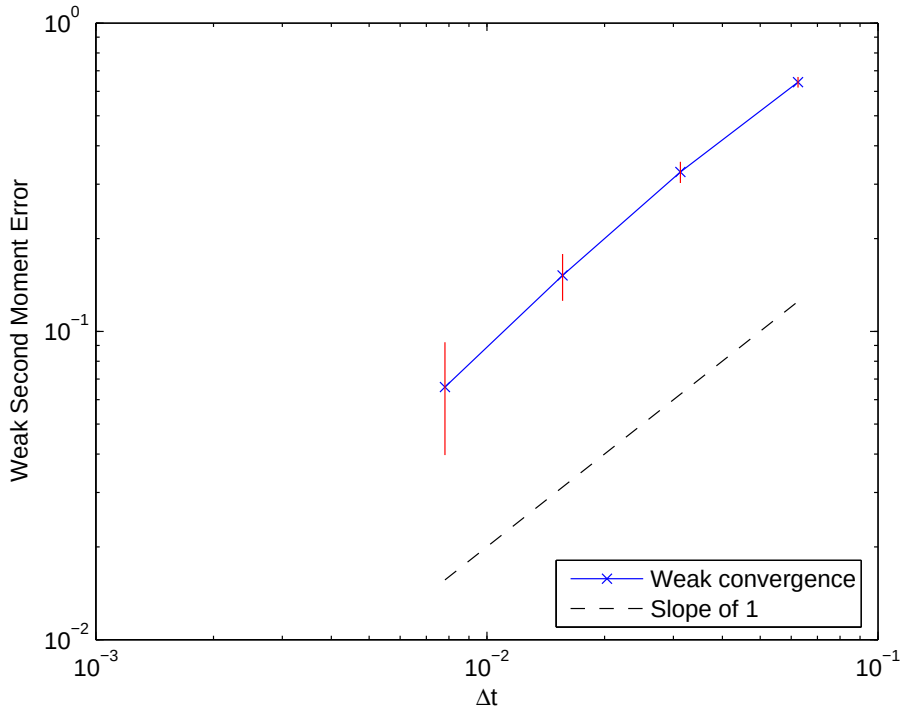

Figure 3: Weak convergence in the 'switch plus CLE' framework using rate constants from [29]. Vertical axis measures the error $|\mathbb{E}\left[P^\star(T)^2\right] - \mathbb{E}\left[\widehat{P^\star}(T)^2\right]|$, for $T = 5$, where $P^\star(t)$ in (10) denotes the protein level and $\widehat{P^\star}(t)$ is the numerical approximation with the method described in the appendix. The quantity $\mathbb{E}\left[\widehat{P^\star}(T)^2\right]$ is evaluated via Monte Carlo, and 95% confidence intervals are shown as vertical lines.

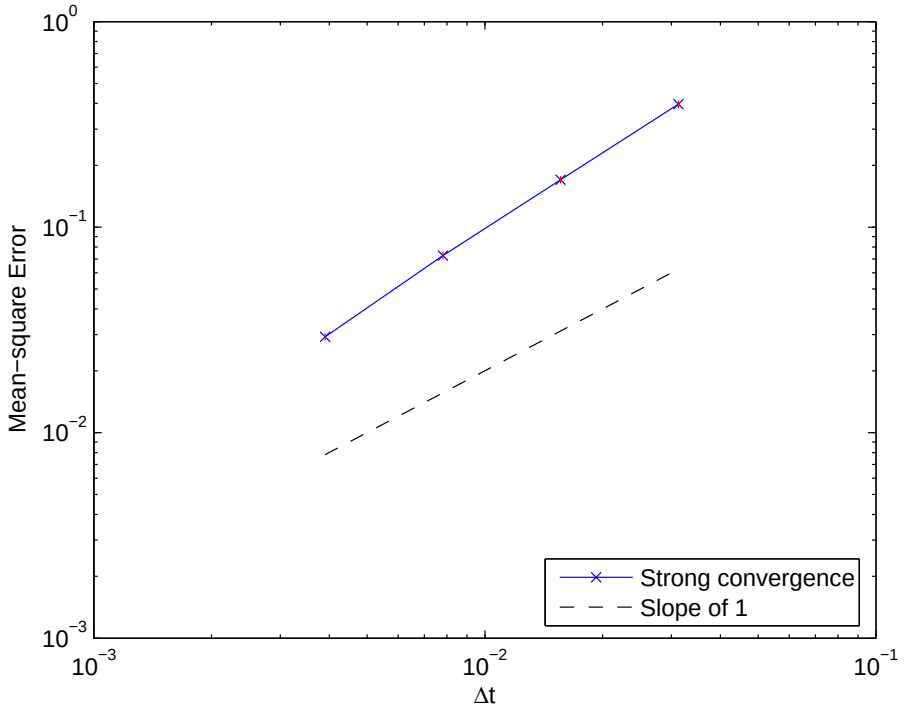

Figure 4: As for Figure 3 except that the strong error $\mathbb{E}\left[\left(P^\star(T)^2 - \widehat{P^\star}(T)^2\right)^2\right]$ is measured. Sample means are shown for 10^4 paths and 95% confidence intervals are negligible.

Paszek [26] considered a hybrid model with the number of active genes forming a discrete-valued stochastic process in the CME regime, and with the levels of mRNA and protein taking real values. He chose mass action ODEs for the reactions involving mRNA and protein, and, as for the simpler system (1)–(4), found that this 'switch plus ODE' hybrid gave a steady-sate variance that does not match the underlying CME. Higham and Khanin [21] showed that a hybrid 'switch plus diffusion' model, where reactions involving mRNA and protein are treated with the CLE approach, reproduces the exact first and second moments for all time. Although the active/inactive model is in a sense more complex than the model in section 2, we emphasize that the number of active genes forms a switch with a finite state space, and hence it is possible to appeal to standard work, such as [24] for existence, uniqueness and simulation theory, and stochastic calculus tools. Our main aim here is to point out that the uniform underestimation of the variance that we established in Theorem 2.1 also applies in this case.

Following [21], if we let $\widehat{M}(t)$ and $\widehat{P}(t)$ denote the mRNA and protein levels arising from the 'switch plus ODE' model, then $\mathbb{E}[M(t)] = \mathbb{E}[\widehat{M}(t)]$, $\mathbb{E}[P(t)] = \mathbb{E}[\widehat{P}(t)]$, $\mathbb{E}[P(t)M(t)] = \mathbb{E}[\widehat{P}(t)\widehat{M}(t)]$, and the discrepancy in the second moments

$$y(t) := \begin{bmatrix} \mathbb{E}[M^2(t)] - \mathbb{E}[\hat{M}^2(t)] \\ \mathbb{E}[P^2(t)] - \mathbb{E}[\hat{P}^2(t)] \end{bmatrix}$$

satisfies

$$\frac{d}{dt}y(t) = -Ay(t) + g(t),$$

where

$$A = \begin{bmatrix} \gamma_r & 0 \\ 0 & \gamma_p \end{bmatrix} \quad \text{and} \quad g(t) = \begin{bmatrix} k_r\mathbb{E}[r] + \gamma_r\mathbb{E}[M] \\ k_p\mathbb{E}[M] + \gamma_p\mathbb{E}[P] \end{bmatrix}.$$

It follows that

$$y(t) = e^{-At}\int_0^t e^{As}g(s)ds.$$

Since $g(t) \geq 0$ for all $t \geq 0$, we conclude that this hybrid model underestimates the true mRNA and protein variances for all time.

We also note that when the reversible reactions $D_i \to D_i^\star$ and $D_i^\star \to D_i$ in (26) are fast compared with the other reactions in the system; that is, both $k_a \gg 1$ and $k_b \gg 1$, with all other rate constants of $O(1)$, then we may introduce a slow-fast decoupling along the lines of [4]. Here, we replace $D_i^\star(t)$ by its steady state in the D_i-$D_i^\star$ subsystem, which effectively reduces (26)–(29) to the fixed-gene system (1)–(4) with the amount of gene equal to $D = D^\star(0)k_a/(k_a + k_b)$. Paszek [26] refers to this as a thermodynamic limit for the full model. Analysis along the lines of that developed above can be used to show that this type of modeling approximation does not have a one-sided effect on the variance; the reduced model may produce a larger or smaller variance depending on the parameter regimes, and the error may change sign over time.

4 Tests with a Second Order Reaction

The results in the previous two sections rely on the first order nature of the reactions. In this section we give some brief numerical evidence that the ideas are relevant more generally, when the first two moments do not form a closed system of ODEs. To do this, we add a protein dimerization stage to the simple gene regulation models.

For the Thattai and van Oudenaarden model (1)–(4), we add the three reactions

$$P + P \quad \stackrel{k_{p2}}{\rightarrow} \quad P_2 \tag{30}$$

$$P_2 \quad \stackrel{k_{-p2}}{\rightarrow} \quad P + P \tag{31}$$

$$P_2 \quad \stackrel{\gamma_{p2}}{\rightarrow} \quad \emptyset. \tag{32}$$

Here, in (30) two protein molecules combine to form a dimer, P_2, and in (31) the process is reversed. In (32) the dimer decays. We note that it has been argued that a difference between the momoner and dimer decay rates can explain the phenomenon of "cooperative stability", which makes a larger spread of protein levels available in vivo [2]. We chose rate constants $k_1 = 0.3$, $k_2 = 0.17$, $k_3 = 0.012$ from [29], $k_4 = 0.0007$, $k_{p2} = 0.025$, $k_{-p2} = 0.5$ from [3], and $\gamma_{p2} = 0.00023$ from [2]. Initial conditions were set to $D(0) = 4, M(0) = 2, P(0) = 4$ and $P_2(0) = 4$, and we record the levels at time $T = 20$.

For the system given by (1)–(4) and (30)–(32), we compared the CME (via Gillespie's algorithm) with the full CLE, 'switch plus diffusion' and 'switch plus ODE' regimes, using an Euler method with stepsize of 0.004. (Comparable results were obtained with a larger stepsize.) Table 1 summarizes the results.

Expected values are estimated with Monte Carlo simulation over 10^5 paths, and approximate 95% confidence intervals are given for each sample mean. In addition to moments and variances for the protein and dimer, we also show their noise strength, $\mathsf{ns}[P]$ and $\mathsf{ns}[P_2]$, respectively, defined as the ratio of variance to mean.

We see from Table 1 that the CME, CLE and 'switch plus diffusion' regimes give comparable results for moments and noise strengths, whereas the 'switch plus ODE' regimes significantly underestimates the variance and noise strength for the protein and dimer.

Table 2 shows the results of an analogous experiment where the Raser and O'Shea system (26)–(29) was augmented with the dimerization reactions (30)–(32). We used $k_a = 0.1$ and $k_b = 0.1$ from [27], $k_1 = 0.3$, $k_2 = 0.17$, $k_3 = 0.012$ from [29], $k_4 = 0.0007$, $k_{p2} = 0.025$, $k_{-p2} = 0.5$ from [3] and $\gamma_{p2} = 0.00023$ from [2]. We see that the conclusions from Table 1 continue to hold.

	CME	CLE	CLE Switch	ODE Switch
$\mathbb{E}[P]$	$[26.23, 26.30]$	$[26.22, 26.28]$	$[26.22, 26.29]$	$[26.54, 26.57]$
$\mathbb{E}[P^2]$	$[717.87, 721.55]$	$[717.28, 720.97]$	$[717.31, 721.00]$	$[712.22, 714.11]$
$\mathbb{E}[P_2]$	$[14.58, 14.63]$	$[14.56, 14.62]$	$[14.55, 14.61]$	$[14.42, 14.46]$
$\mathbb{E}[P_2^2]$	$[231.85, 233.59]$	$[231.20, 232.89]$	$[231.19, 232.91]$	$[217.79, 218.98]$
$\mathsf{var}[P]$	$[29.60, 30.13]$	$[29.84, 30.39]$	$[29.68, 30.22]$	$[7.97, 8.11]$
$\mathsf{ns}[P]$	$[1.125, 1.149]$	$[1.136, 1.159]$	$[1.129, 1.152]$	$[0.300, 0.305]$
$\mathsf{var}[P_2]$	$[19.28, 19.63]$	$[19.06, 19.40]$	$[19.30, 19.66]$	$[9.80, 9.98]$
$\mathsf{ns}[P_2]$	$[1.317, 1.347]$	$[1.304, 1.332]$	$[1.321, 1.351]$	$[0.678, 0.692]$

Table 1: 95% confidence intervals for Monte Carlo sample mean approximations to the first and second moments, variance and noise strength in the CME, CLE, 'switch plus diffusion' and 'switch plus ODE' formulations for (1)–(4) and (30)–(32). Average number of switches per path was 27.

	CME	CLE	CLE Switch	ODE Switch
$\mathbb{E}[P]$	$[19.65, 19.71]$	$[19.67, 19.74]$	$[19.64, 19.70]$	$[20.04, 20.06]$
$\mathbb{E}[P^2]$	$[411.30, 413.91]$	$[412.53, 415.14]$	$[411.15, 413.75]$	$[405.42, 406.41]$
$\mathbb{E}[P_2]$	$[8.47, 8.51]$	$[8.47, 8.52]$	$[8.46, 8.50]$	$[8.31, 8.33]$
$\mathbb{E}[P_2^2]$	$[83.95, 84.81]$	$[84.14, 84.99]$	$[83.62, 84.47]$	$[71.54, 71.88]$
$\mathsf{var}[P]$	$[25.10, 25.56]$	$[25.36, 25.82]$	$[25.20, 25.66]$	$[3.98, 4.05]$
$\mathsf{ns}[P]$	$[1.273, 1.301]$	$[1.285, 1.312]$	$[1.279, 1.306]$	$[0.198, 0.202]$
$\mathsf{var}[P_2]$	$[12.16, 12.40]$	$[12.25, 12.49]$	$[12.04, 12.28]$	$[2.52, 2.56]$
$\mathsf{ns}[P_2]$	$[1.428, 1.464]$	$[1.438, 1.474]$	$[1.417, 1.452]$	$[0.302, 0.308]$

Table 2: 95% confidence intervals for Monte Carlo sample mean approximations to the first and second moments, variance and noise strength in the CME, CLE, 'switch plus diffusion' and 'switch plus ODE' formulations for (26)–(29) and (30)–(32). Average number of switches per path was 8.

5 Discussion and Conclusions

The diffusion approximation to a Markov jump process is useful both analytically and computationally. We have shown here that simple multi-scale diffusion/jump models in gene regulation have advantages over their ODE/jump counterparts. There are many interesting open questions in this area, including:

- How general is the phenomenon shown here that replacing a Langevin component with the reaction rate ODE causes the overall variance to be underestimated?
- Is it possible to develop a theory for state-dependent Markov switches, which arise, for example, when gene activity is regulated by proteins that are created downstream [16]?
- Is there a general existence/uniqueness/numerical convergence theory for diffusion coefficients that involve the square root function?

A Appendix: Theory and Simulation for Infinite State Space Switch

A.1 Set-up

Stochastic differential equations (SDEs) driven by switches are becoming more common as models in science and engineering. A switch typically takes a finite number of possible values, but in this work we need to consider a countably infinite state space, enumerated by the non-negative integers. This requires us to extend the theory for existence, uniqueness and numerical simulation that can be found, for example, in [24], from finite to countably infinite state spaces. We begin by setting up our notation and problem formulation.

Let $r(t)$ be a right-continuous Markov chain on a complete probability space taking values in an infinite state space $\bar{\mathbb{S}} = \{0, 1, 2, \ldots\}$ with generator $\Gamma = (\gamma_{ij})_{i,j\in\bar{\mathbb{S}}}$ given by

$$\mathbb{P}\{r(t+\Delta) = j | r(t) = i\} = \begin{cases} \gamma_{ij}\Delta + o(\Delta) & \text{if } i \neq j, \\ 1 + \gamma_{ii}\Delta + o(\Delta) & \text{if } i = j, \end{cases}$$

where $\gamma_{ij} \geq 0$ is the transition rate from state i to j if $i \neq j$ and

$$\gamma_{ii} = -\sum_{j\neq i} \gamma_{ij}.$$

We assume that the transition rate γ_{ij} satisfies the following condition

$$\max_{i\in\bar{\mathbb{S}}} |\gamma_{ii}| < \infty.$$

Now, consider an autonomous SDE with Markovian switching of the form

$$dx(t) = f(x(t), r(t))dt + g(x(t), r(t))dW(t), \quad 0 \le t \le T, \tag{33}$$

with initial data $x(0) = x_0 \in \mathcal{L}^2_{\mathcal{F}_{t_0}}(\Omega; \mathbb{R}^n)$ and $r(0) = r_0$, where r_0 is an $\bar{\mathbb{S}}$-valued $\mathcal{F}_0$-measurable random variable and

$$f : \mathbb{R}^n \times \bar{\mathbb{S}} \to \mathbb{R}^n \quad \text{and} \quad g : \mathbb{R}^n \times \bar{\mathbb{S}} \to \mathbb{R}^{n \times m}.$$

Here $W(t)$ is an m-dimensional Brownian motion that is independent of the Markov chain.

A.2 Existence and Uniqueness

We begin with an existence, uniqueness and moment bound result, based on the finite state treatment in [24]. We make the traditional global Lipschitz assumptions on the coefficients. For the case where the diffusion coefficients arise through the Chemical Langevin regime, these results apply only up to a stopping time—so that excursions taking population sizes close to zero can be avoided. Deriving more general results that apply directly to non-globally Lipschitz problems is currently an active area [19, 25].

Theorem A.1 *Assume that f and g satisfy a global Lipschitz condition; that is, there exists a positive constant K such that*

$$|f(x,i) - f(y,i)| \vee |g(x,i) - g(y,i)| \le K|x - y| \tag{34}$$

for all $x, y \in \mathbb{R}^n$ and $i \in \bar{\mathbb{S}}$.
Then there exists a unique solution $x(t)$ to equation (33) and, moreover,

$$\mathbb{E}\left(\sup_{0 \le t \le T} |x(t)|^2\right) \le (1 + 3\mathbb{E}|x_0|^2)e^{3KT(T+4)}, \tag{35}$$

so the solution belongs to $\mathcal{M}^2([0,T]; \mathbb{R}^n)$.

Note: $\mathcal{M}^p([a,b]; \mathbb{R}^n)$ means the family of processes $\{f(t)\}_{a \le t \le b}$ in $\mathcal{L}^p([a,b]; \mathbb{R}^n)$ such that $\mathbb{E}\int_a^b |f(t)|^p dt < \infty$; while $\mathcal{L}^p([a,b]; \mathbb{R}^n)$ means the family of $\mathbb{R}^n$-valued $\mathcal{F}_t$-adapted processes $\{f(t)\}_{a \le t \le b}$ such that $\int_a^b |f(t)|^p dt < \infty$ a.s. We also use the notation $[[a,b]]$ to denote a stochastic closed interval, where a or b may be random variables; [24, page 14].

Proof Since almost every sample path of $r(\cdot)$ is a right-continuous step function, there is a sequence $\{\tau_k\}_{k \ge 0}$ of stopping times such that $t_0 = \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_k < \cdots$ and $r(t) = r(\tau_k)$ for $t \in [\tau_k, \tau_{k+1})$.

First, we consider the equation (33) on the interval $t \in [[\tau_0, \tau_1]]$; that is,

$$dx(t) = f(x(t), r_0)dt + g(x(t), r_0)dW(t), \tag{36}$$

with initial data $x(t_0) = x_0$ and $r(t_0) = r_0$. Now, the equation (36) is an SDE without Markovian switching. So, by Mao and Yuan [24, Theorem 3.8], the equation (33) has a unique solution which belongs to $\mathcal{M}^2([[\tau_0, \tau_1]]; \mathbb{R}^n)$. In particular, $x(\tau_1) \in L^2_{\mathcal{F}_{\tau_1}}(\Omega; \mathbb{R})$. After that, we consider the equation (33) on the interval $t \in [[\tau_1, \tau_2]]$ which becomes

$$dx(t) = f(x(t), r(\tau_1))dt + g(x(t), r(\tau_1))dW(t), \tag{37}$$

with initial data $x(\tau_1)$ and $r(\tau_1)$. Again by [24, Theorem 3.8], the equation (33) has a unique solution which belongs to $\mathcal{M}^2([[\tau_1, \tau_2]]; \mathbb{R}^n)$. By repeating this procedure we can see that the equation (33) has a unique solution $x(t)$ on $[0, T]$. Finally, the bound (35) follows by arguing in the same way as [24, Lemma 3.1].

A.3 Numerical Simulation

The natural Euler–Maruyama (EM) method for simulating the switching SDE (33) takes the form

$$X_{k+1} = X_k + f(X_k, r_k^\Delta)\Delta + g(X_k, r_k^\Delta)\Delta W_k. \tag{38}$$

Here, $\Delta > 0$ is a fixed stepsize, X_k is the approximation to $X(t_k)$, with $t_k = k\Delta$, $r_k^\Delta = r(k\Delta)$, $\Delta W_k = W(t_{k+1}) - W(t_k)$ and the initial conditions for the iteration are $X_0 = x_0$ and $r_0^\Delta = r_0$.

For the purpose of analysis, it is convenient to work with a continuous time approximation, $X(t)$, that is defined as

$$X(t) = X_0 + \int_0^t f(\bar{X}(s), \bar{r}(s))ds + \int_0^t g(\bar{X}(s), \bar{r}(s))dW(s), \tag{39}$$

where the 'step processes' $\bar{X}(t)$ and $\bar{r}(t)$ take the form

$$\bar{X}(t) = X_k, \quad \bar{r}(t) = r_k^\Delta \quad \text{for} \quad t \in [t_k, t_{k+1}). \tag{40}$$

Note that $X(t_k) = \bar{X}(t_k) = X_k$, so that $X(t)$ and $\bar{X}(t)$ coincide with the discrete numerical solution at the gridpoints t_k.

The following general moment bounds hold for both the exact and numerical solutions.

Lemma A.1 *Assume that f and g satisfy the linear growth condition; that is, there exists a constant $\bar{K} > 0$ such that*

$$|f(x, i)| \vee |g(x, i)| \quad \leq \quad \bar{K}(1 + |x|) \quad \forall (x, i) \in \mathbb{R}^n \times \bar{\mathbb{S}}. \tag{41}$$

Then for any $p \geq 2$ there is a constant H, which is dependent on only $p, T, \bar{K}, x_0$ but independent of Δ, such that the exact solution $x(t)$ in (33) and the EM approximate solution $X(t)$ in (39) have the property that

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|x(t)|^p\right] \vee \mathbb{E}\left[\sup_{0\leq t\leq T}|X(t)|^p\right] \quad \leq \quad H.$$

Proof Proving this lemma, we can follow the proof in [24, Lemma 4.1].

This result then allows us to establish a strong convergence result for the numerical method.

Theorem A.2 *Assume that f and g satisfy the global Lipschitz condition (34). Then,*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|X(t)-x(t)|^2\right] \leq C\Delta, \tag{42}$$

where C is a positive constant independent of Δ.

Proof It is easy to see that the global Lipschitz condition (34) implies the linear growth condition (41), so that Lemma A.1 applies. Following the proof in [24, Theorem 4.1] and using Lemma A.1, the required assertion follows.

References

[1] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits*, Chapman & Hall, London, 2006.

[2] N. E. Buchler, U. Gerland, and T. Hwa, *Nonlinear protein degradation and the function of genetic circuits*, Proc Natl Acad Sci U S A, 102 (2005), pp. 9559–64.

[3] R. Bundschuh, F. Hayot, and C. Jayaprakash, *The role of dimerization in noise reduction of simple genetic networks*, J Theor Biol, 220 (2003), pp. 261–269.

[4] Y. Cao, D. T. Gillespie, and L. Petzold, *The slow-scale stochastic simulation algorithm*, J. Chem. Phys., 122 (2005), p. 014116.

[5] ——, *Efficient stepsize selection for the Tau-leaping method*, J. Chem. Phys., 124 (2006), p. 044109.

[6] H. de Jong, *Modeling and simulation of genetic regulatory systems: A literature review*, Journal of Computational Biology, 9 (2002), pp. 69–105.

[7] W. E, D. Liu, and E. Vanden-Eijnden, *Nested stochastic simulation algorithms for chemical kinetic systems with multiple time scales*, J. Chem. Phys., 123 (2005), p. 194107.

[8] C. Gadgil, C. H. Lee, and H. G. Othmer, *A stochastic analysis of first-order reaction networks*, Bulletin of Mathematical Biology, 67 (2005), pp. 901–946.

[9] M. Gibson and J. Bruck, *Efficient exact stochastic simulation of chemical systems with many species and many channels*, J. Phys. Chem. A, 104 (2000), pp. 1876–1889.

[10] Ĭ. Ī. Gīhman and A. V. Skorohod, *Stochastic Differential Equations*, Springer-Verlag, New York, 1972. Translated from the Russian by Kenneth Wickwire, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 72.

[11] D. T. Gillespie, *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions*, J. Comp. Phys., 22 (1976), pp. 403—434.

[12] ——, *Exact stochastic simulation of coupled chemical reactions*, J. Phys. Chem., 81 (1977), pp. 2340–2361.

[13] ——, *The chemical Langevin equation*, J. Chem. Phys., 113 (2000), pp. 297–306.

[14] ——, *Approximate accelerated stochastic simulation of chemically reacting systems*, J. Chem. Phys., 115 (2001), pp. 1716–1733.

[15] D. T. Gillespie, *Stochastic simulation of chemical kinetics*, Annual Review of Physical Chemistry, 58 (2007), pp. 35–55.

[16] N. J. Guido, X. Wang, D. Adalsteinsson, D. McMillen, J. Hasty, C. R. Cantor, T. C. Elston, and J. J. Collins, *A bottom-up approach to gene regulation*, Nature, 439 (2006), pp. 856–860.

[17] D. J. Higham, *Modeling and simulating chemical reactions*, SIAM Review, 50 (2008), pp. 347–368.

[18] D. J. Higham and R. Khanin, *Chemical master versus chemical Langevin for first-order reaction networks*, Open Applied Mathematics Journal, 2 (2008), pp. 59–79.

[19] D. J. Higham, X. Mao, and A. Stuart, *Strong convergence of Euler-type methods for nonlinear stochastic differential equations*, SIAM J. Num Anal., 40 (2002), pp. 1041–1063.

[20] P. J. Ingram, M. P. Stumpf, and J. Stark, *Network motifs: structure does not determine function*, BMC Genomics, 7 (2006), p. 108.

[21] R. Khanin and D. J. Higham, *Chemical Master Equation and Langevin regimes for a gene transcription model*, Theoretical Computer Science, 408 (2008), pp. 31–40.

[22] T. G. Kurtz, *Approximation of Population Processes*, SIAM, 1981.

[23] S. MacNamara, K. Burrage, and R. B. Sidje, *Multiscale modeling of chemical kinetics via the master equation*, Multiscale Model. Simul., 6 (2008), pp. 1146–1168.

[24] X. Mao and C. Yuan, *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, London, 2006.

[25] X. Mao, C. Yuan, and G. Yin, *Approximations of Euler-Maruyama type for stochastic differential equations with Markovian switching, under non-Lipschitz conditions*, Journal of Computational and Applied Mathematics, 205 (2007), pp. 936–948.

[26] P. Paszek, *Modeling stochasticity in gene regulation: characterization in the terms of the underlying distribution function*, Bulletin of Mathematical Biology, 69 (2007), pp. 1567–601.

[27] J. Raser and E. O'Shea, *Control of stochasticity in eukaryotic gene expression*, Science, 304 (2004), pp. 1811–4.

[28] E. Renshaw, *Modelling Biological Populations in Space and Time*, Cambridge University Press, 1991.

[29] P. S. Swain, M. Elowitz, and E. Siggia, *Intrinsic and extrinsic contributions to stochasticity in gene expression*, Proc. Natl. Acad. Sci., 99 (2002), pp. 12795–800.

[30] M. Thattai and A. van Oudenaarden, *Intrinsic noise in gene regulatory networks*, Proc. Natl. Acad. Sci., 98 (2001), pp. 8614–19.

[31] T. Tian and K. Burrage, *Stochastic models for regulatory networks of the genetic toggle switch*, Proc. Nat. Acad. Sci, 103 (2006), pp. 8372–8377.

[32] T. E. Turner, S. Schnell, and K. Burrage, *Stochastic approaches for modelling in vivo reactions*, Computational Biology and Chemistry, 28 (2004), pp. 165–178.

[33] M. Ullah, H.Schmidt, K-H.Cho, and O.Wolkenhauer, *Deterministic modelling and stochastic simulation of pathways using MATLAB*, IEE Proc. Systems Biology, 153 (2006), pp. 53–60.

[34] V. Vyshemirsky and M. Girolami, *Bayesian ranking of biochemical system models*, Bioinformatics, 24(6) (2008), pp. 833–839.

[35] J. Wang, J. Zhang, Z. Yuan, and T. Zhou, *Noise-induced switches in network systems of the genetic toggle switch*, BMC Systems Biology, 1:50 (2007).

[36] D. J. Wilkinson, *Stochastic Modelling for Systems Biology*, Chapman & Hall/CRC, 2006.

[37] G. G. Yin and Q. Zhang, *Discrete-Time Markov Chains*, Springer, Berlin, 2005.