Adiabatic reduction of a model of stochastic gene expression with jump Markov process

Romain Yvinec, Changjing Zhuge, Jinzhi Lei & Michael C. Mackey

Journal of Mathematical Biology

ISSN 0303-6812

J. Math. Biol.
DOI 10.1007/s00285-013-0661-y

Springer

J. Math. Biol.
DOI 10.1007/s00285-013-0661-y

Mathematical Biology

Adiabatic reduction of a model of stochastic gene expression with jump Markov process

Romain Yvinec · Changjing Zhuge · Jinzhi Lei · Michael C. Mackey

Received: 20 September 2012 / Revised: 20 February 2013

Abstract This paper considers adiabatic reduction in a model of stochastic gene expression with bursting transcription considered as a jump Markov process. In this model, the process of gene expression with auto-regulation is described by fast/slow dynamics. The production of mRNA is assumed to follow a compound Poisson process occurring at a rate depending on protein levels (the phenomena called bursting in molecular biology) and the production of protein is a linear function of mRNA numbers. When the dynamics of mRNA is assumed to be a fast process (due to faster mRNA degradation than that of protein) we prove that, with appropriate scalings in the burst rate, jump size or translational rate, the bursting phenomena can be transmitted to the slow variable. We show that, depending on the scaling, the reduced equation is either

This research was supported by the Ecole Normale Superieure Lyon (ENS Lyon, France), the National Natural Science Foundation of China, and the Natural Sciences and Engineering Research Council of Canada.

R. Yvinec (✉)
Université de Lyon, CNRS UMR 5208, Université Lyon 1, Institut Camille Jordan,
43 blvd. du 11 novembre 1918, 69622 Villeurbanne, France
e-mail: yvinec@math.univ-lyon1.fr

C. Zhuge · J. Lei
Zhou Pei-Yuan Center for Applied Mathematics, Tsinghua University, Beijing 100084, China
e-mail: zgcj08@mails.tsinghua.edu.cn

J. Lei
e-mail: jzlei@tsinghua.edu.cn

M. C. Mackey
Departments of Physiology, Physics and Mathematics and Centre for Applied Mathematics
in Bioscience and Medicine, McGill University, 3655 Promenade Sir William Osler,
Montreal, QC H3G 1Y6, Canada
e-mail: michael.mackey@mcgill.ca

Published online: 05 March 2013

a stochastic differential equation with a jump Poisson process or a deterministic ordinary differential equation. These results are significant because adiabatic reduction techniques seem to have not been rigorously justified for a stochastic differential system containing a jump Markov process. We expect that the results can be generalized to adiabatic methods in more general stochastic hybrid systems.

Keywords Adiabatic reduction · Piecewise deterministic Markov process · Stochastic bursting gene expression · Quasi-steady state assumption · Scaling limit

Mathematics Subject Classification (2000) 92C45 · 60Fxx · 92C40 · 60J25 · 60J75

1 Introduction

The adiabatic reduction technique is often used to reduce the dimension of a dynamical system when known, or presumptive, fast and slow variables are present. Adiabatic reduction results for deterministic systems of ordinary differential equations have been available since the work by Fenichel (1979) and Tikhonov (1952). This technique has been extended to stochastically perturbed systems when the perturbation is a Gaussian distributed white noise (Berglund and Gentz 2006; Gardiner 1985; Stratonovich 1963; Titular 1978; Wilemski 1976). More recently, separation of time scales in discrete pure jump Markov processes were performed, using a master equation formalism (Santillán and Qian 2011) or a stochastic equation formalism (Kang and Kurtz 2013; Crudu et al. 2012). These papers show that a fast stochastic process can be averaged in the slow time scale, or can induce kicks to the slow variable. However, to the best of our knowledge, this type of approximation has never been extended to the situation in which the (fast) perturbation is a jump Markov process in a piecewise deterministic Markov process.

Jump Markov processes are often used in modelling stochastic gene expressions with explicit bursting in either mRNA or proteins (Friedman et al. 2006; Golding et al. 2005), and have been employed as models for genetic networks (Zeiser et al. 2008) and in the context of excitable membranes (Buckwar and Riedler 2011; Pakdaman et al. 2012; Riedler et al. 2012). Biologically, the 'bursting' of mRNA or protein is simply a process in which there is a production of several molecules within a very short time. In the biological context of modelling stochastic gene expression, explicit models of bursting mRNA and/or protein production have been analyzed recently, either using a discrete (Shahrezaei and Swain 2008) or a continuous formalism (Friedman et al. 2006; Lei 2009; Mackey et al. 2011) as even more experimental evidence from single-molecule visualization techniques has revealed the ubiquitous nature of this phenomenon (Elf et al. 2007; Golding et al. 2005; Ozbudak et al. 2002; Raj and van Oudenaarden 2009; Raj et al. 2006; Suter et al. 2011; Xie et al. 2008).

Traditional models of gene expression are often composed of *at least* two variables (mRNA and protein, and sometimes the promoter state). The use of a reduced one-dimensional model (protein concentration) has been justified so far by an argument concerning the stationary distribution by Shahrezaei and Swain (2008). However, it is clear that the two different models may have the same stationary distribution but very different dynamic behavior (for an example, see Mackey et al. 2011). The adiabatic reduction technique has been used in many studies (Hasty et al. 2000; Mackey et al. 2011) to simplify the analysis of stochastic gene expression dynamics, but without a rigorous mathematical justification.

The present paper gives a theoretical justification of the use of adiabatic reduction in a model of auto-regulation gene expression with a jump Markov process in mRNA transcription. We adopt a formalism based on density evolution (Fokker–Planck like) equations. Our results are of importance since they offer a rigorous justification for the use of adiabatic reduction to jump Markov processes. The model and mathematical results are presented in Sect. 2. Proof of the results are given in Sect. 3, with illustrative simulations in Sect. 4.

2 Model and results

2.1 Continuous-state bursting model

A single round of expression consists of both mRNA transcription and the translation of proteins from mRNA. The mRNA transcription occurs in a burst like fashion depending on the promoter activity. In this study, we assume a simple feedback between the end product (protein) which binds to its own promoter to regulate the transcription activity.

Let X and Y denote the concentrations of mRNA and protein respectively. A simple mathematical model of a single gene expression with feedback regulation and bursting in transcription is given by

$$\frac{dX}{dt} = -\gamma_1 X + \mathring{N}(h, \varphi(Y)), \tag{1}$$

$$\frac{dY}{dt} = -\gamma_2 Y + \lambda_2 X. \tag{2}$$

Here γ_1 and γ_2 are degradation rates for mRNA and proteins respectively, λ_2 is the translational rate, and $\mathring{N}(h, \varphi(Y))$ describes the transcriptional burst that is assumed to be a compound Poisson *white noise* occurring at a rate φ with a non-negative jump size ΔX distributed with density h.

In the model Eqs. (1)–(2), the stochastic transcriptional burst is characterized by the two functions φ and h. We always assume these two functions satisfy

$$\varphi \in \mathcal{C}^\infty(\mathbb{R}^+, \mathbb{R}^+), \ \varphi \text{ and } \varphi' \text{ are bounded, i.e. } \underline{\varphi} \leq \varphi, \varphi' \leq \overline{\varphi} \tag{3}$$

$$h \in \mathcal{C}^\infty(\mathbb{R}^+, \mathbb{R}^+) \quad \text{and} \quad \int_0^\infty x^n h(x)\, dx < \infty, \qquad \forall n \geq 1. \tag{4}$$

For a general density function h, the average burst size is given by

$$b = \int_0^\infty x h(x)\, dx. \tag{5}$$

Remark 1 Hill functions are often used to model self-regulation in gene expression, so that φ is given by

$$\varphi(y) = \varphi_0 \frac{1 + K y^n}{A + B y^n}$$

where φ_0, A, B, K and n are positive parameters (see Mackey et al. 2011 for more details).

An exponential distribution of the burst jump size is often used in modelling gene expression, in agreement with experimental findings by Xie et al. (2008), so that the density function h is given by

$$h(\Delta X) = \frac{1}{b} e^{-\Delta X / b},$$

where b is the average burst size.

The two functions φ and h here satisfy the assumptions (3)–(4).

2.2 Scalings

The Eqs. (1)–(2) are nonlinear, coupled, and analytically not easy to study. This paper provides an analytical understanding of the adiabatic reduction for Eqs. (1)–(2) when mRNA degradation is a fast process, i.e., γ_1 is "large enough" ($\gamma_1 \gg \gamma_2$) but the average protein concentration remains normal. Rapid mRNA degradation has been observed in *E. coli* (and other bacteria), in which mRNA is typically degraded within minutes, whereas most proteins have a lifetime longer than the cell cycle ($\geq$30 min for *E. coli*) (Taniguchi et al. 2010).

In Eqs. (1)–(2), when γ_1 is large, other parameters have to be adjusted accordingly to maintain a normal level of protein. When there is no feedback regulation to the burst rate, the function φ is independent of Y (therefore φ is a constant), and thus the average concentrations of mRNA and protein in a stationary state are

$$X_{\mathrm{eq}} := \lim_{t \to \infty} \mathbb{E}[X(t)] = \frac{b\varphi}{\gamma_1}, \tag{6}$$

$$Y_{\mathrm{eq}} := \lim_{t \to \infty} \mathbb{E}[Y(t)] = \frac{\lambda_2}{\gamma_2} X_{\mathrm{eq}} = \frac{b\varphi\lambda_2}{\gamma_1\gamma_2}. \tag{7}$$

From Eq. (7), when γ_1 is large enough ($\gamma_1 \gg \gamma_2$) and Y_{eq} remains at its normal level, one of the three quantities, b, φ, or λ_2 must be a large number as well. This observation holds even when there is a feedback regulation of the burst rate. Thus, in general, we have three possible scalings (as $\gamma_1 \to \infty$), each of which is biologically observed:

(S1) *Fast promoter activation/deactivation, so that the rate function φ is a large number. In this case, if $\gamma_1 \to \infty$, we assume the ratio φ/γ_1 is independent of γ_1.*

(S2) *Fast transcription, so that the average burst size b is a large number. From Eq.* (5), *this scaling indicates that the density function h changes with the parameter* γ_1 *in a form* $h(\Delta X) = \frac{1}{\gamma_1} h_0(\frac{\Delta X}{\gamma_1})$ *with* $h_0(\cdot)$ *independent of* γ_1.

(S3) *Fast translation, so that the translational rate* λ_2 *is a large number. In this case, if* $\gamma_1 \to \infty$, *we assume the ratio* λ_2/γ_1 *is independent of* γ_1.

These scalings are associated with different types of genes that display different types of kinetics (Schwanhäusser et al. 2011; Suter et al. 2011), and mathematically lead to different forms of the reduced dynamics. In this paper we determine the effective reduced equations for Eqs. (1)–(2) for each of the scaling conditions (S1)–(S3). Our main results are summarized below.

First, under the assumption (S1) (*fast promoter activation/deactivation*), Eqs. (1)–(2) can be approximated by a deterministic ordinary differential equation

$$\frac{dY}{dt} = -\gamma_2 Y + \lambda_2 \psi(Y) \tag{8}$$

where

$$\psi(Y) = b\varphi(Y)/\gamma_1. \tag{9}$$

Next, under the scaling relations (S2) (*fast transcription*) or (S3) (*fast translation*), Eqs. (1)–(2) are reduced to a single stochastic differential equation

$$\frac{dY}{dt} = -\gamma_2 Y + \mathring{N}(\bar{h}(\Delta Y), \varphi(Y)) \tag{10}$$

containing a jump Markov process, and the density $\bar{h}$ for the newly defined process is given by h through

$$\bar{h}(\Delta Y) = \left(\frac{\lambda_2}{\gamma_1}\right)^{-1} h\left(\left(\frac{\lambda_2}{\gamma_1}\right)^{-1} \Delta Y\right). \tag{11}$$

In particular, with the scaling (S2), we have

$$\bar{h}(\Delta Y) = \frac{1}{\lambda_2} h_0\left(\frac{\Delta Y}{\lambda_2}\right). \tag{12}$$

These results can be understood with the following simple arguments. When $\gamma_1 \to \infty$, applying a standard quasi-equilibrium assumption we have

$$\frac{dX}{dt} \approx 0,$$

which yields

$$X(t) \approx \frac{1}{\gamma_1} \mathring{N}(h, \varphi(Y)). \tag{13}$$

In the case of the scaling (S1), the jumps occur with high frequency and an average burst size b. Thus, $X(t)$ approaches the statistical average ($X(t) \approx b\varphi(Y)/\gamma_1$) for a given value Y, which gives Eq. (8). Under scalings (S2) or (S3), substituting Eq. (13) into Eq. (2) yields

$$\begin{aligned}\frac{dY}{dt} &\approx -\gamma_2 Y + \frac{\lambda_2}{\gamma_1}\mathring{N}(h, \varphi(Y))\\ &\approx -\gamma_2 Y + \mathring{N}(\bar{h}, \varphi(Y)).\end{aligned}$$

Exact statements for the results and their mathematical proofs are given below.

2.3 Density evolution equations and main results

The main results are based on the density evolution equations, and show that the evolution equations obtained from Eqs. (1)–(2) and those from Eq. (8) or (10) are consistent with each other when $\gamma_1 \to +\infty$ under the appropriate scaling. The existence of densities for such processes has been studied by Mackey and Tyran-Kamińska (2008), Tyran-Kamińska (2009).

Let $u(t, x, y)$ be the density function of $(X(t), Y(t))$ at time t obtained from the solution of Eqs. (1)–(2). The evolution of the density $u(t, x, y)$ is governed by (Mackey and Tyran-Kamińska 2008)

$$\begin{aligned}\frac{\partial u(t, x, y)}{\partial t} &= \frac{\partial}{\partial x}[\gamma_1 x u(t, x, y)] - \frac{\partial}{\partial y}[(\lambda_2 x - \gamma_2 y)u(t, x, y)]\\ &\quad + \int_0^x \varphi(y)u(t, z, y)h(x - z)\, dz - \varphi(y)u(t, x, y) \qquad (14)\end{aligned}$$

when $(t, x, y) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+$. The corresponding density function of $Y(t)$ is given by

$$u_0(t, y) = \int_0^\infty u(t, x, y)\, dx. \qquad (15)$$

In this paper, we prove that when $\gamma_1 \to \infty$ the density function $u_0(t, y)$ approaches the density $v(t, y)$ for solutions of either the deterministic equation (8) or the stochastic differential equation (10) depending on the scaling. Evolution of the density function for Eq. (8) is given by (Lasota and Mackey 1985)

$$\frac{\partial v(t, y)}{\partial t} = -\frac{\partial}{\partial y}[-\gamma_2 y v(t, y) + \lambda_2 \psi(y) v(t, y)], \qquad (16)$$

where

$$\psi(y) = b\varphi(y)/\gamma_1. \tag{17}$$

Evolution of the density function for Eq. (10) is given by

$$\frac{\partial v(t,y)}{\partial t} = \frac{\partial}{\partial y}[\gamma_2 y v(t,y)] + \int_0^y \varphi(z)v(t,z)\bar{h}(y-z)\,dz - \varphi(y)v(t,y). \tag{18}$$

Here $\bar{h}$ is related to h through

$$\bar{h}(y) = \frac{\gamma_1}{\lambda_2} h\left(\frac{\gamma_1}{\lambda_2} y\right). \tag{19}$$

We note that when φ and h satisfy assumptions (3)–(4), existence of the above densities has been proved by Mackey and Tyran-Kamińska (2008) and Tyran-Kamińska (2009). In particular, for a given initial density function

$$u(0,x,y) = p(x,y), \quad 0 < x, y < +\infty \tag{20}$$

that satisfies

$$p(x,y) \geq 0, \quad \int_0^\infty \int_0^\infty p(x,y)\,dx\,dy = 1, \tag{21}$$

there is a unique solution $u(t,x,y)$ of Eq. (14) that satisfies the initial condition (20) and

$$u(t,x,y) \geq 0, \quad \int_0^\infty \int_0^\infty u(t,x,y)\,dx\,dy = 1 \tag{22}$$

for all $t \in \mathbb{R}^+$.

We can rewrite the Eqs. (16) and (18) in the form

$$\frac{\partial v(t,y)}{\partial t} = \mathcal{T} v(t,y), \tag{23}$$

where $\mathcal{T}$ is a linear operator defined by the right hand side of Eq. (16) or (18).

Definition 1 A smooth function $f : \mathbb{R}^+ \to \mathbb{R}^+$ is a *test function* if $f(y)$ has compact support and $f^{(k)}(0) = 0$ for any $k = 0, 1, 2, \ldots$. An integrable function $v(t,y)$:

$\mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{R}^+$ is said to be a *weak solution* of Eq. (23) if for any test function $f(y)$,

$$\int_0^\infty \left(\frac{\partial v(t,y)}{\partial y} - \mathcal{T} v(t,y) \right) f(y)\, dy = 0, \quad \forall t > 0. \tag{24}$$

Remark 2 It is obvious that any classical solution of Eq. (23) is also a weak solution.

The main result of this section, given below, shows that when γ_1 is large enough, the marginal density of $Y(t)$, $u_0(t, y; \gamma_1)$, as defined below in Eq. (26), gives an approximation of a weak solution of Eq. (16) or (18).

Theorem 1 *Let $u(0, x, y) = p(x, y) \in \mathcal{C}^\infty(\mathbb{R}^{+^2})$ and assume that $p(x, y)$ satisfies*

$$\int_0^\infty x^n p(x,y)\, dx < +\infty, \quad y > 0, \ n = 0, 1, 2, \ldots. \tag{25}$$

For any $\gamma_1 > 0$, let $u(t, x, y; \gamma_1)$ be the associated solution of Eq. (14), and define

$$u_0(t, y; \gamma_1) = \int_0^\infty u(t, x, y; \gamma_1)\, dx. \tag{26}$$

Similarly,

$$p_0(y) = \int_0^\infty p(x,y)\, dx.$$

(1) *Under the scaling (S1), when $\gamma_1 \to \infty$, $u_0(t, y; \gamma_1)$ approaches a weak solution of Eq. (16) $v(t, y)$ with initial condition $v(0, y) = p_0(y)$.*
(2) *Under the scaling (S2) or (S3), when $\gamma_1 \to \infty$, $u_0(t, y; \gamma_1)$ approaches a weak solution of Eq. (18) $v(t, y)$ with initial condition $v(0, y) = p_0(y)$.*

From Definition 1, Theorem 1 means that for any test function $f(y)$,

$$\lim_{\gamma_1 \to \infty} \int_0^\infty \left(\frac{\partial u_0(t,y;\gamma_1)}{\partial t} - \mathcal{T} u_0(t,y;\gamma_1) \right) f(y)\, dy = 0, \quad \forall t > 0. \tag{27}$$

In the next section, we prove Eq. (27) for the three scalings respectively.

3 Proof of the main results

Before proving Theorem 1, we first examine the marginal moments under different scalings.

3.1 Scaling of the marginal moment

Proposition 1 *Let $(X(t), Y(t))$ be the solution of Eqs. (1)–(2), $\mu_k(t) = \mathbb{E}[X(t)^k]$ and $\nu_k(t) = \mathbb{E}[Y(t)X(t)^k]$. Suppose $\mu_k(0) < \infty$ and $\nu_k(0) < \infty$, then $\mu_k(t) < \infty$ and $\nu_k(t) < \infty$ for all t. Moreover, for any fixed $t > 0$:*

1. *If the scaling (S1) holds, both $\mu_k(t)$ and $\nu_k(t)$ are uniformly bounded above and below when γ_1 is large enough.*
2. *If the scaling (S2) holds, when γ_1 is large enough, for $k \geq 1$,*

$$\mu_k(t) \sim \gamma_1^{k-1}, \quad \nu_k(t) \sim \gamma_1^{k-1}, \tag{28}$$

   and $\nu_0(t)$ is uniformly bounded above and below.

3. *If the scaling (S3) holds, when γ_1 is large enough, for $k \geq 1$,*

$$\mu_k(t) \sim \gamma_1^{-1}, \quad \nu_k(t) \sim \gamma_1^{-1}, \tag{29}$$

   and $\nu_0(t)$ is uniformly bounded above and below.

Proof For the two-dimensional stochastic differential equation (1)–(2), the associated infinitesimal generator $\mathcal{A}$ is defined as (Davis 1984, Theorem 5.5)

$$\begin{aligned}\mathcal{A}g(x, y) = -\gamma_1 x \frac{\partial g}{\partial x} + (\lambda_2 x - \gamma_2 y)\frac{\partial g}{\partial y} \\ +\varphi(y)\left(\int_x^\infty h(z-x)g(z, y)\,dz - g(x, y)\right)\end{aligned} \tag{30}$$

for any $g \in \mathcal{C}^1(\mathbb{R}^+ \times \mathbb{R}^+)$. The operator $\mathcal{A}$ is the adjoint of the operator acting on the right hand side of the evolution equation of the density (14). Moreover, for any $g \in \mathcal{C}^1(\mathbb{R}^+ \times \mathbb{R}^+)$, we have

$$\frac{d}{dt}\mathbb{E}g(X_t, Y_t) = \mathbb{E}\mathcal{A}(g(X_t, Y_t)), \tag{31}$$

provided both terms on the right hand side of Eq. (30) are finite. The proposition is proved through calculations of Eq. (31).

To obtain estimations for μ_k, a straightforward calculation from Eq. (30) yields

$$\begin{aligned}\mathcal{A}x^k &= -\gamma_1 k x^k + \varphi(y)\left(\int_x^\infty h(z-x)(z-x+x)^k\,dz - x^k\right)\\ &= -\gamma_1 k x^k + \varphi(y)\sum_{i=0}^{k-1}\binom{k}{i}x^i\int_x^\infty h(z-x)(z-x)^{k-i}\,dz\\ &= -\gamma_1 k x^k + \varphi(y)\sum_{i=0}^{k-1}\binom{k}{i}x^i\mathbb{E}^{k-i}h,\end{aligned}$$

where

$$\mathbb{E}^j h = \int_0^\infty x^j h(x)\,dx.$$

Thus, Eq. (31) yields

$$\frac{d\mu_k(t)}{dt} = -\gamma_1 k\mu_k(t) + \sum_{i=0}^{k-1}\binom{k}{i}\mathbb{E}[\varphi(Y_t)X(t)^i]\mathbb{E}^{k-1}h. \tag{32}$$

We then obtain, with the assumption (3),

$$\underline{\varphi}\sum_{i=0}^{k-1}\binom{k}{i}\mu_i(t)\mathbb{E}^{k-i}h \le \dot{\mu}_k(t) + \gamma_1 k\mu_k(t) \le \overline{\varphi}\sum_{i=0}^{k-1}\binom{k}{i}\mu_i(t)\mathbb{E}^{k-i}h. \tag{33}$$

Now, we can obtain estimations of μ_k for different scalings from inequalities (33).

1. Assume the scaling (S1) so that both $\overline{\varphi}/\gamma_1$ and $\underline{\varphi}/\gamma_1$ are independent of γ_1 when γ_1 is large enough. Applying Gronwall's inequality to inequalities (33) with $k = 1$ yields, for all $t > 0$,

$$\frac{\underline{\varphi}\,b}{\gamma_1} + \left[\mu_1(0) - \frac{\underline{\varphi}\,b}{\gamma_1}\right]e^{-\gamma_1 t} \le \mu_1(t) \le \frac{\overline{\varphi}\,b}{\gamma_1} + \left[\mu_1(0) - \frac{\overline{\varphi}\,b}{\gamma_1}\right]e^{-\gamma_1 t}.$$

Thus, $\mu_1(t)$ is uniformly bounded above and below when γ_1 is large enough.

Iteratively, for all $t > 0$ and $k > 1$, there are constants $\bar{c}_k, \underline{c}_k > 0$ independent of γ_1 such that

$$\frac{\underline{\varphi}\,\underline{c}_k}{k\gamma_1} + \left[\mu_k(0) - \frac{\underline{\varphi}\,\underline{c}_k}{k\gamma_1}\right]e^{-k\gamma_1 t} \le \mu_k(t) \le \frac{\overline{\varphi}\,\bar{c}_k}{k\gamma_1} + \left[\mu_k(0) - \frac{\overline{\varphi}\,\bar{c}_k}{k\gamma_1}\right]e^{-k\gamma_1 t},$$

and hence $\mu_k(t)$ is uniformly bounded above and below when γ_1 is large enough.

2. Assume the scaling (S2) so that $\mathbb{E}^{k-i}h \sim \gamma_1^{k-i}$ when γ_1 is large enough. We note $\mu_0(t) = 1$, and therefore inductively, for any t and $k \geq 1$,

$$\frac{\underline{\varphi}\,\mathbb{E}^k h}{k\gamma_1} + O(\gamma_1^{k-2}) \leq \mu_k(t) \leq \frac{\overline{\varphi}\,\mathbb{E}^k h}{k\gamma_1} + O(\gamma_1^{k-2}).$$

Thus, we have $\mu_k(t) \sim \gamma_1^{k-1}$ when γ_1 is large enough.

3. Assume the scaling (S3) so that λ_2/γ_1 is independent of γ_1 when γ_1 is large enough. Calculations similar to those in case (S1) gives $\mu_k(t) \sim \gamma_1^{-1}$.

Analogous results for $\nu_k(t)$ are obtained with similar calculations with $g(x, y) = x^k y$ in Eq. (30). Namely, we have

$$\mathcal{A}x^k y = -(\gamma_1 k + \gamma_2)x^k y + \lambda_2 x^{k+1} + \varphi(y)\sum_{i=0}^{k-1}\binom{k}{i}x^i y \mathbb{E}^{k-i}h.$$

Thus, when $k = 0$, we have

$$\dot{\nu}_0 = -\gamma_2\nu_0 + \lambda_2\mu_1,$$

and for $k \geq 1$,

$$\begin{aligned}
-(\gamma_1 k + \gamma_2)\nu_k(t) + \lambda_2\mu_{k+1} + \underline{\varphi}\sum_{i=0}^{k-1}\binom{k}{i}\nu_i(t)\mathbb{E}^{k-i}h \\
\leq \dot{\nu}_k(t) \leq -(\gamma_1 k + \gamma_2)\nu_k(t) + \lambda_2\mu_{k+1} + \overline{\varphi}\sum_{i=0}^{k-1}\binom{k}{i}\nu_i(t)\mathbb{E}^{k-i}h.
\end{aligned}$$

Then ν_0 is uniformly bounded for each scaling (S1), (S2), and (S3). Then, iteratively using the inequalities for $\dot{\nu}_k$, the scaling of μ_{k+1} and Gronwall's inequality yields the desired result for each scaling.

Remark 3 Define the marginal moments

$$u_k(t, y) = \int_0^\infty x^k u(t, x, y)\,dx, \tag{34}$$

then

$$\mu_k(t) = \int_0^\infty u_k(t, y)\,dy.$$

Hence the integrals $\int_0^\infty u_k(t, y)dy$ satisfy the same scaling as $\mu_k(t)$ when $\gamma_1 \to \infty$.

Remark 4 From inequalities (33), when $\gamma_1 \to \infty$ the moments $\dot{\mu}_k(t)$ have the same scaling as $\mu_k(t)$. Moreover, the same scalings are valid for the integrals $\int_0^\infty \frac{\partial u_k(t,y)}{\partial t}\,dy$.

3.2 Proof of Theorem 1

Proof Throughout the proof, we omit γ_1 in the solution $u(t, x, y; \gamma_1)$ and in the marginal density $u_0(t, y; \gamma_1)$, and keep in mind that they are dependent on the parameter γ_1 through Eq. (14).

First, from the results in Sect. 3.1 and assumption (25), the marginal moments

$$u_n(t, y) = \int_0^\infty x^n u(t, x, y)\,dx, \tag{35}$$

are well defined for $t > 0$, $y > 0$ and $n \geq 0$. Hence

$$\begin{aligned}\lim_{x\to\infty} x^n u(t, x, y) = 0, \quad \forall t, y, n > 0.\\ \lim_{x\to 0} x^n u(t, x, y) = 0, \quad \forall t, y, n \geq 1.\end{aligned} \tag{36}$$

From Eq. (14), we multiply by x^n and integrate on both sides. By Eq. (36), we have

$$\begin{aligned}\frac{\partial u_n}{\partial t} &= -n\gamma_1 u_n - \lambda_2 \frac{\partial u_{n+1}}{\partial y} + \gamma_2 \frac{\partial (y u_n)}{\partial y}\\ &\quad + \int_0^\infty \int_0^x \varphi(y) x^n u(t, z, y) h(x - z)\,dz\,dx - \varphi(y) u_n.\end{aligned} \tag{37}$$

Since

$$\int_0^\infty \int_0^x \varphi(y) x^n u(t, z, y) h(x - z)\,dz\,dx = \sum_{j=0}^n \binom{n}{j} \varphi(y) u_{n-j} \mathbb{E}^j h,$$

we have

$$\frac{\partial u_n}{\partial t} = -n\gamma_1 u_n - \lambda_2 \frac{\partial u_{n+1}}{\partial y} + \gamma_2 \frac{\partial (y u_n)}{\partial y} + \varphi(y) \sum_{j=1}^n \binom{n}{j} u_{n-j} \mathbb{E}^j h. \tag{38}$$

In particular, when $n = 0$,

$$\frac{\partial u_0}{\partial t} = -\lambda_2 \frac{\partial u_1}{\partial y} + \gamma_2 \frac{\partial (y u_0)}{\partial y}, \tag{39}$$

and when $n \geq 1$,

$$\frac{1}{\gamma_1}\frac{\partial u_n}{\partial t} = -nu_n - \frac{\lambda_2}{\gamma_1}\frac{\partial u_{n+1}}{\partial y} + \frac{\gamma_2}{\gamma_1}\frac{\partial (yu_n)}{\partial y} + \frac{1}{\gamma_1}\varphi(y)\sum_{j=1}^{n}\binom{n}{j}u_{n-j}\mathbb{E}^j h. \quad (40)$$

Thus, for any $n \geq 1$,

$$\begin{aligned} u_n &= -\frac{\lambda_2}{n\gamma_1}\frac{\partial u_{n+1}}{\partial y} + \frac{\gamma_2}{n\gamma_1}\frac{\partial (yu_n)}{\partial y} \\ &\quad + \frac{1}{n\gamma_1}\varphi(y)\sum_{j=1}^{n}\binom{n}{j}u_{n-j}\mathbb{E}^j h - \frac{1}{n\gamma_1}\frac{\partial u_n}{\partial t}. \end{aligned} \quad (41)$$

Now, we are ready to prove the results for the three scalings by iteratively calculating u_1 from Eq. (41).

For the scaling (S1) so $\varphi(y) \sim \gamma_1$, and (here $b = \mathbb{E}h$)

$$u_1 = \frac{b\varphi(y)}{\gamma_1}u_0 + \frac{1}{\gamma_1}\left[\frac{\partial}{\partial y}(\gamma_2 y u_1 - \lambda_2 u_2) - \frac{\partial u_1}{\partial t}\right]. \quad (42)$$

Substituting Eq. (42) into Eq. (39), we obtain

$$\frac{\partial u_0}{\partial t} = \frac{\partial}{\partial y}[\gamma_2 y u_0 - \lambda_2 \psi(y) u_0] - \frac{\lambda_2}{\gamma_1}\frac{\partial}{\partial y}\left[\frac{\partial}{\partial y}(\gamma_2 y u_1 - \lambda_2 u_2) - \frac{\partial u_1}{\partial t}\right], \quad (43)$$

where $\psi(y) = b\varphi(y)/\gamma_1$. Now, we only need to show that for any test function $f(y)$,

$$\lim_{\gamma_1\to\infty}\frac{\lambda_2}{\gamma_1}\int_0^\infty f(y)\frac{\partial}{\partial y}\left[\frac{\partial}{\partial y}(\gamma_2 y u_1 - \lambda_2 u_2) - \frac{\partial u_1}{\partial t}\right]dy = 0, \quad \forall t > 0. \quad (44)$$

We note that the integral

$$\begin{aligned} \int_0^\infty f(y)\frac{\partial}{\partial y}\left[\frac{\partial}{\partial y}(\gamma_2 y u_1 - \lambda_2 u_2) - \frac{\partial u_1}{\partial t}\right]dy &= -\int_0^\infty f'(y)\frac{\partial u_1}{\partial t}\,dy \\ + \int_0^\infty f''(y)(\gamma_2 y u_1 - \lambda_2 u_2)\,dy \end{aligned}$$

is uniformly bounded when γ_1 is large enough, and Eq. (44) is straightforward from the Remarks 3 and 4. Thus, we conclude that $u_0(t, y)$ approaches a weak solution of Eq. (16) and point (1) of Theorem 1 is proved.

For the scaling (S2) so that $\mathbb{E}^j h \sim \gamma_1^j$ when $\gamma_1 \to \infty$, let

$$b_j = \gamma_1^{-j}\mathbb{E}^j h, \quad (j = 0, 1, \ldots) \quad (45)$$

which are independent of γ_1 when $\gamma_1 \to \infty$. Hence, from Eq. (41) and Proposition 1, we have

$$\begin{aligned}
\gamma_1^{-(n-1)}u_n &= -\frac{\lambda_2}{n}\frac{\partial(\gamma_1^{-n}u_{n+1})}{\partial y} + \frac{\gamma_2}{n\gamma_1}\frac{\partial(y\gamma_1^{-(n-1)}u_n)}{\partial y} + \frac{1}{n}\varphi(y)u_0 b_n \\
&\quad + \frac{1}{n\gamma_1}\varphi(y)\sum_{j=1}^{n-1}\binom{n}{j}\gamma_1^{-(n-j-1)}u_{n-j}b_j - \frac{1}{n\gamma_1}\frac{\partial(\gamma_1^{-(n-1)}u_n)}{\partial t} \\
&= \frac{1}{n}b_n\varphi(y)u_0 - \frac{\lambda_2}{n}\frac{\partial(\gamma_1^{-n}u_{n+1})}{\partial y} + \frac{1}{n\gamma_1}C_n(t,y),
\end{aligned}$$

where

$$C_n(t,y) = \gamma_2\frac{\partial(y\gamma_1^{-(n-1)}u_n)}{\partial y} + \varphi(y)\sum_{j=1}^{n-1}\binom{n}{j}\gamma_1^{-(n-j-1)}u_{n-j}b_j - \frac{\partial(\gamma_1^{-(n-1)}u_n)}{\partial t}.$$

Therefore,

$$\begin{aligned}
u_1 &= b_1\varphi(y)u_0 - \lambda_2\frac{\partial}{\partial y}[\gamma_1^{-1}u_2] + \frac{1}{\gamma_1}C_1(t,y) \\
&= b_1\varphi(y)u_0 - \lambda_2\frac{\partial}{\partial y}\left[\frac{1}{2}b_2\varphi(y)u_0 - \frac{\lambda_2}{2}\frac{\partial(\gamma_1^{-2}u_3)}{\partial y} + \frac{1}{2\gamma_1}C_2(t,y)\right] + \frac{1}{\gamma_1}C_1(t,y) \\
&= b_1\varphi(y)u_0 - b_2\frac{\lambda_2}{2!}\frac{\partial}{\partial y}(\varphi(y)u_0) + \frac{\lambda_2^2}{2!}\frac{\partial^2}{\partial y^2}\left[\frac{1}{3}b_3\varphi(y)u_0 - \frac{\lambda_2}{3}\frac{\partial(\gamma_1^{-3}u_4)}{\partial y} + \frac{1}{3\gamma_1}C_3(t,y)\right] \\
&\quad + \frac{1}{\gamma_1}C_1(t,y) - \frac{\lambda_2}{2!\gamma_1}\frac{\partial}{\partial y}C_2(t,y) \\
&\quad \cdots\cdots\cdots\cdots \\
&= \sum_{k=1}^{\infty}\frac{(-\lambda_2)^{k-1}}{k!}b_k\frac{\partial^{k-1}}{\partial y^{k-1}}(\varphi(y)u_0) + \frac{1}{\gamma_1}\sum_{k=1}^{\infty}\frac{(-\lambda_2)^{k-1}}{k!}\frac{\partial^{k-1}}{\partial y^{k-1}}C_k(t,y).
\end{aligned}$$

Thus, denote

$$C(t,y) = -\lambda_2\frac{\partial}{\partial y}\left[\sum_{k=1}^{\infty}\frac{(-\lambda_2)^{k-1}}{k!}\frac{\partial^{k-1}}{\partial y^{k-1}}C_k(t,y)\right] = \sum_{k=1}^{\infty}\frac{(-\lambda_2)^k}{k!}\frac{\partial^k}{\partial y^k}C_k(t,y)$$

and from Eq. (45), we have

$$\begin{aligned}
-\lambda_2\frac{\partial u_1}{\partial y} &= \sum_{k=1}^{\infty}\frac{(-\lambda_2)^k}{k!}(\gamma_1^{-k}\mathbb{E}^k h)\frac{\partial^k}{\partial y^k}(\varphi(y)u_0) + \frac{1}{\gamma_1}C(t,y) \\
&= \sum_{k=1}^{\infty}\frac{1}{k!}\left(-\frac{\lambda_2}{\gamma_1}\right)^k\left(\int_0^{\infty}x^k h(x)\,dx\right)\frac{\partial^k}{\partial y^k}(\varphi(y)u_0) + \frac{1}{\gamma_1}C(t,y)
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^\infty \bar{h}(x)\left[\sum_{k=1}^\infty \frac{1}{k!}(-x)^k \frac{\partial^k}{\partial y^k}(\varphi(y)u_0)\right] dx + \frac{1}{\gamma_1}C(t,y) \\
&= \int_0^\infty \bar{h}(x)(\varphi(y-x)u_0(t,y-x) - \varphi(y)u_0(t,y))\,dx + \frac{1}{\gamma_1}C(t,y) \\
&= \int_0^\infty \bar{h}(x)\varphi(y-x)u_0(t,y-x)\,dx - \varphi(y)u_0(t,y) + \frac{1}{\gamma_1}C(t,y) \\
&= -\int_y^{-\infty} \bar{h}(y-z)\varphi(z)u_0(t,z)\,dz - \varphi(y)u_0(t,y) + \frac{1}{\gamma_1}C(t,y) \\
&= \int_0^y \bar{h}(y-z)\varphi(z)u_0(t,z)\,dz - \varphi(y)u_0(t,y) + \frac{1}{\gamma_1}C(t,y). \qquad (46)
\end{aligned}
$$

Here we note $\varphi(z) = 0$ when $z < 0$.

For any test function $f(y)$, similar to the argument in the scaling (S1), the integral

$$\int_0^\infty C(t,y)f(y)\,dy$$

is uniformly bounded when γ_1 is large enough, and hence

$$\lim_{\gamma_1\to\infty} \frac{1}{\gamma_1}\int_0^\infty C(t,y)f(y)\,dy = 0, \quad \forall t > 0.$$

Therefore, from Eqs. (39) and (46), when $\gamma_1 \to \infty$, u_0 approaches a weak solution of Eq. (18), and point (2) in Theorem 1 is proved.

Now, we consider the scaling (S3) so λ_2/γ_1 is independent of γ_1 when $\gamma_1 \to \infty$. From Eq. (41) and Proposition 1, we have

$$
\begin{aligned}
u_n &= -\frac{1}{n}\frac{\lambda_2}{\gamma_1}\frac{\partial u_{n+1}}{\partial y} + \frac{\gamma_2}{n\gamma_1}\frac{\partial(yu_n)}{\partial y} + \frac{1}{n\gamma_1}\varphi(y)u_0\mathbb{E}^n h \\
&\quad + \frac{1}{n\gamma_1}\varphi(y)\sum_{j=1}^{n-1}\binom{n}{j}u_{n-j}\mathbb{E}^j h - \frac{1}{n\gamma_1}\frac{\partial u_n}{\partial t} \\
&= \frac{1}{n\gamma_1}\varphi(y)u_0\mathbb{E}^n h - \frac{1}{n}\frac{\lambda_2}{\gamma_1}\frac{\partial u_{n+1}}{\partial y} + \frac{1}{n\gamma_1}R_n(t,y),
\end{aligned}
$$

where

$$R_n(t, y) = \gamma_2 \frac{\partial (y u_n)}{\partial y} + \varphi(y) \sum_{j=1}^{n-1} \binom{n}{j} u_{n-j} \mathbb{E}^j h - \frac{\partial u_n}{\partial t}.$$

Therefore,

$$\begin{aligned}
u_1 &= \frac{1}{\gamma_1} \varphi(y) u_0 \mathbb{E}^1 h - \frac{\lambda_2}{\gamma_1} \frac{\partial}{\partial y} u_2 + \frac{1}{\gamma_1} R_1(t, y) \\
&= \frac{1}{\gamma_1} \varphi(y) u_0 \mathbb{E}^1 h - \frac{\lambda_2}{\gamma_1} \frac{\partial}{\partial y} \left[\frac{1}{2\gamma_1} \varphi(y) u_0 \mathbb{E}^2 h - \frac{1}{2} \frac{\lambda_2}{\gamma_1} \frac{\partial}{\partial y} u_3 + \frac{1}{2\gamma_1} R_2(t, y) \right] \\
&\quad + \frac{1}{\gamma_1} R_1(t, y) \\
&= \frac{1}{\gamma_1} \varphi(y) u_0 \mathbb{E}^1 h - \frac{1}{2!} \frac{\lambda_2}{\gamma_1^2} \mathbb{E}^2 h \frac{\partial}{\partial y} [\varphi(y) u_0] \\
&\quad + \frac{1}{2!} \left(\frac{\lambda_2}{\gamma_1} \right)^2 \frac{\partial}{\partial y} \left[\frac{1}{3\gamma_1} \varphi(y) u_0 \mathbb{E}^3 h - \frac{1}{3} \frac{\lambda_2}{\gamma_1} \frac{\partial}{\partial u_4} \right] \\
&\quad + \frac{1}{\gamma_1} \sum_{k=1}^{3} \frac{1}{k!} \left(-\frac{\lambda_2}{\gamma_1} \right)^{k-1} \frac{\partial^{k-1}}{\partial y^{k-1}} R_k(t, y) \\
&\quad \cdots\cdots\cdots \\
&= -\frac{1}{\lambda_2} \sum_{k=1}^{\infty} \frac{1}{k!} \left(-\frac{\lambda_2}{\gamma_1} \right)^k \mathbb{E}^k h \frac{\partial^{k-1}}{\partial y^{k-1}} [\varphi(y) u_0] \\
&\quad + \frac{1}{\gamma_1} \sum_{k=1}^{\infty} \frac{1}{k!} \left(-\frac{\lambda_2}{\gamma_1} \right)^{k-1} \frac{\partial^{k-1}}{\partial y^{k-1}} R_k(t, y).
\end{aligned}$$

Denote

$$R(t, y) = \sum_{k=1}^{\infty} \frac{1}{k!} \left(-\frac{\lambda_2}{\gamma_1} \right)^k \frac{\partial^k}{\partial y^k} R_k(t, y),$$

and in a manner similar to the above argument, we have

$$\begin{aligned}
-\lambda_2 \frac{\partial u_1}{\partial y} &= \sum_{k=1}^{\infty} \frac{1}{k!} \left(-\frac{\lambda_2}{\gamma_1} \right)^k \mathbb{E}^k h \frac{\partial^k}{\partial y^k} [\varphi(y) u_0] + R(t, y) \\
&= \int_0^y \bar{h}(y - z) \varphi(z) u_0(t, z)\, dz - \varphi(y) u_0(t, y) + R(t, y). \qquad (47)
\end{aligned}$$

Finally, we note $\mu_k(t) \sim \gamma_1^{-1}$ in the scaling (S3), hence for any test function $f(y)$,

$$\lim_{\gamma_1\to\infty}\int_0^\infty R(t,y)f(y)=0.$$

Thus, from Eqs. (39) and (47), when $\gamma_1 \to \infty$, u_0 approaches to a weak solution of Eq. (18), and point (3) in Theorem 1 is proved.

4 Illustration

We performed numerical simulations of the stochastic differential Eqs. (1)–(2) to illustrate the results in previous sections. In our simulations, we took parameter values so that γ_1 increases with the scaling (S2). As the intensity of the jumps is bounded, we used an accept/reject numerical scheme to simulate jump times, and used the exact solution of the deterministic part of Eqs. (1)–(2) between the jumps (the equations are linear between jumps). For a given set of parameters, we simulate a trajectory for a sufficiently long time (a bound on the convergence rate can be obtained by the coupling method, as proved by Bardet et al. 2013) so that the stochastic process reaches its stationary state. We then computed its equilibrium density (as well as the first and second moments) by sampling a large number of values (10^6) of the stochastic process at random times. Finally, we compare the marginal density for $Y(t)$ with the analytic steady-state solution of the one-dimensional equation (18). To quantify the differences, we used the L^1, L^2 and L^∞ norms (the parameter values are taken such that the asymptotic density is bounded).

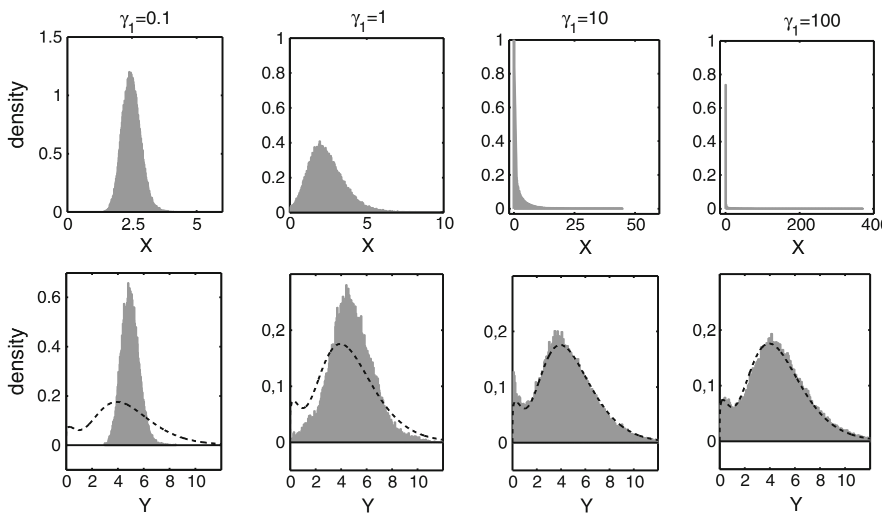

Fig. 1 Adiabatic reduction with the scaling (S2). *Upper panels* show the histograms for the first variable X. *Bottom panels* show the histograms for the second variable Y. *Dashed lines* are obtained from the one-dimensional equation (18). Functions $\varphi(Y)$ and $h(\Delta Y)$ are given by Remark 1, and parameters used are $\varphi_0 = 5, \gamma_2 = 1, \lambda_2 = 2, K = 1, A = 4, B = 1, n = 4, b = \gamma_1/2$ and, from *left to right*, $\gamma_1 = 0.1, 1, 10, 100$

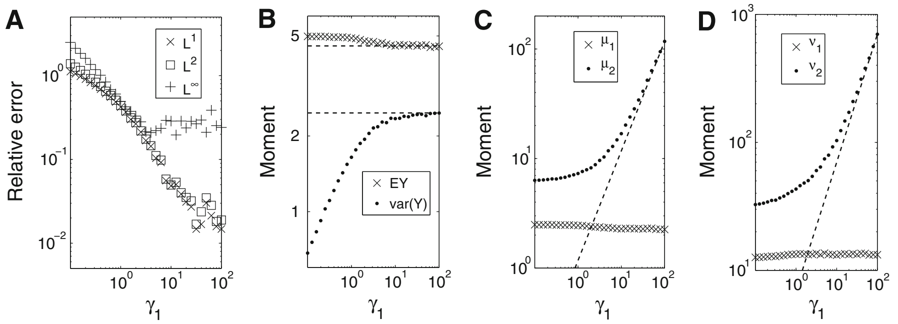

Fig. 2 Adiabatic reduction with the scaling (S2). **a** The norm differences between the numerical marginal density of $Y(t)$ and the analytic steady-state solution of the one-dimensional equation (18). Results for classical L^1, L^2 and L^∞ norms are shown, as indicated in the legend. **b** Asymptotic moment values of the second variable Y, as indicated on the legend. *Dashed lines* are obtained by the analytical asymptotic moment values obtained from the one-dimensional equation (18). **c** The moments μ_1 and μ_2 as functions of γ_1. **d** The moments ν_1 and ν_2 as functions of γ_1. In **c** and **d**, *the dashed lines* have a slope of $+1$. Parameters used are same as in Fig. 1

Results are shown in Figs. 1 and 2. First, Fig. 1 shows that as γ_1 increased, the marginal steady-state distribution approaches the analytical limit. Differences between the distributions are quantified in Fig. 2, where we show norm differences between the numerical and analytic distributions. We also show the behaviour of the moments. Notice that the marginal moment of Y approaches the analytic moment of the one-dimensional stochastic process as $\gamma_1 \to \infty$. Also, we verify the predicted behaviour of the moment involving the first variable X, μ_k and ν_k for $k = 1, 2$, as in Proposition 1. Results show good agreement with our theoretical predictions.

5 Summary

We have considered adiabatic reduction in a model of single gene expression with auto-regulation that is mathematically described by a jump Markov process defined by Eqs. (1)–(2). If mRNA degradation is a fast process, i.e., $\gamma_1 \gg \gamma_2$, we derived reduced forms of the governing equations under the three scaling situations so that the stationary protein level remains fixed when $\gamma_1 \to \infty$: (1) If the promoter activation/deactivation is also a fast process, then the protein concentration dynamics can be approximated by a deterministic ordinary differential equation (8), and the mRNA concentration is approximately given by $X = b\varphi(Y)/\gamma_1$. (2) If either the transcription or the translation is a fast process, then the protein concentration dynamics can be approximated by a single stochastic differential equation (10) with jump Markov process. We expect that these results may be generalized to justify adiabatic reduction methods in more general stochastic hybrid systems of gene regulation network dynamics.

Acknowledgments This work was supported by the Natural Sciences and Engineering Research Council (NSERC, Canada), the Mathematics of Information Technology and Complex Systems (MITACS, Canada), the National Natural Science Foundation of China (NSFC 11272169, China) and the région Rhône-Alpes (mobility fellowship), and carried out in Montréal, Lyon and Beijing. We thank our colleague M. Tyran-Kamińska for valuable discussions.

References

Bardet JB, Christen A, Guillin A, Malrieu F, Zitt PA (2013) Total variation estimates for the TCP process. Electron J Probab 18(10):1–21

Berglund N, Gentz B (2006) Noise-induced phenomena in slow-fast dynamical systems, a sample-paths approach. Springer, Berlin

Buckwar E, Riedler MG (2011) An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution. J Math Biol 63(6):1051–1093

Crudu A, Debussche A, Muller A, Radulescu O (2012) Convergence of stochastic gene networks to hybrid piecewise deterministic processes. Ann Appl Probab 5:1822–1859

Davis MHA (1984) Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models. J R Stat Soc Ser B 46(3):353–388

Elf J, Li GW, Xie XS (2007) Probing transcription factor dynamics at the single-molecule level in a living cell. Science 316(5828):1191–1194. doi:10.1126/science.1141967

Fenichel N (1979) Geometric singular perturbation theory for ordinary differential equations. J Differ Equ 31(1):53–98. doi:10.1016/0022-0396(79)90152-9

Friedman N, Cai L, Xie XS (2006) Linking stochastic dynamics to population distribution: an analytical framework of gene expression. Phys Rev Lett 97(16):168–302. doi:10.1103/PhysRevLett.97.168302

Gardiner CW (1985) Handbook of stochastic methods. Springer, Berlin

Golding I, Paulsson J, Zawilski SM, Cox EC (2005) Real-time kinetics of gene activity in individual bacteria. Cell 123(6):1025–1036

Hasty J, Pradines J, Dolnik M, Collins J (2000) Noise-based switches and amplifiers for gene expression. Proc Natl Acad Sci USA 97(5):2075–2080

Kang HW, Kurtz TG (2013) Separation of time-scales and model reduction for stochastic reaction networks. Ann App Probab 23:529–583

Lasota A, Mackey MC (1985) Probabilistic properties of deterministic systems. Cambridge University Press, Cambridge

Lei J (2009) Stochasticity in single gene expression with both intrinsic noise and fluctuation in kinetic parameters. J Theor Biol 256:485–492

Mackey MC, Tyran-Kamińska M (2008) Dynamics and density evolution in piecewise deterministic growth processes. Ann Pol Math 94(2):111–129. doi:10.4064/ap94-2-2

Mackey MC, Tyran-Kamińska M, Yvinec R (2011) Molecular distributions in gene regulatory dynamics. J Theor Biol 274(1):84–96. doi:10.1016/j.jtbi.2011.01.020

Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A (2002) Regulation of noise in the expression of a single gene. Nat Genet 31(1):69–73. doi:10.1038/ng869

Pakdaman K, Thieullen M, Wainrib G (2012) Fluid limit theorems for stochastic hybrid systems with application to neuron models. Adv Appl Probab 42(3):761–794

Raj A, van Oudenaarden A (2009) Single-molecule approaches to stochastic gene expression. Annu Rev Biophys 38(1):255–270. doi:10.1146/annurev.biophys.37.032807.125928

Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S (2006) Stochastic mRNA synthesis in mammalian cells. PLoS Biol 4(10):e309. doi:10.1371/journal.pbio.0040309

Riedler MG, Thieullen M, Wainrib G (2012) Limit theorems for infinite-dimensional piecewise deterministic Markov processes. Applications to stochastic excitable membrance models. Electron J Probab 17(55): 1–48

Santillán M, Qian H (2011) Irreversible thermodynamics in multiscale stochastic dynamical systems. Phys Rev E 83:1–8

Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M (2011) Global quantification of mammalian gene expression control. Nature 473:337–342

Shahrezaei V, Swain PS (2008) Analytical distributions for stochastic gene expression. Proc Natl Acad Sci USA 105(45):17256–17261. doi:10.1073/pnas.0803850105

Stratonovich R (1963) Topics in the theory of random noise, vol 1. General theory of random processes. Nonlinear transformations of signals and noise, revised English edition. Translated from the Russian by Richard A. Silverman edn. Gordon and Breach Science Publishers, New York

Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F (2011) Mammalian genes are transcribed with widely different bursting kinetics. Science 332(6028):472–474. doi:10.1126/science.1198817

Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M, Hearn J, Emili A, Xie XS (2010) Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. Science 329:533–538

Tikhonov AN (1952) Systems of differential equations containing small parameters in the derivatives. Mat Sb (NS) 31(73):575–586

Titular U (1978) A systematic solution procedure for the Fokker–Planck equation of a Brownian particle in the high-friction case. Phys A 91:321–344

Tyran-Kamińska M (2009) Substochastic semigroups and densities of piecewise deterministic Markov processes. J Math Anal Appl 357(2):385–402

Wilemski G (1976) On the derivation of Smoluchowski equations with corrections in the classical theory of Brownian motion. J Stat Phys 14:153–169

Xie XS, Choi PJ, Li GW, Lee NK, Lia G (2008) Single-molecule approach to molecular biology in living bacterial cells. Annu Rev Biophys 37(1):417–444. doi:10.1146/annurev.biophys.37.092607.174640

Zeiser S, Franz U, Wittich O, Liebscher V (2008) Simulation of genetic networks modelled by piecewise deterministic Markov processes. IET Syst Biol 2(3):113–135