Deterministic-like model reduction for a class of multi-scale stochastic differential equations with application to biomolecular systems (Extended Version)

Narmada Herath, *Student Member, IEEE,* Domitilla Del Vecchio, *Senior Member, IEEE*

Abstract

In this paper, we consider the problem of model order reduction for a class of singularly perturbed stochastic differential equations with linear drift terms. We present a reduced-order model that approximates both slow and fast variable dynamics when the time-scale separation is large. Specifically, we show that, on a finite time interval, the moments of all orders of the slow variables for the reduced-order model become closer to those of the original system as time separation is increased. A similar result holds for the first and second moments of the fast variable. Biomolecular systems with linear propensity functions, modeled by the chemical Langevin equation fit the class of systems considered in this work. Thus, as an application example, we analyze the trade-offs between noise and information transmission in a typical gene regulatory network motif, for which, both slow and fast variables are required. [1]

I. Introduction

The evolution of many dynamical systems takes place on multiple time-scales. Examples include climate systems, electrical systems, and biological systems [4], [5], [6]. The dynamics

This work was funded by AFOSR grant # FA9550-14-1-0060.

Narmada Herath is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 77 Mass. Ave, Cambridge MA nherath@mit.edu

Domitilla Del Vecchio is with the Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Mass. Ave, Cambridge MA ddv@mit.edu

[1]The preliminary results on this work appeared in the conference papers [1], [2], [3].

of such systems can be described using a set of ordinary differential equations (ODE) or stochastic differential equations (SDE) in the standard singular perturbation form, where the system variables are separated into 'slow' and 'fast' categories and a small parameter ϵ is used to represent the separation in the time-scales [7]. The analysis of such systems can be simplified by obtaining a reduced-order model that approximates the dynamics of the original system. To obtain a reduced-order model, one can appeal to the singular perturbation theory [7], [8] or to the averaging principle [9].

In the deterministic setting, the derivation of a reduced-order system is mainly accomplished using the Tikhonov's theorem, where the reduced-order model is obtained by setting ϵ to zero in the original system dynamics [8], [7]. This yields an algebraic equation that approximates the fast variable, which is in turn substituted into the slow variables' differential equation to obtain a reduced-order model for the slow variables' dynamics. The averaging principle can also be used to obtain an approximation of the slow variables' dynamics, in which, the fast dynamics are eliminated by integration of the system functions [9].

In addition to deterministic systems, stochastic models have also gained a lot of interest in many areas such as finance, population biology, and systems and synthetic biology [10], [11]. For example, biomolecular systems are intrinsically stochastic due to the randomness in chemical reactions and the chemical Langevin equation has been widely used to model the stochastic nature of these systems in the form of a stochastic differential equation [12].

Several works have appeared on singular perturbation methods for stochastic differential equations. However, these methods cannot be used when the diffusion terms of the fast variable are state-dependent and are of the order $\sqrt{\epsilon}$, which is the case in the chemical Langevin equation. Therefore, these works cannot be applied to stochastic differential equation models of biomolecular systems. In particular, the work by Kabanov and Pergamenshchikov provides a stochastic version of the Tikhonov's theorem for systems where the diffusion coefficient of the fast variable is $o(\sqrt{\epsilon}/|\sqrt{\ln(\epsilon)}|)$ [13]. Their results show that when the time-scale separation becomes large (ϵ becomes small), the reduced system converges in probability to the original system, under suitable assumptions including the exponential stability of the slow manifold. However, it is discussed that for the class of systems where the diffusion coefficient is $O(\sqrt{\epsilon})$, the fast variable may be oscillatory and may not converge in probability to the slow manifold. A related study by Berglund and Gentz uses a sample-path approach to find the probability of the solution being concentrated around a neighborhood of the slow manifold of the deterministic system [14]. This study includes the case where the diffusion coefficient is of $O(\sqrt{\epsilon})$. However, in this case, the analysis predicts that as ϵ decreases the probability of the trajectory of the fast variable escaping a neighborhood of the slow manifold increases. Their study also provides an approximation for the slow variable, but this approximation is defined for the time interval that the fast variable is within a neighborhood of the slow manifold. Therefore, the length of the time interval for which the slow variable approximation is valid decreases as ϵ decreases.

In [8], Kokotovic et al. developed a singular perturbation approach for linear stochastic systems in which the diffusion coefficient is a constant term. Their analysis also includes systems where the diffusion coefficient is scaled by the singular perturbation parameter ϵ. In both cases, they obtained a reduced system that converges to the original system in the mean squared sense as ϵ becomes small. However, this method cannot be applied to the case where the diffusion term is a function of the state variables. Tang and Basar approached the problem of singularly perturbed stochastic systems using the notion of stochastic input-to-state stability [15]. They obtained stability results for the original system under the assumptions that the reduced fast and slow subsystems are input-to-state stable. However, this method only quantifies the error for the fast variable, and therefore, it cannot be used to approximate the slow variable dynamics.

Aside from singular perturbation based approaches, averaging methods have also been extended for stochastic differential equations. They mostly consider systems with diffusion terms of order $\sqrt{\epsilon}$ [9]. In his pioneering work, R.Z. Khasminskii derived a reduced-order model where it is shown that the slow variables of the original system converge in distribution to the variables of the reduced-order system as the time-scale separation becomes large [16]. More recently, an application of the averaging principle to chemical Langevin equations was presented in [17]. However, the averaging methods require the integration of the system's vector field, which may be undesirable for systems of high dimension. Furthermore, averaging methods obtain approximations only for the slow variables, but do not provide any approximation for the fast variables.

In many applications, it is necessary to approximate both slow and fast variables in order to utilize the reduced-order model for analysis. Particularly, in biomolecular systems, chemical species often participate in both slow and fast reactions and hence the corresponding concentrations are neither slow nor fast variables, but instead are mixed variables. In these systems, a coordinate transformation can be employed to take the system to standard singular perturbation form [18], in which fast and slow variables may not directly correspond to the physical variables of interest. We illustrate this point in the application example of this paper.

In this work, we consider a class of stochastic differential equations with linear drift and nonlinear diffusion terms, including the case where the diffusion term of the fast variable is of the order $\sqrt{\epsilon}$. This class of systems is particularly common on biomolecular processes. We present a reduced-order SDE and an algebraic equation that approximate both slow and fast dynamics, respectively, following a similar approach to deterministic singular perturbation theory. We show that the error between the moments of the original and the reduced-order systems are of $O(\epsilon)$, for moments of all orders for the slow variable and for first and second order moments for the fast variable. We then demonstrate the application of the results on a gene regulatory network motif, where species dynamics typically consist of both slow and fast components. For this system, we derive the reduced-order model and illustrate how both slow and fast variable approximations can be used concurrently in analyzing trade-offs between noise and information transmission.

This paper is organized as follows. In Section II, we introduce the system model under study together with the underlying assumptions. In Section III, we introduce the reduced-order system and present the results on the quantification of the error between the original and reduced-order models. Section IV and Section V include examples that demonstrate the application of the results.

Notation: We use $\mathbb{E}[\cdot]$ to denote the expected value of a random variable. $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. $\mathbb{Z}_{\geq 0}$ denotes the set of nonnegative integers and $\mathbb{Z}_{>0}$ denotes the set of positive integers. Similarly, $\mathbb{Z}^n_{\geq 0}$ and $\mathbb{Z}^n_{>0}$ denote the sets of vectors of length n, with nonnegative and positive integer elements, respectively.

II. System Model

We consider the singularly perturbed stochastic differential equations

$$\dot{x} = f_x(x, z, t) + \sigma_x(x, z, t)\Gamma_x, \qquad x(0) = x_0, \tag{1}$$

$$\epsilon\dot{z} = f_z(x, z, t, \epsilon) + \sqrt{\epsilon}\sigma_z(x, z, t, \epsilon)\Gamma_z, \ z(0) = z_0, \tag{2}$$

where $x \in D_x \subset \mathbb{R}^n$ is the slow variable and $z \in D_z \subset \mathbb{R}^m$ is the fast variable. Γ_x is a d_x-dimensional white noise process. Let Γ_f be a d_f-dimensional white noise process, while Γ_z is a $(d_x + d_f)$-dimensional white noise process. We assume that the system (1)–(2) satisfies the following assumptions.

Assumption 1. The functions $f_x(x,z,t)$ and $f_z(x,z,t,\epsilon)$ are affine functions of the state variables x and z, i.e., we can write $f_x(x,z,t) = A_1x + A_2z + A_3(t)$, where $A_1 \in \mathbb{R}^{n\times n},\ A_2 \in \mathbb{R}^{n\times m}$ and $A_3(t) \in \mathbb{R}^n$, $f_z(x,z,t,\epsilon) = B_1x + B_2z + B_3(t) + \alpha(\epsilon)(B_4x + B_5z + B_6(t))$, where $B_1, B_4 \in \mathbb{R}^{m\times n},\ B_2, B_5 \in \mathbb{R}^{m\times m},\ B_3(t), B_6(t) \in \mathbb{R}^m$, $A_3(t)$ and $B_3(t)$ are continuously differentiable functions, and $\alpha(\epsilon)$ is a continuously differentiable function with $\alpha(0) = 0$.

Assumption 2. Let $\Phi(x,z,t) = \sigma_x(x,z,t)\sigma_x(x,z,t)^T$, $\Lambda(x,z,t,\epsilon) = \sigma_z(x,z,t,\epsilon)\sigma_z(x,z,t,\epsilon)^T$, and $\Theta(x,z,t,\epsilon) = \sigma_z(x,z,t,\epsilon)[\ \sigma_x(x,z,t)\ 0\ ]^T$. Then, we assume that $\Phi(x,z,t)$, $\Lambda(x,z,t,\epsilon)$, and $\Theta(x,z,t,\epsilon)$ are affine functions of x and z, and that $\lim_{\epsilon\to 0}\Lambda(x,z,t,\epsilon) < \infty$ and $\lim_{\epsilon\to 0}\Theta(x,z,t,\epsilon) < \infty$ for all x, z and t. Furthermore, we assume that the functions $\Phi(x,z,t)$, $\Lambda(x,z,t,\epsilon)$, and $\Theta(x,z,t,\epsilon)$ are continuously differentiable in t and ϵ.

Assumption 3. Matrix B_2 is Hurwitz.

We also assume that the system (1)–(2) admits a unique well-defined solution on a finite time interval. Sufficient conditions for the existence and uniqueness of solutions of stochastic differential equations are given by the Lipschitz continuity and bounded growth of system functions [19]. However, the class of systems considered in this work includes systems of the form where the diffusion term is a square-root function of the state variables, as Assumption 2 requires the squared diffusion terms to be linear functions of the state variables. Therefore, such systems may not guarantee the Lipschitz continuity conditions for the diffusion coefficient. For this type of systems, a set of sufficient conditions that guarantee the existence of solutions can be found in [10].

In the next section, we introduce the reduced-order system and present the results on the error quantification between the original and reduced-order systems.

III. Results

A. Reduced-order model

We introduce a reduced-order model by setting $\epsilon = 0$ in the original system (1)–(2), as in the case of deterministic singular perturbation theory. Under Assumption 2, $\epsilon = 0$ leads to the algebraic equation $f_z(x,z,t,0) = B_1x + B_2z + B_3(t) = 0$, for which, Assumption 3 guarantees the existence of a unique global solution $z = \gamma(x,t)$, given by

$$\gamma(x,t) = -B_2^{-1}(B_1x + B_3(t)). \tag{3}$$

Upon substitution of $z = \gamma(x,t)$ into (1), we obtain the *reduced slow system*

$$\dot{\bar{x}} = f_x(\bar{x}, \gamma(\bar{x},t), t) + \sigma_x(\bar{x}, \gamma(\bar{x},t), t)\Gamma_x, \;\; \bar{x}(0) = x_0, \tag{4}$$

which only depends on $\bar{x}$.

We assume that system (4) has a unique well-defined solution on a finite time interval $[0, t_1]$.

Next, we define a candidate approximation for the fast variable dynamics in the form

$$\bar{z}(t) = \gamma(\bar{x}(t), t) + g(\bar{x}(t), t)N, \tag{5}$$

where $N \in \mathbb{R}^d$ is a random vector whose components are independent standard normal random variables, and $g(\bar{x}(t), t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{m \times d}$ is a function that satisfies the Lyapunov equation

$$\begin{aligned} g(\bar{x}(t),t)g(\bar{x}(t),t)^T B_2^T + B_2 g(\bar{x}(t),t)g(\bar{x}(t),t)^T \\ = -\Lambda(\bar{x}, \gamma(\bar{x}(t),t), t, 0). \end{aligned} \tag{6}$$

We call equation (5) the *reduced fast system*.

We now present the results on the error quantification between the original and reduced-order systems. To this end, we first introduce the notation used to denote the moment dynamics (notation adapted from [20], [21]). Consider the vectors $x = [x_1, \ldots, x_n]^T$ and $k = (k_1, \ldots, k_n)$ where $x_i, k_i \in \mathbb{R}$ for $i = 1, \ldots, n$. Let $x^{(k)} = x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$. Then $\mathbb{E}[x^{(k)}]$ denotes the moment of x corresponding to the vector k, where the order of the moment is $\sum_{i=1}^n k_i$. In order to denote the P^{th} order moments for all $P \in \mathbb{Z}_{\geq 0}$, we define the set $\mathcal{G}_r^P = \{(c_1, \ldots, c_r) \in \mathbb{Z}_{\geq 0}^r | \sum_{i=1}^r c_i \leq P\}$. Then, we have the following main result.

Theorem 1. *Consider the original system in (1)–(2) and the reduced system in (4)–(5). Under Assumptions 1 - 3 there exists* $\epsilon^* > 0, \; t_1, \; t_b > 0$ *with* $t_1 > t_b$ *such that for* $\epsilon < \epsilon^*$

$$||\mathbb{E}[\bar{x}^{(k)}] - \mathbb{E}[x^{(k)}]|| = O(\epsilon), \;\; \forall k \in \mathcal{G}_n^{\mathcal{N}}, \;\; \mathcal{N} \in \mathbb{Z}_{>0}, \;\; t \in [0, t_1], \tag{7}$$

$$||\mathbb{E}[\bar{z}^{(l)}] - \mathbb{E}[z^{(l)}]|| = O(\epsilon), \;\; \forall l \in \mathcal{G}_m^2, \qquad t \in [t_b, t_1]. \tag{8}$$

The proof of this theorem utilizes several intermediate results and is presented in Appendix A. The outline of the proof is as follows. First, we show that the moment dynamics of the original system are in the standard singular perturbation form, and that setting $\epsilon = 0$ in the original moment dynamics yields the moment dynamics of the reduced-order system. This holds for moments of all orders for the slow variables and up to second order moments for the fast variables. As the moment dynamics are deterministic, we then apply the Tikhonov's theorem to demonstrate the convergence of the moments of the reduced-order system to the moments of the original system, as ϵ decreases. The stability conditions of the slow manifold of the original moment dynamics required for the application of the Tikhonov's theorem are guaranteed by Assumption 3.

From the reduced-order approximations given in equations (4)–(5), we note the similarity with the reduced-order model obtained by singular perturbation theory for deterministic systems [7]. In particular, the slow variable's dynamics are well approximated by substituting the expression of the slow manifold given by $z = \gamma(x,t)$ in equation (3) into the slow variable's dynamics given in equation (1). This implies that for this class of systems, the slow variable approximation can be obtained in the same manner as in the deterministic singular perturbation method.

By contrast, from expression (5) we note that the fast variable approximation contains the term $g(\bar{x},t)N$, which is in addition to the slow manifold expression $\gamma(\bar{x},t)$ that would be obtained with direct application of deterministic singular perturbation theory. This additional term is required in order to account for the noise of the fast variables. In fact, considering the system in the fast time-scale $\tau = t/\epsilon$, we see that the SDE of the fast variable is given by

$$\frac{dz}{d\tau} = f_z(x,z,t,\epsilon) + \sigma_z(x,z,t,\epsilon)\tilde{\Gamma}_z, \tag{9}$$

where $\tilde{\Gamma}_z$ represents Γ_z in the fast-time scale, i.e, $\tilde{\Gamma}_z(\tau) = \sqrt{\epsilon}\ \Gamma_z(t)$ as shown in [22, p.173]. For the case where the diffusion term is of the order $\sqrt{\epsilon}$, the term $\sigma_z(x,z,t,\epsilon)$ is independent of ϵ and thus $\sigma_z(x,z,t,0) \neq 0$. This shows that the fast variable is subject to noise, given by the diffusion term $\sigma_z(x,z,t,\epsilon)$, and thus the expression $\gamma(x,t)$ dose not provide an adequate approximation for the noise on z.

The noise in the fast variable can be "neglected" in the slow variable approximation because the slow subsystem "filters out" the noise from the fast variable. Such noise must instead be considered to approximate the noise properties of the fast variable, as we illustrate in the following example. Consider the system

$$\dot{x} = -a_1 x + a_2 z + v_1\Gamma_1, \tag{10}$$

$$\epsilon\dot{z} = -z + v_2\sqrt{\epsilon}\Gamma_2, \tag{11}$$

where $a_1, a_2 > 0$.

Setting $\epsilon = 0$, we obtain the system:

$$\dot{\bar{x}} = -a_1\bar{x} + v_1\Gamma_1, \tag{12}$$

$$z = \gamma(\bar{x}, t) = 0. \tag{13}$$

To analyze the error of this approximation, we can directly calculate the steady state moments for both the original and reduced-order systems using their linearity. This yields

$$\mathbb{E}[x^2] = \frac{a_2^2 v_2^2}{2a_1} \frac{\epsilon}{(1 + a_1\epsilon)} + \frac{v_1^2}{2a_1}, \qquad \mathbb{E}[z^2] = \frac{v_2^2}{2},$$

$$\mathbb{E}[\bar{x}^2] = \frac{v_1^2}{2a_1}, \qquad \mathbb{E}[\gamma(\bar{x}, t)^2] = 0.$$

It is seen that $\mathbb{E}[x^2]$ converges to $\mathbb{E}[\bar{x}^2]$ as ϵ approaches zero, however, $\mathbb{E}[z^2]$ remains constant as ϵ goes to zero. That is, the reduced-order system (12)–(13) obtained by setting $\epsilon = 0$ provides a good approximation for the slow variable in terms of the second moment, but it is not a good approximation for the fast variable dynamics. This is due to the fact that the x-subsystem is not affected by the noise Γ_2 as ϵ tends to zero, which can be explained by considering the power spectra and frequency response of the x and z subsystems.

Using the frequency response from input Γ_2 to the output z of the z-subsystem, given by $H_{z\Gamma}(j\omega) = \frac{1}{j\omega + 1/\epsilon}$ we can calculate the power spectrum of z as $S_{zz}(\omega) = \frac{(v_2/\sqrt{\epsilon})^2}{(\omega^2 + (1/\epsilon)^2)}$, which is illustrated in Fig. 1. It can be seen that as ϵ approaches zero, the magnitude of $S_{zz}(\omega)$ decreases at low frequencies but increases at high frequencies, in a way that the variance of z remains constant. However, considering the frequency response from z to x of the x-subsystem, given by $H_{xz}(j\omega) = \frac{a_2}{j\omega + a_1}$, we see that the x-subsystem is a low-pass filter with a cut-off frequency of a_1 that is independent of ϵ (Fig. 1). Therefore, x only selects the low frequency components of signal z, which decrease with ϵ, leading to a decrease in the variance of signal x as ϵ decreases. Thus, the reduced-order system obtained by setting $\epsilon = 0$ provides a good approximation for the slow variable dynamics. However, as the variance of z remains constant as ϵ decreases, the expression $\bar{z} = \gamma(\bar{x}, t)$ by itself does not provide a good approximation for the fast variable stochastic dynamics.

In the next two sections, we consider the application of this theory to an academic example first (Section IV) and then to a biomolecular system (Section V).

IV. Academic Example

We consider the following system, which takes a similar form to the SDEs that appear in affine term structure models in finance [10]:

$$\dot{x} = -2x + z + 10 + \sqrt{2z + 1}\ \Gamma_1, \tag{14}$$

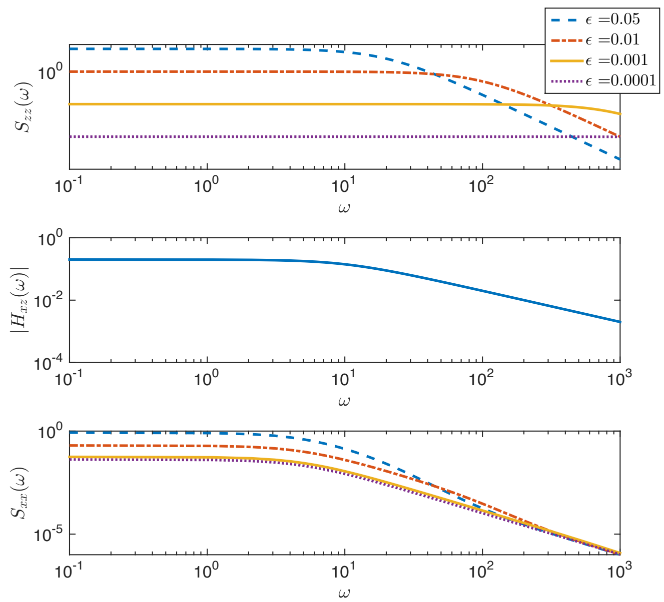

Fig. 1: Power spectrum of z (top). Frequency response from z to x (centre). Power spectrum of x (bottom). The parameters used are $a_1 = 10,\ a_2 = 2$ and $v_1 = 1,\ v_2 = 10$.

$$\epsilon \dot{z} = -z + 15 + \sqrt{\epsilon(2z+1)}\ \Gamma_2. \tag{15}$$

This system satisfies the Assumptions 1 - 3, and using the results of [10] it can be verified that there exists a unique, well-defined solution. Setting $\epsilon = 0$, we obtain the slow manifold $z = 15$. This yields the following reduced-order model for the slow variable:

$$d\bar{x} = -2\bar{x} + 25 + \sqrt{31}\ \Gamma_1. \tag{16}$$

Based on (5), the fast variable approximation for this system is of the form

$$z = 15 + g(x)N, \tag{17}$$

where $g(x)g(x)^T(-1) + (-1)g(x)g(x)^T = 31$ and N is a standard normal random variable. After solving for $g(x)$, the fast variable approximation is given by

$$\bar{z} = 15 + \sqrt{15.5}N. \tag{18}$$

Simulations of the original and the reduced-order systems were performed using the Euler-Maruyama method [23] for stochastic differential equations and the sample means were calculated using 500,000 realizations. Fig. 2 illustrates the second and third order moments of the slow variable and second order moments of the fast variable for the original and reduced-order systems. It can be seen that as ϵ decreases the moments of the original system tends to the moments of the reduced-order system.

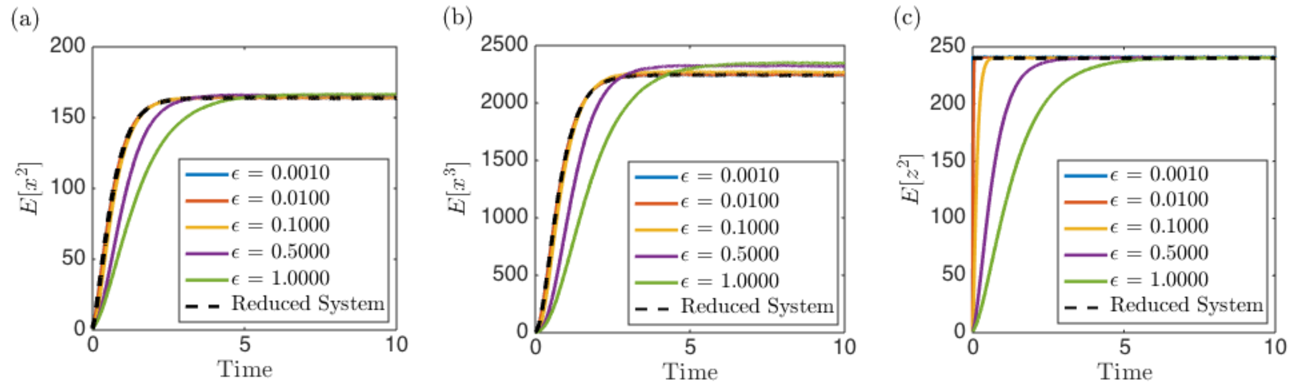

Fig. 2: Moments of the original and reduced systems. (a) Second moments of the slow variable. (b) Third moments of the slow variable. (c) Second moments of the fast variable.

By virtue of Theorem 1, the reduced-order model (16) provides a good approximation of the higher moments for the slow variable, as illustrated in Fig. 2. However, for the fast variable, only the first and second moments are well approximated, and there is no guarantee that the higher order moments are also approximated well, as we show by analyzing the third order moments of the system considered in this section.

To calculate the third order moments, we represent the fast variable dynamics of the above system (14)–(15) in the form

$$\epsilon dz = c_1 z + c_2 + \sqrt{\epsilon(d_1 z + d_2)}\ \Gamma_2.$$

Then, the third order moment dynamics are written as

$$\begin{aligned}
\epsilon\frac{d\mathbb{E}[z]}{dt} &= c_1\mathbb{E}[z] + c_2,\\
\epsilon\frac{d\mathbb{E}[z^2]}{dt} &= 2c_1\mathbb{E}[z^2] + 2c_2\mathbb{E}[z] + d_1\mathbb{E}[z] + d_2,\\
\epsilon\frac{d\mathbb{E}[z^3]}{dt} &= 3c_1\mathbb{E}[z^3] + 3c_2\mathbb{E}[z^2] + 3d_1\mathbb{E}[z^2] + 3d_2\mathbb{E}[z].
\end{aligned}$$

Setting $\epsilon = 0$, we obtain

$$\overline{\mathbb{E}[z]} = \frac{-c_2}{c_1}, \tag{19}$$

$$\overline{\mathbb{E}[z^2]} = \frac{2c_2^2 + d_1 c_2 - d_2 c_1}{2c_1^2}, \tag{20}$$

$$\overline{\mathbb{E}[z^3]} = \frac{3c_1c_2d_2 + c_1d_1d_2 - 2{c_2}^3 - 3{c_2}^2d_1 - c_2{d_1}^2}{2{c_1}^3}. \tag{21}$$

The reduced fast system is given by $\bar{z} = \gamma(\bar{x}, t) + g(\bar{x}, t)N(0, 1)$, where $\gamma(\bar{x}, t) = \frac{-c_2}{c1}$ and $g(\bar{x}, t) = \frac{d_1\gamma(\bar{x},t)+d_2}{-2c_1}$. Calculating the moment dynamics for the reduced fast system we obtain

$$\mathbb{E}[\bar{z}] = \gamma(\mathbb{E}[\bar{x}], t) = \frac{-c_2}{c_1}, \tag{22}$$

$$\mathbb{E}[\bar{z}^2] = \mathbb{E}[\gamma(\bar{x}, t)^2] + \mathbb{E}[g(\bar{x}, t)^2] = \frac{2c_2^2 + d_1c_2 - d_2c_1}{2c_1^2}, \tag{23}$$

$$\mathbb{E}[\bar{z}^3] = \mathbb{E}[\gamma(\bar{x}, t)^3] + 3\mathbb{E}[\gamma(\bar{x}, t)g(\bar{x}, t)^2] = \frac{c_2(3c_1d_2 - 2c_2^2 - 3c_2d_1)}{2c_1^3}. \tag{24}$$

Considering the equations for the slow manifold in (19) - (21) and the moments of the reduced fast system (22)–(24), we have that $\|\overline{\mathbb{E}[z]} - \mathbb{E}[\bar{z}]\| = 0$, $\|\overline{\mathbb{E}[z^2]} - \mathbb{E}[\bar{z}^3]\| = 0$, however, $\|\overline{\mathbb{E}[z^3]} - \mathbb{E}[\bar{z}^3]\| = \frac{d_1(c_1d_2 - c_2d_1)}{2c_1^3}$, which is different from zero. Therefore, it follows that setting $\epsilon = 0$ in the third moments of the fast variable does not yield the third moment of the reduced fast system.

From the general form of the moments in (22)–(24) it follows that the terms $\gamma(x, t)$ and $g(x, t)$ are not sufficient to approximate the third moment. This suggests that approximation of higher order moments of the fast variable would require additional terms in the reduced fast system. However, in many applications, particularly biomolecular systems, the commom measures of noise are coefficient of variation and signal-to-noise ratio, which are functions of only the mean and the variance. Thus, the first and second moments provide sufficient information for analysis of these systems.

V. Application Example

In this section, we demonstrate how the results obtained above can be used to characterize stochastic properties of biological systems. The time-scale separation property has been widely used for model order reduction in the analysis and design of biomolecular systems. More recently, deterministic singular perturbation techniques have been used to quantify impedance-like effects that arise in the design of biomolecular systems. These effects, termed retroactivity, arise at the interconnection of biomolecular components and cause a perturbation in the output signal of the upstream component, similar to loading effects in electrical circuits [24], [25]. Another source of signal perturbation in biological systems is the intrinsic noise due to the randomness in chemical reactions [26], [27]. Therefore, it is important to also account for stochastic effects in the analysis and design of biomolecular systems.

In this example, we consider the interconnection of transcriptional components, typically found in gene regulatory networks appearing both in natural and synthetic biological systems [6]. We model the system dynamics using the chemical Langevin equation and obtain a reduced-order model using the technique developed in this work. The reduced-order model is then used to quantify the errors in the system due to retroactivity and stochasticity. We investigate the interplay between each of these errors and identify trade-offs that arise in signal transmission in biomolecular systems.

A. System Model

We consider the interconnection of two transcriptional components shown in Fig. 3. Each transcriptional component [28] can be viewed as a system that takes as input a transcription factor, that is, a protein that can activate or repress a target gene, and gives as output the target gene's protein product.

The interconnection of Fig. 3, in which transcription factor Y activates the expression of a fluorescent protein G, is ubiquitous in synthetic genetic circuits as an indirect way of measuring the concentration of a transcription factor of interest, Y in this case. In fact, it is reasonable to think that the concentration of the fluorescent protein G should follow that of Y, possibly with some lag due to the process of gene expression encapsulated by the measuring device.

Here, we study how well the concentration of G tracks that of Y in the presence of noise.

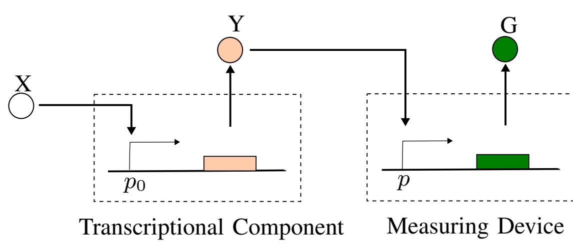

Fig. 3: Protein X acts as an input to the upstream transcriptional component, which produces the output protein Y. The downstream transcriptional components takes protein Y as an input and produces protein G.

The chemical reactions for this system can be written as follows: $\mathrm{X}+\mathrm{p}_0 \underset{\alpha_2}{\overset{\alpha_1}{\rightleftharpoons}} \mathrm{C}_0, \mathrm{C}_0 \xrightarrow{\beta_1} \mathrm{Y}+\mathrm{C}_0, \mathrm{Y} \xrightarrow{\delta_1} \phi$, $\mathrm{Y}+\mathrm{p} \underset{\alpha_4}{\overset{\alpha_3}{\rightleftharpoons}} \mathrm{C}, \mathrm{C} \xrightarrow{\beta_2} \mathrm{G}+\mathrm{C}, \mathrm{G} \xrightarrow{\delta_2} \phi$ [6], [28]. Protein X binds to promoter p_0 and produces complex C_0 where α_1 and α_2 are the association and dissociation rate constants. β_1 is the total production rate constant of protein Y considering both transcription and translation rates. δ_1 is the decay rate constant of protein Y, which includes both degradation and dilution of the protein. Similarly, α_3 and α_4 are the association and dissociation rate constants for protein Y and the promoter p_0, β_2 is the total production rate constant of protein G and δ_2 is the decay rate constant of protein G. Since DNA does not dilute cell growth, the total amount of promoter in the system is conserved giving $p_{T0} = p_0 + C_0$ and $p_T = p + C$ [6]. Denoting the system volume by Ω, the chemical Langevin equations for the system are given by

$$
\begin{aligned}
\frac{dC_0}{dt} &= \alpha_1 X(p_{T0} - C_0) - \alpha_2 C_0 + \sqrt{\frac{\alpha_1 X(p_{T0} - C_0)}{\Omega}}\Gamma_1 - \sqrt{\frac{\alpha_2 C_0}{\Omega}}\Gamma_2, \\
\frac{dY}{dt} &= \beta_1 C_0 - \delta_1 Y + \sqrt{\frac{\beta_1 C_0}{\Omega}}\Gamma_3 - \sqrt{\frac{\delta_1 Y}{\Omega}}\Gamma_4 \\
&\boxed{-\alpha_3 Y(p_T - C) + \alpha_4 C - \sqrt{\frac{\alpha_3 Y(p_T - C)}{\Omega}}\Gamma_5 + \sqrt{\frac{\alpha_4 C}{\Omega}}\Gamma_6}, \\
\frac{dC}{dt} &= \alpha_3 Y(p_T - C) - \alpha_4 C + \sqrt{\frac{\alpha_3 Y(p_T - C)}{\Omega}}\Gamma_5 - \sqrt{\frac{\alpha_4 C}{\Omega}}\Gamma_6, \\
\frac{dG}{dt} &= \beta_2 C - \delta_2 G + \sqrt{\frac{\beta_2 C}{\Omega}}\Gamma_7 - \sqrt{\frac{\delta_2 G}{\Omega}}\Gamma_8,
\end{aligned}
\tag{25}
$$

where Γ_i for $i = 1, ..., 8$ are independent Gaussian white noise processes. The binding of a transcription factor to downstream promoter sites introduces an additional rate of change in the dynamics of the transcription factor, which is represented by the boxed terms in equation (25) for the transcription factor Y. This additional rate of change, known as 'retroactivity', causes a change in the dynamics of the transcription factor's concentration with respect to the isolated case, that is, when the transcription factor is not binding [24], [29]. It was also shown in the works of [24] and [29] that increasing the number of downstream binding sites p_T increases the effect of retroactivity on the transcription factors.

The nominal and perturbed trajectories for Y and G for different amounts of p_T can be seen in Fig. 4. The nominal system dynamics, without perturbation due to retroactivity or noise, are obtained by simulating the ODE model obtained when $\Gamma_i = 0$ for $i = 1, ..., 8$ and the boxed terms are zero in the system (25). The perturbed trajectories are obtained using Gillespie's direct method [30]. For lower values of p_T the signal G closely follows the nominal signal, but the signal is highly perturbed by noise. As p_T increases the noise in the signal G decreases, however, the signal is highly attenuated due to retroactivity. This observation is consistent with the fact that using a high gene copy number (large p_T) is seen as a way of reducing noise in gene expression and protein production [31], [32]. However, the downside of this is that increasing p_T alters the dynamics of the input transcription factor (i.e. the protein Y), as experimentally observed in

[25]. For signal Y, by contrast, both retroactivity and noise increase as p_T is increased. This is consistent with prior observations in [33], where it was shown that increasing the copy number, and consequently increasing retroactivity, leads to a lower signal-to-noise ratio of transcription factors.

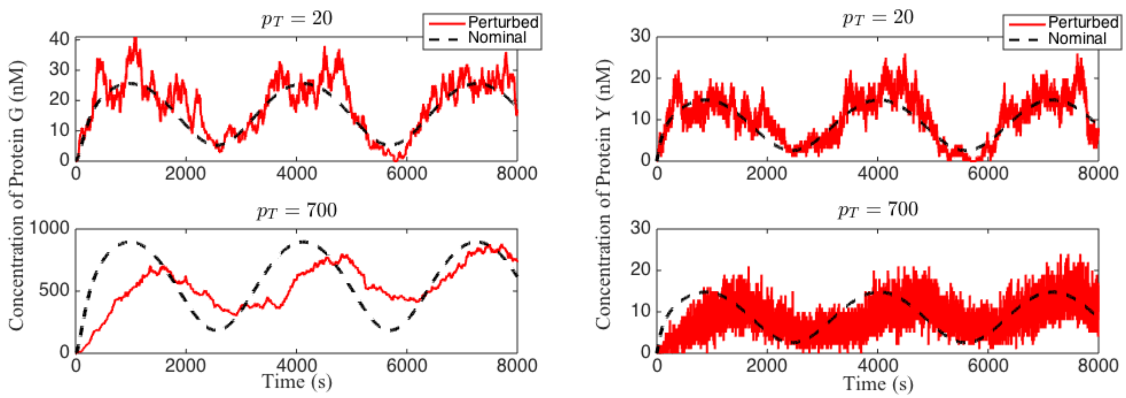

Fig. 4: Nominal and perturbed signals. G is obtained by simulating system (25) using the Gillespie algorithm [30]. The parameter values are $X = 2 + 1.5sin(\omega t)$nM, $\alpha_1 = 1\text{nM}^{-1}\text{s}^{-1}$, $\alpha_2 = 20\text{s}^{-1}$, $\alpha_3 = 1\text{nM}^{-1}\text{s}^{-1}$, $\alpha_4 = 100\text{s}^{-1}, \beta_1 = 0.01\text{s}^{-1}$, $\beta_2 = 0.1\text{s}^{-1}$, $\delta_1 = \delta_2 = 0.01\text{s}^{-1}$, $p_{T0} = 100$nM and ω = 0.002 rad/s.

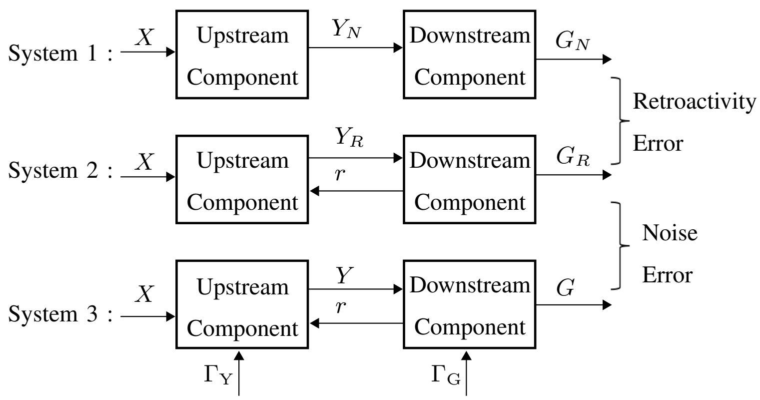

Fig. 5: The signal 'r' denotes the retroactivity to the upstream system. System 1 represents the nominal system in the absence of any perturbations. System 2 represents the system is perturbed only with retroactivity. System 3 represents the system perturbed with both retroactivity and intrinsic noise. Γ_Y encapsulates the noise in the upstream component given by Γ_i for $i = 1, ..., 4$ and Γ_G encapsulates the noise in the downstream component given by Γ_i for $i = 5, ..., 8$.

In the sequel, we mathematically quantify the above trade-offs between retroactivity and noise for proteins Y and G. To this end, we formally introduce System 1 as the nominal system,

System 2 as an intermediate system perturbed only with retroactivity, and System 3 as the perturbed system including both retroactivity and noise, given in Fig. 5. Next, we derive the dynamics for each of these systems. The system (25) exhibits time-scale separation as the binding/unbinding reactions between transcription factors and promoter sites are much faster than protein production/decay [6]. Thus, we can represent the system dynamics in the standard singular perturbation form by defining the small parameter $\epsilon = \delta_1/\alpha_2 \ll 1$. Representing the system variables by the non-dimensional quantities $c_0 = C_0/p_{T0}$, $y = Y/(\beta_1 p_{T0}/\delta_1)$, $c = C/p_T$, $g = G/(\beta_2 p_T/\delta_2)$, and $\bar{t} = t\delta_1$, and defining the dissociation constants $k_{d1} = \alpha_2/\alpha_1$ and $k_{d2} = \alpha_4/\alpha_3$ with $a = \alpha_4/\alpha_2$, we can take the system to the standard singular perturbation form using the change of variable $v = y + \frac{p_T\delta_1}{\beta_1 p_{T0}}c$, which yields

$$\begin{aligned}
\epsilon\frac{dc_0}{d\bar{t}} &= \frac{X}{k_{d1}} - c_0 + \sqrt{\epsilon\frac{X}{k_{d1}p_{T0}\Omega}}\tilde{\Gamma}_1 - \sqrt{\epsilon\frac{c_0}{p_{T0}\Omega}}\tilde{\Gamma}_2, \\
\frac{dv}{d\bar{t}} &= c_0 - (v - \frac{p_T\delta_1}{\beta_1 p_{T0}}c) + \sqrt{\frac{\delta_1 c_0}{\beta_1 p_{T0}\Omega}}\tilde{\Gamma}_3 - \sqrt{\frac{\delta_1(v - \frac{p_T\delta_1}{\beta_1 p_{T0}}c)}{\beta_1 p_{T0}\Omega}}\tilde{\Gamma}_4, \\
\epsilon\frac{dc}{d\bar{t}} &= \frac{a\beta_1 p_{T0}(v - \frac{p_T\delta_1}{\beta_1 p_{T0}}c)}{k_{d2}\delta_1} - ac + \sqrt{\epsilon\frac{a\beta_1 p_{T0}(v - \frac{p_T\delta_1}{\beta_1 p_{T0}}c)}{k_{d2}\delta_1 p_T\Omega}}\tilde{\Gamma}_5 - \sqrt{\epsilon\frac{ac}{p_T\Omega}}\tilde{\Gamma}_6, \\
\frac{dg}{d\bar{t}} &= \frac{\delta_2}{\delta_1}c - \frac{\delta_2}{\delta_1}g + \sqrt{\frac{\delta_2^2}{\delta_1\beta_2 p_T\Omega}c}\tilde{\Gamma}_7 - \sqrt{\frac{\delta_2^2}{\delta_1\beta_2 p_T\Omega}g}\tilde{\Gamma}_8,
\end{aligned} \tag{26}$$

where we have assumed that the binding between the proteins and promoter sites are weak, giving $C_0 \ll p_{T0}$ and $C \ll p_T$, and $\tilde{\Gamma}_i$ for $i = 1, ..., 8$ represent white noise processes in the time-scale $\bar{t}$.

It follows that the system (26) fits the structure of the original system in (1)–(2) with v and g as the slow variables and c_0 and c as the fast variables. We have that the drift terms and the squared diffusion terms are linear in the state variables, satisfying Assumptions 1 - 2. The matrix B_2 defined in Assumption 2 is given by

$$\begin{bmatrix} -1 & 0 \\ 0 & -\frac{ap_T}{k_{d2}} - a \end{bmatrix},$$

where we have that all the parameter constants are positive. Thus, the matrix B_2 is Hurwitz, satisfying Assumption 3. Therefore, the assumptions of Theorem 1 are satisfied.

We note that, due to the square-root form of the diffusion terms, the system (26) does not satisfy the sufficient Lipschitz continuity conditions for the existence of a unique solution for SDEs.

Furthermore, the system parameters do not satisfy the conditions for the existence of a unique solution for affine SDEs in [10]. The existence of a solution for chemical Langevin equations where the arguments of the square-root diffusion terms remain positive is an ongoing research question [34], [35]. However, the validity of the chemical Langevin equation representation for chemical kinetics is based on the assumption that the molecular counts are sufficiently large [12]. In line with this, the work in [36] considers several examples of one-dimensional systems and show that the probability of molecular counts reaching zero decreases as the initial condition increases. Considering higher dimensional models, in [37], we show that the minimum time for the molecular counts to reach a lower bound starting from a given set of initial conditions increases as the initial conditions become appropriately large (as defined in [37]), thereby keeping the argument of the square-root positive for a longer time interval.

Next, setting $\epsilon = 0$, we obtain the reduced-order system

$$\frac{dv}{d\bar{t}} = \frac{X}{k_{d1}} - (1-R)v + \sqrt{\frac{\delta_1 \frac{X}{k_{d1}}}{\beta_1 p_{T0}\Omega}}\tilde{\Gamma}_3 - \sqrt{\frac{\delta_1 (1-R)v}{\beta_1 p_{T0}\Omega}}\tilde{\Gamma}_4, \tag{27}$$

$$\frac{dg}{d\bar{t}} = \frac{\delta_2 \beta_1 p_{T0} v}{\delta_1^2 (p_T + k_{d2})} - \frac{\delta_2}{\delta_1} g + \sqrt{\frac{\delta_2^2 \beta_1 p_{T0} v}{\delta_1^2 \beta_2 p_T (p_T + k_{d2})\Omega}}\tilde{\Gamma}_7 - \sqrt{\frac{\delta_2^2}{\delta_1 \beta_2 p_T \Omega} g}\tilde{\Gamma}_8, \tag{28}$$

$$c_0 = \frac{X}{k_{d1}} + \sqrt{\frac{X}{p_{T0} k_{d1}\Omega}} N_1, \tag{29}$$

$$c = \frac{v\beta_1 p_{T0}}{\delta_1 (p_T + k_{d2})} + \sqrt{\frac{v\beta_1 p_{T0} k_{d2}}{\delta_1 p_T \Omega (p_T + k_{d2})^2}} N_2, \tag{30}$$

where $R = \frac{p_T}{p_T + k_{d2}}$, N_1 and N_2 are standard normal random variables. This system describes the dynamics for the perturbed system denoted by System 3 in Fig. 5 where the dimensionless concentration for protein Y is given by $y = v - \frac{p_T \delta_1}{\beta_1 p_{T0}} c$. Next, the dynamics for System 2, which only includes the error due to retroactivity can be found by taking $\Gamma_i = 0$ for $i = 1, \ldots, 8$ in (27)–(30), which yields

$$\frac{dv_R}{d\bar{t}} = \frac{X}{k_{d1}} - (1-R)v_R, \quad \frac{dg_R}{d\bar{t}} = \frac{\delta_2 \beta_1 p_{T0} v_R}{\delta_1^2 (p_T + k_{d2})} - \frac{\delta_2}{\delta_1} g_R, \tag{31}$$

$$c_{R0} = \frac{X}{k_{d1}}, \quad c_R = \frac{v_R \beta_1 p_{T0}}{\delta_1 (p_T + k_{d2})}. \tag{32}$$

Then, we can use the fast variable approximation for c_R given in (32) to rewrite the system dynamics in the original variable $y_R = v_R - c_R$, to obtain

$$\text{System 2}: \ \dot{y}_R = (1-R)\left(\frac{X}{k_{d1}} - y_R\right), \tag{33}$$

$$\dot{g}_R = \frac{\delta_2 \beta_1 p_{T0} y_R}{\delta_1^2 k_{d2}} - \frac{\delta_2}{\delta_1} g_R. \tag{34}$$

Similarly, the reduced-order dynamics for the nominal system (i.e without the boxed terms that represent retroactivity effects and with $\Gamma_i = 0$ for $i = 1, \ldots, 8$) can be written as

$$\text{System 1}: \quad \dot{y}_N = \frac{X}{k_{d1}} - y_N, \tag{35}$$

$$\dot{g}_N = \frac{\delta_2 \beta_1 p_{T0} y_N}{\delta_1^2 k_{d2}} - \frac{\delta_2}{\delta_1} g_N, \tag{36}$$

Next, using the system definitions in Fig. 5, we define the error due to retroactivity in Y and G as $\frac{|\Delta y_R|}{|y_N|} = \frac{|y_R - y_N|}{|y_N|}$ and $\frac{|\Delta g_R|}{|g_N|} = \frac{|g_R - g_N|}{g_N}$, respectively. Similarly, the error due to noise in the signals Y and G are can be defined as $\frac{|\Delta y_S|}{|y_R|} = \frac{|y - y_R|}{|y_R|}$ and $\frac{|\Delta g_S|}{|g_R|} = \frac{|g - g_R|}{|g_R|}$, respectively. We consider the input X to be of the form $X = k_1 + k_2 sin(\bar{\omega}\bar{t})$ with $k_1 > k_2$ to mimic a typical periodic signal from a clock [38]. As we are interested in the error in the temporal dynamics, we analyze each of the errors arising due to the time-varying component of the input $\tilde{X} = k_2 sin(\bar{\omega}\bar{t})$.

To quantify the error due to retroactivity, we take the ratio of amplitude of the signals Δy_R and Δg_R to the amplitude of the nominal signals Δy_N and Δg_N, respectively. Therefore, the error in y and g due to retroactivity is given by $\frac{|\Delta y_R(j\bar{\omega})|}{|y_N(j\bar{\omega})|}$ and $\frac{|\Delta g_R(j\bar{\omega})|}{|g_N(j\bar{\omega})|}$, respectively.

To quantify the error due to noise we consider the coefficient of variation, which is a standard measure of noise, defined as the ratio of standard deviation to the mean value of a signal. Since the drift functions in the system (27)–(28) are linear, the mean signals of y and g are given by y_R and g_R, respectively. Therefore, the terms $\mathbb{E}[(\Delta y_S)^2]$ and $\mathbb{E}[(\Delta g_S)^2]$ give the variances of signals y and g. Then, to quantify the noise error in Y we take $\frac{\sqrt{|\mathbb{E}[(\Delta y_S)^2](j\bar{\omega})|}}{|y_R(j\bar{\omega})|\sqrt{k_2}}$, where $|\mathbb{E}[(\Delta y_S)^2](j\bar{\omega})|k_2$ gives the amplitude of the signal $\mathbb{E}[(\Delta y_S)^2]$ and $|y_R(j\bar{\omega})|k_2$ gives the amplitude of the signal y_R for the input $\tilde{X} = k_2 sin(\bar{\omega}\bar{t})$. Similarly, the noise error in G can be quantified by the expression $\frac{\sqrt{|\mathbb{E}[(\Delta g_S)^2](j\bar{\omega})|}}{|g_R(j\bar{\omega})|\sqrt{k_2}}$.

B. Retroactivity Error

In order to find the retroactivity error, we consider the System 1 and System 2 in Fig. 5, for which the dynamics are given by (35)–(36) and (33)–(34). We use the linearity of the system (35)–(36) and (33)–(34) to directly evaluate the frequency response with a periodic input of the form $\tilde{X} = k_2 sin(\bar{\omega}\bar{t})$ and calculate the error in Y and G as

$$\frac{|\Delta y_R(j\bar{\omega})|}{|y_N(j\bar{\omega})|} = \frac{R\bar{\omega}}{\sqrt{\bar{\omega}^2 + (1-R)^2}}, \tag{37}$$

$$\frac{|\Delta g_R(j\bar{\omega})|}{|g_N(j\bar{\omega})|} = \frac{R\bar{\omega}}{\sqrt{\bar{\omega}^2 + (1-R)^2}}. \tag{38}$$

Since $R = \frac{p_T}{p_T + k_{d1}}$ monotonically increases with p_T, it follows that the error due to retroactivity in both Y and G increases as p_T increases.

C. Noise Error

Next, we quantify the noise error in Y by considering the dynamics for System 2 and System 3 in Fig. 5. As the drift coefficients of the system (27) - (30) are linear, we have that $\mathbb{E}[y] = \mathbb{E}[v] - \mathbb{E}[c] = y_R$. Therefore, the error $\mathbb{E}[(\Delta y_S)^2]$ is equivalent to the variance of y given by $\mathbb{E}[(y - \mathbb{E}[y])^2]$. Here, we note that the dynamics of the variable y consists of both slow and fast components, and therefore we require both slow and fast variable approximations to represent the dynamics of y using the reduced-order model.

Thus, we use the fast variable approximation for c given in (30) to derive the first and second moment dynamics for the variable y as shown in Appendix B to obtain

$$\frac{d\mathbb{E}[y]}{d\bar{t}} = (1-R)\left(\frac{X}{k_{d1}} - \mathbb{E}[y]\right), \tag{39}$$

$$\frac{d\mathbb{E}[y^2]}{d\bar{t}} = (1-R)\left[2\frac{X}{k_{d1}}\mathbb{E}[y] - 2\mathbb{E}\left[y^2\right] + \frac{\delta_1 X}{k_{d1}\beta_1 p_{T0}\Omega} + \frac{\delta_1 \mathbb{E}[y]}{\beta_1 p_{T0}\Omega}\right]. \tag{40}$$

Then, using the first and second moment dynamics we find the dynamics for the variance of y given by $\mathbb{E}[(y - \mathbb{E}[y])^2] = \mathbb{E}[y^2] - \mathbb{E}[y]^2$, which yields

$$\frac{d\mathbb{E}[(y - \mathbb{E}[y])^2]}{d\bar{t}} = (1-R)\left[\frac{\delta_1 X}{k_{d1}\beta_1 p_{T0}\Omega} + \frac{\delta_1 \mathbb{E}[y]}{\beta_1 p_{T0}\Omega} - 2\delta\mathbb{E}\left[(y - \mathbb{E}[y])^2\right]\right], \tag{41}$$

where $R = \frac{p_T}{p_T + k_{d2}}$ as defined in the derivation of the reduced system (27)–(30).

As system (41) is linear, we can directly evaluate its frequency response with the input $\tilde{X} = k_2 sin(\bar{\omega}\bar{t})$ which, by normalizing by the average signal $|y_R(j\bar{\omega})|k_2 = \frac{(1-R)k_2}{k_{d1}\sqrt{(\bar{\omega}^2 + (1-R)^2)}}$, leads to

$$\frac{\sqrt{|\mathbb{E}[\Delta y_S^2](j\bar{\omega})|}}{|y_R(j\bar{\omega})|\sqrt{k_2}} = \frac{\sqrt{k_{d1}\delta_1}(\bar{\omega}^2 + (1-R)^2)^{(1/4)}}{\sqrt{(1-R)\beta_1 p_{T0}\Omega k_2}}$$

Since the function R monotonically increases with p_T, the noise error in Y increases as p_T increases. Therefore, decreasing the downstream copy number p_T minimizes both retroactivity and noise errors in Y.

Next, we quantify the noise error in G by considering the dynamics for System 2 and System 3 in Fig. 5. Due to the linearity of the drift coefficients, the expression $\mathbb{E}[(\Delta g_S)^2]$, where g_S

was defined as $g_S = g - g_R$, gives the variance of the signal g. Thus, we use the dynamics of the variances of signals v and g to quantify the noise error in G. To this end, denote the variance of the signal v and the covariance between v and g by $\mathbb{E}[(\Delta v)^2] = \mathbb{E}[(v - \mathbb{E}[v])^2]$, $\mathbb{E}[\Delta vg] = (\mathbb{E}[vg] - \mathbb{E}[g]\mathbb{E}[v])$, respectively. Then, using the moment dynamics of the system (27) - (30), the dynamics for the variances are derived as

$$\frac{d\mathbb{E}[v]}{d\tau} = \frac{X}{k_{d1}} - (1-R)\mathbb{E}[v], \tag{42}$$

$$\frac{d\mathbb{E}[g]}{d\tau} = \frac{\delta_2\beta_1 p_{T0}\mathbb{E}[v]}{\delta_1^2(p_T + k_{d2})} - \frac{\delta_2}{\delta_1}\mathbb{E}[g], \tag{43}$$

$$\begin{aligned}
\frac{d\mathbb{E}[(\Delta v)^2]}{d\tau} &= -2(1-R)\mathbb{E}\big[(\Delta v)^2\big] + \frac{\delta_1\frac{X}{k_{d1}}}{\beta_1 p_{T0}\Omega} + \frac{\delta_1(1-R)\mathbb{E}[v]}{\beta_1 p_{T0}\Omega},\\
\frac{d\mathbb{E}[(\Delta vg)]}{d\tau} &= -((1-R) + \frac{\delta_2}{\delta_1})\mathbb{E}[(\Delta vg)] + \frac{\delta_2\beta_1 p_{T0}\mathbb{E}[(\Delta v)^2]}{\delta_1^2(p_T + k_{d2})},\\
\frac{d\mathbb{E}[(\Delta g)^2]}{d\tau} &= 2\frac{\delta_2\beta_1 p_{T0}}{\delta_1^2(p_T + k_{d2})}\mathbb{E}[(\Delta vg)] - 2\frac{\delta_2}{\delta_1}\mathbb{E}\big[(\Delta g)^2\big] + \frac{\delta_2^2\beta_1 p_{T0}\mathbb{E}[v]}{\delta_1^2\beta_2 p_T(p_T + k_{d2})\Omega} + \frac{\delta_2^2}{\delta_1\beta_2 p_T\Omega}\mathbb{E}[g],
\end{aligned} \tag{44}$$

Then, evaluating the frequency response for the system (42)–(44), we can quantify the noise error in G as

$$\frac{\sqrt{|\mathbb{E}[\Delta g_S^2](j\bar{\omega})|}}{|g_R(j\bar{\omega})|\sqrt{k_2}} = \sqrt{\frac{\delta_2^2\delta_1^2 + \delta_1^4\bar{\omega}^2}{k_2\ \Omega}}\sqrt[4]{A(p_T,\bar{\omega})} \tag{45}$$

where the function $A(p_T, \bar{\omega})$ decreases with increasing p_T for sufficiently small $\bar{\omega}$, as shown in Appendix C. Therefore, as we consider an input of the form $\tilde{X} = k_2 sin(\omega t)$, where $\omega = \bar{\omega}\delta_1$, the noise error in G decreases as p_T increases when the input frequency ω is sufficiently smaller than the bandwidth of the nominal system given by δ_1. Thus, in contrast to the noise error in Y, a higher value of p_T should be used to decrease the noise in G. This is due to the fact that increasing the amount of downstream copy number p_T leads to an increase in the amount of protein G, which in turn reduces the amount of relative fluctuations, as observed previously [31], [32].

Furthermore, since the noise error in Y increases with p_T in contrast to that of G, and Y is an input to the downstream component that produces G, we consider how the noise in Y propagates downstream to the signal noise in G. To this end, we observe from Fig. 4 that increasing p_T causes an increase in the high frequency noise of signal Y. However, the downstream component with the output signal G acts as a low-pass filter, which suggests that increasing high frequency noise content in Y will have a minimal effect on the noise of G.

Comparing the results obtained in this section for the noise error in G with the retroactivity error in G given by equation (38) demonstrates a trade-off between stochastic and deterministic perturbation in signal G. Fig. 6, illustrates this trade-off for p_T in the range 1: 1000 nM. Similarly, the expressions for the retroactivity error in (38) and the noise error in (45) can be used to quantify this trade-off for different parameter values and find an optimal value of p_T that would minimize the combined perturbation when designing biological circuits.

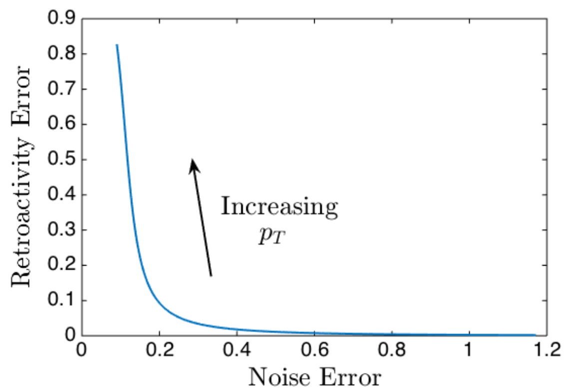

Fig. 6: Trade-off between retroactivity and noise in signal G for p_T in the range 1:1000 nM (obtained using the equations (38) and (45)). The parameter values are as in Figure 4.

VI. Conclusion

In this work, we have considered the problem of model order reduction for a class of stochastic differential equations in singular perturbation form. We introduced a reduced-order model that approximates both the slow and fast dynamics of the original system and can be obtained by solving two algebraic equations. For the slow variable approximation, it was shown that the error between the moments of the reduced system and moments of the original system are of $O(\epsilon)$. For the fast variable approximation, it was shown that the first and the second moments of the reduced system are within an $O(\epsilon)$-neighborhood of the first and second moments of the original system, respectively.

We then illustrated the application of our results with several examples. First, we considered an academic example and demonstrated the derivation of the reduced-order model and verified the results of error convergence through numerical simulations. We then considered an example of a biomolecular system that is typically encountered in both natural and synthetic genetic networks. The system dynamics were modeled using chemical Langevin equations and, by using the reduced-order model of Theorem 1, we analyzed and quantified trade-offs that arise in the design of genetic circuits. In particular, through this example, we demonstrate how both fast and slow variables approximations are required to quantify the noise properties of physical variables, which typically are mixed, i.e., neither slow nor fast. In future work, we will extend these results to systems with non-linear drift terms, which will allow us to consider biomolecular systems with multi-molecular reactions.

Appendix A

Here, we provide the proof of Theorem 1. We begin by presenting the set of intermediate results that will be used in the proof. In Claim 1, we prove that the moment dynamics of the original system can be written in the standard singular perturbation form. In Claim 2, we derive the moment dynamics of the reduced-order system for moments of all orders for the slow variable approximation and for first and second order moments of the fast variable approximation. In Claim 3, we derive the set of reduced-order moment equations obtained by setting $\epsilon = 0$ in the moments of the original system. The above results are used in Lemma 1 to prove that setting $\epsilon = 0$ in the moment dynamics of the original system yields the moment dynamics of the reduced-order system, for all moments of the slow variable and up to second order moments for the fast variable.

In order to derive the moment dynamics, we first define the set $\mathcal{K}_r^P = \{(k_1, \ldots, k_r) \in \mathbb{Z}_{\geq 0}^r | \sum_{i=1}^r k_i = P\}$. Then, considering the original system in (1)–(2), denote the state vectors by $x = [x_1, \ldots, x_n]^T$ and $z = [z_1, \ldots, z_m]^T$. We then have the following claim.

Claim 1. *Under Assumption 1 - 2, the moment dynamics of the original system in (1)–(2) can be written in the singular perturbation form:*

$$\frac{d\mathbb{E}[x^{(k)}]}{dt} = \sum_{i \in \mathcal{G}_n^P} C_{1i}(t)\mathbb{E}[x^{(i)}] + \sum_{l \in \mathcal{G}_m^1} \sum_{j \in \mathcal{G}_n^{P-1}} C_{2jl}(t)\mathbb{E}[z^{(l)}x^{(j)}], \forall k \in \mathcal{K}_n^P \tag{46}$$

$$\epsilon\frac{d\mathbb{E}[z^{(g)}]}{dt} = \sum_{a \in \mathcal{G}_m^P} D_{1a}(t,\epsilon)\mathbb{E}[z^{(i)}] + \sum_{c \in \mathcal{G}_n^1} \sum_{b \in \mathcal{G}_m^{P-1}} D_{2bc}(t,\epsilon)\mathbb{E}[x^{(l)}z^{(j)}], \forall g \in \mathcal{K}_m^P \tag{47}$$

$$\epsilon\frac{d\mathbb{E}[z^{(k_z)}x^{(k_x)}]}{dt} = \sum_{u \in \mathcal{G}_n^P} F_{1u}(t,\epsilon)\mathbb{E}[x^{(u)}] + \sum_{a \in \mathcal{G}_m^P} F_{2a}(t,\epsilon)\mathbb{E}[z^{(i)}] + \sum_{q=2}^{P}\sum_{r=1}^{q}\sum_{k \in \mathcal{G}_m^r}\sum_{s \in \mathcal{G}_m^{q-r}} F_{3qrks}(t,\epsilon)\mathbb{E}[z^{(k)}x^{(s)}], \tag{48}$$

for $k_x \in \mathcal{K}_n^{Q_x}$ and $k_z \in \mathcal{K}_m^{Q_z}$, where $Q_x + Q_z = P$, and for appropriate continuous functions $C_{1i}(t)$, $C_{2jl}(t)$ and continuously differentiable functions $D_{1a}(t)$, $D_{2bc}(t)$, $F_{2a}(t,\epsilon)$, $F_{1u}(t,\epsilon)$, $F_{3qrks}(t,\epsilon)$, for $i \in \mathcal{G}_n^P$, $l \in \mathcal{G}_m^1$ and $j \in \mathcal{G}_n^{P-1}$, $a \in \mathcal{G}_m^P$, $c \in \mathcal{G}_n^1$, $b \in \mathcal{G}_m^{P-1}$, $u \in \mathcal{G}_n^P$, $q = 2,\ldots,P$, $r = 1,\ldots,q$, $k \in \mathcal{G}_m^r$, $s \in \mathcal{G}_n^{q-r}$, $P = \{1,\ldots,\mathcal{N}\}$ where $\mathcal{N} \in \mathbb{Z}_{>0}$.

Proof. In order to derive the dynamics of $\mathbb{E}[x^{(k)}]$, consider the drift and diffusion terms of the slow variable dynamics of the original system (1)–(2), which can be denoted as $f_x(x,z,t) = [f_{x_1}(x,z,t),\ldots,f_{x_n}(x,z,t)]^T$ and $\sigma_x(x,z,t) = [\sigma_x^{ij}(x,z,t),t)]$ for $i = 1,\ldots,n$ and $j = 1,\ldots,d_x$. Then, using the Ito formula as in [20, p. 86] the moment dynamics of $x^{(k)}$ for each $k = (k_1,\ldots,k_n) \in \mathcal{K}_n^P$ can be derived as

$$\begin{aligned}\frac{d\mathbb{E}[x^{(k)}]}{dt} &= \sum_{i=1}^{n} k_i \mathbb{E}[f_{x_i}(x,z,t)\bar{x}_1^{k_1}\ldots x_i^{k_i-1}\ldots x_n^{k_n}] \\ &+ \frac{1}{2}\sum_{p=1}^{n} k_p(k_p-1)\mathbb{E}[\phi_{pp}(x,z,t)x_1^{k_1}\ldots x_p^{k_p-2}\ldots x_n^{k_n}] \\ &+ \sum_{l=2}^{n}\sum_{j=1}^{l-1} k_l k_j \mathbb{E}[\phi_{jl}(x,z,t)\bar{x}_1^{k_1}\ldots x_j^{k_j-1}\ldots x_l^{k_l-1}\ldots x_n^{k_n}],\end{aligned} \tag{49}$$

where $\phi_{ij}(x,z,t)$ for $i,j = 1,\ldots,n$ are the elements of the matrix $\Phi(x,z,t)$ defined in Assumption 2. From Assumptions $1-2$, we have that the functions $f_{x_i}(x,z,t)$ and $\phi_{ij}(x,z,t)$ are affine in x and z. Hence, it follows that the dynamics of P^{th} order moments will depend only on moments of order up to P. Under Assumption 1, we also have that $A_3(t)$ and $B_3(t)$ are continuous functions. Therefore, for appropriate continuous functions $C_{1i}(t)$ and $C_{2jl}(t)$ for $i \in \mathcal{G}_n^P$, $l \in \mathcal{G}_m^1$ and $j \in \mathcal{G}_n^{P-1}$, the moment dynamics in (49) can be written in the form given in (52).

In order to derive the dynamics of $\mathbb{E}[z^{(g)}]$, we consider the drift and diffusion terms of the fast variable dynamics of the original system (1)–(2), which can be denoted as $(1/\epsilon)f_z(x,z,t) = (1/\epsilon)[f_{z_1}(x,z,t,\epsilon),\ldots,f_{z_m}(x,z,t,\epsilon)]^T$ and $(1/\sqrt{\epsilon})\sigma_z(x,z,t,\epsilon) = (1/\sqrt{\epsilon})[\sigma_z^{ij}(x,z,t,\epsilon)]$ for $i = 1,\ldots,n$ and $j = 1,\ldots,(d_x+d_f)$. Then, as in [20, p. 86] the moment dynamics of $z^{(g)}$ for each $g = (g_1,\ldots,g_m) \in \mathcal{K}_m^P$ can be written as

$$\begin{aligned}\frac{d\mathbb{E}[z^{(g)}]}{dt} &= \sum_{i=1}^{m} g_i \mathbb{E}\left[\frac{1}{\epsilon}f_{z_i}(x,z,t,\epsilon)z_1^{g_1}\ldots z_i^{g_i-1}\ldots z_m^{g_m}\right] \\ &+ \frac{1}{2}\sum_{p=1}^{m} g_p(g_p-1)\mathbb{E}\left[\frac{1}{\epsilon}\lambda_{pp}(x,z,t,\epsilon)z_1^{g_1}\ldots z_p^{g_p-2}\ldots z_m^{g_m}\right] \\ &+ \sum_{l=2}^{m}\sum_{p=1}^{l-1} g_l g_p \mathbb{E}\left[\frac{1}{\epsilon}\lambda_{pl}(x,z,t,\epsilon)z_1^{g_1}\ldots z_p^{g_p-1}\ldots z_l^{g_l-1}\ldots z_m^{g_m}\right],\end{aligned} \tag{50}$$

$\lambda_{ij}(x,z,t,\epsilon)$ for $i,j=1,\dots,m$ are the elements of the matrix $\Lambda(x,z,t,\epsilon)$ defined in Assumption 2. Under Assumptions $1-2$, we have that the functions $f_{z_i}(x,z,t,\epsilon)$ and $\lambda_{ij}(x,z,t,\epsilon)$ are affine in x and z and are continuously differentiable in their arguments. Thus, it follows that the dynamics of the moments of order P will depend only on moments of order less than or equal to P. Then, multiplying both sides of the equation (50) by ϵ, we can represent the moments $\mathbb{E}[z^{(g)}]$ for each $g \in \mathcal{K}_m^P$ as in (47) for appropriate continuously differentiable functions $D_{1a}(t)$ and $D_{2bc}(t)$ for $a \in \mathcal{G}_m^P$, $c \in \mathcal{G}_n^1$, $b \in \mathcal{G}_m^{P-1}$.

In order to find the dynamics of $\mathbb{E}[z^{(k_z)}x^{(k_x)}]$, we consider the vector $[z_1,\dots,z_m,x_1,\dots,x_n]^T$, for which the drift and the diffusion terms can be represented by the vector $[f_{z_1}(x,z,t,\epsilon)/\epsilon,\dots,$ $f_{z_m}(x,z,t,\epsilon)/\epsilon, f_{x_1}(x,z,t)\dots f_{x_n}(x,z,t)]^T$ and matrix $[(1/\sqrt{\epsilon})\sigma_z^{ij}(x,z,t,\epsilon);[\sigma_x(x,z,t)\ 0]^{kj}]$ respectively, for $i=1,\dots,m$, $j=1,\dots,(d_x+d_f)$, $k=1,\dots,n$, in which $[\sigma_x(x,z,t)\quad 0]$ denotes a matrix-valued function $\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}\to\mathbb{R}^{n\times(d_x+d_f)}$ and $[\sigma_x(x,z,t)\quad 0]^{kj}$ denotes the elements of the matrix. Then, from [20, p. 86], the moment dynamics of $z^{(k_z)}x^{(k_x)}$ for $k_z=(c_1,\dots,c_m)\in\mathcal{K}_m^{P_z}$ and $k_x=(k_1,\dots,k_m)\in\mathcal{K}_n^{P_x}$ can be written as

$$
\begin{aligned}
&\frac{\mathbb{E}[z^{(k_z)}x^{(k_x)}]}{dt} = \\
&\sum_{i=1}^{m} c_i\mathbb{E}\Big[\frac{1}{\epsilon}f_{z_i}(x,z,t,\epsilon)z_1^{c_1}\dots z_i^{c_i-1}\dots z_m^{c_m}x_1^{k_1}\dots x_n^{k_n}\Big] \\
&+\sum_{i=1}^{n} k_i\mathbb{E}\Big[f_{x_i}(x,z,t)z_1^{c_1}\dots z_m^{c_m}x_1^{k_1}\dots x_i^{k_i-1}\dots x_n^{k_n}\Big] \\
&+\frac{1}{2}\sum_{i=1}^{m} c_i(c_i-1)\Big[\frac{\lambda_{ii}(x,z,t,\epsilon)}{\epsilon}z_1^{c_1}\dots z_i^{c_i-2}\dots z_m^{c_m}x_1^{k_1}\dots x_n^{k_n}\Big] \\
&+\frac{1}{2}\sum_{i=1}^{n} k_i(k_i-1)\Big[\phi_{ii}(x,z,t,\epsilon)z_1^{c_1}\dots z_m^{c_m}x_1^{k_1}\dots x_i^{k_i-2}\dots x_n^{k_n}\Big] \\
&+\sum_{i=2}^{m}\sum_{j=1}^{i-1} c_ic_j\mathbb{E}\Big[\frac{\lambda_{ji}^z(x,z,t,\epsilon)}{\epsilon}z_1^{c_1}\dots z_j^{c_j-1}\dots z_i^{c_i-1}\dots z_m^{c_m}x_1^{k_1} \\
&\qquad\dots x_n^{k_n}\Big] \\
&+\sum_{i=1}^{n}\sum_{j=1}^{m} k_ic_j\mathbb{E}\Big[\frac{\theta_{ji}(x,z,t,\epsilon)}{\sqrt{\epsilon}}z_1^{c_1}\dots z_j^{c_j-1}\dots z_m^{c_m}x_1^{k_1}\dots x_i^{k_i-1} \\
&\qquad\dots x_n^{k_n}\Big] \\
&+\sum_{i=2}^{n}\sum_{j=1}^{i-1} k_ik_j\mathbb{E}\Big[\phi_{ji}(x,z,t)z_1^{c_1}\dots z_m^{c_m}x_1^{k_1}\dots x_i^{k_i-1}\dots x_j^{k_j-1} \\
&\qquad\dots x_n^{k_n}\Big].
\end{aligned}
\tag{51}
$$

where $\lambda_{ij}(x,z,t,\epsilon)$, $\phi_{lk}(x,z,t)$ and $\theta_{ik}(x,z,t,\epsilon)$ for $i,j=1,\dots m$, $l,k=1,\dots n$ are the elements of the matrices $\Lambda(x,z,t,\epsilon)$, $\Phi(x,z,t)$ and $\Theta(x,z,t,\epsilon)$ defined in Assumption 2, respectively. We have that the functions $f_{z_i}(x,z,t,\epsilon)$, $f_{x_i}(x,z,t)$, $\lambda_{ij}(x,z,t,\epsilon)$, $\phi_{lk}(x,z,t)$ and $\theta_{ik}(x,z,t,\epsilon)$ are affine in x and z and are continuously differentiable in their arguments due to Assumption 1 - 2. Thus, for appropriate functions $F_{2a}(t,\epsilon)$, $F_{1u}(t,\epsilon)$, $F_{3qrks}(t,\epsilon)$, for $a\in\mathcal{G}_m^P$, $u\in\mathcal{G}_n^P$, $q=2,\dots,P$, $r=1,\dots,q$, $k\in\mathcal{G}_m^r$, $s\in\mathcal{G}_n^{q-r}$ the dynamics of $\mathbb{E}[z^{(k_z)}x^{(k_x)}]$ can be written in the form of (48). □

Next, we derive the moment dynamics of the reduced-order system (4)–(5). For this, denote the state vector of the reduced slow system by $\bar{x}=[\bar{x}_1,\dots,\bar{x}_n]^T$ and let $\gamma(\bar{x},t)=[\gamma_1(\bar{x},t),\dots,\gamma_m(\bar{x},t)]$, $g(\bar{x},t)=[g_{\bullet 1}(\bar{x},t),\dots,g_{\bullet(d_x+d_f)}(\bar{x},t)]$ where $g_{\bullet i}(\bar{x},t)$ denotes the i^{th} column of $g(\bar{x},t)$. Then, we have the following claim.

Claim 2. *Under Assumptions 1 - 2, the moment dynamics of the reduced slow system in (4) are given by*

$$\frac{d\mathbb{E}[\bar{x}^{(k)}]}{dt}=\sum_{i\in\mathcal{G}_n^P}C_{1i}(t)\mathbb{E}[\bar{x}^{(i)}]+\sum_{l\in\mathcal{G}_m^1}\sum_{j\in\mathcal{G}_n^{P-1}}C_{2jl}(t)\mathbb{E}[\gamma(\bar{x},t)^{(l)}\bar{x}^{(j)}],\qquad\forall k\in\mathcal{K}_n^{P_x},\tag{52}$$

and the dynamics of the first and second moments of the reduced fast system in (5) are given by

$$\mathbb{E}[\bar{z}^{(h)}]=\mathbb{E}[\gamma(\bar{x},t)^{(h)}]+(P_z-1)\mathbb{E}[\sum_{l=1}^{d_x+d_f}g_{\bullet l}(\bar{x},t)^{(h)}],\ \forall h\in\mathcal{K}_m^{P_z},\tag{53}$$

where the functions $C_{1i}(t):\mathbb{R}\to\mathbb{R}$ and $C_{2jl}(t):\mathbb{R}\to\mathbb{R}$ satisfies equation (46) in Claim 1, $P_x=\{1,\dots,\mathcal{N}\}$ where $\mathcal{N}\in\mathbb{Z}_{>0}$ and $P_z=\{1,2\}$.

Proof. To find the moment dynamics $\mathbb{E}[\bar{x}^{(k)}]$, note that the reduced slow system (4) is obtained by taking $z=\gamma(x,t)$ in (1). Thus, the dynamics for $\mathbb{E}[\bar{x}^{(k)}]$ for all $k=(k_1,\dots,k_n)\in\mathcal{K}_n^P$ can be obtained by following the proof of Claim 1 with $x=\bar{x}$ and $z=\gamma(\bar{x},t)$, which yields equation (52).

Next, to derive the moment dynamics of the reduced fast system in (5), we take the expectation of equation (5), which yields

$$\mathbb{E}[\bar{z}]=\mathbb{E}[\gamma(\bar{x},t)+g(\bar{x},t)N].\tag{54}$$

From the definition of the reduced fast system, we have that N is a vector of standard normal random variables independent of the random vector $\bar{x}$. Thus, we have that $\mathbb{E}[N]=0$ and as the expectation operator is linear, we obtain $\mathbb{E}[\bar{z}] = \mathbb{E}[\gamma(\bar{x}, t)]$, which yields equation (53) for the case where $P_z = 1$.

Next, considering the second moment of the reduced fast system we have that

$$\mathbb{E}[z_i z_j] = \mathbb{E}\left[(\gamma_i(\bar{x}, t) + \sum_{l=1}^{d_x+d_f} g_{il}(\bar{x}, t)N_l)(\gamma_j(\bar{x}, t) + \sum_{l=1}^{d_x+d_f} g_{jl}(x, t)N_l)\right].$$

Expanding further, we obtain

$$\mathbb{E}[z_i z_j] = \mathbb{E}[\gamma_i(\bar{x}, t)\gamma_j(x, t)] + \mathbb{E}\left[\gamma_i(\bar{x}, t) \sum_{l=1}^{d_x+d_f} g_{jl}(\bar{x}, t)N_l\right] + \mathbb{E}\left[\sum_{l=1}^{d_x+d_f} g_{il}(\bar{x}, t)N_l\gamma_j(\bar{x}, t)\right]$$
$$+ \mathbb{E}\left[(\sum_{l=1}^{d_x+d_f} g_{il}(\bar{x}, t)N_l)(\sum_{l=1}^{d_x+d_f} g_{jl}(\bar{x}, t)N_l))\right].$$

Since the elements of the vector N are independent standard normal random variables, we have that $\mathbb{E}[N_i] = 0$ for all i, and $\mathbb{E}[N_i N_j] = 0$ for $i \neq j$ and $\mathbb{E}[N_i N_j] = 1$ for $i = j$. Therefore, we obtain $\mathbb{E}[z_i z_j] = \mathbb{E}[\gamma_i(\bar{x}, t)\gamma_j(\bar{x}, t)] + \mathbb{E}\Big[\sum_{l=1}^{d_x+d_f} g_{il}(\bar{x}, t)g_{jl}(\bar{x}, t)\Big]$, which can be written in the form of equation (53) for the case where $P_z = 2$. □

Next, we analyze the set of moment equations obtained by setting $\epsilon = 0$ in the moment dynamics of the original system given in Claim 1.

Claim 3. *Setting $\epsilon = 0$ in the moment dynamics of the original system in (46)–(48) yields the following reduced-order system for all moments of the slow variable and up to second order moments of the fast variable:*

$$\frac{d\mathbb{E}[x^{(k)}]}{dt} = \sum_{i \in \mathcal{G}_n^P} C_{1i}(t)\mathbb{E}[x^{(i)}] + \sum_{l \in \mathcal{G}_m^1} \sum_{j \in \mathcal{G}_n^{P-1}} C_{2jl}(t)\mathbb{E}[\gamma(x, t)^{(l)}x^{(j)}], \qquad \forall k \in \mathcal{K}_n^{P_x}, \tag{55}$$

$$\mathbb{E}[z^{(h)}] = \mathbb{E}[\gamma(x, t)^{(h)}] + (P_z - 1)\mathbb{E}[\sum_{l=1}^{d_x+d_f} g_{\bullet l}(\bar{x}, t)^{(h)}], \qquad \forall h \in \mathcal{K}_m^{P_z}, \tag{56}$$

where the functions $C_{1i}(t) : \mathbb{R} \to \mathbb{R}$ and $C_{2jl}(t) : \mathbb{R} \to \mathbb{R}$ satisfies equation (46) in Claim 1, $P_x = \{1, \ldots, \mathcal{N}\}$ where $\mathcal{N} \in \mathbb{Z}_{>0}$ and $P_z = \{1, 2\}$.

Proof. From Claim 1, we note that the fast variables that appear in the slow variable dynamics in (46) are of the form $\mathbb{E}[z_i x^{(j)}]$ for $j = (k_1, \ldots, k_n) \in \mathcal{G}_n^{P-1}$ and $i = 1, \ldots, m$. Thus, we first consider setting $\epsilon = 0$ in the dynamics of $\mathbb{E}[z_i x^{(j)}]$. The dynamics of $\mathbb{E}[z_i x^{(j)}]$ can be obtained from the derivation of (48) in the proof of Claim 1, with $k_z = (c_1, \ldots, c_m) \in \mathcal{K}_m^1$ which gives $c_i = 1$ and $c_l = 0$ for $l \neq i$ for each $i = 1, \ldots, m$. Then, we have

$$\epsilon \frac{d\mathbb{E}[z_i x^{(j)}]}{dt} = \mathbb{E}\Big[f_{z_i}(x,z,t,\epsilon)x^{(j)}\Big] + \epsilon \sum_{l=1}^{n} k_l \mathbb{E}[f_{x_l}(x,z,t) z_i x_1^{k_1} \dots x_l^{k_l-1} \dots x_n^{k_n}] + \epsilon \frac{1}{2}\sum_{p=1}^{n} k_p(k_p-1)\mathbb{E}[\phi_{pp}(x,z,t) z_i x_1^{k_1} \dots x_p^{k_p-2} \dots x_n^{k_n}] + \sqrt{\epsilon}\sum_{m=1}^{n} k_m \mathbb{E}\Big[\theta_{im}(x,z,t,\epsilon) x_1^{k_1} \dots x_m^{k_m-1} \dots x_n^{k_n}\Big] + \epsilon \sum_{l=2}^{n}\sum_{p=1}^{l-1} k_l k_p \mathbb{E}[\phi_{pl}(x,z,t) z_i x_1^{k_1} \dots x_p^{k_p-1} \dots x_l^{k_l-1} \dots x_n^{k_n}], \tag{57}$$

where $\phi_{lk}(x,z,t)$ and $\theta_{il}(x,z,t,\epsilon)$ are the elements of the matrices $\Phi(x,z,t)$ and $\Theta(x,z,t,\epsilon)$ defined in Assumption 2, where we have that $\lim_{\epsilon\to 0}\theta_{ij}(x,z,t,0) < \infty$ for $i = 1,\dots,m$ and $j = 1,\dots,n$. Thus, setting $\epsilon = 0$ in the dynamics of the vector $\mathbb{E}[zx^{(j)}]$ yields

$$\mathbb{E}[f_z(x,z,t,0)x^{(j)}] = 0. \tag{58}$$

Under Assumption 2 - 3, there exist a unique solution to equation (58), which is given by $\mathbb{E}[zx^{(j)}] = -B_2^{-1}(B_1\mathbb{E}[xx^{(j)}] + B_3(t)\mathbb{E}[x^{(j)}])$. Considering the expression for $\gamma(x,t)$ in equation (3), it follows that $\mathbb{E}[\gamma(x,t)x^{(j)}] = \mathbb{E}[-B_2^{-1}(B_1 x + B_3(t))x^{(j)}] = \mathbb{E}[zx^{(j)}]$. Thus, it can be seen that setting $\epsilon = 0$ in (57), we obtain

$$\mathbb{E}[z_i x^{(j)}] = \mathbb{E}[\gamma_i(x,t)x^{(j)}], \qquad i \in \{1,\dots,m\},\ j = (k_1,\dots,k_n) \in \mathcal{G}_n^{P-1}. \tag{59}$$

Then, substituting (59) in (46) yields the set of equations (57) for the moments of the slow variable x.

Considering the equation (59) with $P = 1$, we also obtain that $\mathbb{E}[z_i] = \mathbb{E}[\gamma_i(x,t)]$, which results in the equation (56) for the case where $P_z = 1$.

Next, we consider setting $\epsilon = 0$ in the second order moment dynamics of z in (47) given by the case where $g \in K_m^2$. We denote these moments by $\mathbb{E}[z_i z_j]$ for $i,j = 1,\dots,m$, for which, the dynamics can be obtained from the derivation of (47) in the proof of Claim 1, taking $g = (g_1,\dots,g_m) \in K_m^2$ with $g_i = 1, g_j = 1$ and $g_l = 0$ for all $l \neq i,j$. Then, representing the second moments of z in matrix form we have

$$\epsilon \frac{d\mathbb{E}[zz^T]}{dt} = \mathbb{E}[zf_z(x,z,t,\epsilon)^T] + \mathbb{E}[f_z(x,z,t,\epsilon)z^T] + \mathbb{E}[\sigma_z(x,z,t,\epsilon)\sigma_z(x,z,t,\epsilon)^T]. \tag{60}$$

Then, setting $\epsilon = 0$ in the equation (60) together with Assumptions 1 - 2, yields

$$\mathbb{E}[zx^T]B_1^T + \mathbb{E}[zz^T]B_2^T + \mathbb{E}[z]B_3(t)^T + B_1\mathbb{E}[xz^T] \\ + B_2\mathbb{E}[zz^T] + B_3(t)\mathbb{E}[z^T] + \Lambda(\mathbb{E}[x], \mathbb{E}[z], t, 0) = 0. \tag{61}$$

From equation (59), we can write $\mathbb{E}[z] = \mathbb{E}[\gamma(x,t)]$ and $\mathbb{E}\big[zx^T\big] = \mathbb{E}\big[\gamma(x,t)x^T\big]$ for the case where $P = 1$ and $P = 2$, which can then be used in (61) to obtain

$$\mathbb{E}[zz^T]B_2^T + B_2\mathbb{E}[zz^T] = -\mathbb{E}\big[\gamma(x,t)x^T\big]B_1^T - \mathbb{E}[\gamma(x,t)]B_3(t)^T \\ - B_1\mathbb{E}\big[\gamma(x,t)x^T\big]^T - B_3(t)\mathbb{E}[\gamma(x,t)]^T - \Lambda(\mathbb{E}[x], \mathbb{E}[\gamma(x,t)], t, 0). \tag{62}$$

The equation (62) is in the form of the Lyapunov equation

$$A^T P + PA = -Q,$$

with

$$P = \mathbb{E}[zz^T],$$
$$Q = -\mathbb{E}\big[\gamma(x,t)x^T\big]B_1^T - \mathbb{E}[\gamma(x,t)]B_3(t)^T - B_1\mathbb{E}\big[\gamma(x,t)x^T\big]^T \\ - B_3(t)\mathbb{E}[\gamma(x,t)]^T - \Lambda(\mathbb{E}[x], \mathbb{E}[\gamma(x,t)], t, 0),$$
$$A = B_2^T.$$

From Assumption 3, we have that the matrix B_2 is Hurwitz, and therefore, there exists a unique solution for $\mathbb{E}\big[zz^T\big]$ in the equation (62). Thus, to prove that the solution to (62) is in the form of

$$\mathbb{E}[zz^T] = \mathbb{E}\big[\gamma(x,t)\gamma(x,t)^T + g(x,t)g(x,t)^T\big] \tag{63}$$

given by (56) for the case where $P_z = 2$ we substitute (63) in (62), which yields

$$\mathbb{E}[\gamma(x,t)\gamma(x,t)^T]B_2^T + \mathbb{E}[g(x,t)g(x,t)^T]B_2^T \\ + B_2\mathbb{E}[\gamma(x,t)\gamma(x,t)^T] + B_2\mathbb{E}[g(x,t)g(x,t)^T] = \\ - \mathbb{E}\big[\gamma(x,t)x^T\big]B_1^T - \mathbb{E}[\gamma(x,t)]B_3(t)^T - B_1\mathbb{E}\big[\gamma(x,t)x^T\big]^T \\ - B_3(t)\mathbb{E}[\gamma(x,t)]^T - \Lambda(\mathbb{E}[x], \mathbb{E}[\gamma(x,t)], t, 0).$$

Simplifying further using the linearity of the expectation operator and the function $\gamma(x,t)$, and noting that $B_1x + B_3(t) = -B_2\gamma(x,t)$ from the expression for $\gamma(x,t)$ in equation (3),we have that

$$\mathbb{E}[\gamma(x,t)\gamma(x,t)^T]B_2^T + \mathbb{E}[g(x,t)g(\bar{x},t)^T]B_2^T$$
$$+ B_2\mathbb{E}[\gamma(x,t)\gamma(x,t)^T] + B_2\mathbb{E}[g(x,t)g(x,t)^T] =$$
$$\mathbb{E}\big[\gamma(x,t)\gamma(x,t)^T\big]B_2^T + B_2\mathbb{E}\big[\gamma(x,t)\gamma(x,t)^T\big]$$
$$- \Lambda(\mathbb{E}[x], \gamma(\mathbb{E}[x],t), t, 0).$$

Canceling the common terms on both sides yields the expression

$$\mathbb{E}[g(x,t)g(x,t)^T]B_2^T + B_2\mathbb{E}[g(x,t)g(x,t)^T] = -\Lambda(\mathbb{E}[x], \gamma(\mathbb{E}[x],t), t, 0),$$

which is satisfied by the definition of the function $g(x,t)$ in (6). Thus, we have that setting $\epsilon = 0$ in the moments of the original system yields the equation (56) for $P_z = 2$, i.e. for the second order moments of z. □

Lemma 1. *Consider the original system in (1)–(2), the reduced system in (4)–(5) , the moment dynamics of the original system in (46)–(48) and the moment dynamics of the reduced system in (52). We have that, under Assumptions 1 - 3, the commutative diagram in Fig. 7 holds.*

Proof. Proof follows from Claim 1, Claim 2 and Claim 3. □

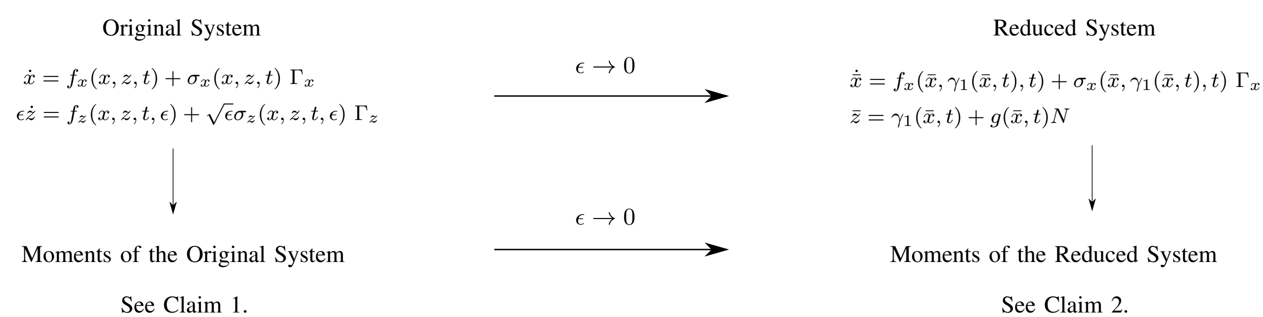

Fig. 7: Setting $\epsilon = 0$ in the moment dynamics of the original system yields the moment dynamics of the reduced system.

Proof of Theorem 1: From the commutative diagram in Lemma 1, it follows that setting $\epsilon = 0$ in the moment dynamics of the original system yields the moment dynamics of the reduced-order system for up to second order moments of the fast variable and for all moments of the slow variable. Therefore, we can apply Tikhonov's theorem to the moment dynamics of the original system in (46)–(48) to obtain the result (7)–(8).

We first prove that the assumptions of the Tikhonov's theorem are satisfied. To this end, we begin by considering the assumption on the global exponential stability of the boundary layer dynamics of the system (46)–(48). From equation (46), it follows that the fast variables that appear in (46) are of the form $\mathbb{E}[z_i x^{(j)}]$ for $j \in \mathcal{G}_n^{P-1}$. Thus, we define the vector $b_j = [b_{1j}, \ldots, b_{mj}]^T$ for the boundary layer variable for $\mathbb{E}[z_i x^{(j)}]$ where $b_{ij} = \mathbb{E}[z_i x^{(j)}] - \mathbb{E}[\gamma_i(x,t)x^{(j)}]$ for $j \in \mathcal{G}_n^{P-1}$ and $i = 1, \ldots, m$. Furthermore, in Theorem 1, we only consider up to the second order moments of the fast variable. Thus, we define the matrix V as the boundary layer variable for $\mathbb{E}[z_l z_k]$ where the elements of V are given by $v_{lk} = \mathbb{E}[z_l z_k] - \mathbb{E}\Big[\gamma_l(x,t)\gamma_k(x,t) + \sum_{h=1}^{d_x+d_f} g(x,t)_{lh} g(x,t)_{kh}\Big]$, for $l, k = 1, \ldots, m$. Then the dynamics of the variable b_{ij} and v_{lk} are given by

$$\frac{db_{ij}}{dt} = \frac{d\mathbb{E}[z_i x^{(j)}]}{dt} - \frac{d\mathbb{E}[\gamma_i(x,t)x^{(j)}]}{dt},$$
$$\frac{dv_{lk}}{dt} = \frac{d\mathbb{E}[z_l z_k]}{dt} - \frac{d\mathbb{E}\Big[\gamma_l(x,t)\gamma_k(x,t) + \sum_{h=1}^{d_x+d_f} g(x,t)_{lh} g(x,t)_{kh}\Big]}{dt}.$$

Let $\tau = t/\epsilon$ be the time variable in the fast time-scale. Then we have that

$$\frac{db_{ij}}{d\tau} = \epsilon\frac{d\mathbb{E}[z_i x^{(j)}]}{dt} - \epsilon\frac{d\mathbb{E}[\gamma_i(x,t)x^{(j)}]}{dt},$$
$$\frac{dv_{lk}}{d\tau} = \epsilon\frac{d\mathbb{E}[z_l z_k]}{dt} - \epsilon\frac{d\mathbb{E}\Big[\gamma_l(x,t)\gamma_k(x,t) + \sum_{h=1}^{d_x+d_f} g(x,t)_{lh} g(x,t)_{kh}\Big]}{dt}.$$

Since from (3) we have that $\gamma_i(x,t)$ is a linear function of x, and since $j \in \mathcal{G}_n^{P-1}$ we have that $x^{(j)}$ contains moments of order up to $P-1$. Therefore, it follows that $\gamma_i(x,t)x^{(j)}$ can be written in terms of P^{th} or lower order moments of x and $\gamma_l(x,t)\gamma_k(x,t)$ consists of up to second order moments of x . Furthermore, from (6) we have that $g(x,t)g(x,t)^T$ is a matrix whose elements are linear functions of x. Therefore, for appropriate functions $Q_k(t)$ for $k \in \mathcal{G}_n^P$, $Z_r(t)$ for $r \in \mathcal{G}_n^2$ and employing the linearity of the differential operator, we can write

$$\frac{db_{ij}}{d\tau} = \epsilon\frac{d\mathbb{E}[z_i x^{(j)}]}{dt} - \epsilon\sum_{k\in\mathcal{G}_n^P} Q_k(t)\frac{d\mathbb{E}[x^{(k)}]}{dt},$$
$$\epsilon\frac{dv_{lk}}{dt} = \epsilon\frac{d\mathbb{E}[z_l z_k]}{dt} - \epsilon\sum_{r\in\mathcal{G}^2} Z_r(t)\frac{d\mathbb{E}[x^{(r)}]}{dt}.$$

Substituting from (46) and using the expansions of $d\mathbb{E}[z_i x^{(j)}]/dt$, $d\mathbb{E}[z_k z_k]/dt$, (see proof of Claim 3), yields

$$\frac{db_{ij}}{d\tau} = \mathbb{E}\Big[f_{z_i}(x,z,t,\epsilon)x^{(j)}\Big]$$

$$+ \epsilon \sum_{l=1}^{n} j_l \mathbb{E}[f_{x_l}(x,z,t) z_i x_1^{j_1} \dots x_l^{j_l - 1} \dots x_n^{j_n}]$$
$$+ \epsilon \frac{1}{2} \sum_{p=1}^{n} j_p (j_p - 1) \mathbb{E}[\phi_{pp}(x,z,t) z_i x_1^{j_1} \dots x_p^{j_p - 2} \dots x_n^{j_n}]$$
$$+ \sqrt{\epsilon} \sum_{m=1}^{n} j_m \mathbb{E}\Big[\theta_{im}(x,z,t,\epsilon) x_1^{j_1} \dots x_m^{j_m - 1} \dots x_n^{j_n}\Big]$$
$$+ \epsilon \sum_{m=2}^{n} \sum_{p=1}^{m-1} j_m j_p \mathbb{E}[\phi_{pm}(x,z,t) z_i x_1^{j_1} \dots x_p^{j_p - 1} \dots x_m^{j_m - 1} \dots x_n^{j_n}]$$
$$- \epsilon \sum_{s=0}^{P} \sum_{k \in \mathcal{K}^s} Q_k(t) \Bigg(\sum_{i \in \mathcal{N}^P} C_{1i}(t) \mathbb{E}[\bar{x}^{(i)}] + \sum_{l \in \mathcal{C}^1} \sum_{j \in \mathcal{N}^{P-1}} C_{2jl}(t) \mathbb{E}[\gamma(\bar{x},t)^{(l)} \bar{x}^{(j)}] \Bigg), \tag{64}$$
$$\frac{dv_{lk}}{dt} = \mathbb{E}[z_l f_{z_k}(x,z,t,\epsilon)] + \mathbb{E}[z_k f_{z_l}(x,z,t,\epsilon)] + \mathbb{E}[\lambda_{lk}(x,z,t,\epsilon)] - \epsilon \sum_{r \in \mathcal{N}^2} Z_r(t) \frac{d\mathbb{E}[x^{(r)}]}{dt}. \tag{65}$$

where $\phi_{ij}(x,z,t)$, $\theta_{lj}(x,z,t,\epsilon)$, $\lambda_{lk}(x,z,t,\epsilon)$ for $i,j = 1,\dots n$ and $l,k = 1,\dots,m$ are the elements of the matrices $\Phi(x,z,t)$, $\Theta(x,z,t,\epsilon)$ and $\Lambda(x,z,t,\epsilon)$ given in Assumption 2.

We obtain the boundary layer system for the vector b_j, by setting $\epsilon = 0$ in the equation (64), with $\mathbb{E}[zx^{(j)}] = b_j + \mathbb{E}[\gamma(x,t)x^{(j)}]$. Under Assumptions 1 - 2, the functions $f_x(x,z,t)$, $f_z(x,z,t,\epsilon)$, $\phi_{ij}(x,z,t)$, $\theta_{lj}(x,z,t,\epsilon)$, $Q_k(t)$, $C_{1i}(t)$, $C_{2jl}(t)$ are continuous and therefore are bounded on any compact interval $t = [0,t_1]$. We also have that $\theta_{lj}(x,z,t,0) < \infty$, from Assumption 2. Additionally, as the system (46)–(48) is linear, the solutions $\mathbb{E}[x^{(i)}]$ and $\mathbb{E}[\gamma(x,t)^{(l)}x^{(j)}]$ exist and are bounded on any compact interval $t = [0,t_1]$. Therefore, the boundary layer dynamics b_j are given by $\frac{db_j}{d\tau} = \mathbb{E}\big[f_z(x,z,t,0)x^{(j)}\big]\Big|_{\mathbb{E}[zx^{(j)}] = b_j + \mathbb{E}[\gamma(x,t)x^{(j)}]}$. Together with Assumption 1 and the definition of $\gamma(x,t)$ in equation (3), we then obtain

$$\begin{aligned}\frac{db_j}{d\tau} &= B_1 \mathbb{E}[xx^{(j)}] + B_2(b_j + \mathbb{E}[\gamma(x,t)x^{(j)}]) + B_3(t)\mathbb{E}[x^{(j)}] \\ &= B_2 b_j.\end{aligned} \tag{66}$$

In order to obtain the boundary layer dynamics for V, we first represent the elements of b_j where $j = (k_1,\dots,k_n) \in \mathcal{G}_n^{P-1}$ for the case $P = 1$ as a vector $d = \mathbb{E}[z] - \mathbb{E}[\gamma(x,t)]$ and for the case $P = 2$ as a matrix $E = \mathbb{E}\big[zx^T\big] - \mathbb{E}[\gamma(x,t)x^T]$. Then, the boundary layer dynamics for the matrix V can be obtained by setting $\epsilon = 0$ in (65) and taking $\mathbb{E}[zx^T] = E + \mathbb{E}[\gamma(x,t)x^T]$ and $\mathbb{E}[z] = d + \mathbb{E}[\gamma(x,t)]$, which yields

$$\frac{dV}{d\tau} = VB_2^T + B_2V^T + EB_1^T + B_1E^T + B_3(t)d^T + dB_3(t)^T. \tag{67}$$

Under Assumption 3, we have that the origin is a globally exponential stable equilibrium point of the boundary layer dynamics b_j in (66). Next, to determine the stability of the boundary layer dynamics V, we consider the solution of (67) for V given by [39]

$$\begin{aligned} V(\tau) =& e^{B_2\tau}V(0)e^{B_2^T\tau} + \int_0^\tau e^{B_2(\tau-v)}(E(v)B_1^T + B_1E(v)^T \\ & + B_3(t)d(v)^T + d(v)B_3(t)^T)(e^{B_2(\tau-v)})^T dv. \end{aligned}$$

Then, considering the solutions for E and d, which can be obtained from (66), and are in the form $E(\tau) = E(0)e^{B_2\tau}$ and $d(\tau) = d(0)e^{B_2\tau}$, and using that B_2 is Hurwitz under Assumption 3, it follows that there exists positive constants C_1 and r_1 such that $\|V(\tau)\|_F \leq C_1(\|d(0)\|_F + \|E(0)\|_F + \|V(0)\|_F)e^{-r_1\tau}$, where $\|.\|_F$ denotes the Frobenius norm. Then, taking $Y = [d \mid E \mid V]$, and considering the exponential stability of E and d, we can write $\|Y(\tau)\|_F \leq C\|Y(0)\|_F e^{-r\tau}$ for positive constants C and r. Thus, we have that the origin is a globally exponentially stable equilibrium point of the boundary layer dynamics V.

Furthermore, we ensure that the additional assumptions of the Tikhonov's theorem also hold. We have that $C_{1i}(t), C_{2lj}(t)$ are continuous functions with respect to time and that the functions $D_{1a}(t,\epsilon), D_{2bc}(t,\epsilon), F_{1i}(t,\epsilon), F_{2a}(t,\epsilon), F_{3qrks}(t,\epsilon)$ in (46)–(48) and their partial derivatives with respect to t and ϵ are continuous, from Claim 1. Due to the linearity of the function $\gamma(x,t)$ and $g(x,t)g(x,t)^T$, we have that the function $\mathbb{E}[\gamma(x,t)x^{(j)}]$ for $j \in \mathcal{G}_n^{P-1}$ has continuous first partial derivatives with respect to its arguments $\mathbb{E}[x^{(k)}]$ for $k \in \mathcal{G}_n^P$ and the function $\mathbb{E}\left[\gamma(x,t)\gamma(x,t)^T + g(x,t)g(x,t)^T\right]$ has continuous first partial derivatives with respect to its arguments $\mathbb{E}[x^{(k)}]$ for $k \in \mathcal{G}_n^2$. Furthermore, we have that the system (52) has a unique solution on a compact time interval $t \in [0, t_1]$, due to its linearity. Hence, the assumptions of the Tikhonov's theorem on a finite time interval are satisfied and applying the theorem to the moment dynamics of the original system in (46)–(48), yields the desired result in (7)–(8).

Appendix B

Here, we show the derivation on the moment dynamics of the variable y using the slow and fast variable approximations. First, we derive the first and second moment dynamics of the variables v and c of the reduced system in (27)–(30) as

$$\frac{d\mathbb{E}[v]}{d\bar{t}} = \frac{X}{k_{d1}} - (1-R)\mathbb{E}[v], \tag{68}$$

$$\mathbb{E}[c] = \frac{\mathbb{E}[v]\beta_1 p_{T0}}{\delta_1(p_T + k_{d2})}, \tag{69}$$

$$\frac{d\mathbb{E}\big[v^2\big]}{d\bar{t}} = 2\frac{X}{k_{d1}}\mathbb{E}[v] - 2(1-R)\mathbb{E}\big[v^2\big] + \frac{\delta_1\frac{X}{k_{d1}}}{\beta_1 p_{T0}\Omega} + \frac{\delta_1(1-R)\mathbb{E}[v]}{\beta_1 p_{T0}\Omega}, \tag{70}$$

$$\mathbb{E}[c^2] = \frac{\mathbb{E}\big[v^2\big]\beta_1^2 p_{T0}^2}{\delta_1^2(p_T + k_{d2})^2} + \frac{\mathbb{E}[v]\beta_1 p_{T0} k_{d2}}{\delta_1 p_T \Omega (p_T + k_{d2})^2}, \tag{71}$$

$$\mathbb{E}[vc] = \frac{\mathbb{E}\big[v^2\big]\beta_1 p_{T0}}{\delta_1(p_T + k_{d2})}. \tag{72}$$

Since $y = v - \frac{p_T\delta_1}{\beta_1 p_{T0}}c$, we can write the dynamics for the first moment of y as $\frac{d\mathbb{E}[y]}{d\bar{t}} = \frac{d\mathbb{E}[v]}{d\bar{t}} - \frac{p_T\delta_1}{\beta_1 p_{T0}}\frac{d\mathbb{E}[c]}{d\bar{t}}$. Then, using the chain rule we obtain $\frac{d\mathbb{E}[y]}{d\bar{t}} = \left(1 - \frac{p_T\delta_1}{\beta_1 p_{T0}}\frac{d\mathbb{E}[c]}{d\mathbb{E}[v]}\right)\frac{d\mathbb{E}[v]}{d\bar{t}}$, and with $\frac{d\mathbb{E}[c]}{d\mathbb{E}[v]} = \frac{\beta_1 p_{T0}}{\delta_1(p_T + k_{d2})}$ from (69), we can write

$$\frac{d\mathbb{E}[y]}{d\bar{t}} = \left(1 - \frac{p_T}{(p_T + k_{d2})}\right)\frac{d\mathbb{E}[v]}{d\bar{t}}. \tag{73}$$

Considering the dynamics for the second moment of y, we have $\frac{d\mathbb{E}[y^2]}{d\bar{t}} = \frac{d\mathbb{E}[(v - \frac{p_T\delta_1}{\beta_1 p_{T0}}c)^2]}{d\bar{t}}$. Substituting the fast variable approximation in (30) we obtain

$$\frac{d\mathbb{E}[y^2]}{d\bar{t}} = \frac{d\mathbb{E}\left[\left(v - \frac{v p_T}{(p_T + k_{d2})} + \sqrt{\frac{v\delta_1 p_T k_{d2}}{\beta_1 p_{T0}\Omega(p_T + k_{d2})^2}}N_2\right)^2\right]}{d\bar{t}}.$$

Since N_2 is a normal random variable independent of y with $\mathbb{E}[N_2] = 0$ and $\mathbb{E}[N_2^2] = 1$, we have that

$$\frac{d\mathbb{E}[y^2]}{d\bar{t}} = \frac{k_{d2}^2}{(p_T + k_{d2})^2}\frac{d\mathbb{E}\big[v^2\big]}{d\bar{t}} + \frac{\delta_1 p_T k_{d2}}{\beta_1 p_{T0}\Omega(p_T + k_{d2})^2}\frac{d\mathbb{E}[v]}{d\bar{t}}. \tag{74}$$

Substituting in (73) and (74) the expressions for the dynamics of $\mathbb{E}[v]$ and $\mathbb{E}[v^2]$ given in (68) and (70), yields the first and second moment dynamics of y as

$$\frac{d\mathbb{E}[y]}{d\bar{t}} = \left(1 - \frac{p_T}{(p_T + k_{d2})}\right)\left(\frac{X}{k_{d1}} - (1-R)\mathbb{E}[v]\right), \tag{75}$$

$$\begin{aligned}\frac{d\mathbb{E}[y^2]}{d\bar{t}} = \frac{k_{d2}^2}{(p_T + k_{d2})^2}\Bigg(2\frac{X}{k_{d1}}\mathbb{E}[v] - 2(1-R)\mathbb{E}\big[v^2\big] + \frac{\delta_1\frac{X}{k_{d1}}}{\beta_1 p_{T0}\Omega} \\ + \frac{\delta_1(1-R)\mathbb{E}[v]}{\beta_1 p_{T0}\Omega}\Bigg) + \frac{\delta_1 p_T k_{d2}}{\beta_1 p_{T0}\Omega(p_T + k_{d2})^2}\left(\frac{X}{k_{d1}} - (1-R)\mathbb{E}[v]\right).\end{aligned} \tag{76}$$

Next, in order to write the equations (75)–(76) in terms of $\mathbb{E}[y]$ and $\mathbb{E}[y^2]$, we use the moments of the fast variable approximation in (69) and (72) to express $\mathbb{E}[v]$ and $\mathbb{E}[v^2]$ in terms of $\mathbb{E}[y]$ and $\mathbb{E}[y^2]$. Towards this end, first consider $\mathbb{E}[v] = \mathbb{E}[y] + \frac{p_T\delta_1}{\beta_1 p_{T0}}\mathbb{E}[c]$. Using the expression for $\mathbb{E}[c]$ from (69) we have

$$\mathbb{E}[v] = \frac{(k_{d2} + p_T)}{k_{d2}} \mathbb{E}[y]. \tag{77}$$

Considering the second moments, we obtain

$$\begin{aligned} \mathbb{E}\big[v^2\big] &= \mathbb{E}\bigg[(y + \frac{p_T \delta_1}{\beta_1 p_{T0}} c)^2\bigg] \\ &= \mathbb{E}\big[y^2\big] + 2\frac{p_T \delta_1}{\beta_1 p_{T0}} \mathbb{E}[yc] + \frac{p_T^2 \delta_1^2}{\beta_1^2 p_{T0}^2} \mathbb{E}\big[c^2\big], \end{aligned} \tag{78}$$

$$\mathbb{E}[vc] = \mathbb{E}\bigg[(y + \frac{p_T \delta_1}{\beta_1 p_{T0}} c)c\bigg] = \mathbb{E}[yc] + \frac{p_T \delta_1}{\beta_1 p_{T0}} \mathbb{E}\big[c^2\big]. \tag{79}$$

Using $\mathbb{E}[yc] = \mathbb{E}[vc] - \frac{p_T\delta_1}{\beta_1 p_{T0}}\mathbb{E}[c^2]$ from (79) in (78) we obtain $\mathbb{E}[v^2] = \mathbb{E}[y^2] + 2\frac{p_T\delta_1}{\beta_1 p_{T0}}\mathbb{E}[vc] - \frac{p_T^2\delta_1^2}{\beta_1^2 p_{T0}^2}\mathbb{E}[c^2]$. Then, substituting for $\mathbb{E}[vc]$ and $\mathbb{E}[c^2]$ from (71)–(72), yields $\mathbb{E}[v^2] = \mathbb{E}[y^2] + 2\frac{\mathbb{E}[v^2]p_T}{(p_T+k_{d2})} - \frac{\mathbb{E}[v^2]p_T^2}{(p_T+k_{d2})^2} - \frac{\mathbb{E}[v]\delta_1 p_T k_{d2}}{\Omega\beta_1 p_{T0}(p_T+k_{d2})^2}$. Simplifying further and using the expression for $\mathbb{E}[v]$ from (77) yields

$$\mathbb{E}\big[v^2\big] = \frac{(k_{d2} + p_T)^2}{k_{d2}^2} \mathbb{E}\big[y^2\big] - \frac{\delta_1 (k_{d2} + p_T) p_T \mathbb{E}[y]}{k_{d2}^2 \Omega \beta_1 p_{T0}}. \tag{80}$$

Then, substituting (77) and (80) in (75)–(76) and simplifying further yields the moment dynamics of the mixed variable given in (39)–(40).

APPENDIX C

Evaluating the magnitude of the frequency response of the system (42) - (44), we find that

$$\frac{\sqrt{|\mathbb{E}[\Delta g_S^2](j\bar{\omega})|}}{|g_R(j\bar{\omega})|\sqrt{k_2}} = \sqrt{\frac{\delta_2^2\delta_1^2 + \delta_1^4\bar{\omega}^2}{k_2\ \Omega}}\sqrt[4]{A(p_T, \bar{\omega})}, \tag{81}$$

where the function $A(p_T, \bar{\omega})$ is in the form

$$A(p_T, \bar{\omega}) = \frac{\big((1-R)^2 + \bar{\omega}^2\big)(N_1 + N_2 + N_3 + N_4)}{(\beta_1 p_{T0}/k_{d1})^2 \beta_2^2 \delta_1^2 p_T^2 D_1 D_2}, \tag{82}$$

with

$$\begin{aligned} N_1 =& -2\delta_1^3(R-1)\bar{\omega}^2(k_{d2}+p_T)(2\beta_2 p_T + \delta_2(k_{d2}+p_T)), \\ N_2 =& \delta_1^2\big(4\beta_2^2 p_T^2\bar{\omega}^2 + 8\beta_2\delta_2 p_T\bar{\omega}^2(k_{d2}+p_T)\big) + \delta_1^2\big(\delta_2^2(k_{d2}+p_T)^2\big(4R^2 - 8R + 5\bar{\omega}^2 + 4\big)\big), \\ N_3 =& -8\delta_1\delta_2^2(R-1)(k_{d2}+p_T)(\beta_2 p_T + \delta_2(k_{d2}+p_T)), \\ N_4 =& 4\delta_2^2(\beta_2 p_T + \delta_2(k_{d2}+p_T))^2 + \delta_1^4\bar{\omega}^2(k_{d2}+p_T)^2\big(R^2 - 2R + \bar{\omega}^2 + 1\big), \\ D_1 =& \big(\delta_1^4\bar{\omega}^4 + 5\delta_1^2\delta_2^2\bar{\omega}^2 + 4\delta_2^4\big), \\ D_2 =& \big(\delta_1^2\big(R^2 - 2R + \bar{\omega}^2 + 1\big) - 2\delta_1\delta_2(R-1) + \delta_2^2\big). \end{aligned}$$

To identify the change in $A(p_T,\bar{\omega})$ with p_T, we consider the derivative of $A(p_T,\bar{\omega})$ with respect to p_T. Evaluating the derivative at $\bar{\omega}=0$, yields

$$\frac{\partial A(p_T,\bar{\omega})}{\partial p_T}\bigg|_{\bar{\omega}=0} = -\frac{2k_{d2}^2\big(N_{d1}+N_{d2}+N_{d3}+\delta_1^3k_{d2}^3\big)}{(\beta_1 p_{T0}/k_{d1})^2\beta_2^2\delta_1^2\delta_2^2p_T^3(\delta_1k_{d2}+\delta_2(k_{d2}+p_T))^3},$$

where

$$\begin{aligned}
N_{d1} &= \delta_2\big(\beta_2^2p_T^3+\beta_2\delta_2p_T\big(k_{d2}^2+3k_{d2}p_T+2p_T^2\big)+\delta_2^2(k_{d2}+p_T)^3\big),\\
N_{d2} &= \delta_1^2k_{d2}^2(\beta_2p_T+3\delta_2(k_{d2}+p_T)),\\
N_{d3} &= \delta_1\delta_2k_{d2}\big(\beta_2p_T(2k_{d2}+3p_T)+3\delta_2(k_{d2}+p_T)^2\big).
\end{aligned}$$

Thus, it can be see that the derivative is negative for all parameter conditions. We note that the function $A(p_T,\bar{\omega})$ is a rational polynomial function in p_T and is continuous with respect to $\bar{\omega}$. Thus, we have that $\frac{\partial A(p_T,\bar{\omega})}{\partial p_T}$ is continuous with respect to $\bar{\omega}$ and thus will remain negative in a neighborhood of $\bar{\omega}=0$. Therefore, the function $A(p_T,\bar{\omega})$ is decreasing with p_T for sufficiently small $\bar{\omega}$.

ACKNOWLEDGMENT

We would like to thank Dr. Mohammad Naghnaeian for helpful discussions and suggestions in writing the paper. We also thank Dr. Abdullah Hamadeh for the contribution to initial work on this project.

REFERENCES

[1] N. Herath, A. Hamadeh, and D. Del Vecchio. Model reduction for a class of singularly perturbed differential equations. In *Proc. of Amer. Control Conf.*, pages 4404 – 4410, 2015.

[2] N. Herath and D. Del Vecchio. Moment convergence in a class of singularly perturbed stochastic differential equations. In *Proc. of Aust. Control Conf.*, pages 43–47, 2015.

[3] N. Herath and D. Del Vecchio. Model reduction for a class of singularly perturbed stochastic differential equations : Fast variable approximation. In *Proc. of Amer. Control Conf.*, pages 3674 – 3679, 2016.

[4] T.P. Barnett. The interaction of multiple time scales in the tropical climate system. *Journal of Climate*, 4(3):269–285, 1991.

[5] J. R. Winkelman, J. H. Chow, J. J. Allemong, and P. V. Kokotovic. Multi-time-scale analysis of a power system. *Automatica*, 16(1):35–43, 1980.

[6] Uri Alon. *An introduction to systems biology: design principles of biological circuits*. CRC press, 2006.

[7] H. K. Khalil. *Nonlinear systems*, volume 3. Prentice hall Upper Saddle River, 2002.

[8] P. Kokotović, H. K. Khalil, and J. O'Reilly. *Singular perturbation methods in control: analysis and design*, volume 25. Siam, 1999.

[9] G. Pavliotis and A. Stuart. *Multiscale methods: averaging and homogenization*, volume 53. Springer Science & Business Media, 2008.

[10] D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical finance*, 6(4):379–406, 1996.

[11] E. Allen. *Modeling with Itô stochastic differential equations*. Springer, 2007.

[12] D. T. Gillespie. The chemical langevin equation. *J. Chem. Phys.*, 113(1):297–306, 2000.

[13] Y. Kabanov and S. Pergamenshchikov. *Two-scale stochastic systems: asymptotic analysis and control*, volume 49. Springer, 2003.

[14] N. Berglund and B. Gentz. *Noise-induced phenomena in slow-fast dynamical systems: a sample-paths approach*. Springer, 2006.

[15] C. Tang and T. Başar. Stochastic stability of singularly perturbed nonlinear systems. In *Proc. of the 40th IEEE Conf. on Decis. and Control*, volume 1, pages 399–404, 2001.

[16] R. Z. Khasminskii. On the principle of averaging the itô stochastic differential equations. *Kybernetika (Prague)*, 4:260–279, 1968.

[17] F. Wu, T. Tian, J. B. Rawlings, and G. Yin. Approximate method for stochastic chemical kinetics with two-time scales by chemical langevin equations. *J. Chem. Phys.*, 144(17):174112, 2016.

[18] M. Contou-Carrere, V. Sotiropoulos, Y. N. Kaznessis, and P. Daoutidis. Model reduction of multi-scale chemical langevin equations. *Systems & Control Letters*, 60(1):75–86, 2011.

[19] B. Øksendal. *Stochastic differential equations*. Springer, 2003.

[20] L. Socha. *Linearization methods for stochastic dynamic systems*, volume 730. Springer Science & Business Media, 2008.

[21] A. Singh and J. P. Hespanha. Approximate moment dynamics for chemically reacting systems. *IEEE Trans. Autom. Control*, 56(2):414–418, 2011.

[22] Wei-Chau Xie. *Dynamic stability of structures*. Cambridge University Press, 2006.

[23] D. J. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM review*, 43(3):525–546, 2001.

[24] D. Del Vecchio, A. J. Ninfa, and E. D. Sontag. Modular cell biology: retroactivity and insulation. *Molecular systems biology*, 4(1), 2008.

[25] S. Jayanthi, K. Nilgiriwala, and D. Del Vecchio. Retroactivity controls the temporal dynamics of gene transcription. *ACS Synth. Biol.*, 2(8):431–441, 2013.

[26] S. S. Andrews, T. Dinh, and A. P. Arkin. Stochastic models of biological processes. In *Encyclopedia of Complexity and Systems Science*, pages 8730–8749. Springer, 2009.

[27] D. A. McQuarrie. Stochastic approach to chemical kinetics. *Journal of applied probability*, 4(3):413–478, 1967.

[28] D. Del Vecchio and R. M. Murray. *Biomolecular Feedback Systems*. Princeton University Press, 2014.

[29] A. Gyorgy and D. Del Vecchio. Modular composition of gene transcription networks. *PLoS Comput Biol*, 10(3):e1003486, 2014.

[30] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434, 1976.

[31] J. M. Raser and E.K. O'Shea. Noise in gene expression: Origins, consequences, and control. *Science*, 309(5743):2010–2013, 2005.

[32] P.S. Swain, M.B. Elowitz, and E.D. Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. In *Proc. Natl Acad. Sci. USA 99*, 2002.

[33] S. Jayanthi and D. Del Vecchio. On the compromise between retroactivity attenuation and noise amplification in gene regulatory networks. In *Proc. of IEEE Conf. on Decis. and Control*, pages 4565 – 4571, 2009.

[34] J. Wilkie and Y. M. Wong. Positivity preserving chemical langevin equations. *Chemical Physics*, 353(1):132–138, 2008.

[35] D. Schnoerr, G. Sanguinetti, and R. Grima. The complex chemical langevin equation. *J. Chem. Phys.*, 141(2):024103, 2014.

[36] L. Szpruch and D. J. Higham. Comparing hitting time behavior of markov jump processes and their diffusion approximations. *Multiscale Modeling & Simulation*, 8(2):605–621, 2010.

[37] N. Herath and D. Del Vecchio. Hitting time behavior for the solution of a stochastic differential equation. http://hdl.handle.net/1721.1/104333, 2016. [Online].

[38] M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, 2000.

[39] H. Kwakernaak and R. Sivan. *Linear optimal control systems*, volume 1. Wiley-interscience New York, 1972.