Model reduction for a class of singularly perturbed stochastic differential equations

Narmada Herath[1], Abdullah Hamadeh[2] and Domitilla Del Vecchio[3]

***Abstract*— A class of singularly perturbed stochastic differential equations (SDE) with linear drift and nonlinear diffusion terms is considered. We prove that, on a finite time interval, the trajectories of the slow variables can be well approximated by those of a system with reduced dimension as the singular perturbation parameter becomes small. In particular, we show that when this parameter becomes small the first and second moments of the reduced system's variables closely approximate the first and second moments, respectively, of the slow variables of the singularly perturbed system. Chemical Langevin equations describing the stochastic dynamics of molecular systems with linear propensity functions including both fast and slow reactions fall within the class of SDEs considered here. We therefore illustrate the goodness of our approximation on a simulation example modeling a well known biomolecular system with fast and slow processes.**

I. INTRODUCTION

Many dynamical systems evolve on multiple timescales. Examples include chemical reactions, climate systems and electrical systems. This separation in time-scales allows the state variables to be categorized into "fast" and "slow" variables. Such systems can be modeled using the singular perturbation framework, where a small parameter ϵ is introduced to capture the separation between the time-scales. The standard approach in analyzing such systems is to approximate the original system by a system with reduced dimension. For deterministic systems, this process is well established, most notably by the Tikhonov's Theorem [1], which quantifies the error between the trajectories of the original and reduced systems in terms of ϵ.

Several works have extended this approach to systems modeled by stochastic differential equations. Kabanov and Pergamenshchikov provide a stochastic version of the Tikhonov's theorem for systems where the diffusion term in the fast variable differential equation is of $o(\sqrt{\epsilon}/|\sqrt{\ln(\epsilon)}|)$ [2]. Their results show that when the time-scale separation becomes large (ϵ becomes small), the reduced system trajectories converge in probability to those of the original system, under the standard singular perturbation assumption that the slow manifold is exponentially stable [1]. Reference [2] also discusses another class of systems where the diffusion term in the fast variable dynamics is on the order of $\sqrt{\epsilon}$. For this type of systems, it is remarked that the fast variable may be oscillatory and it may not converge in probability as ϵ tends to zero. A study by Berglund and Gentz provides an approximation for the slow variable, which is defined only during the time interval in which the fast variable is within a neighborhood of the slow manifold, and the length of this time interval is upper bounded by $O(\epsilon)$ when the diffusion term is on the order of $\sqrt{\epsilon}$ [3]. In [4], Kokotović et al. developed a singular perturbation approach for linear stochastic differential equations in which the diffusion term can be scaled by ϵ but is otherwise constant. In [5], Tang and Başar approach the problem of singularly perturbed stochastic systems using the notion of stochastic input-to-state stability, but only the error for the fast variable is quantified.

In this work, we consider a class of singularly perturbed stochastic differential equations with linear drift and nonlinear diffusion terms, as often found in biomolecular systems modeled by chemical Langevin equations [6]. In these models the diffusion term in the fast variable differential equation is on the order of $\sqrt{\epsilon}$. Hence, the above works cannot be used to approximate the dynamics of the slow variable on a finite time interval. The objective of this paper is to obtain a reduced model that provides an approximation to the slow variable dynamics of the original system including the case where the diffusion term of the fast variable is on the order of $\sqrt{\epsilon}$. To quantify the error in the approximation, we use the first and second moments of the slow variable as they provide a quantification of the mean and the variance of a random variable. We first prove that the moments dynamics of the original singularly perturbed system are also in singular perturbation form. This allows the application of Tikhonov's theorem to the moments dynamics on a finite time interval. This way, we show that the first and second moments of the reduced system are within an $O(\epsilon)$ neighborhood of the first and second moments of the slow variable in the original system. Next, we demonstrate how the results can be applied to biomolecular systems with multiple time-scales modeled using the chemical Langevin equation. A simulation example corresponding to the model of a phosphorylation cycle is also provided.

This paper is organized as follows. In Section II, the system considered is introduced. In Section III, we obtain a reduced model and in Section IV we show that the reduced model is an $O(\epsilon)$-approximation of the original system. In Section V, the results are applied to a biomolecular system modeled by the chemical Langevin equation.

*This work was in part funded by AFOSR grant # FA9550-12-1-0129 and NIGMS grant P50 GMO98792.

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

[2,3]Abdullah Hamadeh and Domitilla Del Vecchio are with the Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Mass. Ave, Cambridge MA ahamadeh@mit.edu, ddv@mit.edu

II. System Model

Consider the following set of stochastic differential equations in the singular perturbation form

$$\Sigma : \begin{cases} \dot{x} = f_x(x,z,t) + \sigma_x(x,z,t)\Gamma_x, & x(0) = x_0 \quad (1) \\ \epsilon\dot{z} = f_z(x,z,t,\epsilon) + \sigma_z(x,z,t,\epsilon)\Gamma_z, & z(0) = z_0 \quad (2) \end{cases}$$

where $x \in \mathbb{R}^n$ is the slow variable, $z \in \mathbb{R}^m$ 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. Then, Γ_z is a (d_x+d_f)-dimensional white noise process that can be expressed in the form $[\ \Gamma_x \quad \Gamma_f\ ]^T$. The functions $f_x : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^n$, $f_z : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^m$, $\sigma_x : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^{n \times d_x}$ and $\sigma_z : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{m \times (d_x+d_f)}$ are continuous functions in their arguments and satisfy the Lipschitz and bounded growth conditions [7] for the existence of a unique solution to system (1) - (2). In this work, we assume that the system Σ satisfies the following assumptions.

Assumption 1. The functions $f_x(x,z,t)$ and $f_z(x,z,t,\epsilon)$ are linear functions of the state variables x and z, i.e., we can write

$$f_x(x,z,t) = A_1 x + A_2 z + A_3(t), \tag{3}$$

where $A_1 \in \mathbb{R}^{n\times n}$, $A_2 \in \mathbb{R}^{n\times m}$ and $A_3(t) \in \mathbb{R}^n$ and there exists a scalar function $\alpha(\epsilon)$ such that,

$$\begin{aligned} f_z(x,z,t,\epsilon) = {} & B_1 x + B_2 z + B_3(t) \\ & + \alpha(\epsilon)(B_4 x + B_5 z + B_6(t)), \end{aligned} \tag{4}$$

where $B_1, B_4 \in \mathbb{R}^{m\times n}$, $B_2, B_5 \in \mathbb{R}^{m\times m}$, $B_3(t), B_6(t) \in \mathbb{R}^n$ and $\alpha(0) = 0$.

Assumption 2. The matrix-valued functions $\sigma_x(x,z,t)$ and $\sigma_z(x,z,t,\epsilon)$ are such that

1) There exists a matrix-valued function $\Phi(x,z,t) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$ that satisfies

$$\sigma_x(x,z,t)\sigma_x(x,z,t)^T = \Phi(x,z,t), \tag{5}$$

where the elements of the matrix-valued function $\Phi(x,z,t)$ are affine in x and z, i.e., we can write $\mathbb{E}[\Phi(x,z,t)] = \Phi(\mathbb{E}[x], \mathbb{E}[z], t)$.

2) There exists a matrix-valued function $\Lambda(x,z,t,\epsilon)$: $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{m\times m}$ that satisfies

$$\sigma_z(x,z,t,\epsilon)\sigma_z(x,z,t,\epsilon)^T = \epsilon\Lambda(x,z,t,\epsilon), \tag{6}$$

where the elements of the matrix-valued function $\Lambda(x,z,t,\epsilon)$ are affine in x and z, i.e., we can write $\mathbb{E}[\Lambda(x,z,t,\epsilon)] = \Lambda(\mathbb{E}[x], \mathbb{E}[z], t, \epsilon)$. Also, we have that $\lim_{\epsilon\to 0}\Lambda(x,z,t,\epsilon)$ is finite for all x, z and t.

3) There exists a matrix-valued function $\Theta(x,z,t,\epsilon)$: $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{m\times n}$, that satisfies

$$\sigma_z(x,z,t,\epsilon)[\ \sigma_x(x,z,t) \quad 0\ ]^T = \Theta(x,z,t,\epsilon), \tag{7}$$

where the elements of the matrix-valued function $\Theta(x,z,t,\epsilon)$ are affine in x and z, i.e., we can write $\mathbb{E}[\Theta(x,z,t,\epsilon)] = \Theta(\mathbb{E}[x], \mathbb{E}[z], t, \epsilon)$. Also, we have that $\lim_{\epsilon\to 0}\Theta(x,z,t,\epsilon) = 0$ for all x, z and t.

Assumption 3. The matrix B_2 is Hurwitz.

Assumption 4. The derivative of $B_3(t)$ with respect to t is continuous.

We will next analyze equations (1) - (2) in two ways. In Section III-A we will first take $\epsilon = 0$ to arrive at a reduced model of (1) - (2) and then derive the moment equations of this reduced model. In Section III-B we derive the moments equations of (1) - (2) and then take $\epsilon = 0$, to obtain reduced moment equations. By comparing the systems obtained via these two procedures and applying Tikhonov's Theorem on the finite time interval, Theorem 1 in Section IV will show that the moments dynamics of the reduced system of (1) - (2) are an $O(\epsilon)$-approximation of the moments dynamics of (1) - (2).

III. Preliminary Results

A. Reduced System

When $\epsilon = 0$ in system (1) - (2) we obtain

$$f_z(x,z,t,0) = B_1 x + B_2 z + B_3(t) = 0 \tag{8}$$

since, from (6), we have that $\sigma_z(x,z,t,0)\sigma_z(x,z,t,0)^T = 0$ and therefore, $\sigma_z(x,z,t,0) = 0$.

Assumption 3 ensures the existence of a unique global solution $z = \gamma_1(x,t)$ to (8), given by

$$\gamma_1(x,t) = -B_2^{-1}(B_1 x + B_3(t)). \tag{9}$$

By substituting $z = \gamma_1(x,t)$ into (1), we obtain the *reduced system*

$$\dot{\bar{x}} = f_x(\bar{x}, \gamma_1(\bar{x},t), t) + \sigma_x(\bar{x}, \gamma_1(\bar{x},t), t)\Gamma_x, \ \ \bar{x}(0) = x_0. \tag{10}$$

Next, we derive the first and second moment equations of the reduced system (10). To this end, let us define the functions γ_2, g_1, and g_2 for $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^{n\times n}$ such that

$$\gamma_2(a,b,t) = -B_2^{-1}(B_1 b + B_3(t)a^T), \tag{11}$$

$$g_1(a, \gamma_1(a,t), t) = A_1 a + A_2\gamma_1(a,t) + A_3(t), \tag{12}$$

$$\begin{aligned} g_2(a,b,\gamma_1(a,t),\gamma_2(a,b,t),t) = {} & A_1 b + A_2\gamma_2(a,b,t) + bA_1^T \\ & + A_3(t)a^T + \gamma_2(a,b,t)^T A_2^T + aA_3(t)^T + \Phi(a,\gamma_1(a,t),t). \end{aligned} \tag{13}$$

We can now make the following claim:

Claim 1. *The first and second moments dynamics of the reduced system (10) can be written in the form*

$$\frac{d\mathbb{E}[\bar{x}]}{dt} = g_1(\mathbb{E}[\bar{x}], \gamma_1(\mathbb{E}[\bar{x}],t), t), \quad \mathbb{E}[\bar{x}(0)] = x_0, \tag{14}$$

$$\begin{aligned} \frac{d\mathbb{E}[\bar{x}\bar{x}^T]}{dt} = {} & \\ & g_2(\mathbb{E}[\bar{x}], \mathbb{E}[\bar{x}\bar{x}^T], \gamma_1(\mathbb{E}[\bar{x}],t), \gamma_2(\mathbb{E}[\bar{x}], \mathbb{E}[\bar{x}\bar{x}^T], t), t), \\ & \mathbb{E}[\bar{x}(0)\bar{x}(0)^T] = x_0 x_0^T. \end{aligned} \tag{15}$$

Proof. As in [8], we can write the moments dynamics of the reduced system (10) as

$$\frac{d\mathbb{E}[\bar{x}]}{dt} = \mathbb{E}[f_x(\bar{x}, \gamma_1(\bar{x}, t), t)], \tag{16}$$

$$\frac{d\mathbb{E}[\bar{x}\bar{x}^T]}{dt} = \mathbb{E}[f_x(\bar{x}, \gamma_1(\bar{x}, t), t)\bar{x}^T] + \mathbb{E}[\bar{x}f_x(\bar{x}, \gamma_1(\bar{x}, t), t)^T] + \mathbb{E}[\sigma_x(x, z, t)\sigma_x(x, z, t)^T]. \tag{17}$$

Using the linearity of the expectation operator and equation (9), we have that

$$\mathbb{E}[\gamma_1(\bar{x}, t)] = -B_2^{-1}(B_1\mathbb{E}[\bar{x}] + B_3(t)) = \gamma_1(\mathbb{E}[\bar{x}], t), \tag{18}$$

$$\begin{aligned}\mathbb{E}[\gamma_1(\bar{x}, t)\bar{x}^T] &= -B_2^{-1}(B_1\mathbb{E}[\bar{x}\bar{x}^T] + B_3(t)\mathbb{E}[x^T]), \\ &= \gamma_2(\mathbb{E}[\bar{x}], \mathbb{E}[\bar{x}\bar{x}^T], t).\end{aligned} \tag{19}$$

Since $\mathbb{E}[\bar{x}\gamma_1(\bar{x}, t)^T] = (\mathbb{E}[\gamma_1(\bar{x}, t)\bar{x}^T])^T$, we can also write $\mathbb{E}[\bar{x}\gamma_1(\bar{x}, t)^T] = \gamma_2(\mathbb{E}[\bar{x}], \mathbb{E}[\bar{x}\bar{x}^T], t)^T$. The initial condition x_0 is deterministic, and therefore, we have that $\mathbb{E}[\bar{x}(0)] = x_0$ and $\mathbb{E}[\bar{x}(0)\bar{x}(0)^T] = x_0x_0^T$. Then, under Assumptions 1 - 2, and the function definitions (12) - (13), the moments dynamics (16) - (17) can be written in the form (14) - (15). □

B. Moments Dynamics of the Original System Σ

In this section we analyze the moments dynamics of the system (1) - (2). To this end, let us define the functions g_3, g_4, g_5 for $a \in \mathbb{R}^n$, $b \in \mathbb{R}^{n\times n}$, $c \in \mathbb{R}^m$, $d \in \mathbb{R}^{m\times m}$, and $e \in \mathbb{R}^{m\times n}$ such that

$$\begin{aligned}&g_3(a, c, t, \epsilon) = B_1a + B_2c + B_3(t) \\ &+ \alpha(\epsilon)(B_4a + B_5c + B_6(t)),\end{aligned} \tag{20}$$

$$\begin{aligned}&g_4(a, b, c, d, e, t, \epsilon) = eB_1^T + dB_2^T + cB_3(t)^T \\ &+ \alpha(\epsilon)(eB_4^T + dB_5^T + cB_6(t)^T) + B_1e^T + \Lambda(a, c, t, \epsilon) \\ &+ B_2d + B_3(t)c^T + \alpha(\epsilon)(B_4e^T + B_5d + B_6(t)c^T),\end{aligned} \tag{21}$$

$$\begin{aligned}&g_5(a, b, c, d, e, t, \epsilon) = \epsilon(eA_1^T + dA_2^T + cA_3(t)^T) + B_1b \\ &+ B_2e + B_3(t)a^T + \Theta(a, c, t, \epsilon) \\ &+ \alpha(\epsilon)(B_4e^T + B_5d + B_6(t)c^T).\end{aligned} \tag{22}$$

Then, we make the following claim :

Claim 2. *The first and second moment equations for Σ in (1) - (2), can be written in the singular perturbation form:*

$$\frac{d\mathbb{E}[x]}{dt} = \mathbb{E}[f_x(x, z, t)] = g_1(\mathbb{E}[x], \mathbb{E}[z], t), \tag{23}$$

$$\begin{aligned}\frac{d\mathbb{E}[xx^T]}{dt} &= \mathbb{E}[f_x(x, z, t)x^T] + \mathbb{E}[xf_x(x, z, t), t)^T] \\ &+ \mathbb{E}[\sigma_x(x, z, t)\sigma_x(x, z, t)^T] \\ &= g_2(\mathbb{E}[x], \mathbb{E}[xx^T], \mathbb{E}[z], \mathbb{E}[zx^T], t),\end{aligned} \tag{24}$$

$$\epsilon\frac{d\mathbb{E}[z]}{dt} = \mathbb{E}[f_z(x, z, t, \epsilon)] = g_3(\mathbb{E}[x], \mathbb{E}[z], t, \epsilon), \tag{25}$$

$$\begin{aligned}\epsilon\frac{d\mathbb{E}[zz^T]}{dt} &= \mathbb{E}[zf_z(x, z, t, \epsilon)^T] + \mathbb{E}[f_z(x, z, t, \epsilon)z^T] \\ &+ \frac{1}{\epsilon}\mathbb{E}[\sigma_z(x, z, t, \epsilon)\sigma_z(x, z, t, \epsilon)^T] \\ &= g_4(\mathbb{E}[x], \mathbb{E}[xx^T], \mathbb{E}[z], \mathbb{E}[zz^T], \mathbb{E}[zx^T], t, \epsilon),\end{aligned} \tag{26}$$

$$\begin{aligned}\epsilon\frac{d\mathbb{E}[zx^T]}{dt} &= \epsilon\mathbb{E}[zf_x(x, z, t)^T] + \mathbb{E}[f_z(x, z, t, \epsilon)x^T] \\ &+ \mathbb{E}[\sigma_z(x, z, t, \epsilon)[\ \sigma_x(x, z, t) \quad 0\ ]^T] \\ &= g_5(\mathbb{E}[x], \mathbb{E}[xx^T], \mathbb{E}[z], \mathbb{E}[zz^T], \mathbb{E}[zx^T], t, \epsilon).\end{aligned} \tag{27}$$

where g_1 and g_2 are defined in (12) and (13), respectively, and g_3, g_4 and g_5 are defined in (20) - (22).

Proof. Note that system Σ in (1) - (2) can be written in the form

$$\dot{x} = f_x(x, z, t) + [\ \sigma_x(x, z, t) \quad 0\ ]\Gamma_z, \quad x(0) = x_0 \tag{28}$$

$$\epsilon\dot{z} = f_z(x, z, t, \epsilon) + \sigma_z(x, z, t, \epsilon)\Gamma_z, \qquad z(0) = z_0 \tag{29}$$

where $[\ \sigma_x(x, z, t) \quad 0\ ]$ is a matrix-valued function $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^{n\times(d_x+d_f)}$. Then, we write the dynamics for the first moments [8] as

$$\frac{d\mathbb{E}[x]}{dt} = \mathbb{E}[f_x(x, z, t)], \quad \frac{d\mathbb{E}[z]}{dt} = \frac{1}{\epsilon}\mathbb{E}[f_z(x, z, t, \epsilon)]. \tag{30}$$

Similarly, from Proposition III.1 in [8], the second moments dynamics are given by

$$\begin{aligned}\frac{d}{dt}\mathbb{E}\begin{bmatrix} xx^T & xz^T \\ zx^T & zz^T \end{bmatrix} &= \begin{bmatrix} xf_x(x, z, t)^T & \frac{1}{\epsilon}xf_z(x, z, t, \epsilon)^T \\ zf_x(x, z, t)^T & \frac{1}{\epsilon}zf_z(x, z, t, \epsilon)^T \end{bmatrix} \\ &+ \begin{bmatrix} f_x(x, z, t)x^T & f_x(x, z, t)z^T \\ \frac{1}{\epsilon}f_z(x, z, t, \epsilon)x^T & \frac{1}{\epsilon}f_z(x, z, t, \epsilon)z^T \end{bmatrix} \\ &+ \begin{bmatrix} \sigma_x(x, z, t)\sigma_x(x, z, t)^T & \frac{1}{\epsilon}[\ \sigma_x(x, z, t) \quad 0\ ]\sigma_z(x, z, t, \epsilon)^T \\ \frac{1}{\epsilon}\sigma_z(x, z, t, \epsilon)[\ \sigma_x(x, z, t) \quad 0\ ]^T & \frac{1}{\epsilon^2}\sigma_z(x, z, t, \epsilon)\sigma_z(x, z, t, \epsilon)^T \end{bmatrix}.\end{aligned} \tag{31}$$

Summing the corresponding entries of the matrices in (31), multiplying both sides by ϵ, using Assumptions 1 - 2 and the linearity of the expectation operator, we obtain the expressions

$$\frac{d\mathbb{E}[x]}{dt} = A_1\mathbb{E}[x] + A_2\mathbb{E}[z] + A_3(t), \tag{32}$$

$$\begin{aligned}\frac{d\mathbb{E}[xx^T]}{dt} &= A_1\mathbb{E}[xx^T] + A_2\mathbb{E}[zx^T] + A_3(t)\mathbb{E}[x^T] \\ &+ \mathbb{E}[xx^T]A_1^T + \mathbb{E}[xz^T]A_2^T + \mathbb{E}[x]A_3(t)^T \\ &+ \Phi(\mathbb{E}[x], \mathbb{E}[z], t),\end{aligned} \tag{33}$$

$$\begin{aligned}\epsilon\frac{d\mathbb{E}[z]}{dt} &= B_1\mathbb{E}[x] + B_2\mathbb{E}[z] + B_3(t) \\ &+ \alpha(\epsilon)(B_4\mathbb{E}[x] + B_5\mathbb{E}[z] + B_6(t)),\end{aligned} \tag{34}$$

$$\begin{aligned}\epsilon\frac{d\mathbb{E}[zz^T]}{dt} &= \mathbb{E}[zx^T]B_1^T + \mathbb{E}[zz^T]B_2^T + \mathbb{E}[z]B_3(t)^T \\ &+ \alpha(\epsilon)(\mathbb{E}[zx^T]B_4^T + \mathbb{E}[zz^T]B_5^T + \mathbb{E}[z]B_6(t)^T) \\ &+ B_1\mathbb{E}[xz^T] + B_2\mathbb{E}[zz^T] + B_3(t)\mathbb{E}[z^T] \\ &+ \alpha(\epsilon)(B_4\mathbb{E}[xz^T] + B_5\mathbb{E}[zz^T] + B_6(t)\mathbb{E}[z^T]) + \\ &\frac{1}{\epsilon}\epsilon\Lambda(\mathbb{E}[x], \mathbb{E}[z], t, \epsilon),\end{aligned} \tag{35}$$

$$\begin{aligned}\epsilon\frac{d\mathbb{E}[xz^T]}{dt} &= \epsilon(A_1\mathbb{E}[xz^T] + A_2\mathbb{E}[zz^T] + A_3(t)\mathbb{E}[z^T]) \\ &+ \mathbb{E}[xx^T]B_1^T + \mathbb{E}[xz^T]B_2^T + \mathbb{E}[x]B_3(t)^T\end{aligned}$$

$$
\begin{aligned}
&+\alpha(\epsilon)(\mathbb{E}[xx^T]B_4^T+\mathbb{E}[xz^T]B_5^T+\mathbb{E}[x]B_6(t)^T) \\
&+\Theta(\mathbb{E}[x],\mathbb{E}[z],t,\epsilon)^T, \qquad (36)
\end{aligned}
$$

$$
\begin{aligned}
\epsilon\frac{d\mathbb{E}[zx^T]}{dt} &= \epsilon(\mathbb{E}[zx^T]A_1^T+\mathbb{E}[zz^T]A_2^T+\mathbb{E}[z]A_3(t)^T) \\
&+B_1\mathbb{E}[xx^T]+B_2\mathbb{E}[zx^T]+B_3(t)\mathbb{E}[x^T] \\
&+\alpha(\epsilon)(B_4\mathbb{E}[xx^T]+B_5\mathbb{E}[zx^T]+B_6(t)\mathbb{E}[x^T]) \\
&+\Theta(\mathbb{E}[x],\mathbb{E}[z],t,\epsilon). \qquad (37)
\end{aligned}
$$

Note that $\mathbb{E}[xz^T]=(\mathbb{E}[zx^T])^T$. Therefore, eliminating equation (36), we can write the system (32) - (37) in the singular perturbation form in (23) - (27). □

Next, we derive the reduced model for the system (23) - (27), when $\epsilon=0$.

Claim 3. *The reduced system of moments equations (23) - (27), obtained when $\epsilon=0$, can be written in the form*

$$
\frac{d\overline{\mathbb{E}[x]}}{dt}=g_1(\overline{\mathbb{E}[x]},\gamma_1(\overline{\mathbb{E}[x]},t),t), \quad \overline{\mathbb{E}[x(0)]}=x_0, \qquad (38)
$$

$$
\begin{aligned}
&\frac{d\overline{\mathbb{E}[xx^T]}}{dt}= \\
&g_2(\overline{\mathbb{E}[x]},\overline{\mathbb{E}[xx^T]},\gamma_1(\overline{\mathbb{E}[x]},t),\gamma_2(\overline{\mathbb{E}[x]},\overline{\mathbb{E}[xx^T]},t),t), \\
&\overline{\mathbb{E}[x(0)x(0)^T]}=x_0x_0^T, \qquad (39)
\end{aligned}
$$

where g_1 is defined in (12) and g_2 is defined in (13).

Proof. From Claim 2, setting $\epsilon=0$ in system (25) - (27), under Assumption 1, we obtain the equations

$$
\begin{aligned}
&B_1\mathbb{E}[x]+B_2\mathbb{E}[z]+B_3(t)=0, \qquad (40)\\
&B_1\mathbb{E}[xx^T]+B_2\mathbb{E}[zx^T]+B_3(t)\mathbb{E}[x^T]=0. \qquad (41)
\end{aligned}
$$

We do not consider the dynamics of the fast variable $\mathbb{E}[zz^T]$ as they do not appear in the slow variable dynamics. Under Assumption 3, we have that $\mathbb{E}[z]=h_1(\mathbb{E}[x],t)$ and $\mathbb{E}[zx^T]=h_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)$ are the unique global solutions to (40) and (41), given by

$$
\begin{aligned}
&h_1(\mathbb{E}[x],t)=-B_2^{-1}[B_1\mathbb{E}[x]+B_3(t)], \qquad (42)\\
&h_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)=-B_2^{-1}[B_1\mathbb{E}[xx^T]+B_3(t)\mathbb{E}[x^T]]. \qquad (43)
\end{aligned}
$$

By substituting $\mathbb{E}[z]=h_1(\mathbb{E}[x],t)$ and $\mathbb{E}[zx^T]=h_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)$ in (23) - (24) we obtain the reduced system

$$
\frac{d\overline{\mathbb{E}[x]}}{dt}=g_1(\overline{\mathbb{E}[x]},h_1(\overline{\mathbb{E}[x]},t),t), \qquad (44)
$$

$$
\begin{aligned}
&\frac{d\overline{\mathbb{E}[xx^T]}}{dt}= \\
&g_2(\overline{\mathbb{E}[x]},\overline{\mathbb{E}[xx^T]},h_1(\overline{\mathbb{E}[x]},t),h_2(\overline{\mathbb{E}[x]},\overline{\mathbb{E}[xx^T]},t),t). \qquad (45)
\end{aligned}
$$

From equation (9), we have that $\gamma_1(\mathbb{E}[x],t)=-B_2^{-1}(B_1\mathbb{E}[x]+B_3(t))$, and comparing this expression with (42) it follows that $h_1(\overline{\mathbb{E}[x]},t)=\gamma_1(\overline{\mathbb{E}[x]},t)$. Therefore, equation (44) takes the form of (38). Comparison of (11) and (43) shows that $h_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)=\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)$, and therefore equation (45) takes the form of (39).

□

IV. Main Results

Lemma 1. *Consider the original system Σ given in (1) - (2), the reduced system in (10), the moments dynamics of the original system in (23) - (27), and moments dynamics of the reduced system in (14) - (15). Then, under Assumptions 1 - 3, the commutative diagram in Fig. 1 holds.*

Proof. Setting $\epsilon=0$ in the original system given in (1) - (2), we obtain the reduced system (10). Then taking the moment dynamics of the reduced system we obtain the system (14) - (15).

Using Claim 2, we have that the moments dynamics of the original system can be written in the singular perturbation form given in (23) - (27). Then from Claim 3, it follows that setting $\epsilon=0$ in the moments dynamics of the original system in (23) - (27), we obtain the system given in (38) - (39), which is equal to the system of moments dynamics of the reduced system given by (14) - (15).

□

Definition 1. Consider the original system Σ defined in (1) - (2) and the reduced system in (10). We say that the reduced system (10) is an $O(\epsilon)$-approximation of the original system (1) - (2) if there exists $t_1\geq 0$ such that

$$
\begin{aligned}
&\|\mathbb{E}[\bar{x}(t)]-\mathbb{E}[x(t)]\|=O(\epsilon),\ t\in[0,t_1],\\
&\|\mathbb{E}[\bar{x}(t)\bar{x}(t)^T]-\mathbb{E}[x(t)x(t)^T]\|_F=O(\epsilon),\ t\in[0,t_1],
\end{aligned} \qquad (46)
$$

where $\|.\|_F$ is the Frobenius norm.

From Lemma 1, we see that setting $\epsilon=0$ in the moment dynamics of Σ leads to the moments dynamics of the reduced system. However, it does not guarantee that the trajectories of the moments $\mathbb{E}[x]$, $\mathbb{E}[xx^T]$ will approach the trajectories of $\mathbb{E}[\bar{x}]$, $\mathbb{E}[\bar{x}\bar{x}^T]$ as $\epsilon\to 0$. Therefore, the following theorem proves that under Assumptions 1 - 4, the moments $\mathbb{E}[x]$ and $\mathbb{E}[xx^T]$ approach $\mathbb{E}[\bar{x}]$ and $\mathbb{E}[\bar{x}\bar{x}^T]$, respectively, as $\epsilon\to 0$, and therefore the reduced system is a good approximation of the slow variable dynamics of the original system.

Theorem 1. *Under Assumptions 1 - 4, the reduced system (10) is an $O(\epsilon)$-approximation of the original system (1) - (2).*

Proof. Consider the commutative diagram in Lemma 2. We see that the moments dynamics of the reduced system can be obtained by setting $\epsilon=0$ in the moments dynamics of the original system. As the moments dynamics are deterministic, under Assumption 3, we can apply Tikhonov's theorem on the finite time interval [1] to the moments dynamics of the original system given in (23) - (27) to obtain (46). To do so, we prove that under Assumption 3 the boundary layer dynamics for the system (23) - (27) are globally exponentially stable. To this end, define the variables

$$
\begin{aligned}
&b_1:=\mathbb{E}[z]-\gamma_1(\mathbb{E}[x],t),\\
&b_2:=\mathbb{E}[zx^T]-\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t),
\end{aligned}
$$

where $\gamma_1(\mathbb{E}[x],t)$ and $\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)$ are solutions to the equations $g_3(\mathbb{E}[x],\mathbb{E}[z],t,0)=0$

Original System Σ

$$\begin{aligned}\dot{x} &= f_x(x,z,t) + \sigma_x(x,z,t)\ \Gamma_x \\ \epsilon\dot{z} &= f_z(x,z,t,\epsilon) + \sigma_z(x,z,t,\epsilon)\ \Gamma_z\end{aligned} \xrightarrow{\epsilon \to 0}$$

Reduced System

$$\dot{\bar{x}} = f_x(\bar{x},\gamma_1(\bar{x},t),t) + \sigma_x(\bar{x},\gamma_1(\bar{x},t),t)\ \Gamma_x$$

$\downarrow$ Moments of the Original System Σ $\xrightarrow{\epsilon \to 0}$ Moments of the Reduced System $\downarrow$

$$\frac{d}{dt}\begin{pmatrix} \mathbb{E}[x] \\ \mathbb{E}[xx^T] \\ \epsilon\mathbb{E}[z] \\ \epsilon\mathbb{E}[zz^T] \\ \epsilon\mathbb{E}[zx^T] \end{pmatrix} = \begin{pmatrix} g_1(\mathbb{E}[x],\mathbb{E}[z],t) \\ g_2(\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z],\mathbb{E}[zx^T],t) \\ g_3(\mathbb{E}[x],\mathbb{E}[z],t,\epsilon) \\ g_4(\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z],\mathbb{E}[zz^T],\mathbb{E}[zx^T],t,\epsilon) \\ g_5(\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z],\mathbb{E}[zz^T],\mathbb{E}[zx^T],t,\epsilon) \end{pmatrix}$$

$$\frac{d}{dt}\begin{pmatrix} \mathbb{E}[\bar{x}] \\ \mathbb{E}[\bar{x}\bar{x}^T] \end{pmatrix} = \begin{pmatrix} g_1(\mathbb{E}[\bar{x}],\gamma_1(\mathbb{E}[\bar{x}],t),t) \\ g_2(\mathbb{E}[\bar{x}],\mathbb{E}[\bar{x}\bar{x}^T],\gamma_1(\mathbb{E}[\bar{x}],t),\gamma_2(\mathbb{E}[\bar{x}],\mathbb{E}[\bar{x}\bar{x}^T],t),t) \end{pmatrix}$$

Fig. 1: Commutative Diagram.

and $g_5(\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z],\mathbb{E}[zz^T],\mathbb{E}[zx^T],t,0) = 0$, respectively. The dynamics of the variables b_1 and b_2 are given by

$$\epsilon\frac{db_1}{dt} = \epsilon\frac{d\mathbb{E}[z]}{dt} - \epsilon\frac{\partial\gamma_1(\mathbb{E}[x],t)}{\partial t} - \epsilon\frac{\partial\gamma_1(\mathbb{E}[x],t)}{\partial\mathbb{E}[x]}\frac{d\mathbb{E}[x]}{dt}, \tag{47}$$

$$\begin{aligned}\epsilon\frac{db_2}{dt} &= \epsilon\frac{d\mathbb{E}[zx^T]}{dt} - \epsilon\frac{\partial\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial t} \\ &\quad - \epsilon\frac{\partial\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\mathbb{E}[x]}\frac{d\mathbb{E}[x]}{dt} \\ &\quad - \epsilon\frac{\partial\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\mathbb{E}[xx^T]}\frac{d\mathbb{E}[xx^T]}{dt},\end{aligned} \tag{48}$$

where $\frac{\partial\gamma_1(\mathbb{E}[x],t)}{\partial\mathbb{E}[x]}$ is the Jacobian of the function $\gamma_1(\mathbb{E}[x],t)$, $\frac{\partial\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\mathbb{E}[x]}$ is a third order tensor, and $\frac{\partial\gamma_2(\mathbb{E}[x]\mathbb{E}[xx^T],t)}{\partial\mathbb{E}[xx^T]}$ is a fourth order tensor.

Let $\tau := t/\epsilon$, be the time variable in the fast time scale. Then, using equation (25) in (47) and equation (27) in (48), the dynamics of b_1 and b_2 in the τ timescale are given by

$$\frac{db_1}{d\tau} = \mathbb{E}[f_z(x,z,t,\epsilon)] - \frac{\partial\gamma_1(\mathbb{E}[x],t)}{\partial\tau} - \frac{\partial\gamma_1(\mathbb{E}[x],t)}{\partial\mathbb{E}[x]}\frac{d\mathbb{E}[x]}{d\tau}, \tag{49}$$

$$\begin{aligned}\frac{db_2}{d\tau} &= \epsilon\mathbb{E}[zf_x(x,z,t)^T] + \mathbb{E}[f_z(x,z,t,\epsilon)x^T] \\ &\quad + \mathbb{E}[\sigma_z(x,z,t,\epsilon)[\ \sigma_x(x,z,t) \quad 0\ ]^T] \\ &\quad - \frac{\partial\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\mathbb{E}[x]}\frac{d\mathbb{E}[x]}{d\tau} - \frac{\partial\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\tau} \\ &\quad - \frac{\partial\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\mathbb{E}[xx^T]}\frac{d\mathbb{E}[xx^T]}{d\tau},\end{aligned} \tag{50}$$

where

$$\frac{d\mathbb{E}[x]}{d\tau} = \epsilon g_1(\mathbb{E}[x],\mathbb{E}[z],t), \tag{51}$$

$$\frac{d\mathbb{E}[xx^T]}{d\tau} = \epsilon g_2(\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z],\mathbb{E}[zx^T],t), \tag{52}$$

from equations (23) - (24), and using equations (9) and (11) we have

$$\frac{\partial\gamma_1(\mathbb{E}[x],t)}{\partial\tau} = -\epsilon B_2^{-1}\frac{dB_3(t)}{dt}, \tag{53}$$

$$\frac{\partial\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\tau} = -\epsilon B_2^{-1}\frac{dB_3(t)}{dt}\mathbb{E}[x^T]. \tag{54}$$

The boundary layer system is obtained by setting $\epsilon = 0$ in (49) - (50). Due to the linearity of the system (23) - (27), the solutions $\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z]$, and $\mathbb{E}[zx^T]$ exist and are bounded on any compact time interval $t \in [0,t_1]$. Therefore, when $\epsilon = 0$, we have that $d\mathbb{E}[x]/d\tau = 0$ and $d\mathbb{E}[xx^T]/d\tau = 0$ from equations (51) - (52). Using Assumption 4, we have that $dB_3(t)/dt$ is continuous on t, and therefore, it is bounded on any compact interval $t \in [0,t_1]$. Thus, it follows that setting $\epsilon = 0$ in (53) - (54), yields $\partial\gamma_1/\partial\tau = 0$ and $\partial\gamma_2/\partial\tau = 0$. Therefore under Assumption 1 - 2, we obtain the following boundary layer dynamics for b_1:

$$\begin{aligned}\frac{db_1}{d\tau} &= B_1\mathbb{E}[x] + B_2b_1 + B_2(-B_2^{-1}(B_1\mathbb{E}[x] + B_3(t))) \\ &\quad + B_3(t) = B_2b_1.\end{aligned} \tag{55}$$

The dynamics for b_2 in the boundary layer system are given by

$$\begin{aligned}\frac{db_2}{d\tau} &= B_1\mathbb{E}[xx^T] + B_2b_2 + B_2(-B_2^{-1}(B_1\mathbb{E}[xx^T] \\ &\quad + B_3(t)\mathbb{E}[x^T])) + B_3(t)\mathbb{E}[x^T] = B_2b_2.\end{aligned}$$

Let $b_2 = [\ b_{21} \quad \ldots \quad b_{2n}\ ]$. Then, we can write the dynamics of b_2 as n systems given by

$$\frac{db_{2i}}{d\tau} = B_2b_{2i}, \qquad \text{for } i = 1,\ldots,n. \tag{56}$$

From (55), (56), and Assumption 3 it follows that the boundary layer system is globally exponential stable. Since the functions g_1 and g_2 are linear, the solution of the system (14) - (15) exists and is unique on any compact time interval $t \in [0,t_1]$. We also have that the functions $g_1,\ g_2,\ g_3,\ g_4$, and g_5 in Lemma 1 are continuous, and from Assumption 2 and Assumption 4, the first and the second partial derivatives of $g_1,\ g_2,\ g_3,\ g_4$, and g_5 with respect to their arguments are continuous. The functions $\gamma_1(\mathbb{E}[x],t)$ and $\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)$ also have continuous first partial derivatives with respect to their arguments. Therefore the assumptions of the Tikhonov's theorem on the finite time interval are satisfied, and relationship (46) holds. $\square$

A. Illustrative Example

From the reduced system (10) we observe that the noise in the fast variable does not affect the slow variable. To provide some insight into why this occurs, consider the system

$$\dot{x} = a_1 x + a_2 z,$$
$$\epsilon \dot{z} = -a_3 z + v_1 \sqrt{\epsilon}\Gamma,$$

where $a_1, a_2, a_3 > 0$. As the system is linear, we can calculate the second moments at the steady state using the autocorrelation function. To this end, let $S_{xx}(\omega)$, $S_{zz}(\omega)$ and $S_\Gamma(j\omega)$ be the power spectra of x, z and Γ, respectively. We first calculate $S_{zz}(\omega)$. Let $H_{z\Gamma}(j\omega)$ be the frequency response from Γ to z. Then, we have that

$$S_{zz}(\omega) = |H_{z\Gamma}(j\omega)|^2 S_\Gamma(\omega) = \frac{(v_1/\sqrt{\epsilon})^2}{(\omega^2 + (1/\epsilon)^2)},$$

where $S_\Gamma(j\omega) = 1$ since Γ is a white noise process. Then, the autocorrelation of z, $R_{zz}(\tau)$, is given by

$$R_{zz}(\tau) = \frac{(v_1/\sqrt{\epsilon})^2}{2(1/\epsilon)} e^{-(1/\epsilon)|\tau|},$$

so that the steady state second moment of z is given by

$$\mathbb{E}[z^2] = R_{zz}(0) = \frac{(v_1/\sqrt{\epsilon})^2}{2(1/\epsilon)} = \frac{v_1^2}{2}.$$

To calculate $S_{xx}(\omega)$, note that z appears as an input to the slow variable. Then, letting $H_{xz}(j\omega)$ be the frequency response from z to x, we have that

$$S_{xx}(\omega) = |H_{xz}(j\omega)|^2 S_{zz}(\omega),$$
$$= \frac{(a_2^2(v_1^2/\epsilon))/(-a_1^2 + 1/\epsilon^2)}{(\omega^2 + a_1^2)} + \frac{(a_2^2(v_1^2/\epsilon))(a_1^2 - 1/\epsilon^2)}{(\omega^2 + (1/\epsilon)^2)}.$$

Thus, the autocorrelation of x is given by

$$R_{xx}(\tau) = \frac{a_2^2(v_1^2/\epsilon)}{(-a_1^2 + \frac{1}{\epsilon^2})(2a_1)} e^{-a_1|\tau|} + \frac{a_2^2(v_1^2/\epsilon)}{(a_1^2 - \frac{1}{\epsilon^2})(\frac{2}{\epsilon})} e^{-\frac{1}{\epsilon}|\tau|},$$

so that the second moment of x is given by

$$\mathbb{E}[x^2] = R_{xx}(0) = \frac{a_2^2 v_1^2}{2a_1} \frac{\epsilon}{(1 + a_1\epsilon)}.$$

Therefore as $\epsilon \to 0$, $\mathbb{E}[x^2] \to 0$, that is, noise Γ does not affect the x-subsystem as ϵ tends to zero. This can be explained by visualizing the power spectrum $S_{zz}(\omega)$ and the frequency response $|H_{xz}(j\omega)|$. Consider $S_{zz}(\omega)$, illustrated in Fig. 2. We see that as $\epsilon \to 0$, $S_{zz}(\omega)$ at low frequencies decreases while at high frequencies it increases. However, $H_{xz}(j\omega)$ is a low-pass filter with a cut-off frequency of a_1 that does not change with ϵ (Fig. 2). Therefore, the x-subsystem selects only the low frequency components of z, which decrease with ϵ, leading to a decrease in power of signal x as ϵ decreases.

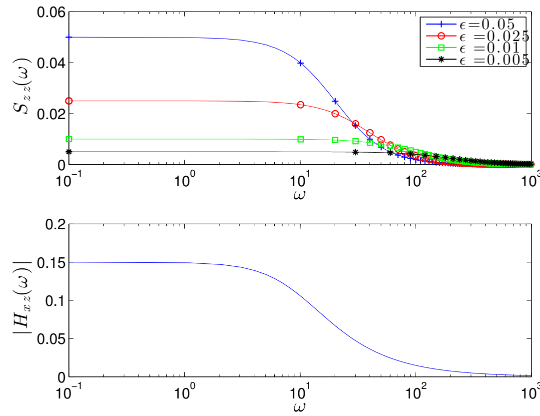

Fig. 2: Power spectrum of z and frequency response from z to x. The parameters used are $a_1 = 10$, $a_2 = 1.5$ and $v_1 = 1$.

V. Application to the Chemical Langevin Equation

In this section, we demonstrate how the results developed in this paper can be applied to analyze the stochastic properties of a biomolecular system. Biological systems are inherently stochastic due to randomness in chemical reactions. There are several tools that have been developed to capture the stochastic behavior of biological systems, such as the chemical master equation, the Fokker-Planck equation and the chemical Langevin equation [9], [6]. The chemical master equation and the Fokker-Planck equation provide descriptions for the time-evolution of the probability density function of the system state while the chemical Langevin equation is a stochastic differential equation that provides an approximation to the chemical master equation under the assumption that the number of molecules in the system is large [9], [6].

Time-scale separation is a common feature in biomolecular systems where reactions can be categorized as slow reactions and fast reactions according to the reaction rate. A biomolecular system with such time-scale separation can be represented by a chemical Langevin equation given by

$$\begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} S_S & 0 \\ S_{F1} & S_{F2} \end{bmatrix} \begin{bmatrix} w_S(x,z,t) \\ \frac{1}{\epsilon} w_F(x,z,t) \end{bmatrix} + \begin{bmatrix} S_S & 0 \\ S_{F1} & S_{F2} \end{bmatrix} \begin{bmatrix} \text{diag}(\sqrt{w_S(x,z,t)})\Gamma_s \\ \text{diag}(\frac{1}{\sqrt{\epsilon}}\sqrt{w_F(x,z,t)})\Gamma_f \end{bmatrix}, \quad (57)$$

where $x \in \mathbb{R}^{N_S}$ and $z \in \mathbb{R}^{N_F}$ are the slow and fast variables, respectively with $x(0) = x_0$ and $z(0) = z_0$, $S = \begin{bmatrix} S_S & 0 \\ S_{F1} & S_{F2} \end{bmatrix}$ is the stoichiometry matrix with $S_S \in \mathbb{R}^{N_S \times M_S}$, $S_{F1} \in \mathbb{R}^{N_F \times M_S}$, $S_{F2} \in \mathbb{R}^{N_F \times M_F}$, $W(x,z) = \begin{bmatrix} w_S(x,z,t) \\ \frac{1}{\epsilon} w_F(x,z,t) \end{bmatrix}$ is the vector of reactions with the slow reaction rates $w_S : \mathbb{R}^{N_S} \times \mathbb{R}^{N_F} \to \mathbb{R}^{M_S}$, the fast reaction rates $w_F : \mathbb{R}^{N_S} \times \mathbb{R}^{N_F} \to \mathbb{R}^{M_F}$, Γ_s and Γ_f are white noise processes.

This system has the structure of system (1) - (2) with

$$\begin{aligned} f_x(x,z,t) &= S_S w_S(x,z,t), \\ f_z(x,z,t,\epsilon) &= \epsilon S_{F1} w_S(x,z,t) + S_{F2} w_F(x,z,t), \\ \sigma_x(x,z,t) &= S_S \mathrm{diag}(\sqrt{w_S(x,z,t)}), \\ \sigma_z(x,z,t,\epsilon) &= [\ \epsilon S_{F1}\mathrm{diag}(\sqrt{w_S(x,z,t)}) \\ &\qquad\qquad S_{F2}\mathrm{diag}(\sqrt{\epsilon}\sqrt{w_F(x,z,t)})\ ], \\ \Gamma_x = \Gamma_s, &\qquad \Gamma_z = [\ \Gamma_s \quad \Gamma_f\ ]^T. \end{aligned}$$

System (57) satisfies Assumptions 1 - 2, when the functions $w_S(x,z,t)$ and $w_F(x,z,t)$ are linear functions of the state variables x and z, which corresponds to a system in which all reactions are well approximated by uni-molecular reactions. Then, we can write

$$\begin{aligned} S_S w_S(x,z,t) &= A_1 x + A_2 z + A_3(t), \\ S_{F2} w_F(x,z,t) &= B_1 x + B_2 z + B_3(t), \\ \Phi(x,z,t) &= S_S \mathrm{diag}(w_S(x,z,t)), \\ \Lambda(x,z,t,\epsilon) &= \epsilon S_{F1}\mathrm{diag}(w_S(x,z,t)) \\ &\quad + S_{F2}\mathrm{diag}(w_F(x,z,t)), \\ \Theta(x,z,t,\epsilon) &= \epsilon S_{F1}\mathrm{diag}(w_S(x,z,t)) S_S^T. \end{aligned}$$

Next, we demonstrate this reduction process using a one-step reaction model for a phophorylation cycle as an example.

A. Application Example

Phosphorylation is an important mechanism in the regulation of many intracellular events such as cell cycle, cell growth and signal transduction. It was further demonstrated that phosphorylation cycles can function as insulation devices in biological circuits by attenuating the deleterious effects of load arising from connection to downstream targets, such as protein substrates or promoter sites controlling gene expression [10], [11], [12]. Tradeoffs between load mitigation and noise were further studied in [13]. Here, we apply the results of this paper to a model of a phosphorylation cycle with load and validate through numerical simulations that indeed the reduced system is an $O(\epsilon)$-approximation of the original system. We consider a phosphorylation cycle in which the phosphorylated protein X* binds to promoter p. Binding/unbinding reactions are much faster than phosphorylation/dephosphorylation reactions [11]. As a result, this system exhibits a separation of time-scales.

We can write the following chemical reactions for the system considered: $\mathrm{X} + \mathrm{Z} \xrightarrow{k_1} \mathrm{X}^* + \mathrm{Z}$, $\mathrm{X}^* + \mathrm{Y} \xrightarrow{k_2} \mathrm{X} + \mathrm{Y}$, $\mathrm{X}^* + \mathrm{p} \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} \mathrm{C}$. Here, protein X is phosphorylated by the kinase Z at rate k_1 and dephosphorylated by phosphotase Y at rate k_2. Phosphorylated protein X* binds to promoter p producing the complex C. We consider the input Z to be of the form $Z(t) = k + Asin(\omega t)$ where k is a constant bias and A and ω are the amplitude and the frequency of the input signal, respectively, as in [10]. The total concentration of X and promoter p, are conserved and therefore we can write $X + X^* + C = X_{tot}$ and $p + C = p_{tot}$. Let Γ_i be white noise processes and Ω be the cell volume. Letting $\Omega = 1$ for simplicity, the chemical Langevin equation for the system is given by

$$\begin{aligned} \frac{dX^*}{dt} &= k_1 Z(t) X_{tot}\left(1 - \frac{X^*}{X_{tot}} - \frac{C}{X_{tot}}\right) - k_2 Y X^* \\ &\quad - k_{\mathrm{on}} X^*(p_{tot} - C) + k_{\mathrm{off}} C \\ &\quad + \sqrt{k_1 Z(t) X_{tot}\left(1 - \frac{X^*}{X_{tot}} - \frac{C}{X_{tot}}\right)}\Gamma_3 \\ &\quad - \sqrt{k_2 Y X^*}\Gamma_4 - \sqrt{k_{\mathrm{on}} X^*(p_{tot} - C)}\Gamma_5 + \sqrt{k_{\mathrm{off}} C}\Gamma_6, \end{aligned} \tag{58}$$

$$\begin{aligned} \frac{dC}{dt} &= k_{\mathrm{on}} X^*(p_{tot} - C) - k_{\mathrm{off}} C + \sqrt{k_{\mathrm{on}} X^*(p_{tot} - C)}\Gamma_5 \\ &\quad - \sqrt{k_{\mathrm{off}} C}\Gamma_6. \end{aligned} \tag{59}$$

We assume that the pathway is weakly activated [14], which enables us to write $X^* \ll X_{tot}$, and that the concentration of X is much larger than that of the downstream binding site, giving $X_{tot} \gg p_{tot}$. Also, assuming weak binding of X* to promoter p, which gives $p_{tot} \gg C$, we can write equations (58) - (59) as

$$\begin{aligned} \frac{dX^*}{dt} &= k_1 Z(t) X_{tot} - k_2 Y X^* - k_{\mathrm{on}} X^* p_{tot} + k_{\mathrm{off}} C \\ &\quad + \sqrt{k_1 Z(t) X_{tot}}\Gamma_3 - \sqrt{k_2 Y X^*}\Gamma_4 - \sqrt{k_{\mathrm{on}} X^* p_{tot}}\Gamma_5 \\ &\quad + \sqrt{k_{\mathrm{off}} C}\Gamma_6, \end{aligned} \tag{60}$$

$$\frac{dC}{dt} = k_{\mathrm{on}} X^* p_{tot} - k_{\mathrm{off}} C + \sqrt{k_{\mathrm{on}} X^* p_{tot}}\Gamma_5 - \sqrt{k_{\mathrm{off}} C}\Gamma_6. \tag{61}$$

Since the binding reactions are much faster than phosphorylation/dephosphorylation, we can write $k_{\mathrm{off}} \gg k_2 Y$. Let $k_d = k_{\mathrm{off}}/k_{\mathrm{on}}$ be the dissociation constant and write $\epsilon = k_2 Y / k_{\mathrm{off}}$ where $\epsilon \ll 1$. Then, writing $a = k_2/k_1$ and letting $y = X^* + C$, the system (60) - (61) can be written in the form

$$\begin{aligned} \frac{dy}{dt} &= k_1 Z(t) X_{tot} - k_2 Y(y - C) + \sqrt{k_1 Z(t) X_{tot}}\Gamma_3 \\ &\quad - \sqrt{k_2 Y (y - C)}\Gamma_4, \end{aligned} \tag{62}$$

$$\begin{aligned} \epsilon\frac{dC}{dt} &= \frac{k_2 Y}{k_d} p_{tot}(y - C) - k_2 Y C + \sqrt{\frac{\epsilon k_2 Y}{k_d} p_{tot}(y - C)}\Gamma_5 \\ &\quad - \sqrt{\epsilon k_2 Y C}\Gamma_6, \end{aligned} \tag{63}$$

which has the structure of the system (57), with the state variables $x = y$, $z = C$ and

$$S = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \tag{64}$$

$$w_S(x,z,t) = \begin{bmatrix} k_1 Z(t) X_{tot} \\ k_2 Y (x - z) \end{bmatrix}, \tag{65}$$

$$w_F(x,z,t) = \begin{bmatrix} \frac{k_2 Y}{k_d} p_{tot}(x - z) \\ k_2 Y z \end{bmatrix}. \tag{66}$$

From (65) - (66), we have that $w_S(x,z,t)$ and $w_F(x,z,t)$ are linear functions of the state variables with $A_1 = -k_2 Y$, $A_2 = k_2 Y$, $A_3(t) = k_1 Z(t) X_{tot}$, $B_1 = k_2 Y p_{tot}/k_d$, $B_2 = -(k_2 Y p_{tot}/k_d + k_2 Y)$ and $B_3(t) = 0$. Therefore Assumptions 1 - 2 are satisfied. We also have $B_2 = -(k_2 Y p_{tot}/k_d +$

$k_2 Y$), satisfying Assumptions 3 and $B_3(t) = 0$ satisfying Assumption 4. We assume that there exist $X_0^* > 0$ and $C_0 > 0$ such that $X^* \geq X_0^*$ and $C \geq C_0$, which will guarantee the necessary Lipschitz continuity properties to ensure the existence of a unique solution to system (62) - (63). Then, we can proceed to find a reduced model for (62) - (63).

Setting $\epsilon = 0$ in (63), we obtain the slow manifold $\gamma_1(y) = y p_{tot}/(p_{tot} + k_d)$. Then the reduced system is given by

$$\frac{d\bar{y}}{dt} = k_1 Z(t) X_{tot} - k_2 Y \bar{y}\left(\frac{k_d}{k_d + p_{tot}}\right) + \sqrt{k_1 Z X_{tot} + k_2 Y \bar{y}\left(\frac{k_d}{k_d + p_{tot}}\right)}\Gamma_y,$$

where we have used the fact that Γ_i are independent Gaussian white noise processes to write $v_1\Gamma_3 - v_2\Gamma_4 = \sqrt{v_1^2 + v_2^2}\Gamma_y$.

Then applying Theorem 1, we have that

$$\|\mathbb{E}[\bar{y}] - \mathbb{E}[y]\| = O(\epsilon),\ t \in [0, t_1]$$
$$\|\mathbb{E}[\bar{y}^2] - \mathbb{E}[y^2]\| = O(\epsilon)\ t \in [0, t_1].$$

Fig. 3 shows the errors in the first and second moments. The simulations are performed using the Euler-Maruyama method [15] for stochastic differential equations and takes the average of 1000 simulation runs.

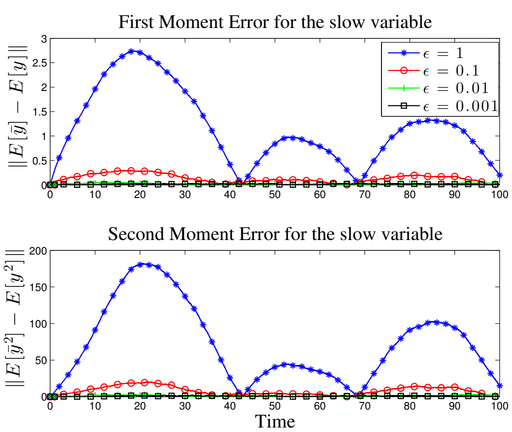

Fig. 3: Errors in the first and the second moments. The parameters used are $Z(t) = 0.1 + 0.05sin(0.1t),\ k_1 = 0.1,\ k_2 = 1,\ k_d = 100,\ X_{tot} = 200,\ Y = 0.2,\ p_{tot} = 200,\ y(0) = 10$ and $C(0) = 5$.

VI. Conclusions and Future Work

In this paper, we considered singularly perturbed stochastic differential equations with linear drift and nonlinear diffusion terms. We obtained a reduced model that approximates the dynamics of the slow variable of the original system when the time-scale separation is large (ϵ is small). Therefore, these results allow the slow dynamics of the original system to be approximated by a system with lower dimensions, which will be useful in analysis and simulations of stochastic system. In particular, benefits in simulations include reduced computation time and cost, especially for stochastic differential equations in the chemical Langevin form, modeling large biomolecular systems.

In future work, we will extend this result to the infinite time interval, consider systems where the drift coefficient is nonlinear and determine an approximation to the fast variables of the original system.

References

[1] H. K. Khalil. *Nonlinear systems*, volume 3. Prentice hall Upper Saddle River, 2002.
[2] Y. Kabanov and S. Pergamenshchikov. *Two-scale stochastic systems: asymptotic analysis and control*, volume 49. Springer, 2003.
[3] N. Berglund and B. Gentz. *Noise-induced phenomena in slow-fast dynamical systems: a sample-paths approach*. Springer, 2006.
[4] P. Kokotović, H. K. Khali, and J. O'reilly. *Singular perturbation methods in control: analysis and design*, volume 25. Siam, 1999.
[5] C. Tang and T. Başar. Stochastic stability of singularly perturbed nonlinear systems. In *Decision and Control, 2001. Proceedings of the 40th IEEE Conference on*, volume 1, pages 399–404. IEEE, 2001.
[6] D. T. Gillespie. The chemical langevin equation. *The Journal of Chemical Physics*, 113(1):297–306, 2000.
[7] L. Arnold. *Stochastic Differential Equations: Theory and Applications*. John Wiley and Sons, 1974.
[8] B. Mélykúti, K. Burrage, and K. C. Zygalakis. Fast stochastic simulation of biochemical reaction systems by alternative formulations of the chemical langevin equation. *The Journal of chemical physics*, 132(16):164109, 2010.
[9] N. G. Van Kampen. Stochastic processes in physics and chemistry, 1981.
[10] D. Del Vecchio, A. J. Ninfa, and E. D. Sontag. Modular cell biology: retroactivity and insulation. *Molecular systems biology*, 4(1), 2008.
[11] S. Jayanthi and D. Del Vecchio. Retroactivity attenuation in bio-molecular systems based on timescale separation. *Automatic Control, IEEE Transactions on*, 56(4):748–761, 2011.
[12] P. M. Rivera and D. Del Vecchio. Optimal design of phosphorylation-based insulation devices. In *American Control Conference (ACC), 2013*, pages 3783–3789. IEEE, 2013.
[13] S. Jayanthi and D. Del Vecchio. On the compromise between retroactivity attenuation and noise amplification in gene regulatory networks. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 4565–4571. IEEE, 2009.
[14] R. Heinrich, B. G. Neel, and T. A. Rapoport. Mathematical models of protein kinase signal transduction. *Molecular cell*, 9(5):957–970, 2002.
[15] D. J. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM review*, 43(3):525–546, 2001.