Moment convergence in a class of singularly perturbed stochastic differential equations

Narmada Herath[1] and Domitilla Del Vecchio[2]

Abstract— **We consider a class of singularly perturbed stochastic differential equations with linear drift and nonlinear diffusion terms. We obtain a reduced-order model that approximates the slow variable dynamics of the original system when the singular perturbation parameter ϵ is small. In our previous work, it was shown that, on a finite time interval, the first and the second moments of the slow variable dynamics of the original system are within an $O(\epsilon)$-neighborhood of the first and the second moments of the reduced-order system. In this paper, we extend this result to show that all moments of the slow variable dynamics of the original system are within an $O(\epsilon)$-neighborhood of the moments of the reduced-order system. We illustrate the application of this approach on a biomolecular system modeled by the chemical Langevin equation.**

I. INTRODUCTION

Time-scale separation is a common feature of many physical systems. It allows the system dynamics to be separated into slow and fast, employing a small parameter ϵ to capture the separation of the time-scales. The first step in analyzing such systems is to obtain a reduced-order model that approximates the dynamics of the slow variables. There are two main methods for obtaining such a reduced-order model: singular perturbation and averaging [1], [2]. The singular perturbation approach involves setting $\epsilon = 0$ in the fast variable dynamics to obtain an algebraic equation describing the slow manifold, which can then be used to derive an approximation for the slow variable dynamics. In the averaging method, the fast dynamics are averaged out by taking the time average of system functions.

Multi-scale stochastic differential equations arise in many applications, such as chemical reactions, population models and financial models. In particular, chemical reactions with multiple time-scales modeled by the chemical Langevin equation take the form of an SDE, where the diffusion term is of the order $\sqrt{\epsilon}$. For this type of systems, the existing singular perturbation approaches for stochastic systems in [1], [3], [4], [5] cannot be applied. The averaging method in [6] can be applied in this case, and it provides a reduced-order model to which the slow dynamics of the original system converge in distribution as ϵ becomes small. However, averaging methods require explicit integration of the system vector field, while singular perturbation methods only require

the local solution of an algebraic equation. Here, we therefore pursue the singular perturbation approach.

In our previous work, we considered a class of singularly perturbed systems with linear drift and nonlinear diffusion terms, including the case where the diffusion term is of order $\sqrt{\epsilon}$ and obtained a reduced-order model via singular perturbation. The first and the second moments of the reduced system were proven to be within an $O(\epsilon)$-neighborhood of the first and second moments of the original system [7]. In this paper, we show that, for the same class of systems, *all* moments of the reduced system are within an $O(\epsilon)$-neighborhood of the corresponding moments of the original system. We then apply the results to a biomolecular system modeled using the chemical Langevin equation.

This paper is organized as follows. In Section II, we introduce the system considered. In Section III, we present the reduced system and derive the moment dynamics of both the original and reduced systems. In Section IV, we prove the convergence result. Section V includes an application example.

II. SYSTEM MODEL

Consider the system

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

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

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 make the following assumptions on system (1) - (2).

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 $f_x(x, z, t) = A_1 x + A_2 z + A_3(t)$, where $A_1 \in \mathbb{R}^{n \times n}$, $A_2 \in \mathbb{R}^{n \times m}$ and $A_3(t) \in \mathbb{R}^n$ and $f_z(x, z, t, \epsilon) = B_1 x + B_2 z + B_3(t) + \alpha(\epsilon)(B_4 x + B_5 z + B_6(t))$, where $B_1, B_4 \in \mathbb{R}^{m \times n}$, $B_2, B_5 \in \mathbb{R}^{m \times m}$, $B_3(t), B_6(t) \in \mathbb{R}^n$, $A_3(t)$ and $B_3(t)$ are continuously differentiable functions, and $\alpha(\epsilon)$ is a continuously differentiable function with $\alpha(0) = 0$.

Assumption 2: The matrix-valued functions $\sigma_x(x, z, t)$ and $\sigma_z(x, z, t, \epsilon)$ are such that, there exist continuously differentiable functions $\Phi(x, z, t) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^{n \times n}$, $\Lambda(x, z, t, \epsilon) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{m \times m}$, $\Theta(x, z, t, \epsilon) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{m \times n}$, that satisfy

$$\sigma_x(x, z, t)\sigma_x(x, z, t)^T = \Phi(x, z, t), \qquad (3)$$

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

$$\sigma_z(x,z,t,\epsilon)\sigma_z(x,z,t,\epsilon)^T = \epsilon\Lambda(x,z,t,\epsilon), \qquad (4)$$

$$\sigma_z(x,z,t,\epsilon)[\ \sigma_x(x,z,t)\quad 0\ ]^T = \Theta(x,z,t,\epsilon), \qquad (5)$$

where the elements of $\Phi(x,z,t)$, $\Lambda(x,z,t,\epsilon)$, $\Theta(x,z,t,\epsilon)$ are affine functions of x and z, i.e., we can write $\mathbb{E}[\Phi(x,z,t)] = \Phi(\mathbb{E}[x],\mathbb{E}[z],t)$, $\mathbb{E}[\Lambda(x,z,t,\epsilon)] = \Lambda(\mathbb{E}[x],\mathbb{E}[z],t,\epsilon)$, and $\mathbb{E}[\Theta(x,z,t,\epsilon)] = \Theta(\mathbb{E}[x],\mathbb{E}[z],t,\epsilon)$. Also, we have that $\lim_{\epsilon\to 0}\Lambda(x,z,t,\epsilon) < \infty$ and $\lim_{\epsilon\to 0}\Theta(x,z,t,\epsilon) = 0$ for all x, z and t.

Assumption 3: Matrix B_2 is Hurwitz.

We assume that there exists a unique well-defined solution to system (1) - (2), on a finite time interval. Sufficient conditions, i.e., Lipschitz continuity and bounded growth of system functions, for the existence and uniqueness of solutions to stochastic differential equations can be found in [8]. The class of systems defined by Assumptions 1 - 2 includes the case where the diffusion term is given by a square-root function of the state, which does not satisfy the Lipschitz continuity condition. For this case, the work in [9] gives conditions that guarantee the existence of a unique solution.

First, we define the reduced system by setting $\epsilon = 0$ in (2). Next, we derive the moment dynamics of both the reduced system and original system, and show that setting $\epsilon = 0$ in the moment dynamics of the original system yields the moment dynamics of the reduced system. Then, by applying the Tikhonov's theorem to the moment dynamics of the original system, we show that all moments of the reduced system are within an $O(\epsilon)$-neighborhood of the moments of the original system.

III. PRELIMINARY RESULTS

Setting $\epsilon = 0$ in system (1) - (2) under Assumption 2, yields the algebraic equation

$$f_z(x,z,t,0) = B_1 x + B_2 z + B_3(t) = 0. \qquad (6)$$

Assumption 3 ensures the existence of a unique global solution $z = \gamma(x,t)$ to (6), given by

$$\gamma(x,t) = -B_2^{-1}(B_1 x + B_3(t)). \qquad (7)$$

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

$$\dot{\bar{x}} = f_x(\bar{x},\gamma(\bar{x},t),t) + \sigma_x(\bar{x},\gamma(\bar{x},t),t)\Gamma_x,\ \bar{x}(0) = x_0. \qquad (8)$$

We assume that there exists a unique solution to system (8), on a finite time interval.

In order to derive the moment dynamics of this system, we first introduce the notation used to represent the moments (notation adapted from [10], [11]). For given vectors $x = [x_1,\ldots,x_n]^T$ and $k = (k_1,\ldots,k_n)$ of length n, let $x^{(k)} = x_1^{k_1} x_2^{k_2}\ldots x_n^{k_n}$. Then, the moment of x corresponding to the vector k, is given by $\mathbb{E}[x^{(k)}]$, where the order of the moment is $\sum_{i=1}^n k_i$.

Similarly, we can write the Pth order moments for all $P \in \mathbb{N}$. To this end, we also define the sets $\mathcal{K}^P = \{(k_1,\ldots,k_n) \in \mathbb{N}^n | \sum_{i=1}^n k_i = P\}$, $\mathcal{G}^P = \{(g_1,\ldots,g_m) \in \mathbb{N}^m | \sum_{i=1}^n g_i = $

$P\}$, $\mathcal{N}^P = \{(l_1,\ldots,l_n) \in \mathbb{N}^n | \sum_{i=1}^n l_i \leq P\}$ and $\mathcal{C}^P = \{(c_1,\ldots,c_m) \in \mathbb{N}^m | \sum_{i=1}^m c_i \leq P\}$ for $P \in \mathbb{N}$.

Next, we proceed to derive the moment dynamics of the reduced system in (8). Denote the state vector by $\bar{x} = [\bar{x}_1,\ldots,\bar{x}_n]^T$ and let $\gamma(\bar{x},t) = [\gamma_1(\bar{x},t),\ldots,\gamma_m(\bar{x},t)]^T$.

Claim 1: There exists continuous functions $C_{1i}(t) : \mathbb{R} \to \mathbb{R}$ and $C_{2jl}(t) : \mathbb{R} \to \mathbb{R}$ for $i \in \mathcal{N}^P$, $l \in \mathcal{C}^1$, $j \in \mathcal{N}^{P-1}$ and $P = 1,\ldots,N$ where $N \in \mathbb{N}_{>0}$, such that under Assumptions 1 - 2, the moment dynamics of the reduced system in (8) are given by

$$\frac{d\mathbb{E}[\bar{x}^{(k)}]}{dt} = \sum_{i\in\mathcal{N}^P} C_{1i}(t)\mathbb{E}[\bar{x}^{(i)}]$$
$$+ \sum_{l\in\mathcal{C}^1}\sum_{j\in\mathcal{N}^{P-1}} C_{2jl}(t)\mathbb{E}[\gamma(\bar{x},t)^{(l)}\bar{x}^{(j)}],\quad \forall k \in \mathcal{K}^P.$$
$$(9)$$

Proof: See Appendix A-1. ∎

Next, we derive the moment dynamics of the original system in (1) - (2). Denote the state vectors by $x = [x_1,\ldots,x_n]^T$ and $z = [z_1,\ldots,z_m]^T$. Then, we make the following claim.

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

$$\frac{d\mathbb{E}[x^{(k)}]}{dt} = \sum_{i\in\mathcal{N}^P} C_{1i}(t)\mathbb{E}[x^{(i)}]$$
$$+ \sum_{l\in\mathcal{C}^1}\sum_{j\in\mathcal{N}^{P-1}} C_{2jl}(t)\mathbb{E}[z^{(l)}x^{(j)}], \forall k \in \mathcal{K}^P \quad (10)$$

$$\epsilon\frac{d\mathbb{E}[z^{(g)}]}{dt} = \sum_{i\in\mathcal{C}^P} D_{1i}(t,\epsilon)\mathbb{E}[z^{(i)}]$$
$$+ \sum_{l\in\mathcal{N}^1}\sum_{j\in\mathcal{C}^{P-1}} D_{2jl}(t,\epsilon)\mathbb{E}[x^{(l)}z^{(j)}], \forall g \in \mathcal{G}^P$$
$$(11)$$

$$\epsilon\frac{d\mathbb{E}[z^{(k_z)}x^{(k_x)}]}{dt} = \sum_{u\in\mathcal{N}^P} F_{1u}(t,\epsilon)\mathbb{E}[x^{(u)}] + \sum_{i\in\mathcal{C}^P} F_{2i}(t,\epsilon)\mathbb{E}[z^{(i)}]$$
$$+ \sum_{q=2}^P\sum_{r=1}^q\sum_{k\in\mathcal{C}^r}\sum_{s\in\mathcal{N}^{q-r}} F_{3qrks}(t,\epsilon)\mathbb{E}[z^{(k)}x^{(s)}], \quad (12)$$

for $k_x \in \mathcal{K}^{P_x}$ and $k_z \in \mathcal{G}^{P_z}$, where $P_x + P_z = P$, and for appropriate continuously differentiable functions $D_{1i}(t)$, $D_{2jl}(t)$, $F_{2i}(t,\epsilon)$, $F_{1u}(t,\epsilon)$, $F_{3qrks}(t,\epsilon)$, for $i \in \mathcal{C}^P$, $l \in \mathcal{N}^1$, $j \in \mathcal{C}^{P-1}$, $u \in \mathcal{N}^P$, $q = 2,\ldots,P$, $r = 1,\ldots,q$, $k \in \mathcal{C}^r$, $s \in \mathcal{N}^{q-r}$, $P = 1,\ldots,N$ where $N \in \mathbb{N}_{>0}$. Furthermore, setting $\epsilon = 0$ in (12) implies

$$\mathbb{E}[z^{(l)}x^{(j)}] = \mathbb{E}[\gamma(x,t)^{(l)}x^{(j)}] \qquad (13)$$

for $l \in \mathcal{C}^1$, $j \in \mathcal{N}^{P-1}$.

Proof: See Appendix A-2. ∎

IV. MAIN RESULTS

Lemma 1: Consider the original system in (1) - (2), the reduced system in (8), the moment dynamics of the reduced system in (9) and the moment dynamics of the original system in (10) - (12). We have that, under Assumptions 1 - 3, the commutative diagram in Fig. 1 holds.

$$\dot{x} = f_x(x,z,t) + \sigma_x(x,z,t)\,\Gamma_x$$
$$\epsilon\dot{z} = f_z(x,z,t,\epsilon) + \sigma_z(x,z,t,\epsilon)\,\Gamma_z$$

$$\epsilon \to 0$$

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

$$\epsilon \to 0$$

Moments of the Original System

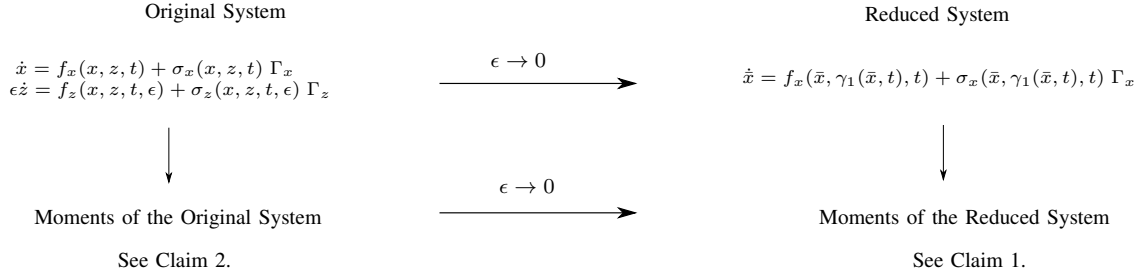Moments of the Reduced System

See Claim 2.

See Claim 1.

Fig. 1: Commutative Diagram.

Proof: In Section 2, the reduced system (8) is defined to be the system obtained when $\epsilon = 0$ in the original system (1) - (2). From Claim 1 and 2, we have that setting $\epsilon = 0$ in the moment dynamics of the original system in (10) - (12), yields the moment dynamics of the reduced system in (9). ∎

Although, Lemma 1 shows that setting $\epsilon = 0$ in the moment dynamics of the original system yields the moment dynamics of the reduced system, it does not guarantee that the trajectories of the moments of the original system approach the trajectories of the corresponding moments of the reduced system as ϵ becomes small. Therefore, next we apply the Tikhonov's theorem and prove that the moments of the reduced system are an $O(\epsilon)$ approximation of the moments of the original system.

Theorem 1: Consider the original system in (1) - (2) and the reduced system in (8). Then, under Assumptions 1 - 3, there exists $t_1 \geq 0$ such that

$$\|\mathbb{E}[\bar{x}^{(k)}] - \mathbb{E}[x^{(k)}]\| = O(\epsilon), \quad \forall k \in \mathcal{K}^N, \quad t \in [0, t_1]. \tag{14}$$

where $\mathcal{K}^N = \{(k_1, \ldots, k_n) \in \mathbb{N}^n \mid \sum_{j=1}^n k_j = N\}$ and $N \in \mathbb{N}_{>0}$.

Proof: From the commutative diagram in Lemma 1, it follows that setting $\epsilon = 0$ in the moments dynamics of the original system yields the moment dynamics of the reduced system. Therefore, we can apply Tikhonov's theorem to the moment dynamics of the original system in (10) - (12) to obtain the result (14). To this end, we prove that the boundary layer dynamics of the system (10) - (12) are globally exponentially stable. From equation (10), it follows that the fast variables that appear in (10) are of the form $\mathbb{E}[z_i x^{(j)}]$ for $j \in \mathcal{N}^{P-1}$. Thus, we define the vector $b_j = [b_{1j}, \ldots, b_{mj}]^T$ where $b_{ij} = \mathbb{E}[z_i x^{(j)}] - \mathbb{E}[\gamma_i(x,t)x^{(j)}]$ for $j \in \mathcal{N}^{P-1}$ and $i = 1, \ldots, m$. The dynamics of the variable b_{ij} for $j \in \mathcal{N}^{P-1}$ and $i = 1, \ldots, m$ are given by

$$\frac{db_{ij}}{dt} = \frac{d\mathbb{E}[z_i x^{(j)}]}{dt} - \frac{d\mathbb{E}[\gamma_i(x,t)x^{(j)}]}{dt}.$$

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

$$\frac{db_{ij}}{d\tau} = \epsilon\left(\frac{d\mathbb{E}[z_i x^{(j)}]}{dt} - \frac{d\mathbb{E}[\gamma_i(x,t)x^{(j)}]}{dt}\right).$$

Note that $\gamma_i(x,t)x^{(j)}$ can be written in terms of Pth or lower order moments of x, since $x^{(j)}$ contains moments of order up to $P - 1$ and from (6) we have that $\gamma_i(x,t)$ is a linear function of x. Therefore, for an appropriate function $Q_k(t)$ for $k \in \mathcal{N}^P$, and employing the linearity of the differential operator, we can write

$$\frac{db_{ij}}{d\tau} = \epsilon\left(\frac{d\mathbb{E}[z_i x^{(j)}]}{dt} - \sum_{k \in \mathcal{N}^P} Q_k(t)\frac{d\mathbb{E}[x^{(k)}]}{dt}\right).$$

Substituting from (10) and using the expansion of $d\mathbb{E}[z_i x^{(j)}]/dt$, (see Appendix A-2), yields

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

$$+ \epsilon\sum_{l=1}^n j_l \mathbb{E}[f_{x_l}(x,z,t)z_i x_1^{j_1}\ldots x_l^{j_l-1}\ldots x_n^{j_n}]$$

$$+ \epsilon\frac{1}{2}\sum_{p=1}^n j_p(j_p - 1)\mathbb{E}[b_{pp}(x,z,t)z_i x_1^{j_1}\ldots x_p^{j_p-2}\ldots x_n^{j_n}]$$

$$+ \sum_{m=1}^n j_m \mathbb{E}\left[q_{im}(x,z,t,\epsilon)x_1^{j_1}\ldots x_m^{j_m-1}\ldots x_n^{j_n}\right]$$

$$+ \epsilon\sum_{m=2}^n\sum_{p=1}^{m-1} j_m j_p \mathbb{E}[b_{pm}(x,z,t)z_i x_1^{j_1}\ldots x_p^{j_p-1}\ldots x_m^{j_m-1}\ldots x_n^{j_n}]$$

$$- \epsilon\sum_{s=0}^P\sum_{k \in \mathcal{K}^s} Q_k(t)\left(\sum_{i \in \mathcal{N}^P} C_{1i}(t)\mathbb{E}[\bar{x}^{(i)}]\right.$$

$$\left. + \sum_{l \in \mathcal{C}^1}\sum_{j \in \mathcal{N}^{P-1}} C_{2jl}(t)\mathbb{E}[\gamma(\bar{x},t)^{(l)}\bar{x}^{(j)}]\right), \tag{15}$$

where $b_{ij}(x,z,t) = \sum_{k=1}^{d_x} \sigma_x^{ik}(x,z,t)\sigma_x^{jk}(x,z,t)$, and $q_{ij}(x,z,t,\epsilon) = \sum_{k=1}^{(d_x+d_f)} \sigma_z^{ik}(x,z,t,\epsilon)[\sigma_x(x,z,t)\quad 0]^{jk}$, in which $[\sigma_x(x,z,t)\quad 0]$ denotes a matrix-valued function $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^{n \times (d_x+d_f)}$. To obtain the boundary layer system for the vector b_j, we set $\epsilon = 0$ in the equation (15). From Assumptions 1 - 3, we have that the functions $f_x(x,z,t)$, $f_z(x,z,t,\epsilon)$, $b_{ij}(x,z,t,\epsilon)$, $Q_k(t)$, $C_{1i}(t)$, $C_{2jl}(t)$ are continuous and therefore are bounded on any compact interval $t = [0, t_1]$. Under Assumption 2, we also have that $q_{im}(x,z,t,0) = 0$. Furthermore, due to the linearity of the system (10) - (12), 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 system is given by

$$\frac{db_j}{d\tau} = \mathbb{E}\left[f_z(x,z,t,0)x^{(j)}\right]\bigg|_{\mathbb{E}[zx^{(j)}]=b_j+\mathbb{E}[\gamma_i(x,t)x^{(j)}]}.$$

Under Assumption 1 and using equation (6), we have that

$$
\begin{aligned}
\frac{db_j}{d\tau} &= B_1 \mathbb{E}[xx^{(j)}] + B_2(b_j + \mathbb{E}[\gamma(x,t)x^{(j)}]) + B_3(t)\mathbb{E}[x^{(j)}], \\
&= B_2 b_j + B_1 \mathbb{E}[xx^{(j)}] + B_2(B_2^{-1}(B_1\mathbb{E}[xx^{(j)}] \\
&\quad + B_3(t)\mathbb{E}[x^{(j)}])) + B_3(t)\mathbb{E}[x^{(j)}], \\
&= B_2 b_j.
\end{aligned}
\tag{16}
$$

Therefore, under Assumption 3, we have that the boundary layer system in (16) is globally exponentially stable. We next verify that the remaining assumptions of the Tikhonov's theorem are satisfied. Note that the system (9) has a unique solution on a compact time interval $t \in [0, t_1]$, due to its linearity. From Claim 1, we have that $C_{1i}(t), C_{2lj}(t)$ are continuous functions with respect to time, and from Claim 2, we have that the functions $D_{1i}(t,\epsilon), D_{2jl}(t,\epsilon), F_{1i}(t,\epsilon), F_{2i}(t,\epsilon), F_{3qrks}(t,\epsilon)$ in (10) - (12) and their partial derivatives with respect to t and ϵ are continuous. Furthermore, we have that the function $\mathbb{E}[\gamma(x,t)x^{(j)}]$ for $j \in \mathcal{N}^{P-1}$ has continuous first partial derivatives with respect to its arguments $\mathbb{E}[x^{(k)}]$ for $k \in \mathcal{N}^P$ and the elements of the Jacobian $\frac{\partial \mathbb{E}[f_z(x,z,t,0)x^{(j)}]}{\partial \mathbb{E}[zx^{(j)}]}$ have continuous first partial derivatives with respect to their arguments $\mathbb{E}[zx^{(j)}]$, $\mathbb{E}[x^{(k)}]$ and t. Therefore 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 (10) - (12), we obtain the result (14). ∎

Remark: For the class of systems in Assumption 1 where $\sigma_z(x,z,t,\epsilon)$ is of order $\sqrt{\epsilon}$ and $\alpha(\epsilon) = 0$ for all ϵ, we can also prove that the diffusion process described by the slow variable $x(t)$ of original system (1) - (2) converges in distribution to the diffusion process described by reduced system (8), as ϵ tends to zero, by using the averaging principle developed by Khasminskii in [6]. We omit the proof here due to space limitations.

V. APPLICATION EXAMPLE

In this section, we consider an example of a biomolecular system with multiple time-scales and apply our results to obtain a reduced model.
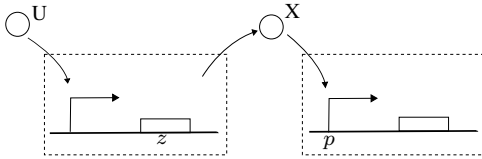

Fig. 2: The upstream transcriptional component takes protein U as the input, and produces the output protein X. The downstream transcriptional component takes protein X as the input.

Consider the system in Fig. 2, where an interconnection of two transcriptional components are shown. Transcriptional components make up gene transcriptional networks, which carry out fundamental processes in a cell for cell survival and growth [12]. The processes in transcriptional components

exhibit time-scale separation and previously singular perturbation methods have been employed to analyze deterministic models of transcriptional components [13]. In Fig. 2, Protein X, the output of the first component binds to the promoter p in the downstream component. We denote by $k(t)$ the production rate of X, and by δ the decay rate constant of X, which includes both degradation and dilution. The chemical reactions for this system can be written as follows : $\phi \underset{\delta}{\overset{k(t)}{\rightleftharpoons}} X, X + p \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C$, where the total concentration of promoter is conserved and is given by $p_t = p + C$. Denoting by Ω the cell volume, the chemical Langevin equations for the system are given by

$$
\begin{aligned}
\dot{X} &= k(t) - \delta X - k_{\text{on}}X(p_t - C) + k_{\text{off}}C + \frac{1}{\sqrt{\Omega}}(\sqrt{k(t)}\Gamma_1 \\
&\quad - \sqrt{\delta X}\Gamma_2 - \sqrt{k_{\text{on}}X(p_t - C)}\Gamma_3 + \sqrt{k_{\text{off}}C}\Gamma_4),
\end{aligned}
\tag{17}
$$

$$
\begin{aligned}
\dot{C} &= k_{\text{on}}X(p_t - C) - k_{\text{off}}C \\
&\quad + \frac{1}{\sqrt{\Omega}}(\sqrt{k_{\text{on}}X(p_t - C)}\Gamma_3 - \sqrt{k_{\text{off}}C}\Gamma_4),
\end{aligned}
\tag{18}
$$

where Γ_i are white noise processes.

As binding and unbinding reactions are much faster that protein production and decay, we have $k_{\text{off}} \gg \delta$. Denote by $k_d = k_{\text{off}}/k_{\text{on}}$, the dissociation constant between the binding and unbinding reaction rates. Then, we can write $\epsilon = \delta/k_{\text{off}}$, where $\epsilon \ll 1$. We assume that the binding is weak, which gives $p_t \gg C$. Then, with $k_{\text{off}} = \delta/\epsilon$, $k_{\text{on}} = \delta/(k_d\epsilon)$ and the change of variable $y = X + C$ we can write the system (17) - (18) in the singular perturbation form as

$$
\dot{y} = k(t) - \delta(y - C) + \frac{1}{\sqrt{\Omega}}\sqrt{k(t)}\Gamma_1 - \frac{1}{\sqrt{\Omega}}\sqrt{\delta(y-C)}\Gamma_2,
\tag{19}
$$

$$
\epsilon\dot{C} = \frac{\delta}{k_d}(y - C)p_t - \delta C + \frac{1}{\sqrt{\Omega}}\sqrt{\frac{\delta}{k_d}(y-C)p_t}\Gamma_3 \\
- \frac{1}{\sqrt{\Omega}}\sqrt{\delta C}\Gamma_4.
\tag{20}
$$

It should be noted that the system (19) - (20) does not satisfy the sufficient conditions for existence of a unique and well-defined solution given by [9]. The problem of existence of well-defined solutions for chemical Langevin equations is an ongoing area of study [14], [15]. In this example, we choose the parameter values that give sufficiently high concentrations of X and C, such that the argument of the square-root terms are always positive. Note that sufficiently high molecular concentrations is also a necessary requirement for the chemical Langevin equation to be a good approximation of the stochastic dynamics of a biomolecular system [16].

To obtain the reduced system, we set $\epsilon = 0$ in equation (20) which yields $\gamma(y) = p_t/(p_t + k_d)$. Then, the reduced system is given by

$$
\begin{aligned}
\dot{\bar{y}} &= k(t) - \delta\left(\frac{k_d}{p_t + k_d}\right)\bar{y} + \frac{1}{\sqrt{\Omega}}\sqrt{k(t)}\Gamma_1 \\
&\quad - \frac{1}{\sqrt{\Omega}}\sqrt{\delta\left(\frac{k_d}{p_t + k_d}\right)\bar{y}}\,\Gamma_2.
\end{aligned}
\tag{21}
$$

Applying Theorem 1 to systems (21) and (19) - (20), we have that

$$\|\mathbb{E}[\bar{y}^p] - \mathbb{E}[y^p]\| = O(\epsilon), \ t \in [0, t_1], \ \forall p \in \mathbb{N}_{>0}. \quad (22)$$

The simulation results performed using the Euler-Maruyama method [17] for $p = 1, 2, 3, 4$ are given in Fig. 3.
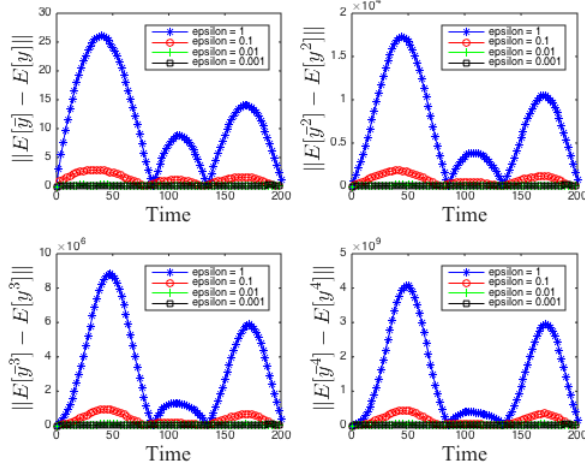

Fig. 3: Errors in the first, second, third and fourth order moments. The parameters used are $k(t) = 10 + 0.5\sin(0.05t)$, $\delta = 0.1$, $k_d = 1000$, $p_t = 2000$, $\Omega = 1$, $y(0) = 100$ and $C(0) = 50$. Moments are calculated using 100,000 simulation runs.

VI. CONCLUSION

We considered a class of singularly perturbed stochastic differential equations with linear drift and nonlinear diffusion terms and obtained a reduced-order model that approximates the slow variable dynamics of the original system. In [7], it was shown that the first and second moments of the reduced system are within an $O(\epsilon)$-neighborhood of the first and second moments of the original system, where ϵ is the singular perturbation parameter. In this work, we extended this result to show that *all* the moments of the reduced system are within an $O(\epsilon)$-neighborhood of the corresponding moments of the original system. We also note that, by using the averaging principle in [6], we can prove convergence in distribution. In future work, we aim at extending this analysis to obtain a reduced-order model for the fast variable and we will consider systems with nonlinear drift terms.

APPENDIX

A-1: Denote the drift and the diffusion terms of the reduced system in (8) by the vector $f_x(\bar{x}, \gamma(\bar{x}, t), t) = [f_{x_1}(\bar{x}, \gamma(\bar{x}, t), t), \ldots, f_{x_n}(\bar{x}, \gamma(\bar{x}, t), t)]^T$ and matrix $\sigma_x(\bar{x}, \gamma(\bar{x}, t), t) = [\sigma_x^{ij}(\bar{x}, \gamma(\bar{x}, t), t)]$ for $i = 1, \ldots, n$ and $j = 1, \ldots, d_x$, respectively. Then, from [10, p. 86] we have that the moment dynamics for each $k \in \mathcal{K}^P$ are given by

$$\frac{d\mathbb{E}[\bar{x}^{(k)}]}{dt} = \sum_{i=1}^{n} k_i \mathbb{E}[f_{x_i}(\bar{x}, \gamma(\bar{x}, t), t)\bar{x}_1^{k_1} \ldots \bar{x}_i^{k_i-1} \ldots \bar{x}_n^{k_n}]$$

$$+ \frac{1}{2}\sum_{p=1}^{n} k_p(k_p - 1)\mathbb{E}[b_{pp}(\bar{x}, \gamma(\bar{x}, t), t)\bar{x}_1^{k_1} \ldots \bar{x}_p^{k_p-2} \ldots \bar{x}_n^{k_n}]$$

$$+ \sum_{l=2}^{n}\sum_{j=1}^{l-1} k_l k_j \mathbb{E}[b_{jl}(\bar{x}, \gamma(\bar{x}, t), t)\bar{x}_1^{k_1} \ldots \bar{x}_j^{k_j-1} \ldots \bar{x}_l^{k_l-1} \ldots \bar{x}_n^{k_n}], \quad (23)$$

where $b_{ij}(\bar{x}, \gamma(\bar{x}, t), t) = \sum_{k=1}^{d_x} \sigma_x^{ik}(\bar{x}, \gamma(\bar{x}, t), t)\sigma_x^{jk}(\bar{x}, \gamma(\bar{x}, t), t)$.

Using Assumptions $1 - 2$ we have that the functions $f_{x_i}(\bar{x}, \gamma(\bar{x}, t), t)$ and $b_{ij}(\bar{x}, \gamma(\bar{x}, t), t)$ are affine functions of $\bar{x}$ and $\gamma(\bar{x}, t)$. Therefore, we observe that the dynamics of the Pth order moments will depend only on moments of order less than or equal to P. Furthermore, from Assumption 1, $A_3(t)$ and $B_3(t)$ are continuous. Thus, for appropriate continuous functions $C_{1i}(t)$ and $C_{2jl}(t)$ for $i \in \mathcal{N}^P$, $l \in \mathcal{C}^1$ and $j \in \mathcal{N}^{P-1}$, equation (23) can be written in the form of (9).

A-2: To find the moment dynamics $\mathbb{E}[x^{(k)}]$, we use Claim 1 with $x = \bar{x}$ and $z = \gamma(\bar{x}, t)$. Then, the moment dynamics for $\mathbb{E}[x^{(k)}]$ for all $k \in \mathcal{K}^P$ takes the form

$$\frac{d\mathbb{E}[x^{(k)}]}{dt} = \sum_{i \in \mathcal{N}^P} C_{1i}(t)\mathbb{E}[x^{(i)}] + \sum_{l \in \mathcal{C}^{(1)}}\sum_{j \in \mathcal{N}^{P-1}} C_{2j}(t)\mathbb{E}[z^{(l)}x^{(j)}].$$

To find the the dynamics of the moment $\mathbb{E}[z^{(g)}]$, denote the drift and the diffusion terms of equation (2) by the vector $(1/\epsilon)f_z(x, z, t) = (1/\epsilon)[f_{z_1}(x, z, t, \epsilon), \ldots, f_{z_m}(x, z, t, \epsilon)]^T$ and matrix $(1/\epsilon)\sigma_z(x, z, t, \epsilon) = (1/\epsilon)[\sigma_z^{ij}(x, z, t, \epsilon)]$ for $i = 1, \ldots, n$ and $j = 1, \ldots, (d_x + d_f)$, respectively. Then, from [10, p. 86] we have that the moment dynamics for $z^{(g)}$ for each $g \in \mathcal{G}^P$ are given by

$$\frac{d\mathbb{E}[z^{(g)}]}{dt} = \sum_{i=1}^{m} g_i \mathbb{E}\left[\frac{1}{\epsilon}f_{z_i}(x, z, t, \epsilon)z_1^{g_1} \ldots z_i^{g_i-1} \ldots z_m^{g_m}\right]$$

$$+ \frac{1}{2}\sum_{p=1}^{m} g_p(g_p - 1)\mathbb{E}\left[\frac{1}{\epsilon^2}d_{pp}(x, z, t, \epsilon)z_1^{g_1} \ldots z_p^{g_p-2} \ldots z_m^{g_m}\right]$$

$$+ \sum_{l=2}^{m}\sum_{p=1}^{l-1} g_l g_p \mathbb{E}\left[\frac{1}{\epsilon^2}d_{pl}(x, z, t, \epsilon)z_1^{g_1} \ldots z_p^{g_p-1} \ldots z_l^{g_l-1} \ldots z_m^{g_m}\right],$$

where $d_{ij}(x, z, t, \epsilon) = \sum_{k=1}^{d} \sigma_z^{ik}(x, z, t, \epsilon)\sigma_z^{jk}(x, z, t, \epsilon)$. From Assumption 2 we have that there exists $\lambda_{ij}(x, z, t, \epsilon) \in \mathbb{R}$ such that $d_{ij}(x, z, t, \epsilon) = \epsilon\lambda_{ij}(x, z, t, \epsilon)$ where $\lambda_{ij}(x, z, t, \epsilon)$ is affine in x and z and $\lim_{\epsilon \to 0} \lambda_{ij}(x, z, t, \epsilon) < \infty$. Using Assumptions 1 we also have that the function $f_{z_i}(x, z, t, \epsilon)$ is an affine function of x and z. Therefore, we observe that the dynamics of the Pth order moments do not depend on moment of order higher than P. Furthermore, from Assumptions 1 - 3, we have that the functions $f_{z_i}(x, z, t, \epsilon)$ and $\lambda_{ij}(x, z, t, \epsilon)$ are continuously differentiable in their arguments. Thus, multiplying by ϵ, we can write the dynamics of the moments $\mathbb{E}[z^{(g)}]$ for each $g \in \mathcal{G}^P$ in the form of (11) for appropriate continuously differentiable functions $D_{1i}(t)$ and $D_{2jl}(t)$ for $i \in \mathcal{C}^P$, $l \in \mathcal{N}^1$, $j \in \mathcal{C}^{P-1}$.

To derive the moment dynamics of $\mathbb{E}[z^{(k_z)}x^{(k_x)}]$, consider the dynamics of the vector $[z_1, \ldots, z_m, x_1, \ldots, x_n]^T$ for which the drift and the diffusion terms can be written as the vector

$[f_{z_1}(x,z,t,\epsilon)/\epsilon, \ldots f_{z_m}(x,z,t,\epsilon)/\epsilon, f_{x_1}(x,z,t) \ldots f_{x_n}(x,z,t)]^T$ and matrix $[(1/\epsilon)\sigma_z^{ij}(x,z,t,\epsilon), [\sigma_x(x,z,t)\ 0]^{kj}$ respectively, for $i = 1,\ldots,m$, $j = 1,\ldots,(d_x+d_f)$, $k = 1,\ldots,n$, in which $[\sigma_x(x,z,t)\ \ 0]$ denotes a matrix-valued function $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}^{n \times (d_x+d_f)}$ and $[\sigma_x(x,z,t)\ \ 0]^{kj}$ denotes the elements of the matrix. Then, from [10, p. 86], we can derive the moment dynamics of $z^{(k_z)}x^{(k_x)}$ for $k_z = (c_1,\ldots,c_m) \in \mathcal{G}^{P_z}$ and $k_x = (k_1,\ldots,k_m) \in \mathcal{K}^{P_x}$ as

$$\frac{\mathbb{E}[z^{(k_z)}x^{(k_x)}]}{dt} =$$

$$\sum_{i=1}^{m} c_i \mathbb{E}\left[\frac{1}{\epsilon} f_{z_i}(x,z,t,\epsilon)z_1^{c_1} \ldots z_i^{c_i-1} \ldots z_m^{c_m}x_1^{k_1} \ldots x_n^{k_n}\right]$$

$$+ \sum_{i=1}^{n} k_i \mathbb{E}\left[f_{x_i}(x,z,t)z_1^{c_1} \ldots z_m^{c_m}x_1^{k_1} \ldots x_i^{k_i-1} \ldots x_n^{k_n}\right]$$

$$+ \frac{1}{2}\sum_{i=1}^{m} c_i(c_i-1)\left[\frac{b_{ii}^z(x,z,t,\epsilon)}{\epsilon^2}z_1^{c_1} \ldots z_i^{c_i-2} \ldots z_m^{c_m}x_1^{k_1} \ldots x_n^{k_n}\right]$$

$$+ \frac{1}{2}\sum_{i=1}^{n} k_i(k_i-1)\left[b_{ii}^x(x,z,t,\epsilon)z_1^{c_1} \ldots z_m^{c_m}x_1^{k_1} \ldots x_i^{k_i-2} \ldots x_n^{k_n}\right]$$

$$+ \sum_{i=2}^{m}\sum_{j=1}^{i-1} c_i c_j \mathbb{E}\left[\frac{b_{ji}^z(x,z,t,\epsilon)}{\epsilon^2}z_1^{c_1} \ldots z_j^{c_j-1} \ldots z_i^{c_i-1} \ldots z_m^{c_m}x_1^{k_1}\right.$$
$$\left. \ldots x_n^{k_n}\right]$$

$$+ \sum_{i=1}^{n}\sum_{j=1}^{m} k_i c_j \mathbb{E}\left[\frac{b_{ji}^{(x,z)}(x,z,t,\epsilon)}{\epsilon}z_1^{c_1} \ldots z_j^{c_j-1} \ldots z_m^{c_m}x_1^{k_1} \ldots\right.$$
$$\left. x_i^{k_i-1} \ldots x_n^{k_n}\right]$$

$$+ \sum_{i=2}^{n}\sum_{j=1}^{i-1} k_i k_j \mathbb{E}\left[b_{ji}^x(x,z,t)z_1^{c_1} \ldots z_m^{c_m}x_1^{k_1} \ldots x_i^{k_i-1} \ldots x_j^{k_j-1}\right.$$
$$\left. \ldots x_n^{k_n}\right]. \quad (24)$$

where $b_{ij}^z(x,z,t,\epsilon) = \sum_{k=1}^{(d_x+d_f)} \sigma_z^{ik}(x,z,t,\epsilon)\sigma_z^{jk}(x,z,t,\epsilon)$, $b_{ij}^x(x,z,t) = \sum_{k=1}^{d_x} \sigma_x^{ik}(x,z,t)\sigma_x^{jk}(x,z,t)$, and $b_{ij}^{(x,z)}(x,z,t,\epsilon) = \sum_{k=1}^{(d_x+d_f)} \sigma_z^{ik}(x,z,t,\epsilon)[\sigma_x(x,z,t)\ \ 0]^{jk}$.

Under Assumption 1 - 2, we have that the functions $f_{z_i}(x,z,t,\epsilon)$, $f_{x_i}(x,z,t)$, $b_{ij}^z(x,z,t,\epsilon)$, $b_{ij}^x(x,z,t)$ and $b_{ij}^{(x,z)}(x,z,t,\epsilon)$ are linear functions of x and z and are continuously differentiable in their arguments. Therefore, we can write the moment dynamics $\mathbb{E}[z^{(k_z)}x^{(k_x)}]$ in the form of (12), for appropriate functions $F_{2i}(t,\epsilon)$, $F_{1u}(t,\epsilon)$, $F_{3qrks}(t,\epsilon)$, for $i \in \mathcal{C}^P$, $u \in \mathcal{N}^P$, $q = 2,\ldots,P$, $r = 1,\ldots,q$, $k \in \mathcal{C}^r$, $s \in \mathcal{N}^{q-r}$.

To prove (13) we consider the dynamics of $\mathbb{E}[z_i x^{(j)}]$ for $j \in \mathcal{N}^{P-1}$ and $i = 1,\ldots,m$, which can be derived from (24), by considering the vector $k_z = (c_1,\ldots,c_m) \in \mathcal{G}^1$ which gives $c_i = 1$ and $c_l = 0$ for $l \neq i$ for each $i = 1,\ldots,m$. Then, the dynamics for $\mathbb{E}[z_i x^{(j)}]$ can be written in the form

$$\epsilon\frac{d\mathbb{E}[z_i x^{(j)}]}{dt} = \mathbb{E}\left[f_{z_i}(x,z,t,\epsilon)x^{(j)}\right]$$
$$+ \epsilon\sum_{l=1}^{n} k_l \mathbb{E}[f_{x_l}(x,z,t)z_i x_1^{k_1} \ldots x_l^{k_l-1} \ldots x_n^{k_n}]$$

$$+ \epsilon\frac{1}{2}\sum_{p=1}^{n} k_p(k_p-1)\mathbb{E}[b_{pp}^x(x,z,t)z_i x_1^{k_1} \ldots x_p^{k_p-2} \ldots x_n^{k_n}]$$

$$+ \sum_{m=1}^{n} k_m \mathbb{E}\left[b_{im}^{(x,z)}(x,z,t,\epsilon)x_1^{k_1} \ldots x_m^{k_m-1} \ldots x_n^{k_n}\right]$$

$$+ \epsilon\sum_{l=2}^{n}\sum_{p=1}^{l-1} k_l k_p \mathbb{E}[b_{pl}^x(x,z,t)z_i x_1^{k_1} \ldots x_p^{k_p-1} \ldots x_l^{k_l-1} \ldots x_n^{k_n}].$$
$$(25)$$

Using Assumption 2, we have that $b_{ij}^{(x,z)}(x,z,t,0) = 0$ for $i = 1,\ldots,m$ and $j = 1,\ldots,n$. Therefore, setting $\epsilon = 0$ in the dynamics of the vector $\mathbb{E}[zx^{(j)}]$ yields

$$\mathbb{E}[f_z(x,z,t,0)x^{(j)}] = 0. \quad (26)$$

Using Assumption 2 - 3, we obtain the unique solution to (26) as $\mathbb{E}[zx^{(j)}] = -B_2^{-1}(B_1\mathbb{E}[xx^{(j)}]+B_3(t)\mathbb{E}[x^{(j)}])$. From equation (7) we have that $\mathbb{E}[\gamma(x,t)x^{(j)}] = \mathbb{E}[-B_2^{-1}(B_1x + B_3(t))x^{(j)}] = \mathbb{E}[zx^{(j)}]$. Therefore, it follows that setting $\epsilon = 0$ in (25), yields $\mathbb{E}[z_i x^{(j)}] = \mathbb{E}[\gamma_i(x,t)x^{(j)}]$ for $i = 1,\ldots,m$, which results in (13).

REFERENCES

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

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

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

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

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

[7] N. Herath, A. Hamadeh, and D. Del Vecchio. Model reduction for a class of singularly perturbed stochastic differential equations. In *American Control Conference (ACC), 2015*. IEEE, (To Appear) 2015.

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

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

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

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

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

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

[14] Joshua Wilkie and Yin Mei Wong. Positivity preserving chemical langevin equations. *Chemical Physics*, 353(1):132–138, 2008.

[15] David Schnoerr, Guido Sanguinetti, and Ramon Grima. The complex chemical langevin equation. *The Journal of chemical physics*, 141(2):024103, 2014.

[16] D. T. Gillespie. The chemical langevin equation. *The Journal of Chemical Physics*, 113(1):297–306, 2000.

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