MIT Open Access Articles

Deterministic-Like Model Reduction for a Class of Multi-Scale Stochastic Differential Equations with Application to Biomolecular Systems

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

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

Abstract—We consider a class of singularly perturbed stochastic differential equations with linear drift terms, and present a reduced-order model that approximates both slow and fast variable dynamics when the time-scale separation is large. We show that, on a finite time interval, moments of all orders of the slow variables for the original system become closer to those of the reduced-order model as time-scale separation is increased. A similar result holds for the first and second moments of the fast variable approximation. 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

Many physical systems evolve on multiple time-scales. Examples include climate systems, electrical systems, and biological systems. The dynamics 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 time-scales [4]. The analysis of such systems can be simplified by obtaining a reduced-order model that approximates the dynamics of the original system.

In the deterministic setting, the derivation of a reduced-order system is mainly accomplished using the singular perturbation method, formalized by Tikhonov's theorem, where the reduced-order model is obtained by setting ϵ to zero in the original system dynamics [5], [4]. 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. Another method that can be used to obtain a reduced-order model is the averaging principle, which eliminates the fast dynamics via integration of system functions to give an approximation for the slow variables' dynamics [6].

In addition to deterministic systems, stochastic models also arise in many areas such as finance, population biology, and

Narmada Herath is with the Electrical Engineering and Computer Science Department, MIT, 77 Mass. Ave, Cambridge MA `nherath@mit.edu`

Domitilla Del Vecchio is with the Mechanical Engineering Department, MIT, 77 Mass. Ave, Cambridge MA `ddv@mit.edu`

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

chemical kinetics. For example, biomolecular systems are intrinsically stochastic due to the randomness in chemical reactions and the chemical Langevin equation (CLE) has been widely used to model the stochastic nature of these systems in the form of an SDE [7].

Several works have extended singular perturbation methods to SDEs [8], [9], [5], [10]. 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. Aside from singular perturbation based approaches, averaging methods have also been extended for SDEs. These methods consider diffusion terms of the order $\sqrt{\epsilon}$ [6], [11], and have recently been applied in approximating the slow variables of CLEs [12]. However, averaging methods require the integration of the system's vector field, which may be undesirable for systems of high dimension. Furthermore, averaging only provides an approximation for the slow variables, and does not approximate 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 [13], 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 in 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

II. SYSTEM MODEL

We consider the singularly perturbed stochastic differential equations

$$\dot{x} = f_x(x,z,t) + \sigma_x(x,z,t)\Gamma_x, \qquad x(0) = x_0, \quad (1)$$
$$\epsilon\dot{z} = f_z(x,z,t,\epsilon) + \sqrt{\epsilon}\sigma_z(x,z,t,\epsilon)\Gamma_z, \ z(0) = z_0, \quad (2)$$

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 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 $f_x(x,z,t) = A_1 x + A_2 z + A_3(t)$, where $A_1 \in \mathbb{R}^{n \times n}$, $A_2 \in \mathbb{R}^{n \times m}$ and $A_3(t) \in \mathbb{R}^n$, $f_z(x,z,t,\epsilon) = B_1 x + B_2 z + B_3(t) + \alpha(\epsilon)(B_4 x + B_5 z + B_6(t))$, where $B_1, B_4 \in \mathbb{R}^{m \times n}$, $B_2, B_5 \in \mathbb{R}^{m \times m}$, $B_3(t), B_6(t) \in \mathbb{R}^m$, $A_3(t)$ 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 $\lim_{\epsilon\to0}\Lambda(x,z,t,\epsilon) < \infty$ and $\lim_{\epsilon\to0}\Theta(x,z,t,\epsilon) < \infty$ 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 SDEs are given by the Lipschitz continuity and bounded growth of system functions [14]. 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 [15].

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 $f_z(x,z,t,0) = B_1 x + B_2 z + B_3(t) = 0$,

for which, Assumption 3 guarantees the existence of a unique global solution $z = \gamma(x,t)$, given by

$$\gamma(x,t) = -B_2^{-1}(B_1 x + B_3(t)). \quad (3)$$

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

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

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

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

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

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

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 (adapted from [16]) used to denote the moment dynamics. 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 (expected value) 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 Pth order moments for all non-negative integers $P \in \mathbb{Z}_{\geq 0}$, we define the set $\mathcal{G}_r^P = \{(c_1, \ldots, c_r) \in \mathbb{Z}_{\geq 0}^r | \sum_{i=1}^r c_i \leq P\}$, where $\mathbb{Z}_{>0}^r$ denotes the sets of vectors of length r with nonnegative integer elements. 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 exist $\epsilon^* > 0$, t_1, $t_b > 0$ with $t_1 > t_b$ such that for $\epsilon < \epsilon^*$*

$$\|\mathbb{E}[\bar{x}^{(k)}] - \mathbb{E}[x^{(k)}]\| = O(\epsilon), \ \forall k \in \mathcal{G}_n^{\mathcal{N}}, \ \mathcal{N} \in \mathbb{Z}_{>0}, \ t \in [0, t_1],$$
$$(7)$$
$$\|\mathbb{E}[\bar{z}^{(l)}] - \mathbb{E}[z^{(l)}]\| = O(\epsilon), \ \forall l \in \mathcal{G}_m^2, \qquad t \in [t_b, t_1], \quad (8)$$

where $\mathbb{Z}_{>0}$ denotes the set of positive integers.

The proof of Theorem 1 is based on applying Tikhonov's theorem to the moment dynamics of the original and reduced-order system, and is similar to the results in our conference papers [2] and [3]. We provide an outline of the proof strategy here, while a complete proof is provided in [17]. 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 Tikhonov's theorem to demonstrate the convergence of the moments of the original system to the moments of the reduced-order 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 [4]. 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 [4].

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 given by the diffusion terms $\sigma_z(x,z,t,\epsilon)$. 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

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

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 [18, 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 $\gamma(\bar{x},t)$ 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. This noise must instead be considered in approximating the noise properties of the fast variable, as we illustrate in the following example. Consider the system

$$\dot{x} = -a_1 x + a_2 z + v_1 \Gamma_1, \qquad \epsilon\dot{z} = -z + v_2\sqrt{\epsilon}\Gamma_2, \quad (10)$$

where $a_1, a_2 > 0$. Setting $\epsilon = 0$, we obtain the system:

$$\dot{\bar{x}} = -a_1\bar{x} + v_1\Gamma_1, \qquad z = \gamma(\bar{x},t) = 0. \qquad (11)$$

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 system (11) 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. 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, 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, 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 (11) obtained by setting $\epsilon = 0$ is a good approximation for the slow variable dynamics, but it does not provide a good approximation for the fast variable.
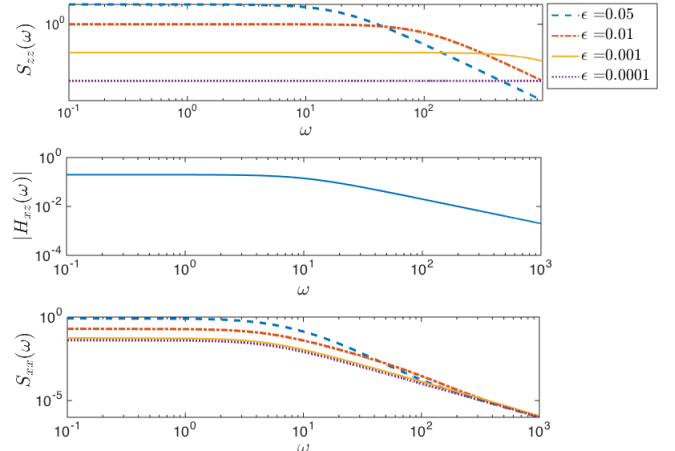

Fig. 1: Power spectrum of z (top). Frequency response from z to x (centre). Power spectrum of x (bottom). The parameters used are $a_1 = 10$, $a_2 = 2$ and $v_1 = 1$, $v_2 = 10$.

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 [15]:

$$\dot{x} = -2x + z + 10 + \sqrt{2z+1}\,\Gamma_1, \qquad (12)$$
$$\epsilon\dot{z} = -z + 15 + \sqrt{\epsilon(2z+1)}\,\Gamma_2. \qquad (13)$$

This system satisfies the Assumptions 1 - 3, and using the results of [15] it can be verified that there exists a unique solution where the arguments of the square-root diffusion terms remain positive at all times.

Setting $\epsilon = 0$, we obtain the slow manifold $z = 15$. This yields the following reduced-order model for the slow variable:

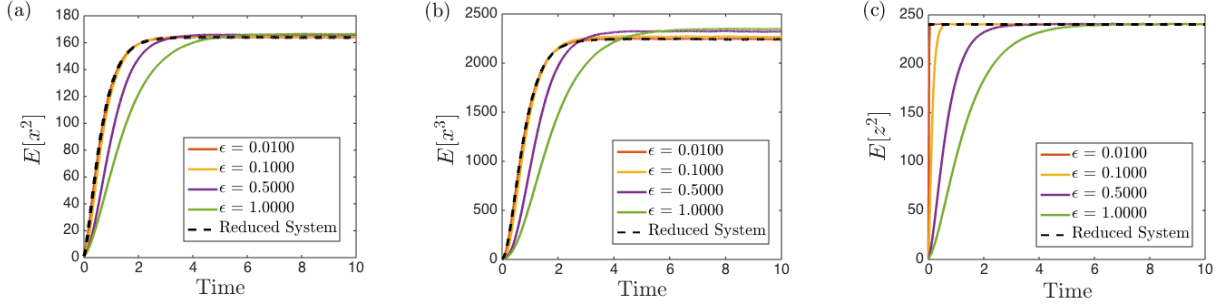$$\dot{\bar{x}} = -2\bar{x} + 25 + \sqrt{31}\,\Gamma_1. \qquad (14)$$

Fig. 2: Moments of the original and reduced systems. Sample means were calculated using 500,000 realizations. (a) Second moments of the slow variable. (b) Third moments of the slow variable. (c) Second moments of the fast variable.

Based on (5), the fast variable approximation for this system is of the form $z = 15 + g(x)N$ 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

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

Simulations of the original and the reduced-order systems were performed using the Euler-Maruyama method [19] for SDEs and were used to calculate the second and third order moments of the slow variable and second order moments of the fast variable, as illustrated in Fig. 2. It can be seen that as ϵ decreases the moments of the original system tend to the moments of the reduced-order system.

By virtue of Theorem 1, the reduced-order model (14) provides a good approximation of the higher moments for the slow variable, as illustrated in Fig. 2. For the fast variable, only the first and second moments are well approximated, while there is no guarantee that the higher order moments are also approximated well. We illustrate this by analyzing the third order moments of the fast variable in (13). To calculate the third order moments, we first represent the fast variable dynamics of the system (12)–(13) in the form $\epsilon\dot{z} = c_1 z + c_2 + \sqrt{\epsilon(d_1 z + d_2)}\,\Gamma_2$.

Then, computing the third order moment dynamics and setting $\epsilon = 0$, we obtain

$$\overline{\mathbb{E}[z]} = -c_2/c_1, \qquad \overline{\mathbb{E}[z^2]} = (2c_2^2 + d_1 c_2 - d_2 c_1)/(2c_1^2) \tag{16}$$

$$\overline{\mathbb{E}[z^3]} = (3c_1 c_2 d_2 + c_1 d_1 d_2 - 2c_2^3 - 3c_2^2 d_1 - c_2 d_1^2)/(2c_1^3). \tag{17}$$

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

$$\mathbb{E}[\bar{z}] = \gamma(\mathbb{E}[\bar{x}], t) = -c_2/c_1, \tag{18}$$

$$\mathbb{E}[\bar{z}^2] = \mathbb{E}[\gamma(\bar{x}, t)^2] + \mathbb{E}[g(\bar{x}, t)^2]$$
$$= (2c_2^2 + d_1 c_2 - d_2 c_1)/(2c_1^2), \tag{19}$$

$$\mathbb{E}[\bar{z}^3] = \mathbb{E}[\gamma(\bar{x}, t)^3] + 3\mathbb{E}[\gamma(\bar{x}, t)g(\bar{x}, t)^2]$$
$$= c_2(3c_1 d_2 - 2c_2^2 - 3c_2 d_1)/(2c_1^3). \tag{20}$$

Considering the equations for the slow manifold in (16) - (17) and the moments of the reduced fast system (18)–(20), we have that $\|\overline{\mathbb{E}[z]} - \mathbb{E}[\bar{z}]\| = 0$, $\|\overline{\mathbb{E}[z^2]} - \mathbb{E}[\bar{z}^3]\| = 0$, however, $\|\overline{\mathbb{E}[z^3]} - \mathbb{E}[\bar{z}^3]\| = \frac{d_1(c_1 d_2 - c_2 d_1)}{2c_1^3}$, 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 (18)–(20) 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 common measures of noise are coefficient of variation and signal-to-noise ratio, which are functions of only the mean and the variance. Therefore, 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 [20], [21]. Another source of signal perturbation in biological systems is the intrinsic noise due to the randomness in chemical reactions [22], [23]. 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 [24]. 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 [25]

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.
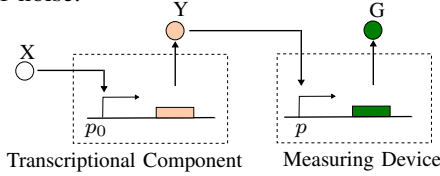

Fig. 3: Protein X acts as an input to the upstream transcriptional component, which produces the output protein Y. The downstream transcriptional component takes protein Y as an input and produces protein G.

The chemical reactions for the system in Fig. 3, can be written as follows: $X + p_0 \underset{\alpha_2}{\overset{\alpha_1}{\rightleftharpoons}} C_0, C_0 \xrightarrow{\beta_1} Y + C_0, Y \xrightarrow{\delta_1} \phi$, $Y + p \underset{\alpha_4}{\overset{\alpha_3}{\rightleftharpoons}} C, C \xrightarrow{\beta_2} G + C, G \xrightarrow{\delta_2} \phi$ [24], [25]. 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 $p_T = p + C$ [24]. Denoting the system volume by Ω, the chemical Langevin equations for the system are given by

$$\frac{dC_0}{dt} = \alpha_1 X(p_{T0} - C_0) - \alpha_2 C_0 + \sqrt{\frac{\alpha_1 X(p_{T0} - C_0)}{\Omega}} \Gamma_1$$
$$- \sqrt{\frac{\alpha_2 C_0}{\Omega}} \Gamma_2,$$

$$\frac{dY}{dt} = \beta_1 C_0 - \delta_1 Y + \sqrt{\frac{\beta_1 C_0}{\Omega}} \Gamma_3 - \sqrt{\frac{\delta_1 Y}{\Omega}} \Gamma_4$$

$$\boxed{-\alpha_3 Y(p_T - C) + \alpha_4 C - \sqrt{\frac{\alpha_3 Y(p_T - C)}{\Omega}} \Gamma_5 + \sqrt{\frac{\alpha_4 C}{\Omega}} \Gamma_6},$$

$$\dot{C} = \alpha_3 Y(p_T - C) - \alpha_4 C + \sqrt{\frac{\alpha_3 Y(p_T - C)}{\Omega}} \Gamma_5 - \sqrt{\frac{\alpha_4 C}{\Omega}} \Gamma_6,$$

$$\frac{dG}{dt} = \beta_2 C - \delta_2 G + \sqrt{\frac{\beta_2 C}{\Omega}} \Gamma_7 - \sqrt{\frac{\delta_2 G}{\Omega}} \Gamma_8,$$

$$(21)$$

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 (21) 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 [20], [25]. It was also shown, both theoretically [20] and experimentally [21], that increasing the number of downstream binding sites p_T increases the effect of retroactivity.

The nominal and perturbed trajectories for Y and G for different amounts of p_T can be seen in Fig. 4(a) and Fig. 4(b). 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 (21). The perturbed trajectories are obtained using Gillespie's direct method [26]. 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 [22], [23]. However, the downside of this is that increasing p_T alters the dynamics of the input transcription factor (i.e. protein Y), due to retroactivity, as previously discussed. For signal Y, by contrast, both retroactivity and noise increase as p_T is increased, which is consistent with prior observations in [27].
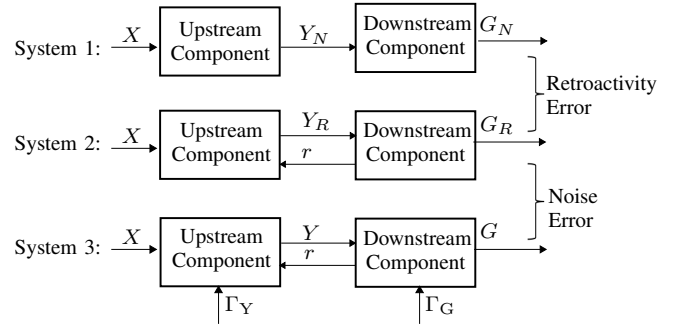

Fig. 5: The signal 'r' denotes the retroactivity to the upstream system. Γ_Y encapsulates the noise in the upstream component given by Γ_i for $i = 1, ..., 4$ and Γ_G encapsulates the noise in the downstream component given by Γ_i for $i = 5, ..., 8$.

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 where the perturbations due to retroactive and noise are absent, 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 (21) exhibits time-scale separation as the binding/unbinding reactions between transcription factors and promoter sites are much faster than protein production/decay [24]. 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 $c_0 = C_0/p_{T0}$, $y = Y/(\beta_1 p_{T0}/\delta_1)$, $c = C/p_T$,
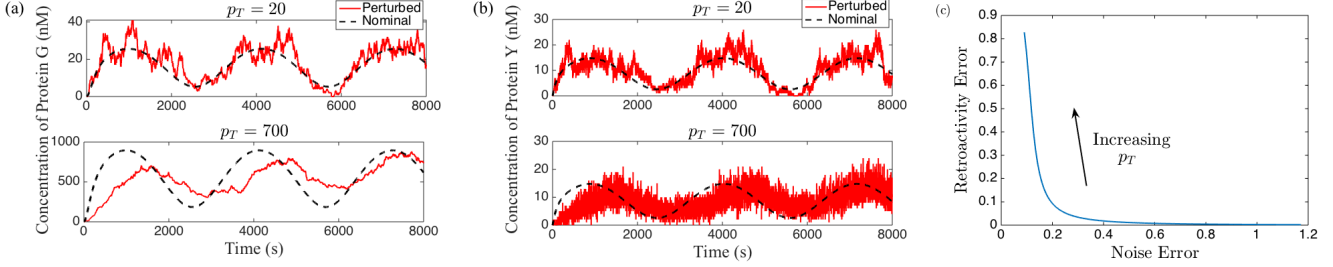
Fig. 4: (a) Nominal and perturbed trajectories for G. (b) Nominal and perturbed trajectories for Y. (c) Trade-off between retroactivity and noise in signal G for p_T in the range 1:1000 nM (obtained using the equations (31) and (32)). The parameter values are $X = 2 + 1.5sin(\omega t)$nM, $\alpha_1 = 1$nM^{-1}s^{-1}, $\alpha_2 = 20$s^{-1}, $\alpha_3 = 1$nM^{-1}s^{-1}, $\alpha_4 = 100$s^{-1}, $\beta_1 = 0.01$s^{-1}, $\beta_2 = 0.1$s^{-1}, $\delta_1 = \delta_2 = 0.01$s^{-1}, $p_{T0} = 100$nM and $\omega = 0.002$ rad/s.

$g = G/(\beta_2 p_T/\delta_2)$, 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

$$\epsilon \frac{dc_0}{d\bar{t}} = \frac{X}{k_{d1}} - c_0 + \sqrt{\epsilon \frac{X}{k_{d1}p_{T0}\Omega}}\tilde{\Gamma}_1 - \sqrt{\epsilon \frac{c_0}{p_{T0}\Omega}}\tilde{\Gamma}_2,$$

$$\frac{dv}{d\bar{t}} = c_0 - v + \frac{p_T \delta_1}{\beta_1 p_{T0}}c + \sqrt{\frac{\delta_1 c_0}{\beta_1 p_{T0}\Omega}}\tilde{\Gamma}_3 - \sqrt{\frac{\delta_1(v - \frac{p_T \delta_1}{\beta_1 p_{T0}}c)}{\beta_1 p_{T0}\Omega}}\tilde{\Gamma}_4,$$

$$\epsilon \frac{dc}{d\bar{t}} = \frac{a\beta_1 p_{T0}(v - \frac{p_T \delta_1}{\beta_1 p_{T0}}c)}{k_{d2}\delta_1} - ac$$
$$+ \sqrt{\epsilon \frac{a\beta_1 p_{T0}(v - \frac{p_T \delta_1}{\beta_1 p_{T0}}c)}{k_{d2}\delta_1 p_T \Omega}}\tilde{\Gamma}_5 - \sqrt{\epsilon \frac{ac}{p_T \Omega}}\tilde{\Gamma}_6,$$

$$\frac{dg}{d\bar{t}} = \frac{\delta_2}{\delta_1}c - \frac{\delta_2}{\delta_1}g + \sqrt{\frac{\delta_2^2}{\delta_1 \beta_2 p_T \Omega}}c\tilde{\Gamma}_7 - \sqrt{\frac{\delta_2^2}{\delta_1 \beta_2 p_T \Omega}}g\tilde{\Gamma}_8, \quad (22)$$

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 (22) 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 the system (22) does not satisfy the sufficient conditions for existence and uniqueness of solution given in [15], and that the existence of a solution for CLEs where the arguments of the square-root diffusion terms remain positive is an ongoing research question [28], [29]. However, the validity of the CLE representation for chemical kinetics is based on the assumption that the molecular counts are sufficiently large [7]. In line with this, the work in [30] considers several one-dimensional models and show that the probability of molecular counts reaching zero decreases as the initial condition increases. Considering higher dimensional models, in [31], 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 [31]), thereby keeping the argument of the square-root positive for a longer time interval.

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

$$\frac{dv}{d\bar{t}} = \frac{X}{k_{d1}} - (1 - R)v + \sqrt{\frac{\delta_1 \frac{X}{k_{d1}}}{\beta_1 p_{T0}\Omega}}\tilde{\Gamma}_3 - \sqrt{\frac{\delta_1(1 - R)v}{\beta_1 p_{T0}\Omega}}\tilde{\Gamma}_4, \quad (23)$$

$$\frac{dg}{d\bar{t}} = \frac{\delta_2 \beta_1 p_{T0}v}{\delta_1^2(p_T + k_{d2})} - \frac{\delta_2}{\delta_1}g + \sqrt{\frac{\delta_2^2 \beta_1 p_{T0}v}{\delta_1^2 \beta_2 p_T(p_T + k_{d2})\Omega}}\tilde{\Gamma}_7$$
$$- \sqrt{\frac{\delta_2^2}{\delta_1 \beta_2 p_T \Omega}}g\tilde{\Gamma}_8, \quad (24)$$

$$c_0 = \frac{X}{k_{d1}} + \sqrt{\frac{X}{p_{T0}k_{d1}\Omega}}N_1, \quad (25)$$

$$c = \frac{v\beta_1 p_{T0}}{\delta_1(p_T + k_{d2})} + \sqrt{\frac{v\beta_1 p_{T0}k_{d2}}{\delta_1 p_T \Omega(p_T + k_{d2})^2}}N_2, \quad (26)$$

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 (23)–(26), which yields

$$\frac{dv_R}{d\bar{t}} = \frac{X}{k_{d1}} - (1 - R)v_R, \quad \frac{dg_R}{d\bar{t}} = \frac{\delta_2 \beta_1 p_{T0}v_R}{\delta_1^2(p_T + k_{d2})} - \frac{\delta_2}{\delta_1}g_R, \quad (27)$$

$$c_{R0} = \frac{X}{k_{d1}}, \quad c_R = \frac{v_R \beta_1 p_{T0}}{\delta_1(p_T + k_{d2})}. \quad (28)$$

Then, we can use the fast variable approximation for c_R given in (28) to rewrite the system dynamics in the original variable $y_R = v_R - c_R$, to obtain

$$\text{System 2}: \quad \dot{y}_R = (1 - R)\left(\frac{X}{k_{d1}} - y_R\right),$$
$$\dot{g}_R = \frac{\delta_2 \beta_1 p_{T0}y_R}{\delta_1^2 k_{d2}} - \frac{\delta_2}{\delta_1}g_R. \quad (29)$$

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

$$\text{System 1}: \quad \dot{y}_N = \frac{X}{k_{d1}} - y_N,$$
$$\dot{g}_N = \frac{\delta_2 \beta_1 p_{T0}y_N}{\delta_1^2 k_{d2}} - \frac{\delta_2}{\delta_1}g_N, \quad (30)$$

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 [25]. 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. Due to the linearity of drift terms in system (23)–(24), 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 (30) and (29). Using the linearity of the systems (30) and (29) to directly evaluate the frequency response with a periodic input of the form $\tilde{X} = k_2 sin(\bar{\omega}\bar{t})$, we calculate the error in Y and G as

$$\frac{|\Delta y_R(j\bar{\omega})|}{|y_N(j\bar{\omega})|} = \frac{|\Delta g_R(j\bar{\omega})|}{|g_N(j\bar{\omega})|} = \frac{R\bar{\omega}}{\sqrt{\bar{\omega}^2 + (1 - R)^2}}. \quad (31)$$

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 (23)–(26) 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 y is a mixed variable whose dynamics consist of both slow and fast components. 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 (26) to derive the first and second moment dynamics for the variable y (see [17] for details) to obtain

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

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

Then, using the first and second moment dynamics we derive the dynamics for the variance of y given by $\mathbb{E}[(\Delta y_S)^2] = \mathbb{E}[(y - \mathbb{E}[y])^2] = \mathbb{E}[y^2] - \mathbb{E}[y]^2$ (see [17] for details). As the moment dynamics are linear, we then directly evaluate the frequency response of $\mathbb{E}[(\Delta y_S)^2]$ with the input $\tilde{X} = k_2 sin(\bar{\omega}\bar{t})$. Next, by normalizing the magnitude of the frequency response by the average signal $|y_R(j\bar{\omega})|k_2 = \frac{(1-R)k_2}{k_{d1}\sqrt{(\bar{\omega}^2 + (1-R)^2)}}$, we obtain

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

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 terms, the expression $\mathbb{E}[(\Delta g_S)^2]$ 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. Then, using the moment dynamics of the system (23) to derive the the variance dynamics for v and g (see [17] for details) and evaluating their frequency response we can quantify the noise error in G as

$$\frac{\sqrt{|\mathbb{E}[\Delta g_S^2](j\bar{\omega})|}}{|g_R(j\bar{\omega})|\sqrt{k_2}} = \sqrt{\frac{\delta_2^2\delta_1^2 + \delta_1^4\bar{\omega}^2}{k_2\ \Omega}}\sqrt[4]{A(p_T, \bar{\omega})} \quad (32)$$

where the function $A(p_T, \bar{\omega})$ decreases with increasing p_T for sufficiently small $\bar{\omega}$, as shown in the Appendix. 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, as previously observed in literature [22], [23], a higher value of p_T should be used to decrease the noise in G.

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(b) 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 for the noise error in G given by equation (32) with the retroactivity error in G given by equation (31) demonstrates a trade-off between the stochastic and deterministic perturbations in the signal G. Fig. 4(c) illustrates this trade-off for p_T in the range 1:1000 nM. Similarly, the expressions for the retroactivity error in (31) and the noise error in (32) 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.

VI. Conclusion

In this technical note, we have considered the problem of model order reduction for a class of SDEs in singular perturbation form, as found in biomolecular systems modeled by CLEs. We presented a reduced-order model that approximates both the slow and fast variables of the original system, which can be obtained by simply solving two algebraic equations. We then illustrated the application of our results with several examples. In particular, we considered a biomolecular system application, in which we quantified a trade-off between deterministic and stochastic signal transmission properties. Through this example, we demonstrate how both fast and slow variable 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

Deriving the dynamics for the means $\mathbb{E}[v]$, $\mathbb{E}[g]$ and variances $\mathbb{E}[(\Delta v)^2]$, $\mathbb{E}[\Delta vg]$, $\mathbb{E}[(\Delta g_S)^2]$ (see [17] for details), and evaluating the magnitude of the frequency response, we find that the noise error is given by equation (32) where the function $A(p_T, \bar{\omega})$ is in the form $A(p_T, \bar{\omega}) = \frac{\left((1-R)^2 + \bar{\omega}^2\right)(N_1 + N_2 + N_3 + N_4)}{(\beta_1 p_{T0}/k_{d1})^2 \beta_2^2 \delta_1^2 p_T^2 D_1 D_2}$, with

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

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

$N_{d1} = \delta_2\left(\beta_2^2 p_T^3 + \beta_2\delta_2 p_T\left(k_{d2}^2 + 3k_{d2}p_T + 2p_T^2\right) + \delta_2^2(k_{d2} + p_T)^3\right)$,
$N_{d2} = \delta_1^2 k_{d2}^2(\beta_2 p_T + 3\delta_2(k_{d2} + p_T))$,
$N_{d3} = \delta_1\delta_2 k_{d2}\left(\beta_2 p_T(2k_{d2} + 3p_T) + 3\delta_2(k_{d2} + p_T)^2\right)$.

It follows that the derivative is negative for all parameter conditions. 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 $\frac{\partial A(p_T, \bar{\omega})}{\partial p_T}$ is continuous with respect to $\bar{\omega}$ and 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

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

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

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

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

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

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

[10] 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.

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

[12] 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.

[13] 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.

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

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

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

[17] N. Herath and D. Del Vecchio. Deterministic-like model reduction for a class of multi-scale stochastic differential equations with application to biomolecular systems (extended version). http://hdl.handle.net/1721.1/110896, 2017. [Online].

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

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

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

[21] S. Jayanthi, K. Nilgiriwala, and D. Del Vecchio. Retroactivity controls the temporal dynamics of gene transcription. *ACS Synth. Biol.*, 2013.

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

[23] 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.

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

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

[26] 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.

[27] 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.

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

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

[30] 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.

[31] 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].