Reduced Linear Noise Approximation for Biochemical Reaction Networks with Time-scale Separation : The stochastic tQSSA$^+$

Narmada Herath[a]

Department of Electrical Engineering and Computer Science,

Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge,

MA 02139, USA

Domitilla Del Vecchio[b]

Department of Mechanical Engineering, Massachusetts Institute of Technology,

77 Massachusetts Avenue, Cambridge, MA 02139, USA

(Dated: 14 February 2018)

Biochemical reaction networks often involve reactions that take place on different time-scales, giving rise to 'slow' and 'fast' system variables. This property is widely used in the analysis of systems to obtain dynamical models with reduced dimensions. In this paper, we consider stochastic dynamics of biochemical reaction networks modeled using the Linear Noise Approximation (LNA). Under time-scale separation conditions, we obtain a reduced-order LNA that approximates both the slow and fast variables in the system. We mathematically prove that the first and second moments of this reduced-order model converge to those of the full system as the time-scale separation becomes large. These mathematical results, in particular, provide a rigorous justification to the accuracy of LNA models derived using the stochastic total quasi-steady state approximation (tQSSA). Since, in contrast to the stochastic tQSSA, our reduced-order model also provides approximations for the fast variable stochastic properties, we term our method the 'stochastic tQSSA$^+$'. Finally, we demonstrate the application of our approach on two biochemical network motifs found in gene-regulatory and signal transduction networks.

[a]nherath@mit.edu

[b]ddv@mit.edu

I. INTRODUCTION

Many biochemical processes involve reactions that occur on different time-scales. For example, in bacterial cells, the binding of transcription factor to DNA takes place on the time-scale of seconds, while protein production and dilution are on the order of hours[1]. Such a separation in time-scales allows the system variables to be separated into slow and fast groups, and this property can be exploited to reduce the complexity of dynamical models. In particular, for deterministic systems, the quasi-steady state approximation gives a reduced-order model for the slow variables, assuming that the fast variables rapidly reach a steady state[2,3]. In the mathematical literature, a system of ordinary differential equations (ODEs) with multiple time-scales is represented as slow and fast subsystems by using a small parameter ϵ to capture the separation in time-scales. The mathematical treatment of such systems is given by two main methods: singular perturbation and averaging[4,5]. The singular perturbation approach, formalized by Tikhonov's theorem, involves setting $\epsilon = 0$ in the system dynamics to obtain an algebraic equation that approximates the fast variables, which is in turn used to derive an approximation for the slow variable dynamics[4,6,7]. In the averaging method, a reduced-order model for the slow variables is obtained by the elimination of the fast dynamics via integration of the system functions[5].

As opposed to deterministic models, employing time-scale separation for model order reduction remains an ongoing area of research for stochastic models of biological systems. Yet, obtaining reduced descriptions of stochastic dynamics is even more critical than for deterministic dynamics in order to increase the speed of simulation and aid analytical studies. Furthermore, accurate reduced-order models are important for precise parameter estimations[8]. The most prominent model used to capture the stochasticity in biological systems is the Chemical Master Equation (CME), which considers the species counts as a set of discrete random variables and describes the time evolution of their probability distributions using a set of ordinary differential equations[9,10]. There have been several works that obtain reduced-order representations of the CME under time-scale separation conditions[11–25]. Among these, a common approach used to reduce the complexity in the simulations of the CME is to approximate the fast variables by their deterministic quasi-steady state expressions[19]. However, validity of this method still remains under investigation[20,22,25–27].

The theoretical analysis of the CME is challenging due to the large system size and the

lack of appropriate analytical tools. Therefore, several approximations to the Chemical Master Equation have been developed under the assumptions that the system's volume and the molecular counts are sufficiently large. One such approximation is the Fokker-Planck equation (FPE) where partial differential equations are used to describe the evolution of the probability distribution of the species counts[9]. Another approximation, equivalent to the FPE, is the chemical Langevin equation (CLE) where the dynamics of species counts are described by stochastic differential equations[28].

There have been several recent works that consider the problem of model order reduction for biochemical reaction networks modeled by the CLE. The work by Contou-Carrere et al. obtains a reduced-order system for the slow variables of the CLE by adiabatic elimination of the fast variable dynamics[29]. They provide a numerical analysis on the error between the full and reduced-order systems; however, an analytical error quantification is not provided. Furthermore, their work does not provide an approximation for the fast variables of the full system. In our previous work[30–32], we considered CLE models of biochemical reaction networks with linear reaction rates and obtained a reduced-order system for both slow and fast variables, following a similar approach to the deterministic singular perturbation technique. It was mathematically demonstrated that the moments of the reduced system are within an $O(\epsilon)$-neighborhood of the moments of the full system.

In addition to the above methods, the mathematical literature also offers model order reduction techniques for multi-scale stochastic differential equations via averaging methods[5,33]. Recently, these methods have been applied in the analysis of systems modeled by the chemical Langevin equation[34]. However, averaging methods require integration of system functions, which may be challenging for systems that are of high dimension or are nonlinear. Moreover, the averaging methods also do not provide an approximation to the fast variable dynamics. In the case of biochemical reaction networks, it is typically important to approximate the fast variable dynamics, because many species are mixed - that is, their concentrations are given by the combination of slow and fast variable concentrations. Therefore, we often require both slow and fast components in the reduced-order model to correctly approximate such species dynamics.

The Linear Noise Approximation (LNA) is another approximate model for the CME, where stochasticity is represented as random fluctuations around a deterministic trajectory using stochastic differential equations or partial differential equations[10,35]. Recently, model

order reduction methods for the LNA have been developed using projection operators[27,36] or singular perturbation analysis[37]. However, in these works, the error between the full and reduced-order models are not analytically quantified. The work by Sootla and Anderson[38] provides an error quantification for model order reduction of LNA developed by Thomal et al.[27,36], but to do so they assume Lipschitz continuity of system functions, which are not Lipschitz-continuous for general systems. Furthermore, the above works only provide an approximation for the slow variable dynamics and do not approximate the fast variables.

In this work, we consider biochemical reaction networks modeled using the LNA. We consider systems with separation of time-scales where the dynamics can be represented in the singular perturbation form with ϵ as the singular perturbation parameter. We then obtain a reduced-order model that approximates both slow and fast variables. We mathematically demonstrate that first and second moments of the reduced system variables are within an $O(\epsilon)$-neighborhood of the first and second moments of the full system variables. These results, in turn, provide a rigorous justification for the accuracy of LNA models derived using the stochastic tQSSA in comparison to the standard quasi-steady state approximation (QSSA). Furthermore, in contrast to the stochastic tQSSA, our reduced system also provides approximations for the fast variables stochastic properties. Hence, we term our method the stochastic tQSSA$^+$. The application of our approach is then demonstrated on two biochemical network motifs found in gene-regulatory networks and signal transduction cascades. Through these examples, we illustrate the practical applications of the reduced-order models and the necessity of both slow and fast variable approximations for analysis. In particular, using the reduced-order model for the gene-regulatory network motif, we further investigate the parameter conditions under which the standard QSSA provides accurate results in the stochastic setting.

This paper is organized as follows. In Section II, we present the LNA model considered in this paper. In Section III, we introduce the reduced-order system and present our results on the error quantification between the full and reduced dynamics. In Section V, we illustrate the application of the results with two examples of biochemical network motifs.

Notation: $\mathbb{E}[\cdot]$ denotes the expected value of a random variable. $\|\cdot\|$ denotes the Euclidean norm for vectors and $\|\cdot\|_F$ denotes the Frobenius norm for matrices.

II. THE LINEAR NOISE APPROXIMATION WITH TIME-SCALE SEPARATION

A. Linear Noise Approximation

Consider a biochemical reaction network with n species $Y_1, \ldots, Y_n$, in a volume Ω, interacting through m reactions of the form:

$$p_{i1}Y_1 + \ldots + p_{in}Y_n \xrightarrow{k_i} r_{i1}Y_1 + \ldots + r_{in}Y_n, \qquad i = 1, \ldots, m,$$
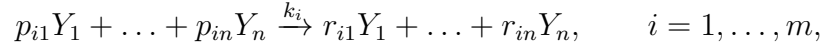
where k_i denotes the rate constant of reaction i and $p_{il} - r_{il}$ is the change in the number of molecules of Y_l due to the reaction i. Let $y(t) = [y_1, \ldots, y_n]^T$ be the state of the system at a given time t where each component y_i represents the molecular count for each species as a discrete random variable. Then, the Chemical Master Equation for this system is of the form

$$\frac{\partial P(y,t)}{\partial t} = \sum_{i=1}^{m} [a_i(y - q_i, t)P(y - q_i, t) - a_i(y, t)P(y, t)], \tag{1}$$

where $a_i(y, t)$ is the microscopic reaction rate proportional to k_i with $a_i(y, t)dt$ being the probability that the reaction i will take place in an infinitesimal time step dt. The variable $q_i = r_i - p_i$ is the stoichiometry vector where $p_i = [p_{i1}, \ldots, p_{in}]^T$ and $r_i = [r_{i1}, \ldots, r_{i1n}]^T$ for $i = 1, \ldots, m$ [39].

The Linear Noise Approximation (LNA) is an approximation to the CME, where the molecular counts are represented by continuous variables under the assumption that the system volume and the molecular counts are sufficiently large. As shown in the work of van Kampen[10], the LNA is derived by taking $y = \Omega v + \sqrt{\Omega}\xi$ in the CME, where Ω is the system volume, v is a vector of deterministic variables and ξ is a vector of random variables that represents the stochastic fluctuations. Then, performing a Taylor series expansion about the deterministic variable Ωv and equating the terms of order $\Omega^{1/2}$ and Ω^0, it is shown that v gives the macroscopic concentrations and the elements of ξ are Gaussian random variables with the dynamics

$$\dot{v} = f(v, t), \tag{2}$$

$$\dot{\xi} = A(v, t)\xi + \sigma(v, t)\Gamma, \tag{3}$$

in which Γ is an m-dimensional white noise process, $f(v,t) = \sum_{i=1}^{m} q_i \tilde{a}_i(v,t)$, $A(v,t) = \frac{\partial f(v,t)}{\partial v}$ and $\sigma(v,t) = [q_1 \sqrt{\tilde{a}_1(v,t)}, \ldots, q_m \sqrt{\tilde{a}_m(v,t)}]$. The function $\tilde{a}_i(v,t)$ is the macroscopic reaction rate which can be approximated by $\tilde{a}_i(v,t) = \frac{1}{\Omega} a_i(\Omega v, t)$ as $\Omega \to \infty$ and $y \to \infty$ such that the concentration $v = y/\Omega$ remains constant[40].

B. System model with time-scale separation

In this work, we consider biochemical reaction networks in which the chemical reactions take place on two well-separated time-scales. For the system (2)–(3), let m_s be the number of slow reactions and m_f be the number of fast reactions where $m_s + m_f = m$. Then, by using a small parameter ϵ, the reaction rate vector can be arranged in the form $\tilde{a}(v,t) = [\hat{a}_s(v,t), (1/\epsilon)\hat{a}_f(v,t)]^T$ where $\hat{a}_s(v,t) \in \mathbb{R}^{m_s}$ represents the reaction rates of slow reactions and $(1/\epsilon)\hat{a}_f(v,t) \in \mathbb{R}^{m_f}$ represents the reaction rates of fast reactions. The corresponding stoichiometry vectors q_i can be written in the form $q = [q_1, \ldots, q_{m_s}, q_{m_s+1}, \ldots, q_{m_s+m_f}]$ where q_i for $i = 1, \ldots, m_s$ represent the change in the molecular counts given by the slow reactions, and q_i for $i = m_s + 1, \ldots, m_s + m_f$ represent the change in the molecular counts given by the fast reactions. Because, chemical species often take part in both slow and fast reactions, the above separation in reaction rates does not necessarily correspond to a partitioning of the system's species into fast and slow. Often, a coordinate change is necessary to identify the slow and fast variables in the system and write it in the standard singular perturbation form[29,41]. Therefore, here we consider systems in which the species can be partitioned into n_s slow variables and n_f fast variables with $n_s + n_f = n$, according to the following claim:

Claim 1. *Assume that there is an invertible matrix $T = [T_x^T, T_z^T]^T$ with $T_x \in \mathbb{R}^{n_s \times n}$ and $T_z \in \mathbb{R}^{n_f \times n}$ such that the change of variables $x = T_x v$, $z = T_z v$, takes the system (2) into the singular perturbation form*

$$\dot{x} = f_x(x, z, t), \tag{4}$$

$$\epsilon \dot{z} = f_z(x, z, t, \epsilon). \tag{5}$$

Then, the change of variables $\psi_x = T_x \xi$, $\psi_z = T_z \xi$ transforms system (3), into the singular perturbation form

$$\dot{\psi}_x = S_x(x, z, t)\psi_x + S_z(x, z, t)\psi_z + \sigma_x(x, z, t)\Gamma_x, \tag{6}$$

$$\epsilon \dot{\psi}_z = F_x(x,z,t,\epsilon)\psi_x + F_z(x,z,t,\epsilon)\psi_z + \sigma_z(x,z,t,\epsilon)\Gamma_z, \qquad (7)$$

where Γ_x is an m_s-dimensional white noise process, $\Gamma_z = [\Gamma_x^T, \Gamma_f^T]^T$, where Γ_f is an m_f-dimensional white noise process and

$$S_x(x,z,t) = \frac{\partial f_x(x,z,t)}{\partial x}, \qquad\qquad S_z(x,z,t) = \frac{\partial f_x(x,z,t)}{\partial z},$$

$$F_x(x,z,t,\epsilon) = \frac{\partial f_z(x,z,t,\epsilon)}{\partial x}, \qquad\qquad F_z(x,z,t,\epsilon) = \frac{\partial f_z(x,z,t,\epsilon)}{\partial z},$$

$$\sigma_x(x,z,t) = T_x\left[q_1\sqrt{\hat{a}_{s1}(T^{-1}[x^T,z^T]^T,t)}, \ldots, q_{m_s}\sqrt{\hat{a}_{sm_s}(T^{-1}[x^T,z^T]^T,t)} \right],$$

$$\sigma_z(x,z,t,\epsilon) = \left[\begin{matrix} \epsilon\left[q_1\sqrt{\hat{a}_{s1}(T^{-1}[x^T,z^T]^T,t)}, \ldots, q_{m_s}\sqrt{\hat{a}_{sm_s}(T^{-1}[x^T,z^T]^T,t)} \right]^T T_z^T \\ \left[q_{m_s+1}\sqrt{\epsilon\hat{a}_{f1}(T^{-1}[x^T,z^T]^T,t)}, \ldots, q_{m_s+m_f}\sqrt{\epsilon\hat{a}_{fm_f}(T^{-1}[x^T,z^T]^T,t)} \right]^T T_z^T \end{matrix} \right]^T.$$

Proof. See Appendix A.

$\square$

Note, that the noise Γ_x on the slow variables is m_s-dimensional, while the noise Γ_z is $(m_s + m_f)$-dimensional. This is due to the fact that the fast reactions do not affect the slow variable dynamics in the above form, i.e., ϵ does not appear in the slow variable dynamics (4), but the fast dynamics can be dependent on both slow and fast reactions (see Appendix A for details).

There are several works in the literature that have investigated existence of such coordinate transformations that allow the separation of slow and fast variables for deterministic systems[42] and also for chemical Langevin equation models[29].

Following the results of Claim 1, we consider biochemical reaction networks where the Linear Noise Approximation model can be written in the standard singular perturbation form

$$\dot{x} = f_x(x,z,t), \qquad\qquad\qquad x(0) = x_0, \qquad (8)$$

$$\epsilon\dot{z} = f_z(x,z,t,\epsilon), \qquad\qquad\qquad z(0) = z_0, \qquad (9)$$

$$\dot{\psi}_x = S_x(x,z,t)\psi_x + S_z(x,z,t)\psi_z + \sigma_x(x,z,t)\Gamma_x, \qquad \psi_x(0) = \psi_{x0}, \qquad (10)$$

$$\epsilon\dot{\psi}_z = F_x(x,z,t,\epsilon)\psi_x + F_z(x,z,t,\epsilon)\psi_z + \sigma_z(x,z,t,\epsilon)\Gamma_z, \qquad \psi_z(0) = \psi_{z0}, \qquad (11)$$

where $x \in \mathbb{R}^{n_s}$, $\psi_x \in \mathbb{R}^{n_s}$ are the slow variables and $z \in \mathbb{R}^{n_f}$, $\psi_z \in \mathbb{R}^{n_f}$ are the fast variables. Γ_x is an m_s-dimensional white noise process. Then, $\Gamma_z = [\Gamma_x^T, \Gamma_f^T]^T$, where Γ_f is an

m_f-dimensional white noise process.

We refer to the system (8) - (11) as the full system and we first make the following assumptions for $x \in \mathbb{R}^{n_s}$ and $z \in \mathbb{R}^{n_f}$.

Assumption 1. The functions $f_x(x, z, t)$, $f_z(x, z, t, \epsilon)$ are twice continuously differentiable. The Jacobian $\frac{\partial f_z(x,z,t,0)}{\partial z}$ has continuous first and second partial derivatives with respect to its arguments.

Assumption 2. The matrix-valued functions $\sigma_x(x, z, t)\sigma_x(x, z, t)^T$, $\sigma_z(x, z, t, \epsilon)[\sigma_x(x, z, t) \ 0]^T$ and $\sigma_z(x, z, t, \epsilon)\sigma_z(x, z, t, \epsilon)^T$ are continuously differentiable. Furthermore, let $\Lambda(x, z, t, \epsilon) = \frac{\sigma_z(x,z,t,\epsilon)\sigma_z(x,z,t,\epsilon)^T}{\epsilon}$. We have that $\sigma_z(x, z, t, 0) = 0$ and $\Lambda(x, z, t, 0)$ is bounded for given x, z, t.

Assumption 3. There exists an isolated real root $z = \gamma_1(x, t)$, for the equation $f_z(x, z, t, 0) = 0$, for which, the matrix $\frac{\partial f_z(x,z,t,0)}{\partial z}\big|_{z=\gamma_1(x,t)}$ is Hurwitz (i.e. eigenvalues of the matrix have strictly negative real parts), uniformly in x and t. Furthermore, we have that the first partial derivative of $\gamma_1(x, t)$ is continuous with respect to its arguments. Also, the initial condition z_0 is in the region of attraction of the equilibrium point $z = \gamma_1(x_0, 0)$ for the system $\frac{dz}{d\tau} = f_z(x_0, z, 0, 0)$.

Assumption 4. The system $\dot{x} = f_x(x, \gamma_1(x, t), t)$ has a unique bounded solution for $t \in [0, t_1]$.

In the next section, we present the reduced-order model that we define to approximate the slow and fast variables when ϵ is small in the full system (8) - (11), and quantify the error between the moment dynamics of the full and the reduced systems.

III.   RESULTS

To define the reduced system we follow a similar approach to the singular perturbation theory[4] by setting $\epsilon = 0$ in the full system (8) - (11). This yields

$$f_z(x, z, t, 0) = 0, \tag{12}$$

$$F_x(x, z, t, 0)\psi_x + F_z(x, z, t, 0)\psi_z = 0. \tag{13}$$

Let $z = \gamma_1(x, t)$ be an isolated root of equation (12). Then, it follows that the unique solution to equation (13) is

$$\psi_z = -F_z(x, \gamma_1(x, t), t, 0)^{-1} F_x(x, \gamma_1(x, t), t, 0)\psi_x.$$

Note that the invertibility of matrix F_z follows from Assumption 3 where it is assumed that $F_z(x, \gamma_1(x, t), t, 0) = \frac{\partial f_z(x, z, t, 0)}{\partial z}\big|_{z = \gamma_1(x, t)}$ is Hurwitz. Let $\gamma_2(x, t) = -F_z(x, \gamma_1(x, t), t, 0)^{-1} F_x(x, \gamma_1(x, t), t, 0)$. Substituting $z = \gamma_1(x, t)$ and $\psi_z = \gamma_2(x, t)\psi_x$ in equations (8) and (10), we obtain the following candidate approximation for the slow variable dynamics:

$$\dot{\bar{x}} = f_x(\bar{x}, \gamma_1(\bar{x}, t), t), \qquad\qquad \bar{x}(0) = x_0, \qquad\qquad (14)$$

$$\dot{\bar{\psi}}_x = S(\bar{x}, t)\bar{\psi}_x + \sigma_x(\bar{x}, \gamma_1(\bar{x}, t), t)\Gamma_x, \quad \bar{\psi}_x(0) = \psi_{x0}, \qquad (15)$$

where $S(\bar{x}, t) = S_x(\bar{x}, \gamma_1(\bar{x}, t), t) + S_z(\bar{x}, \gamma_1(\bar{x}, t), t)\gamma_2(\bar{x}, t)$.

Next, to introduce a candidate approximation for the fast variable dynamics we define:

$$\bar{z} = \gamma_1(\bar{x}, t), \qquad\qquad\qquad\qquad (16)$$

$$\bar{\psi}_z = \gamma_2(\bar{x}, t)\bar{\psi}_x + g(\bar{x}, t)N(0, 1), \qquad\qquad (17)$$

where $N(0, 1) \in \mathbb{R}^d$ is a vector of standard normal random variables and $g(x, t) : \mathbb{R}^{n_s} \times \mathbb{R} \to \mathbb{R}^{n_f \times d}$ is a function that satisfies the Lyapunov equation

$$F_z(x, \gamma_1(x, t), t, 0)g(x, t)g(x, t)^T$$
$$+ g(x, t)g(x, t)^T F_z(x, \gamma_1(x, t), t, 0)^T$$
$$= -\Lambda(x, \gamma_1(x, t), t, 0), \qquad (18)$$

where $\Lambda(x, z, t, \epsilon)$ is defined in Assumption 2.

We refer to the equations (14)–(17) as the *reduced system*. Next, we show that the reduced system (14)–(17) is a good approximation of the full system (8)–(11) as ϵ tends to zero. In particular, we demonstrate that the error in the first and second moments between the full and the reduced system variables are $O(\epsilon)$. To this end, denote by $\mathbb{E}[\cdot]$ the expected value of a random variable. Then, we have the following theorem:

Theorem 1. *Consider the full system (8)–(11) and the reduced system in (14)–(17). Then, under Assumptions 1 - 4, there exist $\epsilon^* \geq 0$ such that for $0 < \epsilon < \epsilon^*$, $t \in [0, t_1]$ we have*

$$\|x(t) - \bar{x}(t)\| = O(\epsilon), \qquad\qquad (19)$$

9

$$\|\mathbb{E}[\psi_x(t)] - \mathbb{E}[\bar{\psi}_x(t)]\|\| = O(\epsilon), \qquad (20)$$

$$\|\mathbb{E}[\psi_x(t)\psi_x(t)^T] - \mathbb{E}[\bar{\psi}_x(t)\bar{\psi}_x(t)^T]\|_F = O(\epsilon), \qquad (21)$$

and for any $0 < t_b < t_1$ there exists $\epsilon^{**} \leq \epsilon^*$ such that for $\epsilon < \epsilon^{**}$ and $t \in [t_b, t_1]$ we have

$$\|z(t) - \bar{z}(t)\| = O(\epsilon), \qquad (22)$$

$$\|\mathbb{E}[\psi_z(t)] - \mathbb{E}[\bar{\psi}_z(t)]\|\| = O(\epsilon), \qquad (23)$$

$$\|\mathbb{E}[\psi_z(t)\psi_x(t)^T] - \mathbb{E}[\bar{\psi}_z(t)\bar{\psi}_x(t)^T]\|_F = O(\epsilon), \qquad (24)$$

$$\|\mathbb{E}[\psi_z(t)\psi_z(t)^T] - \mathbb{E}[\bar{\psi}_z(t)\bar{\psi}_z(t)^T]\|_F = O(\epsilon). \qquad (25)$$

The proof of this theorem is presented in Appendix B together with several intermediate results that are used in the proof. We provide an outline of the proof here. First, we derive the moment dynamics of the full system (8) - (11) and show that they are also in the standard singular perturbation form (Lemma 1, Appendix B). Then, we derive the moment dynamics of the reduced system (14)–(17) in Lemma 2. Next, we show that setting $\epsilon = 0$ in the moment dynamics of the full system yields the moment dynamics of the reduced system (Lemma 3 and 4, Appendix B). As the moment dynamics are deterministic we then apply the Tikhonov's theorem[4] using the stability conditions provided by Assumption 3 to prove Theorem 1.

Theorem 1 shows that the reduced system (14)–(17) provides a good approximation to the slow and fast variables of the full system (8) - (11) in terms of the first and second moments of the stochastic fluctuations ψ_x and ψ_z. Thus, we have that, as ϵ tends to zero, the mean, variance and the covariance of the stochastic fluctuations ψ_x and ψ_z are well approximated by those of the reduced system variables $\bar{\psi}_x$ and $\bar{\psi}_z$. Therefore, in the case where the full system in the singular perturbation form is obtained via a coordinate change as in Claim 1, we have that the reduced system provides a good approximation for the variables v and ξ that consists of both slow and fast dynamics (see Appendix B). This is illustrated in the diagram of Figure 1. We also note that these results hold under the stability assumption of the Tikhonov's theorem required for the deterministic dynamics, given by Assumption 3, and that there are no additional stability assumptions.

From the definition of the reduced system (14)–(15), we note that the slow variable dynamics can be approximated by setting $\epsilon = 0$, similar to singular perturbation in deterministic systems[4]. However, the fast variable approximation $\bar{\psi}_z$ requires an additional term
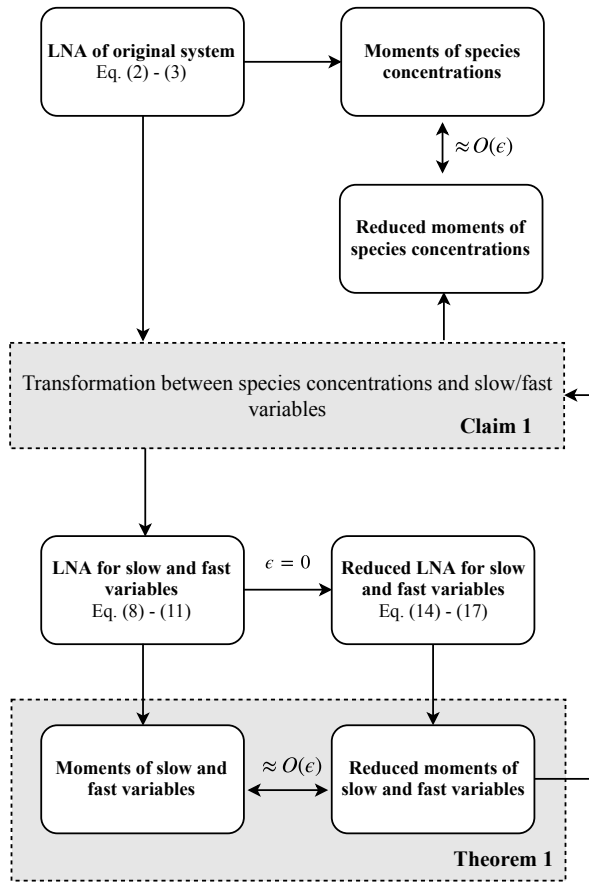
FIG. 1. Schematic diagram illustrating the model reduction approach given by Theorem 1. The $O(\epsilon)$-closeness of the moments of the original species concentrations and the moments obtained via the reduced system is proved in Appendix C.

$g(\bar{x}, t)$. This additional term is required to capture the noise properties of the fast variable, which are not captured by $\gamma_2(\bar{x}, t)$ alone, since setting $\epsilon = 0$ in equation (11) eliminates the diffusion term $\sigma_z(x, z, t, \epsilon)$, which contributes to the second moment of the fast variable. To illustrate this point, we provide a simple example. Consider the system

$$\frac{d\psi_x}{dt} = a_1\psi_x + a_2\psi_z, \tag{26}$$

$$\epsilon\frac{d\psi_z}{dt} = a_3\psi_z + \sqrt{\epsilon}v_1\Gamma, \tag{27}$$

in which referring to system (10)–(11) we have $\sigma_x = 0$ and $\sigma_z = v_1$. Setting $\epsilon = 0$ in this system gives

$$\frac{d\bar{\psi}_x}{dt} = a_1\bar{\psi}_x, \quad \bar{\psi}_z = 0. \tag{28}$$

11

Computing the steady state second moments of system (26)–(27) and system (28) obtained by setting $\epsilon = 0$ we have

$$\mathbb{E}[\psi_x^2] = \frac{a_2^2 v_1^2}{2a_1} \frac{\epsilon}{(1 + a_1 \epsilon)}, \qquad\qquad \mathbb{E}[\bar{\psi}_x^2] = 0,$$

$$\mathbb{E}[\psi_z^2] = \frac{v_1^2}{2a_3}, \qquad\qquad\qquad\quad \mathbb{E}[\bar{\psi}_z^2] = 0.$$

From here, it can be seen that as ϵ tends to zero the second moment of ψ_x tends to the moment of the approximation $\bar{\psi}_x$, however, the moment of ψ_z does not converge to the moment of $\bar{\psi}_z$. Thus, the reduced system obtained by setting $\epsilon = 0$ provides a good approximation for the moments of the slow variable, but not for those of the fast variable.

As it can be seen from equations (27)–(28), the reason for the poor approximation of the fast variable is due to the elimination of the noise term v_1 when ϵ is set to zero in (26)–(27).

This can be further analyzed by representing the full system in the fast time-scale $\tau = t/\epsilon$:

$$\frac{d\psi_x}{d\tau} = \epsilon(a_1 \psi_x + a_2 \psi_z), \tag{29}$$

$$\frac{d\psi_z}{d\tau} = a_3 \psi_z + v_1 \tilde{\Gamma}, \tag{30}$$

where $\tilde{\Gamma} = \sqrt{\epsilon}\, \Gamma$ is a white noise process in the fast time-scale. Here, we observe that setting $\epsilon = 0$ does not eliminate the noise term of the fast variable dynamics, and thus, this noise term should be taken into account in the fast variable approximation. The term $g(\bar{x}, t)$ in the reduced fast system captures this fast variable noise. This can be seen by computing the values of $\gamma_2(\bar{x}, t)$ and $g(\bar{x}, t)^2$ for the system (26)-(27). From (27) and (13), we have that $\gamma_2(\bar{x}, t) = 0$. Writing the Lyapunov equation (18) for the system (27), we have

$$a_3 g(\bar{x}, t)^2 + g(\bar{x}, t)^2 a_3 = v_1^2,$$

which yields $g(\bar{x}, t)^2 = v_1^2/2a_3$. This gives the steady state moment $\mathbb{E}[\bar{\psi}_z^2] = v_1^2/2a_3$ for the reduced fast system, which is the same as $\mathbb{E}[\psi_z^2]$. The dynamics of the slow variable ψ_x vary at a much slower rate than the dynamics of ψ_z, and therefore, the noise of the fast variable ψ_z can be neglected in the slow variable approximation as it is essentially 'filtered out'[30,32].

Remark 1: The stochastic fluctuations ψ_x and ψ_z are multivariate Gaussian random variables and thus their probability distributions are fully characterized by the mean and the covariance[10]. From Theorem 1, we have that the first and second moments of $\psi_x(t)$

converge to the moments of the vector $\bar{\psi}_x(t)$, and the first and second moments of $\psi_z(t)$ converge to the moments of $\bar{\psi}_z(t)$, as ϵ tends to zero. Thus, we further have that the vectors $\psi_x(t)$ and $\psi_z(t)$ converge in distribution to the vectors $\bar{\psi}_x(t)$ and $\bar{\psi}_z(t)$, respectively.

Next, we investigate how the reduced system derived in this section relates to the commonly used total quasi-steady state approximation in stochastic analysis.

IV. VALIDITY OF LNA MODELS DERIVED USING STOCHASTIC TQSSA

The quasi-steady state approximation is widely used in the deterministic setting to obtain reduced-order models of biochemical reaction networks under time-scale separation conditions. Recently, the QSSA has also been extended to stochastic systems, where the fast variables are approximated by their deterministic quasi-steady state expressions in stochastic simulations. This method is termed stochastic QSSA[20,43].

Several studies in the deterministic setting have shown that the validity of the standard QSSA is limited to certain parameter conditions. Thus, the total quasi-steady state approximation (tQSSA) has been introduced as an alternate approximation with increased accuracy[44,45]. The tQSSA, first proposed for enzymatic reactions, considers a coordinate change to identify the slow variables in the system. This has then been adapted in the stochastic setting where stochastic simulations of the CME are performed using the deterministic quasi-steady state expressions of the fast variables obtained using the tQSSA[20,43,46]. There are several studies that investigate the validity of these stochastic quasi-steady state approximations. Particularly, the work by J.K Kim et al. demonstrates, via two-dimensional systems modeled by the CME and LNA, that the tQSSA provides a better approximation to the original system in comparison to the standard QSSA[20,22]. Here, we use the results derived in Theorem 1 to provide a rigorous mathematical justification for the accuracy of the stochastic tQSSA for general reaction networks modeled through the LNA. The inaccuracy of the standard QSSA in the deterministic setting results from treating the species that consists of both slow and fast dynamics as purely slow variables. In contrast, the tQSSA involves defining the slow variables of the system via a coordinate change, in which the slow variable dynamics do not depend on the fast reactions. This corresponds to Claim 1, where

a coordinate transformation is used to identify the slow and fast variables in the system. Thus, the deterministic counterpart of the reduced LNA given by equation (14) is equivalent to the reduced system obtained by using the tQSSA. We next derive the dynamics of the stochastic fluctuations under the stochastic tQSSA. To this end, let $\hat{x}$ and $\hat{\psi}_x$ represent the variables in the LNA model obtained using the stochastic tQSSA. Then, we have that the deterministic dynamics are given by

$$\dot{\hat{x}} = f_x(\hat{x}, \gamma_1(\hat{x}, t), t). \tag{31}$$

Next, deriving the corresponding dynamics for the stochastic fluctuation using the definition of the LNA in Section II we have

$$\dot{\hat{\psi}}_x = \frac{\partial f_x(\hat{x}, \gamma_1(\hat{x}, t), t)}{\partial \hat{x}} \hat{\psi}_x + \sigma_x(\hat{x}, \gamma_1(\hat{x}, t), t).$$

Using the chain rule we can write this system as

$$\dot{\hat{\psi}}_x = \left(\frac{\partial f_x(\hat{x}, z, t)}{\partial \hat{x}} + \frac{\partial f_x(\hat{x}, z, t)}{\partial z} \frac{\partial \gamma_1(\hat{x}, t)}{\partial \hat{x}} \right)\Bigg|_{z=\gamma_1(\hat{x}, t)} \hat{\psi}_x$$
$$+ \sigma_x(\hat{x}, \gamma_1(\hat{x}, t), t). \tag{32}$$

Then, by the implicit function theorem[47] we have that

$$\frac{\partial \gamma_1(\hat{x}, t)}{\partial \hat{x}} = -\frac{\partial f_z(x, z, t, 0)}{\partial z}\Bigg|_{z=\gamma_1(\hat{x}, t)}^{-1} \frac{\partial f_z(x, z, t, 0)}{\partial x}\Bigg|_{z=\gamma_1(\hat{x}, t)},$$
$$= -F_z(\hat{x}, \gamma_1(\hat{x}, t), t, 0)^{-1} F_x(\hat{x}, \gamma_1(\hat{x}, t), t, 0),$$

where we have from Section III that $-F_z(\hat{x}, \gamma_1(\hat{x}, t), t, 0)^{-1} F_x(\hat{x}, \gamma_1(\hat{x}, t), t, 0) = \gamma_2(\hat{x}, t)$.

Substituting $\partial \gamma_1(\hat{x}, t)/\partial \hat{x} = \gamma_2(\hat{x}, t)$ in (32), we obtain

$$\dot{\hat{\psi}}_x = (S_x(\hat{x}, \gamma_1(\hat{x}, t), t) + S_z(\hat{x}, \gamma_1(\hat{x}, t), t)\gamma_2(\hat{x}, t))\hat{\psi}_x$$
$$+ \sigma_x(\hat{x}, \gamma_1(\hat{x}, t), t), \tag{33}$$

which is equivalent to the dynamics of $\bar{\psi}_x$ in the reduced system (8) and (11) derived through singular perturbation. Thus, we have that the LNA model obtained under stochastic tQSSA is equivalent to the slow variable approximations (14)–(15) in our reduced system. This is illustrated in the schematic in Figure 2. Then, from Theorem 1 we have that the moments of the LNA model obtained using stochastic tQSSA provide a good approximation
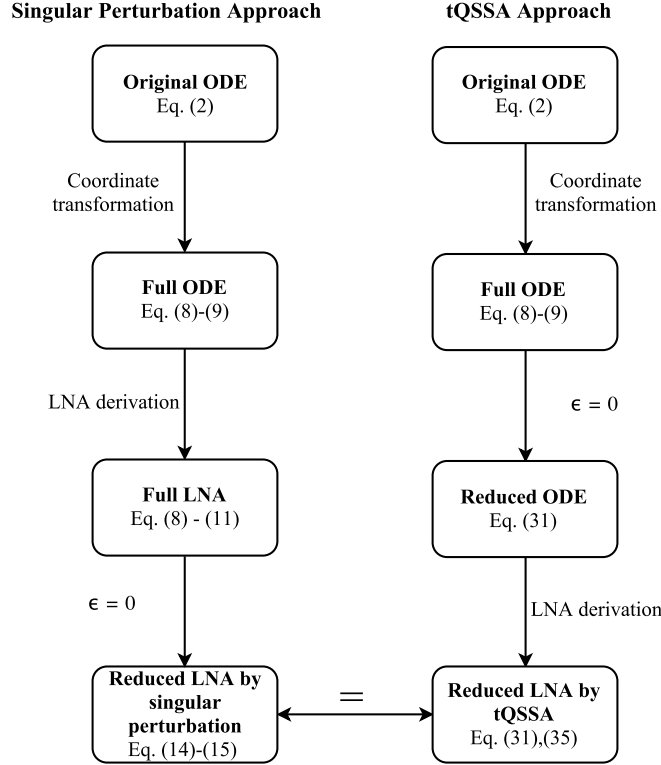
14

FIG. 2. Schematic diagram illustrating the relationship between the singular perturbation approach and tQSSA approach.

for the moments of slow variables (8) and (11) in the full system. This demonstrates that, similar to the deterministic setting, the LNA model obtained using the tQSSA provides a good approximation for the dynamics of the slow variables of the full LNA model, under all parameter conditions. Furthermore, this indicates that when the molecular counts are sufficiently large, the stochastic tQSSA yield a valid approximation for CME under time-scale separation, similar to the observations in previous studies[20,22].

Therefore, our results provide a rigorous justification for the validity of stochastic tQSSA models, which has been absent in previous work. However, unlike the tQSSA, the reduced model derived in (14)–(17) also provide approximations for the fast variable stochastic properties. Therefore, we term our method stochastic tQSSA$^+$. These fast variables stochastic properties can be used to analyze the reduced dynamics of the original species concentrations. Similar considerations in the literature suggests using the prefactor QSSA[48] method to study the dynamics of the fast variables in the reduced setting[20]. The prefactor QSSA method, first proposed for deterministic models of gene-regulatory networks, involves first

using the coordinate transformation to identify the slow and fast variables to perform the model reduction and then transforming the reduced system variables back into the original form[48]. This is equivalent to the singular perturbation approach in deterministic setting and thus produces accurate results in deterministic models. However, J.K. Kim et al. has shown that although this provides a good approximation in deterministic models, using the reactions rates obtained through the prefactor QSSA in stochastic simulations does not produce accurate results[20,22]. By contrast, our results provide a method to obtain accurate approximations of the original system variables.

In the next section, we demonstrate how the results developed in this section can be applied to biochemical reaction networks with two time-scales. Furthermore, we use an application example to investigate the parameter conditions under which the standard quasi-steady state provides a good approximation of the original system dynamics.

V. APPLICATION EXAMPLES

In this section, we illustrate the application of our results to two biochemical network motifs that exhibit time-scale separation. First, we consider a gene-regulatory network motif. We derive the reduced LNA and validate it through numerical simulations. We then use the reduced system to investigate the validity of quasi-steady state approximations that are often performed in the stochastic setting. As a second example, we consider a cellular signaling motif.

A. Example I: Gene-regulatory network motif

We consider the gene-regulatory network motif shown in Figure 3, in which two genes are regulating each other in a negative feedback loop. Gene-regulatory networks describe the interactions between genes and the proteins that regulate the expressions of these genes. The regulatory proteins, known as transcription factors, can activate or repress the expression of a gene by binding to its promoter site and by either aiding or interfering with the transcription process by RNAP[49]. Through these activation and repression processes, the cell regulates the levels of proteins in the cell, which are responsible for a vast majority of cellular functions[49].

In deterministic models of gene-regulatory networks, activation and repression processes

are often modeled by Hill functions[49]. These are derived using the QSSA where it is assumed that the binding reactions between the DNA promoter sites and transcription factors equilibrate faster than transcription and translation. Similarly, it has become common practice to use the deterministic quasi-steady state expressions (e.g. Michaelis-Menten/Hill functions) in the simulation of stochastic biochemical reaction networks as a convenient way to eliminate the fast dynamics and reduce simulation time[19]. However, the validity of such approximations in the stochastic setting remains under investigation[20,27]. In the previous section, we demonstrated that the reduced LNA model obtained using the stochastic tQSSA is equivalent to the slow variable dynamics in the reduced system (14)-(17), and thus provides a good approximation for the full system variables under all parameter conditions. In this example, we derive the reduced LNA for the system in Figure 3, and use it to investigate the validity of reduced-order LNA models obtained through the stochastic standard QSSA.

Consider the gene-regulatory network motif in Figure 3. We have protein M activating the production of protein G, which in turn represses the production of protein M. The chemical
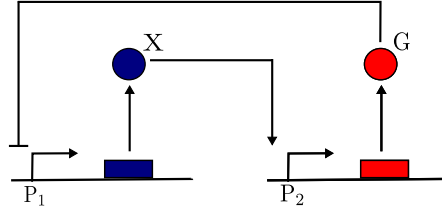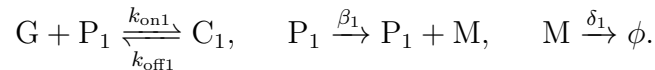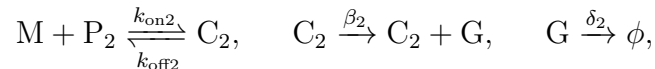

FIG. 3. Protein M activates the production of protein G, which in turn represses the production of protein M.

reactions are given by

$$G + P_1 \underset{k_{\text{off1}}}{\overset{k_{\text{on1}}}{\rightleftharpoons}} C_1, \qquad P_1 \xrightarrow{\beta_1} P_1 + M, \qquad M \xrightarrow{\delta_1} \phi.$$

Protein G represses the production of M by binding to the promoter P_1 leading to the inactive complex C_1. The production of protein M is given by the second reaction, in which we have lumped together transcription and translation. Protein M decays at rate δ_1, which accounts for degradation and dilution.

Similarly, the chemical reactions that describe the expression and decay of protein G can be written as

$$M + P_2 \underset{k_{\text{off2}}}{\overset{k_{\text{on2}}}{\rightleftharpoons}} C_2, \qquad C_2 \xrightarrow{\beta_2} C_2 + G, \qquad G \xrightarrow{\delta_2} \phi,$$

where M activates the production of protein G by first binding to promoter P_2 to produce the transcriptionally active complex C_2. We assume that the total concentration of promoters are conserved, giving $p_{T1} = p_1 + c_1$, $p_{T2} = p_2 + c_2$ where the lower-case letters represent macroscopic concentrations. Then, the macroscopic reaction rate equations can be written as

$$\begin{aligned}
\frac{dm}{dt} &= \beta_1(p_{T1} - c_1) - \delta_1 m - k_{on2}m(p_{T2} - c_2) + k_{off2}c_2, \\
\frac{dc_1}{dt} &= k_{on1}g(p_{T1} - c_1) - k_{off1}c_1, \\
\frac{dc_2}{dt} &= k_{on2}m(p_{T2} - c_2) - k_{off2}c_2, \\
\frac{dg}{dt} &= \beta_2 c_2 - \delta_2 g - k_{on1}g(p_{T1} - c_1) + k_{off1}c_1.
\end{aligned} \tag{34}$$

Since binding and unbinding reactions are much faster than production and decay, we can define the small parameter $\epsilon = \delta_1/k_{off1}$ where $\epsilon \ll 1$. Let $k_{d1} = k_{off1}/k_{on1}$ and $k_{d2} = k_{off2}/k_{on2}$ with $a = k_{off2}/k_{off1}$. Then, considering the species vector $v = [m, g, c_1, c_2]$, we can partition the reaction rate vector into slow and fast groups as $\tilde{a}(v, t) = [\hat{a}_s(y, t), (1/\epsilon)\hat{a}_f(v, t)]$ where $\hat{a}_s(v, t) = [\beta_1(p_{T1} - c_1), \delta_1 m, \beta_2 c_2, \delta_2 g]$ and $\hat{a}_f(y, t) = [\frac{\delta_1}{k_{d1}}g(p_{T1} - c_1), \delta_1 c_1, \frac{a\delta_1}{k_{d2}}m(p_{T2} - c_2), a\delta_1 c_2]$, and the corresponding stoichiometry matrix is given by

$$q = \begin{bmatrix}
1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 \\
0 & 0 & 1 & -1 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & -1
\end{bmatrix}.$$

The variables m and g are not fast nor slow because M and G participate in both slow and fast reactions. Thus, we consider the coordinate change $w = m + c_2$, $u = g + c_1$, which takes the system (34) into singular perturbation form:

$$\begin{aligned}
\frac{dw}{dt} &= \beta_1(p_{T1} - c_1) - \delta_1(w - c_2), \\
\frac{du}{dt} &= \beta_2 c_2 - \delta_2(u - c_1), \\
\epsilon \frac{dc_1}{dt} &= \frac{\delta_1}{k_{d1}}(u - c_1)(p_{T1} - c_1) - \delta_1 c_1, \\
\epsilon \frac{dc_2}{dt} &= \frac{a\delta_1}{k_{d2}}(w - c_2)(p_{T2} - c_2) - a\delta_1 c_2,
\end{aligned} \tag{35}$$

where the slow variables are given by w and u, and the fast variables are given by c_1 and c_2. Considering Claim 1, the change of coordinates $w = m + c_2$, $u = g + c_1$ corresponds to

$T_x = [1\ 0\ 0\ 1,\ 0\ 1\ 1\ 0]^T$, $T_z = [0\ 0\ 1\ 0,\ 0\ 0\ 0\ 1]^T$, where $x = [w, u]^T$ and $z = [c_1, c_2]^T$. Then, as described in Section II, the dynamics of the stochastic fluctuations of the LNA can be written using (35) as

$$
\begin{aligned}
\frac{d\psi_w}{dt} &= -\delta_1 \psi_w - \beta_1 \psi_{c_1} + \delta_1 \psi_{c_2} + \sqrt{\beta_1(p_{T1} - c_1)}\Gamma_1 - \sqrt{\delta_1(w - c_2)}\Gamma_2 \\
\frac{d\psi_u}{dt} &= -\delta_2 \psi_u + \delta_2 \psi_{c_1} + \beta_2 \psi_{c_2} + \sqrt{\beta_2 c_2}\Gamma_3 - \sqrt{\delta_2(u - c_1)}\Gamma_4, \\
\epsilon \frac{d\psi_{c_1}}{dt} &= \frac{\delta_1}{k_{d1}}(p_{T1} - c_1)\psi_u - (\frac{\delta_1}{k_{d1}}(p_{T1} + u - 2c_1) + \delta_1)\psi_{c_1} \\
&\quad + \sqrt{\frac{\epsilon \delta_1}{k_{d1}}(u - c_1)(p_{T1} - c_1)}\Gamma_5 - \sqrt{\epsilon \delta_1 c_1}\Gamma_6, \\
\epsilon \frac{d\psi_{c_2}}{dt} &= \frac{a\delta_1(p_{T2} - c_2)}{k_{d2}}\psi_w - (\frac{a\delta_1}{k_{d2}}(p_{T2} + w - 2c_2) + a\delta_1)\psi_{c_2} \\
&\quad + \sqrt{\frac{\epsilon a\delta_1}{k_{d2}}(w - c_2)(p_{T2} - c_2)}\Gamma_7 - \sqrt{\epsilon a\delta_1 c_2}\Gamma_8.
\end{aligned}
\tag{36}
$$

We note that the system functions are given by polynomials of the system variables, and thus Assumptions 1 and 2 are satisfied. Considering Assumption 3, we have that $\frac{\partial f_z}{\partial z}$ where $f_z(x, z, t, \epsilon) = [\frac{\delta_1}{k_{d1}}(u - c_1)(p_{T1} - c_1) - \delta_1 c_1, \frac{a\delta_1}{k_{d2}}(w - c_2)(p_{T2} - c_2) - a\delta_1 c_2]^T$ is Hurwitz for all parameter values, and setting $\epsilon = 0$ in the fast variable dynamics f_z yields the unique solution $[c_1, c_2]^T = [\frac{1}{2}(u + p_{T1} + k_{d1}) - \frac{1}{2}\sqrt{(u + p_{T1} + k_{d1})^2 - 4up_{T1}}, \frac{1}{2}(w + p_{T2} + k_{d2}) - \frac{1}{2}\sqrt{(w + p_{T2} + k_{d2})^2 - 4wp_{T2}}]^T$ that is feasible under conditions $0 \le c_1 \le p_{T1}$ and $0 \le c_2 \le p_{T2}$. Next, in order to derive the reduced system, we set $\epsilon = 0$ in the system (35)-(36), which yields

$$
\begin{aligned}
c_1 &= \frac{(u + p_{T1} + k_{d1})}{2} - \frac{\sqrt{(u + p_{T1} + k_{d1})^2 - 4up_{T1}}}{2}, \\
c_2 &= \frac{(w + p_{T2} + k_{d2})}{2} - \frac{\sqrt{(w + p_{T2} + k_{d2})^2 - 4wp_{T2}}}{2}, \\
\psi_{c1} &= \frac{(p_{T1} - c_1)\psi_u}{((p_{T1} + u - 2c_1) + k_{d1})}, \\
\psi_{c2} &= \frac{(p_{T2} - \bar{c}_2)\psi_w}{((p_{T2} + w - 2\bar{c}_2) + k_{d2})}.
\end{aligned}
\tag{37}
$$

Next, in order to determine the additional terms in the fast variable approximations for the stochastic variables ψ_{c1} and ψ_{c2} we write the Lyapunov equation (18) as

$$
\begin{bmatrix} -(\frac{\delta_1}{k_{d1}}(p_{T1} + u - 2c_1) + \delta_1) & 0 \\ 0 & -(\frac{a\delta_1}{k_{d2}}(p_{T2} + w - 2c_2) + a\delta_1) \end{bmatrix} g(w)g(w)^T
$$

$$+g(w)g(w)^T \begin{bmatrix} -(\frac{\delta_1}{k_{d1}}(p_{T1} + u - 2c_1) + \delta_1) & 0 \\ 0 & -(\frac{a\delta_1}{k_{d2}}(p_{T2} + v - 2c_2) + a\delta_1) \end{bmatrix}$$

$$= -\begin{bmatrix} \frac{\delta_1}{k_{d1}}(u - c_1)(p_{T1} - c_1) + \delta_1 c_1 & 0 \\ 0 & \frac{a\delta_1}{k_{d2}}(w - c_2)(p_{T2} - c_2) + a\delta_1 c_2 \end{bmatrix},$$

which yields

$$g(w)$$
$$= \begin{bmatrix} \sqrt{\frac{(\bar{u}-\bar{c}_1)(p_{T1}-\bar{c}_1)+k_{d1}\bar{c}_1}{2\sqrt{(\bar{u}+p_{T1}+k_{d1})^2-4\bar{u}p_{T1}}}} & 0 \\ 0 & \sqrt{\frac{(\bar{w}-\bar{c}_2)(p_{T2}-\bar{c}_2)+k_{d2}\bar{c}_2}{2\sqrt{(\bar{w}+p_{T2}+k_{d2})^2-4\bar{w}p_{T2}}}} \end{bmatrix}.$$

Then, substituting the terms (37) in the system (35)-(36), and using $g(w)$ and a vector of normal random variables $[N_1(0,1), N_2(0,1)]^T$ according to (17) we obtain the reduced system

$$\frac{d\bar{w}}{dt} = \beta_1(p_{T1} - \bar{c}_1) - \delta_1(\bar{w} - \bar{c}_2),$$

$$\frac{d\bar{u}}{dt} = \beta_2\bar{c}_2 - \delta_2(\bar{u} - \bar{c}_1),$$

$$\frac{d\bar{\psi}_w}{dt} = -\delta_1\bar{\psi}_w - \frac{\beta_1(p_{T1} - \bar{c}_1)\bar{\psi}_u}{p_{T1} + \bar{u} - 2\bar{c}_1 + k_{d1}} + \delta_1 \frac{(p_{T2} - \bar{c}_2)\bar{\psi}_w}{p_{T2} + \bar{w} - 2\bar{c}_2 + k_{d2}}$$
$$+ \sqrt{\beta_1(p_{T1} - \bar{c}_1)}\Gamma_1 - \sqrt{\delta_1(\bar{w} - \bar{c}_2)}\Gamma_2,$$

$$\frac{d\bar{\psi}_u}{dt} = -\delta_2\bar{\psi}_u + \frac{\delta_2(p_{T1} - \bar{c}_1)\bar{\psi}_u}{p_{T1} + \bar{u} - 2\bar{c}_1 + k_{d1}} + \frac{\beta_2(p_{T2} - \bar{c}_2)\bar{\psi}_w}{p_{T2} + \bar{w} - 2\bar{c}_2 + k_{d2}}$$
$$+ \sqrt{\beta_2\bar{c}_2}\Gamma_3 - \sqrt{\delta_2(\bar{u} - \bar{c}_1)}\Gamma_4, \tag{38}$$

$$\bar{c}_1 = \frac{(\bar{u} + p_{T1} + k_{d1})}{2} - \frac{\sqrt{(\bar{u} + p_{T1} + k_{d1})^2 - 4\bar{u}p_{T1}}}{2},$$

$$\bar{c}_2 = \frac{(\bar{w} + p_{T2} + k_{d2})}{2} - \frac{\sqrt{(\bar{w} + p_{T2} + k_{d2})^2 - 4\bar{w}p_{T2}}}{2},$$

$$\bar{\psi}_{c1} = \frac{(p_{T1} - \bar{c}_1)\psi_u}{((p_{T1} + \bar{u} - 2\bar{c}_1) + k_{d1})} + \sqrt{\frac{(\bar{u} - \bar{c}_1)(p_{T1} - \bar{c}_1) + k_{d1}\bar{c}_1}{2\sqrt{(\bar{u} + p_{T1} + k_{d1})^2 - 4\bar{u}p_{T1}}}}N_1(0,1),$$

$$\bar{\psi}_{c2} = \frac{(p_{T2} - \bar{c}_2)\bar{\psi}_w}{((p_{T2} + \bar{w} - 2\bar{c}_2) + k_{d2})} + \sqrt{\frac{(\bar{w} - \bar{c}_2)(p_{T2} - \bar{c}_2) + k_{d2}\bar{c}_2}{2\sqrt{(\bar{w} + p_{T2} + k_{d2})^2 - 4\bar{w}p_{T2}}}}N_2(0,1).$$

Figure 4 illustrates simulation results for the second moments of $\bar{\psi}_w$, $\bar{\psi}_u$, $\bar{\psi}_{c1}$ and $\bar{\psi}_{c2}$. As ϵ tends to zero the moments of the full system become closer to the moments of the reduced system. This confirms our results in Theorem 1.
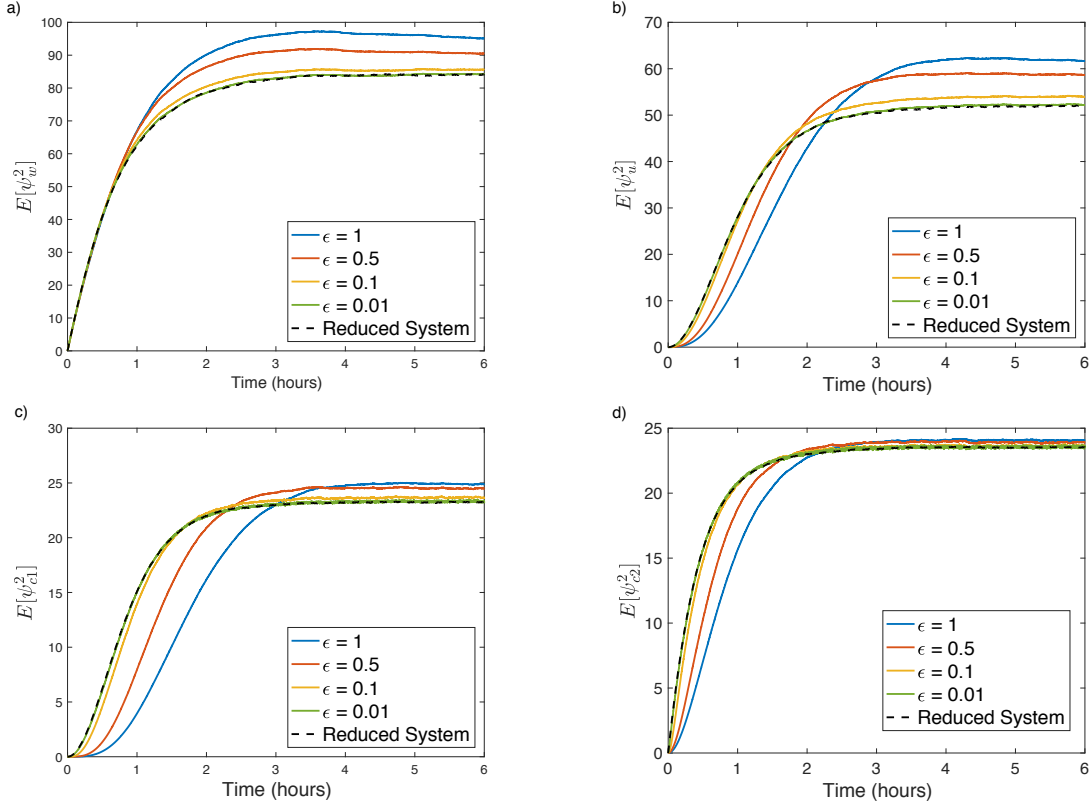
20

FIG. 4. Sample moments of the full and reduced systems obtained by numerically simulating the systems (35)–(36) and (38). The simulations were performed using the Euler-Maruyama method for the stochastic differential equations and the moments are computed using the average of 500,000 simulation runs. (a) Second moments of ψ_w and $\bar{\psi}_w$. (b) Second moments of ψ_u and $\bar{\psi}_u$. (c) Second moments of ψ_{c1} and $\bar{\psi}_{c1}$. (d) Second moments of ψ_{c2} and $\bar{\psi}_{c2}$. The parameter values are $\beta_1 = 1 \text{ hr}^{-1}$, $\beta_2 = 2 \text{ hr}^{-1}$, $\delta_1 = 1 \text{ hr}^{-1}$, $\delta_2 = 3 \text{ hr}^{-1}$, $k_{d1} = 50$ nM, $k_{d2} = 100$ nM, $p_{T1} = 100$ nM, $p_{T1} = 100$ nM, $w(0) = 0$, $c_1(0) = 0$, $c_2(0) = 0$, $u(0) = 0$, $\psi_w(0) = 0$, $\psi_u(0) = 0$, $\psi_{c1}(0) = 0$, $\psi_{c2}(0) = 0$.

Next we use the reduced system (38) to analyze the validity of stochastic models conveniently obtained by using the quasi-steady state approximation.

1. Validity of LNA models based on the Michaelis-Menten function

In this section, we use the reduced system (38) to investigate the validity of the stochastic models obtained using the standard QSSA. Using the deterministic QSSA to reduce the deterministic system is a convenient way to obtain a reduced LNA. {However, the valid-

ity of this reduction remains elusive. Here, we first derive the LNA model based on the Michaelis-Menten function obtained using the stochastic standard QSSA, similar to several works in the literature[20,27,50]. Next, we present the corresponding moment dynamics of the system. Then, we use the reduced system (38) proposed in our work and compute the corresponding moment dynamics. Comparing these two sets of moment dynamics we identify conditions under which the LNA model based on the Michaelis-Menten function provides a good approximation.

As performed in standard references[49], we derive the Michaelis-Menten function based model by setting the time derivatives $\frac{dc_1}{dt}$ and $\frac{dc_2}{dt}$ in (34) to zero obtaining

$$c_1 = \frac{gp_{T1}}{g + k_{d1}}, \qquad c_2 = \frac{mp_{T2}}{m + k_{d2}}.$$

Then, substituting these in system (34) we obtain the reduced deterministic dynamics:

$$\frac{d\tilde{m}}{dt} = \beta_1 \frac{p_{T1}k_{d1}}{\tilde{g} + k_{d1}} - \delta_1 \tilde{m}, \tag{39}$$

$$\frac{d\tilde{g}}{dt} = \beta_2 \frac{\tilde{m}p_{T2}}{\tilde{m} + k_{d2}} - \delta_2 \tilde{g}. \tag{40}$$

Next, we derive the dynamics for the corresponding stochastic fluctuations as explained in Section II, which yields

$$\frac{d\tilde{\psi}_m}{dt} = -\delta_1 \tilde{\psi}_m - \beta_1 \frac{p_{T1}k_{d1}}{(\tilde{g} + k_{d1})^2} \tilde{\psi}_g + \sqrt{\beta_1 \frac{p_{T1}k_{d1}}{\tilde{g} + k_{d1}}} \Gamma_1$$

$$+ \sqrt{\delta_1 \tilde{m}} \Gamma_2, \tag{41}$$

$$\frac{d\tilde{\psi}_g}{dt} = -\delta_2 \tilde{\psi}_g + \beta_2 \frac{p_{T2}k_{d2}}{(\tilde{m} + k_{d2})^2} \tilde{\psi}_m + \sqrt{\beta_2 \frac{\tilde{m}p_{T2}}{\tilde{m} + k_{d2}}} \Gamma_3$$

$$+ \sqrt{\delta_2 \tilde{g}} \Gamma_4. \tag{42}$$

Thus, the Michaelis-Menten function based LNA model is given by system (39)–(42). The dynamics for the second moments of $\tilde{\psi}_m$ and $\tilde{\psi}_g$ are given by

$$\frac{d\mathbb{E}\left[\tilde{\psi}_m^2\right]}{dt} = -2\delta_1 \mathbb{E}\left[\tilde{\psi}_m^2\right] - 2R_1(\tilde{g})\beta_1 \mathbb{E}\left[\tilde{\psi}_m \tilde{\psi}_g\right]$$

$$+ \beta_1 \frac{p_{T1}k_{d1}}{\tilde{g} + k_{d1}} + \delta_1 \tilde{m},$$

$$\frac{d\mathbb{E}\left[\psi_g^2\right]}{dt} = -2\delta_2 \mathbb{E}\left[\tilde{\psi}_g^2\right] + 2R_2(\tilde{m})\beta_2 \mathbb{E}\left[\tilde{\psi}_m \tilde{\psi}_g\right] \tag{43}$$

$$+ \beta_2 \frac{\tilde{m} p_{T1}}{\tilde{m} + k_{d2}} + \delta_2 \tilde{g},$$

$$\frac{d\mathbb{E}\left[\tilde{\psi}_m \tilde{\psi}_g\right]}{dt} = (-\delta_1 - \delta_2)\mathbb{E}\left[\tilde{\psi}_m \tilde{\psi}_g\right] - R_1(\tilde{g})\beta_1 \mathbb{E}\left[\tilde{\psi}_g^2\right]$$

$$+ R_2(\tilde{m})\beta_2 \mathbb{E}\left[\tilde{\psi}_m^2\right],$$

where we have defined $R_1(\tilde{g}) = \frac{p_{T1}/k_{d1}}{(\tilde{g}/k_{d1}+1)^2}$ and $R_2(\tilde{m}) = \frac{p_{T2}/k_{d2}}{(\tilde{m}/k_{d2}+1)^2}$.

Next, we use the reduced-order model developed in our work to precisely identify the conditions under which the above Michaelis-Menten function based LNA model provides a good approximation for the statistical properties of the original system. To this end, we use the reduced system (38) to derive the moment dynamics (see Appendix D for details) as

$$\frac{d\bar{m}}{dt} = \frac{1}{1 + \boxed{R_2(\bar{m})}}\left(\beta_1 \frac{p_{T1} k_{d1}}{\bar{g} + k_{d1}} - \delta_1 \bar{m}\right), \tag{44}$$

$$\frac{d\bar{g}}{dt} = \frac{1}{1 + \boxed{R_1(\bar{g})}}\left(\beta_2 \bar{m} \frac{p_{T2}}{\bar{m} + k_{d2}} - \delta_2 g\right), \tag{45}$$

$$\frac{d\mathbb{E}\left[\bar{\psi}_m^2\right]}{dt} = \frac{1}{1 + \boxed{R_2(\bar{m})}}\left[\left(-2\delta_1 + \boxed{F_2(\bar{m})}\right)\mathbb{E}\left[\bar{\psi}_m^2\right] + 2R_1(\bar{g})\beta_1 \mathbb{E}\left[\bar{\psi}_m \bar{\psi}_g\right]\right.$$

$$\left. + \left(\boxed{-G_2(\bar{m})} + 1\right)\beta_1 \frac{p_{T1} k_{d1}}{\bar{g} + k_{d1}} + \left(\boxed{G_2(\bar{m})} + 1\right)\delta_1 \bar{m}\right], \tag{46}$$

$$\frac{d\mathbb{E}\left[\bar{\psi}_g^2\right]}{dt} = \frac{1}{1 + \boxed{R_1(\bar{g})}}\left[\left(-2\delta_2 + \boxed{F_1(\bar{g})}\right)\mathbb{E}\left[\bar{\psi}_g^2\right] + 2R_2(\bar{m})\beta_1 \mathbb{E}\left[\bar{\psi}_m \bar{\psi}_g\right]\right.$$

$$\left. + \left(\boxed{-G_1(\bar{g})} + 1\right)\beta_2 \frac{p_{T1} \bar{m}}{\bar{m} + k_{d2}} + \left(\boxed{G_1(\bar{g})} + 1\right)\delta_2 g\right], \tag{47}$$

$$\frac{d\mathbb{E}\left[\bar{\psi}_m \bar{\psi}_g\right]}{dt} = \left(-\frac{\delta_1}{1 + \boxed{R_2(\bar{m})}} - \frac{\delta_2}{1 + \boxed{R_1(\bar{g})}} + \boxed{H_2(\bar{m})} + \boxed{H_1(\bar{g})}\right)\mathbb{E}\left[\bar{\psi}_m \bar{\psi}_g\right]$$

$$- \frac{R_1(\bar{g})\beta_1 \mathbb{E}\left[\bar{\psi}_g^2\right]}{1 + \boxed{R_2(\bar{m})}} + \frac{R_2(\bar{m})\beta_2 \mathbb{E}\left[\bar{\psi}_m^2\right]}{1 + \boxed{R_1(\bar{g})}} - \boxed{I(\bar{m}, \bar{g})}, \tag{48}$$

where

$$F_i(y) = 4\frac{R_i(y)}{(y + k_{di})}\frac{dy}{dt},$$

$$G_i(y) = \frac{2y R_i(y)(1 + 2R_i(y))}{(y + k_{di})(1 + R_i(y))^2},$$

$$H_i(y) = \frac{2R_i(y)}{(y + k_{di})(1 + R_i(y))}\frac{dy}{dt},$$

$$I(\bar{m}, \bar{g}) = \frac{1}{1 + R_2(\bar{m})}\frac{1}{1 + R_1(\bar{g})}(\beta_1 \bar{g} R_1(\bar{g})^2 + \beta_2 \bar{m} R_2(\bar{m})^2).$$

The boxed terms in (44)–(48) highlight the extra terms compared to the Michaelis-Menten function based moment dynamics (39)–(40) and (43). From this, it follows that the Michaelis-Menten function based model of (39)–(42) becomes a good approximation of the full LNA model only when the boxed terms are sufficiently small. In order to analyze the conditions under which the boxed terms become small, we first note from the equations (44)–(45) that

$$\bar{m}(t) \le \max\{\bar{m}(0), \beta_1 p_{T1}/\delta_1\} := b_{\bar{m}}$$

and

$$\bar{g}(t) \le \max\{\bar{g}(0), \beta_2 p_{T2}/\delta_2\} := b_{\bar{g}}$$

for all t. Thus, we have that

$$\left| \frac{d\bar{m}}{dt} \right| \le \beta_1 p_{T1} + \delta_1 \max\{\bar{m}(0), \beta_1 p_{T1}/\delta_1\} := b_{d\bar{m}}$$

and

$$\left| \frac{d\bar{g}}{dt} \right| \le \beta_2 p_{T2} + \delta_2 \max\{\bar{g}(0), \beta_2 p_{T2}/\delta_2\} := b_{d\bar{g}}.$$

Then, considering the terms $F_1(\bar{g})$ and $F_2(\bar{m})$ we have that

$$\left| F_1(\bar{g}) \right| \le 2 \left| \frac{dR_1(\bar{g})}{d\bar{g}} \right| (\beta_2 p_{T2} + \delta_2 \max\{\bar{g}(0), \beta_2 p_{T2}/\delta_2\}),$$

$$\left| F_2(\bar{m}) \right| \le 2 \left| \frac{dR_2(\bar{m})}{d\bar{m}} \right| (\beta_1 p_{T1} + \delta_1 \max\{\bar{m}(0), \beta_1 p_{T1}/\delta_1\}).$$

From this we have that for given bounds $b_{d\bar{m}}$ and $b_{d\bar{g}}$, $|F_i(y)|$ terms become small when the terms $|\frac{dR_i(y)}{dy}|$ are sufficiently small. Similarly, from the expressions of $R_i(y)$, $G_i(y)$, $H_i(y)$ and $I(\bar{m}, \bar{g})$, we have that for given bounds $b_{\bar{m}}$, $b_{\bar{g}}$, $b_{d\bar{m}}$ and $b_{d\bar{g}}$, the terms $|H_i(y)|$ become small when the terms $|\frac{dR_i(y)}{dy}|$ are sufficiently small and the terms $G_i(y)$, $I(\bar{m}, \bar{g})$ become small when $R_i(y)$ terms are sufficiently small. Thus, in this case, we can conclude that when $R_i(y)$ and $|\frac{dR_i(y)}{dy}|$ are sufficiently small, the Michaelis-Menten function based moment dynamics (39)–(40) and (43) become close to the moment dynamics (44)–(48), which is in turn a good approximation of the moments of the original system when ϵ is small. Thus, it follows that the Michaelis-Menten function based models becomes a good approximation of the original system when ϵ becomes small and when $R_i(y)$ and $|\frac{dR_i(y)}{dy}|$ are also small. We illustrate this via simulations in Figure 4, which shows that the moments of the Michaelis-Menten function based model become close to those of the singular perturbation-based model when the dissociation constants k_{d1} and k_{d2} are large (making $R_i(y)$ and $|\frac{dR_i(y)}{dy}|$ small).
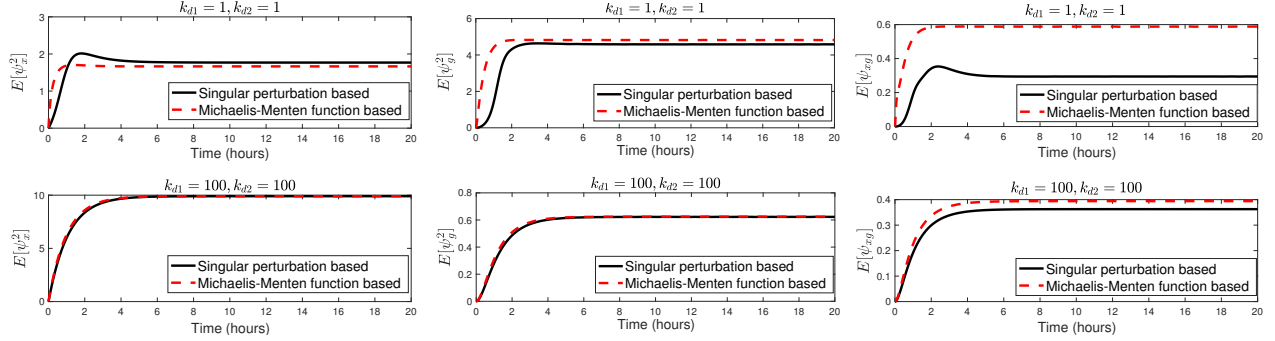
FIG. 5. Moments of the Michaelis-Menten function based LNA model and the singular perturbation based LNA model. The parameter values are $\beta_1 = 1 \text{ hr}^{-1}$, $\beta_2 = 2 \text{ hr}^{-1}, \delta_1 = 1 \text{ hr}^{-1}, \delta_2 = 3 \text{ hr}^{-1}$, $p_{T1} = 10 \text{ nM}$, $p_{T1} = 10 \text{ nM}$.

The terms $R_1(\bar{g})$ and $R_2(\bar{m})$ in the deterministic dynamics (44)–(45) have been previously studied in the context of modularity in gene-regulatory networks. In particular, Del Vecchio et al.[49,51,52] showed that the terms of the form $R_1(\bar{g})$ and $R_2(\bar{m})$ physically arise due to the "loading" that promoter binding sites apply to their transcription factor regulators. These effects, termed retroactivity, cause a change in the dynamics of transcription factors upon binding to the DNA promoter sites and are not captured by Michaelis-Menten function based models alone. Thus, the results obtained in this section indicate that, similar to deterministic systems, retroactivity effects also impact the stochastic dynamics of the system.

As previously described in Section III, the use of the deterministic QSSA in stochastic models has been previously studied by several researchers in the context of CME and LNA. In particular, the study by J.K. Kim et al. considers two dimensional systems (with one slow variable and one fast variable) modeled by the CME, and points to two main sources of inaccuracy for the stochastic models obtained through standard QSSA[20]. The first source of inaccuracy arises from treating species that contain both slow and fast variables as purely slow variables (such as the free proteins M and G in this example). This is also a source of inaccuracy in the deterministic QSSA as noted by several studies[48,51]. The second source of inaccuracy is disregarding the noise of the fast variable when taking the QSSA. In line with this, J.K. Kim et al. analyzed the steady state distribution and the variance of the fast variable and determine that the stochastic QSSA becomes accurate when the sensitivity of the quasi-steady state expression to changes of the slow variable is small. This sensitivity term is equivalent to the retroactivity terms $R_1(\bar{g})$ and $R_2(\bar{m})$ in our model. Thus, our results recapitulate the findings of Kim's steady state analysis on systems with two species,

but, being based on general and rigorous derivation of the moment dynamics, our results are applicable to systems with more than two species and can also be used to study transient dynamics, as we have shown in this example.

Additionally, a similar study has been performed by Thomas et al., where a reduced-order approximation for LNA is proposed.[27]. They analyze several enzyme and gene network motifs and determine that the discrepancy between the quasi-steady state LNA models and the reduced-order LNA models proposed in their work arises mainly due to disregarding the noise of the fast variables. However, the reduced-order LNA developed in their work also regards some species concentrations that are the combination of fast and slow variables as purely slow variables. Therefore, their transient dynamics may not be accurately approximated by their reduction. In contrast to this, here, we identify conditions for which both the transient dynamics and the steady state are accurate.

The analysis in this section, being based on mathematical derivations as opposed to simulations, therefore provides a general criterion to determine when the quasi-steady state LNA of gene-regulatory networks is close to the original system. The fast variable approximations can also be used to analyze how the noise of the transcription factor-DNA complexes change with retroactivity effects, which can give insights into the interplay between modularity and noise in gene-regulatory networks.

B.   Example II: Signal transduction network motif

As another example, we consider the dynamics of species involved in signal transmission via phosphorylation cycles. Phosphorylation cycles are network motifs that form signal transduction cascades in cells. They play a central role in cell physiology for transmitting signals from outside the cell to initiate cellular responses by activating target gene expression. Furthermore, engineered phosphorylation systems have also been used in synthetic biology applications to transmit signals while buffering from loading effects that were discussed in the previous example[41,51,53,54]. Here, we consider the system shown in Figure 6. Protein M is phosphorylated by kinase Z and produces protein M*, which is dephosphorylated by phosphatase Y. The phosphorylated protein M* acts as a transcription factor that activates protein G.

Phosphorylation and dephosphorylation processes can be modeled by the one-step
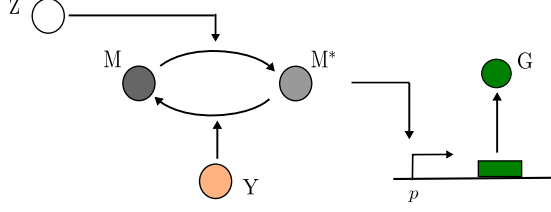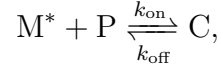
FIG. 6. Protein M is phosphorylated by kinase Z and dephosphorylated by phosphatase Y. Phosphorylated protein M* binds to the downstream promoter P.

reactions[49]

$$M + Z \xrightarrow{k_1} M^* + Z, \qquad M^* + Y \xrightarrow{k_2} M + Y,$$

where k_1 and k_2 are the rate constants. The binding between protein M* and promoter P produces a complex C, giving

$$M^* + P \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} C,$$

where k_{on} and k_{off} are the binding and unbinding rate constants. Then, the production and decay of protein G is given by

$$C \xrightarrow{\beta} C + G, \quad G \xrightarrow{\delta} \phi,$$

where β is the production rate constant, and δ is the decay rate constant. The total concentration of protein M and promoter P are conserved, giving $M_T = m + m^* + c$ and $p_T = p + c$, where we use lower-case letters to denote the corresponding macroscopic concentrations. Then, the macroscopic reaction rate equations for this system can be written as

$$\frac{dm^*}{dt} = k_1 Z(t)(M_T - m^* - c) - k_2 Y m^* - k_{\mathrm{on}} m^* (p_T - c) + k_{\mathrm{off}} c,$$
$$\frac{dc}{dt} = k_{\mathrm{on}} m^* (p_T - c) - k_{\mathrm{off}} c,$$
$$\frac{dg}{dt} = \beta c - \delta g,$$

where we consider the concentration of the kinase given by $Z(t)$ to be a deterministic input to the system.

As binding and unbinding reactions are much faster than phosphorylation and dephosphorylation reactions, we have that $k_{\mathrm{off}} \gg k_2 Y$. Thus, we can define the small parameter $\epsilon = k_2 Y / k_{\mathrm{off}}$. Let $k_d = k_{\mathrm{off}} / k_{\mathrm{on}}$ be the dissociation constant for the binding reaction between M* and P. Then, as described in Section II, we can consider the species vector

$v = [m^*, \ g, \ c]^T$ and partition the reaction rate vector into slow and fast groups in the form $\tilde{a}(y,t) = [\hat{a}_s(v,t), (1/\epsilon)\hat{a}_f(y,t)]$ where $\hat{a}_s(v,t) = [k_1 Z(t)(M_T - m^* - c), \ k_2 Y m^*, \ \beta c, \ \delta g]$ and $\hat{a}_f(v,t) = [\frac{k_2 Y}{k_d} m^*(p_T - c), \ k_2 Y c]$, with the corresponding stoichiometry matrix given by

$$
q = \begin{bmatrix} 1 & -1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}.
$$

This yields the system dynamics

$$
\begin{aligned}
\frac{dm^*}{dt} &= k_1 Z(t)(M_T - m^* - c) - k_2 Y m^* - \frac{k_2 Y}{\epsilon k_d} m^*(p_{tot} - c) + \frac{k_2 Y}{\epsilon} c, \\
\frac{dg}{dt} &= \beta c - \delta g, \\
\frac{dc}{dt} &= \frac{k_2 Y}{\epsilon k_d} m^*(p_{tot} - c) - \frac{k_2 Y}{\epsilon} c.
\end{aligned}
\tag{49}
$$

Although, the singular perturbation parameter appears in the system of equations (49), we note that the slow and fast dynamics are not well separated and the system is not in the standard singular perturbation form given in (8)–(9). Thus, we consider the change of coordinates $w = m^* + c$, which yields

$$
\begin{aligned}
\frac{dw}{dt} &= k_1 Z(t)(M_T - w) - k_2 Y(w - c), \\
\frac{dg}{dt} &= \beta c - \delta g, \\
\epsilon \frac{dc}{dt} &= \frac{k_2 Y}{k_d}(w - c)(p_T - c) - k_2 Y c.
\end{aligned}
\tag{50}
$$

where we have that the slow variables are w and g, and the fast variable is given by c. Referring to Claim 1, we have that the coordinate change $w = m^* + c$ corresponds to $T_x = [1 \ 0 \ 1, \ 0 \ 1 \ 0]^T$, $T_z = [0 \ 0 \ 1]$, where we have that $x = [w, g]^T$ and $z = c$. Therefore, from (50) we can write the following equations for the dynamics of the stochastic fluctuations as explained in Section II:

$$
\begin{aligned}
\frac{d\psi_w}{dt} &= (-k_1 Z(t) - k_2 Y)\psi_w + k_2 Y \psi_c \\
&\quad + \sqrt{k_1 Z(t)(M_T - w + c)}\Gamma_1 - \sqrt{k_2 Y(w - c)}\Gamma_2, \\
\frac{d\psi_g}{dt} &= \beta \psi_c - \delta \psi_g + \sqrt{\beta c}\Gamma_3 - \sqrt{\delta g}\Gamma_4, \\
\epsilon \frac{d\psi_c}{dt} &= \frac{k_2 Y(p_T - c)}{k_d}\psi_w + \left(-\frac{k_2 Y p_T}{k_d} - \frac{k_2 Y v}{k_d} + \frac{k_2 Y}{k_d} 2c - k_2 Y\right)\psi_c \\
&\quad + \sqrt{\epsilon \frac{k_2 Y}{k_d}(w - c)(p_T - c)}\Gamma_5 - \sqrt{\epsilon k_2 Y c}\Gamma_6.
\end{aligned}
\tag{51}
$$

28

Next, we derive the reduced-order dynamics of system (50)–(51) using Theorem 1. From (50), it follows that the system functions are polynomials of the state variables. Therefore, Assumptions 1 and 2 are satisfied. Solving $f_z = \frac{k_2 Y}{k_d}(w-z)(p_T-z) - k_2 Y z = 0$, yields the unique solution $\gamma_1(w) = \frac{1}{2}(w + p_T + k_d) - \frac{1}{2}\sqrt{(w + p_T + k_d)^2 - 4vp_T}$, feasible under the physical constraints $0 \le c \le p_T$. We note that $\frac{\partial f_z}{\partial z}$ is negative for all parameter values and thus Assumption 3 is satisfied. Next, to determine the fast variable approximation for ψ_c in the form of equation (17), we write

$$\left(-\frac{k_2 Y p_T}{k_d} - \frac{k_2 Y v}{k_d} + \frac{k_2 Y}{k_d} 2c - k_2 Y\right) g(w)g(w)^T$$
$$+ g(w)g(w)^T \left(-\frac{k_2 Y p_T}{k_d} - \frac{k_2 Y v}{k_d} + \frac{k_2 Y}{k_d} 2c - k_2 Y\right)$$
$$= -\frac{k_2 Y}{k_d}(w-c)(p_T-c) - k_2 Y c,$$

which yields

$$g(w) = \sqrt{\frac{\frac{k_2 Y}{k_d}(w-c)(p_{tot}-c) + k_2 Y c}{2\frac{k_2 Y}{k_d}\sqrt{(w + p_{tot} + k_d)^2 - 4wp_T}}}.$$

Then, the reduced system is given by

$$\frac{d\bar{w}}{dt} = k_1 Z(t)(M_T - \bar{w}) - k_2 Y(\bar{w} - c),$$
$$\frac{d\bar{g}}{dt} = \beta\bar{c} - \delta\bar{g},$$
$$\frac{d\bar{\psi}_w}{dt} = (-k_1 Z(t) - k_2 Y)\bar{\psi}_w + \frac{k_2 Y(p_T - \bar{c})}{(p_T + \bar{w} - 2\bar{c} + k_d)}\bar{\psi}_w$$
$$+ \sqrt{k_1 Z(t)(M_T - \bar{w} + \bar{c})}\Gamma_1 - \sqrt{k_2 Y(\bar{w} - \bar{c})}\Gamma_2,$$
$$\frac{d\bar{\psi}_g}{dt} = \beta\frac{(p_T - \bar{c})}{(p_T + \bar{w} - 2\bar{c} + k_d)}\bar{\psi}_w - \delta\bar{\psi}_g + \sqrt{\beta\bar{c}}\Gamma_3 - \sqrt{\delta\bar{g}}\Gamma_4,$$
$$\bar{c} = \frac{1}{2}(\bar{w} + p_T + k_d) - \frac{1}{2}\sqrt{(\bar{w} + p_T + k_d)^2 - 4p_T\bar{w}},$$
$$\bar{\psi}_c = \frac{(p_T - \bar{c})}{(p_T + \bar{w} - 2\bar{c} + k_d)}\bar{\psi}_w + \sqrt{\frac{\frac{k_2 Y}{k_d}(\bar{w} - \bar{c})(p_T - \bar{c}) + k_2 Y\bar{c}}{2\frac{k_2 Y}{k_d}\sqrt{(\bar{w} + p_T + k_d)^2 - 4\bar{w}p_T}}}N(0,1). \qquad (52)$$

Figure 7 illustrates simulation results for the second moments of ψ_w, ψ_g and ψ_c.

As ϵ tends to zero the moments of the full system tend to the moments of the reduced system in accordance to Theorem 1.

The reduced system derived in this section can be used to mathematically analyze the noise properties of signals transmitted through phosphorylation cycles. A common measure
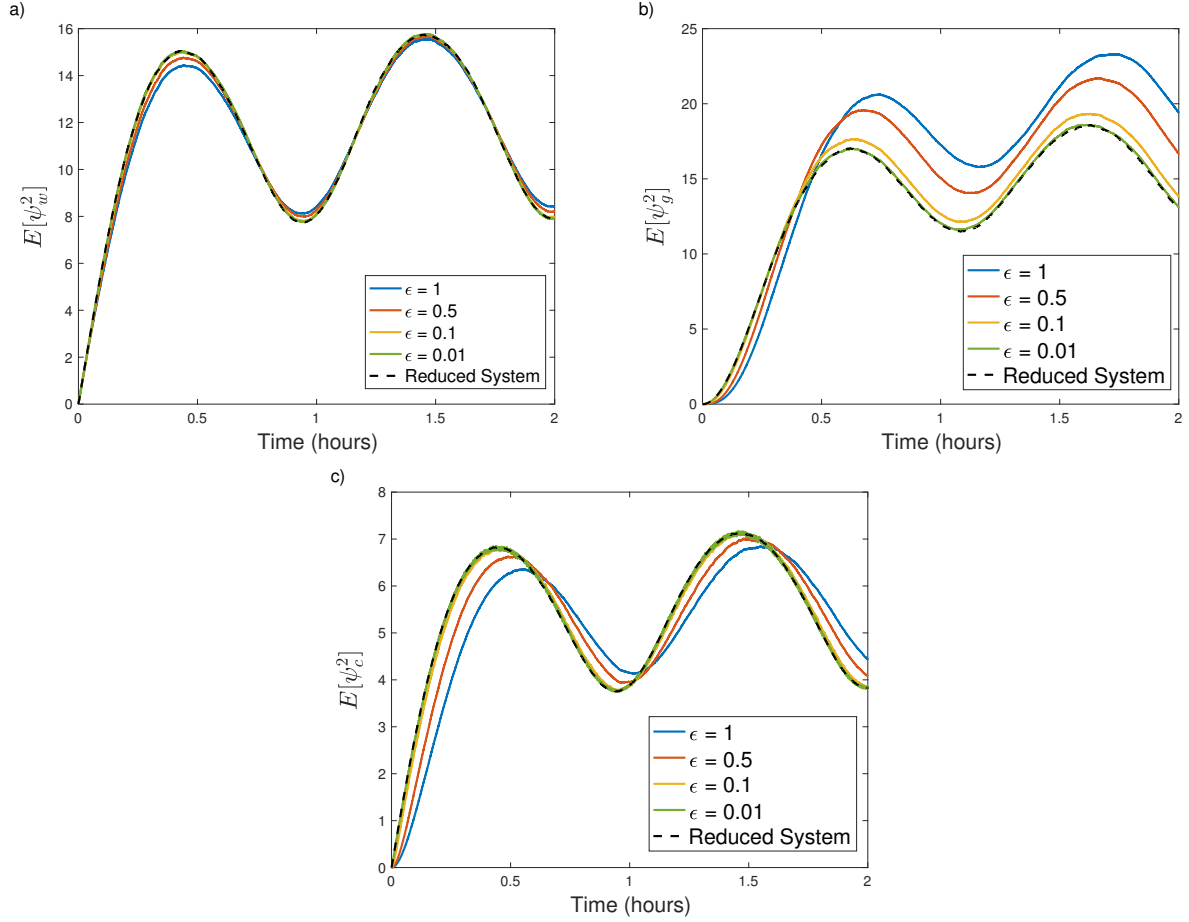
FIG. 7. Sample moments of the full and reduced systems obtained by numerically simulating the systems (50)–(51) and (52). The simulations were performed using the Euler-Maruyama method for the stochastic differential equations and the moments are computed using the average of 300,000 simulation runs. The parameter values are $Z(t) = 10 + 8.5\sin(6t)$ nM, $k_1 = 0.06$ nMhr^{-1}, $k_2 = 0.6$ nMhr^{-1}, $k_d = 50$ nM, $M_T = 100$ nM, $Y = 10$ nM, $p_T = 50$ nM, $\delta = 6$ hr^{-1}, $\beta = 10$ hr^{-1}, $w(0) = 0$, $c(0) = 0$, $g(0) = 0$, $\psi_w(0) = 0$, $\psi_g(0) = 0$, $\psi_c(0) = 0$.

of noise is the coefficient of variation, which is defined as the ratio of standard deviation to the mean of a random variable. The moment dynamics of the reduced system can therefore be used in the computation of such noise measures. Finally, note that since the species that carry the signal M* is a mixed (fast and slow) species, both fast and slow variable approximations are necessary to analyze noise propagation in these signal transduction systems.

VI. CONCLUSION

In this work, we addressed the problem of model order reduction for biochemical reaction networks with time-scale separation, where the system dynamics are modeled with the LNA. After transforming the system into standard singular perturbation form, we developed a reduced-order model that approximates the slow and fast dynamics of the full system when the time-scale separation is large. In particular, we showed that the error between the moment dynamics of the full system and the reduced systems is $O(\epsilon)$, where ϵ is the singular perturbation parameter that captures the time-scale separation. This error quantification only requires stability of the fast variables boundary layer dynamics, which is a simple condition to check and the same condition required in deterministic singular perturbation. Different from existing work, we have also presented an approximation of the fast variables. Furthermore, we show that the slow variable dynamics in our reduced-order model are equivalent to the reduced model obtained using the stochastic tQSSA approximation. Thus, our results also provide a rigorous justification for the validity of the LNA models obtained using the stochastic tQSSA. When the molecular counts are sufficiently large, this further provides a justification for the validity of the stochastic tQSSA in the CME. Since our reduction, differently from the stochastic tQSSA, also provides an approximation of the fast variable stochastic properties, we have termed it the stochastic tQSSA$^+$.

We have detailed the application of our work to two examples of biochemical reaction networks: a gene-regulatory network motif and a signal transduction module. For these examples, we derived the reduced-order LNA and verified the analytical results through numerical simulations. Through these examples we highlighted the necessity of both slow and fast variable approximations for practical application of the model reduction framework. Finally, for the gene-regulatory network motif, we identified conditions under which commonly used stochastic quasi-steady state models provide a good approximation. Our results can be used to substantially reduce the dimensionality of stochastic models of biochemical reaction networks, thus aiding analytical quantification of noise and simulation time. Furthermore, our results would also be useful in parameter estimations where accurate reduced-order models are required to obtain accurate and precise parameter estimations[8].

ACKNOWLEDGMENTS

Appendix A: Proof of Claim 1

Applying the coordinate transformation $x = T_x v$, $z = T_z v$ to equation (2), with $\tilde{a}(v,t) = [\hat{a}_s(v,t), (1/\epsilon)\hat{a}_f(v,t)]^T$ and $q = [q_1, \ldots, q_{m_s}, q_{m_s+1}, \ldots, q_{m_s+m_f}]$, with $v = T^{-1}[x^T, z^T]^T$ we have

$$\dot{x} = T_x f(T^{-1}[x^T, z^T]^T, t) = T_x \sum_{i=1}^{m_s} q_i \hat{a}_{si}(T^{-1}[x^T, z^T]^T, t)$$

$$+ T_x \sum_{i=q_{m_s}+1}^{m_s+m_f} q_i (1/\epsilon) \hat{a}_{fi}(T^{-1}[x^T, z^T]^T, t)$$

$$= f_x(x, z, t), \tag{A1}$$

$$\dot{z} = T_z f(T^{-1}[x^T, z^T]^T, t) = T_z \sum_{i=1}^{m_s} q_i \hat{a}_{si}(T^{-1}[x^T, z^T]^T, t)$$

$$+ T_z \sum_{i=q_{m_s}+1}^{m_s+m_f} q_i (1/\epsilon) \hat{a}_{fi}(T^{-1}[x^T, z^T]^T, t)$$

$$= \frac{1}{\epsilon} f_z(x, z, t, \epsilon). \tag{A2}$$

Thus, from equation (A1), if follows that $T_x q_i = 0$ for $i = m_s + 1, \ldots, m_s + m_f$. Thus, the fast reactions do not appear in the slow dynamics, however, the slow reactions can appear in the fast dynamics.

Applying the coordinate transformation $\psi_x = T_x \xi$, $\psi_z = T_z \xi$, to equation (3), we have that

$$\dot{\psi}_x = T_x[A(v,t)\xi] + T_x \sigma(v,t)\Gamma,$$

$$\dot{\psi}_z = T_z[A(v,t)\xi] + T_z \sigma(v,t)\Gamma.$$

Since $A(v,t) = \frac{\partial f(v,t)}{\partial v}$ and $v = T^{-1}[x^T, z^T]^T$, using the chain rule we can write

$$\dot{\psi}_x = T_x \left[\frac{\partial f(T^{-1}[x^T, z^T]^T, t)}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f(T^{-1}[x^T, z^T]^T, t)}{\partial z} \frac{\partial z}{\partial v} \right] \xi$$

$$+ T_x \left[q_1 \sqrt{\tilde{a}_1(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_m \sqrt{\tilde{a}_m(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma,$$

$$\dot{\psi}_z = T_z \left[\frac{\partial f(T^{-1}[x^T, z^T]^T, t)}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f(T^{-1}[x^T, z^T]^T, t)}{\partial z} \frac{\partial z}{\partial v} \right] \xi$$

$$+ T_z \left[q_1 \sqrt{\tilde{a}_1(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_m \sqrt{\tilde{a}_m(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma.$$

Using the linearity of the differentiation operator and the transformation $x = T_x v$, $z = T_z v$, we obtain

$$\dot{\psi}_x = \left[\frac{\partial T_x f(T^{-1}[x^T, z^T]^T, t)}{\partial x} T_x + \frac{\partial T_x f(T^{-1}[x^T, z^T]^T, t)}{\partial z} T_z \right] \xi$$
$$+ T_x \left[q_1 \sqrt{\tilde{a}_1(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_m \sqrt{\tilde{a}_m(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma,$$

$$\dot{\psi}_z = \left[\frac{\partial T_z f(T^{-1}[x^T, z^T]^T, t)}{\partial x} T_x + \frac{\partial T_z f(T^{-1}[x^T, z^T]^T, t)}{\partial z} T_z \right] \xi$$
$$+ T_z \left[q_1 \sqrt{\tilde{a}_1(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_m \sqrt{\tilde{a}_m(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma.$$

From (A1)–(A2), we have that $T_x f(T^{-1}[x^T, z^T]^T, t) = f_x(x, z, t)$ and $T_z f(T^{-1}[x^T, z^T]^T, t) = \frac{1}{\epsilon} f_z(x, z, t, \epsilon)$. Furthermore, substituting for $\tilde{a}(T^{-1}[x^T, z^T]^T, t) = [\hat{a}_s(T^{-1}[x^T, z^T]^T, t), (1/\epsilon)\hat{a}_f(T^{-1}[x^T, z^T]^T, t)]^T$, we have

$$\dot{\psi}_x = \frac{\partial f_x(x, z, t)}{\partial x} \psi_x + \frac{\partial f_x(x, z, t)}{\partial z} \psi_z +$$
$$T_x \left[q_1 \sqrt{\hat{a}_{s1}(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_{m_s} \sqrt{\hat{a}_{sm_s}(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma_x$$
$$+ T_x \left[q_{m_s+1} \sqrt{\frac{1}{\epsilon} \hat{a}_{f1}(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_{m_s+m_f} \sqrt{\frac{1}{\epsilon} \hat{a}_{fm_f}(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma_f, \qquad (A3)$$
$$\dot{\psi}_z = \frac{\partial \frac{1}{\epsilon} f_z(x, z, t, \epsilon)}{\partial x} \psi_x + \frac{\partial \frac{1}{\epsilon} f_z(x, z, t, \epsilon)}{\partial z} \psi_z +$$
$$T_z \left[q_1 \sqrt{\hat{a}_{s1}(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_{m_s} \sqrt{\hat{a}_{sm_s}(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma_x$$
$$+ T_z \left[q_{m_s+1} \sqrt{\frac{1}{\epsilon} \hat{a}_{f1}(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_{m_s+m_f} \sqrt{\frac{1}{\epsilon} \hat{a}_{fm_f}(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma_f, \qquad (A4)$$

where $\Gamma = [\Gamma_x^T, \Gamma_f^T]^T$. From (A1) we have that, $T_x q_i = 0$ for $i = m_s + 1, \ldots, m_s + m_f$. Then, multiplying (A4) by ϵ, and taking $\Gamma_z = [\Gamma_x^T, \Gamma_f^T]^T$, we can write the system (A3)–(A4) in the form of system (6)–(7).

Appendix B: Proof of Theorem 1

In this section, we present the proof of Theorem 1. We first derive a set of intermediate results, given in Claim 2 - 4 and Lemma 1, that will be used in the proof. In Claim 2, we derive the moment dynamics of the full system and show that these moment dynamics are also in the singular perturbation form. In Claim 3, we derive the moment dynamics of the reduced system. Next, in Claim 4, we show that setting $\epsilon = 0$ in the moment dynamics of

the full system, yields the moment dynamics of the reduced system. Then, these results are summarized in Lemma 1, which is then used to prove Theorem 1.

Lemma 1. *The first and second moment dynamics for the variables ψ_x and ψ_z of the full system (8) - (11) can be expressed in the singular perturbation form*

$$\frac{d\mathbb{E}[\psi_x]}{dt} = S_x(x, z, t)\mathbb{E}[\psi_x] + S_z(x, z, t)\mathbb{E}[\psi_z], \tag{B1}$$

$$\frac{d\mathbb{E}[\psi_x\psi_x^T]}{dt} = S_x(x, z, t)\mathbb{E}[\psi_x\psi_x^T] + S_z(x, z, t)\mathbb{E}[\psi_z\psi_x^T]$$

$$+ \mathbb{E}[\psi_x\psi_x^T]S_x(x, z, t)^T + (\mathbb{E}[\psi_z\psi_x^T])^T S_z(x, z, t)^T$$

$$+ \sigma_x(x, z, t)\sigma_x(x, z, t)^T, \tag{B2}$$

$$\epsilon\frac{d\mathbb{E}[\psi_z]}{dt} = F_x(x, z, t, \epsilon)\mathbb{E}[\psi_x] + F_z(x, z, t, \epsilon)\mathbb{E}[\psi_z], \tag{B3}$$

$$\epsilon\frac{d\mathbb{E}[\psi_z\psi_x^T]}{dt} = \epsilon\mathbb{E}[\psi_z\psi_x^T]S_x(x, z, t)^T + \epsilon\mathbb{E}[\psi_z\psi_z^T]S_z(x, z, t)^T$$

$$+ F_x(x, z, t, \epsilon)\mathbb{E}[\psi_x\psi_x^T] + F_z(x, z, t, \epsilon)\mathbb{E}[\psi_z\psi_x^T]$$

$$+ \sigma_z(x, z, t, \epsilon)[\sigma_x(x, z, t) \ 0]^T, \tag{B4}$$

$$\epsilon\frac{d\mathbb{E}[\psi_z\psi_z^T]}{dt} = F_x(x, z, t, \epsilon)\mathbb{E}[\psi_x\psi_z^T] + F_z(x, z, t, \epsilon)\mathbb{E}[\psi_z\psi_z^T]$$

$$+ \mathbb{E}[\psi_z\psi_x^T]F_x(x, z, t, \epsilon)^T + \mathbb{E}[\psi_z\psi_z^T]F_z(x, z, t, \epsilon)^T$$

$$+ \frac{1}{\epsilon}\sigma_z(x, z, t, \epsilon)\sigma_z(x, z, t, \epsilon)^T, \tag{B5}$$

where x and z are the solutions of the equations (8) - (9), and the initial conditions are given by $\mathbb{E}[\psi_x(0)] = \psi_{x0}$, $\mathbb{E}[\psi_x\psi_x^T(0)] = \psi_{x0}\psi_{x0}^T$, $\mathbb{E}[\psi_z(0)] = \psi_{z0}$, $\mathbb{E}[\psi_z\psi_x^T(0)] = \psi_{z0}\psi_{x0}^T$, $\mathbb{E}[\psi_z\psi_z^T(0)] = \psi_{z0}\psi_{z0}^T$.

Proof. We express the equations (10) - (11) in the form

$$\dot{\psi}_x = S_x(x, z, t)\psi_x + S_z(x, z, t)\psi_z + [\sigma_x(x, z, t) \ 0]\Gamma_z,$$

$$\epsilon\dot{\psi}_z = F_x(x, z, t, \epsilon)\psi_x + F_z(x, z, t, \epsilon)\psi_z + \sigma_z(x, z, t, \epsilon)\Gamma_z,$$

where $[\sigma_x(x, z, t) \ 0] \in \mathbb{R}^{n\times(m_s+m_f)}$. Then, as the x and z are deterministic we use the linearity of the expectation operator to derive the dynamics for the first moments as

$$\frac{d\mathbb{E}[\psi_x]}{dt} = S_x(x, z, t)\mathbb{E}[\psi_x] + S_z(x, z, t)\mathbb{E}[\psi_z], \tag{B6}$$

$$\frac{d\mathbb{E}[\psi_z]}{dt} = \frac{1}{\epsilon}F_x(x, z, t, \epsilon)\mathbb{E}[\psi_x] + \frac{1}{\epsilon}F_z(x, z, t, \epsilon)\mathbb{E}[\psi_z]. \tag{B7}$$

Similarly, Proposition III.1 in Bence et al.[55] can be used to write the second moment dynamics as

$$
\frac{d}{dt}\mathbb{E}\begin{bmatrix} \psi_x \psi_x^T & \psi_x \psi_z^T \\ \psi_z \psi_x^T & \psi_z \psi_z^T \end{bmatrix} =
$$

$$
\begin{bmatrix} \psi_x (S_x(x,z,t)\psi_x + S_z(x,z,t)\psi_z)^T & \frac{1}{\epsilon}\psi_x (F_x(x,z,t,\epsilon)\psi_x + F_z(x,z,t,\epsilon)\psi_z)^T \\ \psi_z (S_x(x,z,t)\psi_x + S_z(x,z,t)\psi_z)^T & \frac{1}{\epsilon}\psi_z (F_x(x,z,t,\epsilon)\psi_x + F_z(x,z,t,\epsilon)\psi_z)^T \end{bmatrix}
$$

$$
+ \begin{bmatrix} (S_x(x,z,t)\psi_x + S_z(x,z,t)\psi_z)\psi_x^T & (S_x(x,z,t)\psi_x + S_z(x,z,t)\psi_z)\psi_z^T \\ \frac{1}{\epsilon}(F_x(x,z,t,\epsilon)\psi_x + F_z(x,z,t,\epsilon)\psi_z)\psi_x^T & \frac{1}{\epsilon}(F_x(x,z,t,\epsilon)\psi_x + F_z(x,z,t,\epsilon)\psi_z)\psi_z^T \end{bmatrix}
$$

$$
+ \begin{bmatrix} \sigma_x(x,z,t)\sigma_x(x,z,t)^T & \frac{1}{\epsilon}[\ \sigma_x(x,z,t)\ \ 0\ ]\sigma_z(x,z,t,\epsilon)^T \\ \frac{1}{\epsilon}\sigma_z(x,z,t,\epsilon)[\ \sigma_x(x,z,t)\ \ 0\ ]^T & \frac{1}{\epsilon^2}\sigma_z(x,z,t,\epsilon)\sigma_z(x,z,t,\epsilon)^T \end{bmatrix}. \tag{B8}
$$

Then, summing the corresponding entries of the matrices in equation (B8) and using the linearity of the expectation operator, the equations (B6)–(B8) can be written in the form (B1)–(B5). We have that $\mathbb{E}[\psi_x \psi_z^T] = (\mathbb{E}[\psi_z \psi_x^T])^T$, and thus, we do not consider the dynamics of the variable $\mathbb{E}[\psi_x \psi_z^T]$ in the equations (B1)–(B5). Furthermore, the initial conditions ψ_{x0} and ψ_{z0} are deterministic, which yields $\mathbb{E}[\psi_x(0)] = \psi_{x0}$, $\mathbb{E}[\psi_x \psi_x^T(0)] = \psi_{x0}\psi_{x0}^T$, $\mathbb{E}[\psi_z(0)] = \psi_{z0}$, $\mathbb{E}[\psi_z \psi_x^T(0)] = \psi_{z0}\psi_{x0}^T$, $\mathbb{E}[\psi_z \psi_z^T(0)] = \psi_{z0}\psi_{z0}^T$. $\qquad\square$

Next, we derive the moment dynamics of the reduced system (14)–(17).

Lemma 2. *The first and second moment dynamics for the variable $\bar{\psi}_x$ of the reduced system (14)–(15) can be written in the form*

$$
\frac{d\mathbb{E}[\bar{\psi}_x]}{dt} = S(\bar{x},t)\mathbb{E}[\bar{\psi}_x], \qquad \mathbb{E}[\bar{\psi}_x(0)] = \psi_{x0}, \tag{B9}
$$

$$
\frac{d\mathbb{E}[\bar{\psi}_x \bar{\psi}_x^T]}{dt} = S(\bar{x},t)\mathbb{E}[\bar{\psi}_x \bar{\psi}_x^T] + \mathbb{E}[\bar{\psi}_x \bar{\psi}_x^T]S(\bar{x},t)^T
$$

$$
+ \sigma_x(\bar{x},\gamma_1(\bar{x},t),t),t)\sigma_x(\bar{x},\gamma_1(\bar{x},t),t),t)^T,
$$

$$
\mathbb{E}[\bar{\psi}_x(0)\bar{\psi}_x(0)^T] = \psi_{x0}\psi_{x0}^T, \tag{B10}
$$

and the first and second moments for the variable $\bar{\psi}_z$ of the reduced system (14)–(15) can be written in the form

$$
\mathbb{E}[\bar{\psi}_z] = \gamma_2(\bar{x},t)\mathbb{E}[\bar{\psi}_x] \tag{B11}
$$

35

$$\mathbb{E}[\bar{\psi}_z \bar{\psi}_z^T] = \gamma_2(\bar{x}, t) \mathbb{E}[\bar{\psi}_x \bar{\psi}_x^T] \gamma_2(\bar{x}, t)^T + g(\bar{x}, t) g(\bar{x}, t)^T. \tag{B12}$$

Proof. Using the Proposition III.1 in Bence et al.[55], the first and second moment dynamics of $\bar{\psi}_x$ in (15) can be written in the form

$$\frac{d\mathbb{E}[\bar{\psi}_x]}{dt} = \mathbb{E}[S(\bar{x}, t) \bar{\psi}_x],$$

$$\frac{d\mathbb{E}[\bar{\psi}_x \bar{\psi}_x^T]}{dt} = \mathbb{E}[S(\bar{x}, t) \bar{\psi}_x \bar{\psi}_x^T] + \mathbb{E}[\bar{\psi}_x (\bar{\psi}_x^T S(\bar{x}, t)^T)]$$

$$+ \sigma_x(\bar{x}, \gamma_1(\bar{x}, t), t), t) \sigma_x(\bar{x}, \gamma_1(\bar{x}, t), t), t)^T.$$

We have that the dynamics of $\bar{x}$ in (14) are deterministic. Therefore, using the linearity of the expectation operator, the moment dynamics of $\bar{\psi}_x$ in (15) can be written in the form of (B9)–(B10).

In order to derive the moments of the variable $\bar{\psi}_z$, we take the expected value of equation (16), which yields $\mathbb{E}[\bar{\psi}_z] = \mathbb{E}[\gamma_2(\bar{x}, t) \bar{\psi}_x]$ as the elements of the vector $N(0, 1)$ are normal random variable with zero mean. Since $\bar{x}$ is deterministic, we have that $\mathbb{E}[\bar{\psi}_z] = \gamma_2(\bar{x}, t) \mathbb{E}[\bar{\psi}_x]$. Calculating the second moment of $\bar{\psi}_z$, we obtain

$$\mathbb{E}[\bar{\psi}_z \bar{\psi}_z^T] = \mathbb{E}[(\gamma_2(\bar{x}, t) \psi_x + g(\bar{x}, t) N(0, 1))(\gamma_2(\bar{x}, t) \psi_x + g(\bar{x}, t) N(0, 1))^T].$$

Expanding further and using the fact that $N(0, 1)$ is independent of $\bar{\psi}_x$, we have

$$\mathbb{E}[\bar{\psi}_z \bar{\psi}_z^T] = \gamma_2(\bar{x}, t) \mathbb{E}[\psi_x \psi_x^T] \gamma_2(\bar{x}, t)^T$$

$$+ g(\bar{x}, t) \mathbb{E}[N(0, 1)] \psi_x^T \gamma_2(\bar{x}, t)^T$$

$$+ \gamma_2(\bar{x}, t) \psi_x \mathbb{E}[N(0, 1)^T] g(\bar{x}, t)^T$$

$$+ g(\bar{x}, t) \mathbb{E}[N(0, 1) N(0, 1)^T] g(\bar{x}, t)^T.$$

Since $N(0, 1)$ is a vector of standard normal random variables, we have that $\mathbb{E}[N(0, 1)] = 0$ and $\mathbb{E}[N(0, 1) N(0, 1)^T] = I_{d \times d}$ where $I_{d \times d}$ is an $d \times d$ identity matrix. Thus, we have that $\mathbb{E}[\bar{\psi}_z \bar{\psi}_z^T] = \gamma_2(\bar{x}, t) \mathbb{E}[\bar{\psi}_x \bar{\psi}_x^T] \gamma_2(\bar{x}, t)^T + g(\bar{x}, t) g(\bar{x}, t)^T.$ $\square$

Next, we derive the set of reduced-order moments obtained by setting $\epsilon = 0$ in the moment dynamics of the full system (B1)–(B5).

Lemma 3. *Setting $\epsilon = 0$ in the system of moment dynamics (B1)–(B5) and the dynamics of x and z given by (8)–(9), yields the moment dynamics of the reduced system (B9)–(B12) where the dynamics of $\bar{x}$ and $\bar{z}$ are given by (14) and (16), respectively.*

Proof. Setting $\epsilon = 0$, we have

$$0 = f_z(x, z, t, 0), \tag{B13}$$

$$0 = F_x(x, z, t, 0)\mathbb{E}[\psi_x] + F_z(x, z, t, 0)\mathbb{E}[\psi_z], \tag{B14}$$

$$0 = F_x(x, z, t, 0)\mathbb{E}[\psi_x\psi_x^T] + F_z(x, z, t, 0)\mathbb{E}[\psi_z\psi_x^T], \tag{B15}$$

$$0 = F_x(x, z, t, 0)\mathbb{E}[\psi_x\psi_z^T] + F_z(x, z, t, 0)\mathbb{E}[\psi_z\psi_z^T]$$
$$+ \mathbb{E}[\psi_z\psi_x^T]F_x(x, z, t, \epsilon)^T + \mathbb{E}[\psi_z\psi_z^T]F_z(x, z, t, 0)^T$$
$$+ \Lambda(x, z, t, 0). \tag{B16}$$

Under Assumption 3, there exists an isolated real root $z = \gamma_1(x, t)$ for equation (B13). Thus, the corresponding unique solutions to equations (B14)–(B15) are given by

$$\mathbb{E}[\psi_z] = -F_z(x, \gamma_1(x, t), t, 0)^{-1}(F_x(x, \gamma_1(x, t), t, 0)\mathbb{E}[\psi_x])$$
$$= \gamma_2(x, t)\mathbb{E}[\psi_x], \tag{B17}$$

$$\mathbb{E}[\psi_z\psi_x^T] = -F_z(x, \gamma_1(x, t), t, 0)^{-1}(F_x(x, \gamma_1(x, t), t, 0)\mathbb{E}[\psi_x\psi_x^T])$$
$$= \gamma_2(x, t)\mathbb{E}[\psi_x\psi_x^T]. \tag{B18}$$

Substituting $z = \gamma_1(x, t)$ and equations (B17)–(B18) in (8) and the moment equations (B1)–(B5), yields the dynamics of $\bar{x}$ given by (14) and the moment dynamics of the variable $\bar{\psi}_x$ reduced system given by (B9)–(B10).

Furthermore, we have that the equation (B17) is equivalent to the first moment of the variable $\bar{\psi}_z$ of the reduced system given by (B11). Next, in order to solve equation (B16), we substitute (B17)–(B18) in (B16), which yields

$$F_z(x, \gamma_1(x, t), t, 0)\mathbb{E}[\psi_z\psi_z^T] + \mathbb{E}[\psi_z\psi_z^T]F_z(x, \gamma_1(x, t), t, 0)^T$$
$$= -F_x(x, \gamma_1(x, t), t, 0)\mathbb{E}[\psi_x\psi_x^T]\gamma_2(x, t)^T$$
$$- \gamma_2(x, t)\mathbb{E}[\psi_x\psi_x^T]F_x(x, \gamma_1(x, t), t, \epsilon)^T$$
$$- \Lambda(x, \gamma_1(x, t), t, 0), \tag{B19}$$

which is in the form of a Lyapunov equation $AP + PA^T = -Q$ with

$$A = F_z(x, \gamma_1(x, t), t, 0),$$
$$P = \mathbb{E}[\psi_z\psi_z^T],$$
$$Q = F_x(x, \gamma_1(x, t), t, 0)\mathbb{E}[\psi_x\psi_x^T]\gamma_2(x, t)^T$$

37

$$+ \gamma_2(x,t)\mathbb{E}[\psi_x \psi_x^T]F_x(x, \gamma_1(x,t), t, \epsilon)^T + \Lambda(x, \gamma_1(x,t), t, 0)$$

Under Assumption 3, the matrix $F_z(x, \gamma_1(x,t), t, 0)$ is Hurwitz for all x and t and thus, the equation (B19) has a unique solution $\mathbb{E}[\psi_z \psi_z^T] = h(x, \mathbb{E}[\psi_x \psi_x^T], t)$. In order to prove that $h(x, \mathbb{E}[\psi_x \psi_x^T], t) = \gamma_2(x,t)\mathbb{E}[\psi_x \psi_x^T]\gamma_2(x,t)^T + g(x,t)g(x,t)^T$ given by the second moment of the $\bar{\psi}_z$ of the reduced system in (B12), we substitute $\mathbb{E}[\psi_z \psi_z^T] = \gamma_2(x,t)\mathbb{E}[\psi_x \psi_x^T]\gamma_2(x,t)^T + g(x,t)g(x,t)^T$, which yields

$$F_z(x, \gamma_1(x,t), t, 0)\gamma_2(x,t)\mathbb{E}[\psi_x \psi_x^T]\gamma_2(x,t)^T + F_z(x, \gamma_1(x,t), t, 0)g(x,t)g(x,t)^T$$
$$+ \gamma_2(x,t)\mathbb{E}[\psi_x \psi_x^T]\gamma_2(x,t)^T F_z(x, \gamma_1(x,t), t, 0)^T + g(x,t)g(x,t)^T F_z(x, \gamma_1(x,t), t, 0)^T$$
$$= -F_x(x, \gamma_1(x,t), t, 0)\mathbb{E}[\psi_x \psi_x^T]\gamma_2(x,t)^T - \gamma_2(x,t)\mathbb{E}[\psi_x \psi_x^T]F_x(x, \gamma_1(x,t), t, \epsilon)^T$$
$$- \Lambda(x, \gamma_1(x,t), t, 0)$$

Then, canceling the common terms on both sides we obtain

$$F_z(x, \gamma_1(x,t), t, 0)g(x,t)g(x,t)^T + g(x,t)g(x,t)^T F_z(x, \gamma_1(x,t), t, 0)^T = -\Lambda(x, \gamma_1(x,t), t, 0)$$

which satisfies the equation (18) in the definition of the reduced fast system. $\square$

We then have the following result.

Lemma 4. *Consider the full system in (8)–(11), the reduced system in (14)–(15), and the moment dynamics for the full and reduced systems in (B1)–(B5), (B9)–(B12) respectively. We have that, under Assumptions 1 - 3, the commutative diagram in Figure 8 holds.*

Proof. Proof follows from Lemmas 1, 2 and 3. $\square$

Although, Lemma 1 shows that the setting $\epsilon = 0$ in the moment dynamics of the full system yields the moment dynamics of the reduced system, it does not guarantee that the trajectories of the moments become close to each other as ϵ decreases. Therefore, we next use the Tikhonov's theorem to prove that the moments of the reduced system are within an $O(\epsilon)$-neighborhood of the moments of the full system.

Proof of Theorem 1:

It can be seen from the commutative diagram in Lemma 1 that setting $\epsilon = 0$ in the moment equations of the full system together with the dynamics of x and z yields the

Original System $\longrightarrow$ Reduced System

$$\dot{x} = f_x(x, z, t),$$
$$\epsilon \dot{z} = f_z(x, z, t, \epsilon),$$
$$\dot{\psi}_x = S_x(x, z, t)\psi_x + S_z(x, z, t)\psi_z + \sigma_x(x, z, t)\Gamma_x,$$
$$\epsilon \dot{\psi}_z = F_x(x, z, t, \epsilon)\psi_x + F_z(x, z, t, \epsilon)\psi_z + \sigma_z(x, z, t, \epsilon)\Gamma_z.$$

$$\dot{\bar{x}} = f_x(\bar{x}, \gamma_1(\bar{x}, t), t),$$
$$\bar{z} = \gamma_1(\bar{x}, t),$$
$$\dot{\bar{\psi}}_x = [S_x(\bar{x}, \gamma_1(\bar{x}, t), t) + S_z(\bar{x}, \gamma_1(\bar{x}, t), t)\gamma_2(\bar{x}, t)]\bar{\psi}_x$$
$$\qquad\qquad + \sigma_x(\bar{x}, \gamma_1(\bar{x}, t), t)\Gamma_x,$$
$$\bar{\psi}_z = \gamma_2(\bar{x}, t)\psi_x + g(\bar{x}, t)N(0, 1).$$

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

$$\dot{x} = f_x(x, z, t),$$
$$\epsilon \dot{z} = f_z(x, z, t, \epsilon),$$
$$\frac{d}{dt}\begin{pmatrix} \mathbb{E}[\psi_x] \\ \mathbb{E}[\psi_x\psi_x^T] \\ \epsilon\mathbb{E}[\psi_z] \\ \epsilon\mathbb{E}[\psi_z\psi_x^T] \\ \epsilon\mathbb{E}[\psi_z\psi_z^T] \end{pmatrix} = \dots.$$

$$\dot{\bar{x}} = f_x(\bar{x}, \gamma_1(\bar{x}, t), t),$$
$$\frac{d}{dt}\begin{pmatrix} \mathbb{E}[\bar{\psi}_x] \\ \mathbb{E}[\bar{\psi}_x\bar{\psi}_x^T] \end{pmatrix} = \dots,$$
$$\bar{z} = \gamma_1(\bar{x}, t),$$
$$\begin{pmatrix} \mathbb{E}[\bar{\psi}_z] \\ \mathbb{E}[\bar{\psi}_z\bar{\psi}_z^T] \end{pmatrix} = \dots.$$
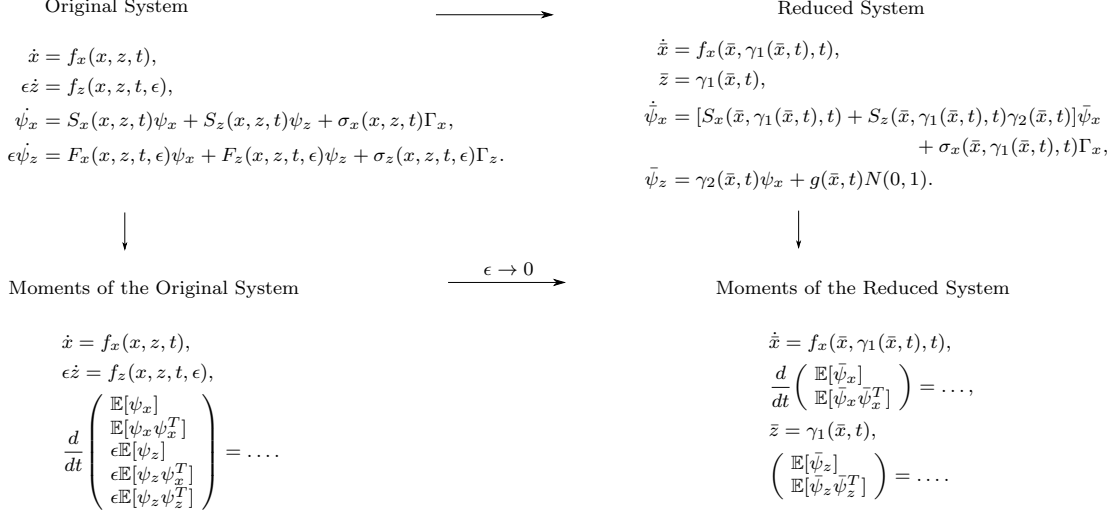
FIG. 8. Setting $\epsilon = 0$ in the moment dynamics of the full system yields the moment dynamics of the reduced system.

moment equations of the reduced system. Thus, as the moment dynamics are deterministic, we can apply the Tikhonov's theorem to the moments of the full system together with the dynamics of x and z to prove the results given in Theorem 1. Towards this end, we first prove that the assumptions of the Tikhonov's theorem are satisfied. We first consider the boundary layer dynamics of the moment dynamics (B1)–(B5), where we define the boundary layer variables as

$$b_1 = z - \gamma_1(x, t), \tag{B20}$$

$$b_2 = \mathbb{E}[\psi_z] - \gamma_2(x, t)\mathbb{E}[\psi_x], \tag{B21}$$

$$b_3 = \mathbb{E}[\psi_z\psi_x^T] - \gamma_2(x, t)\mathbb{E}[\psi_x\psi_x^T], \tag{B22}$$

$$b_4 = \mathbb{E}[\psi_z\psi_z^T] - (\gamma_2(x, t)\mathbb{E}[\psi_x\psi_x^T]\gamma_2(x, t)^T + g(x, t)g(x, t)^T). \tag{B23}$$

Define $\gamma_3(x, \mathbb{E}[\psi_x\psi_x^T], t) = \gamma_2(x, t)\mathbb{E}[\psi_x\psi_x^T]\gamma_2(x, t)^T + g(x, t)g(x, t)^T$. Then, we have that the dynamics of the boundary layer variables are given by

$$\frac{db_1}{dt} = \frac{dz}{dt} - \frac{d\gamma_1(x, t)}{dt},$$
$$\frac{db_2}{dt} = \frac{d\mathbb{E}[\psi_z]}{dt} - \frac{d\gamma_2(x, t)\mathbb{E}[\psi_x]}{dt},$$
$$\frac{db_3}{dt} = \frac{d\mathbb{E}[\psi_z\psi_x]}{dt} - \frac{d\gamma_2(x, t)\mathbb{E}[\psi_x\psi_x^T]}{dt},$$
$$\frac{db_4}{dt} = \frac{d\mathbb{E}[\psi_z\psi_z^T]}{dt} - \frac{d\gamma_3(x, \mathbb{E}[\psi_x\psi_x^T], t)}{dt}.$$

39

Taking $\tau = t/\epsilon$ to be the time variable in the fast time-scale we have that

$$\frac{db_1}{d\tau} = \epsilon\frac{dz}{dt} - \epsilon\frac{\partial\gamma_1(x,t)}{\partial t} - \epsilon\frac{\partial\gamma_1(x,t)}{\partial x}\frac{dx}{dt},$$

$$\frac{db_2}{d\tau} = \epsilon\frac{d\mathbb{E}[\psi_z]}{dt} - \epsilon\mathbb{E}[\psi_x]\frac{\partial\gamma_2(x,t)}{\partial t} - \epsilon\mathbb{E}[\psi_x]\frac{\partial\gamma_2(x,t)}{\partial x}\frac{dx}{dt}$$
$$- \epsilon\gamma_2(x,t)\frac{d\mathbb{E}[\psi_x]}{dt},$$

$$\frac{db_3}{d\tau} = \epsilon\frac{d\mathbb{E}[\psi_z\psi_x^T]}{dt} - \epsilon\mathbb{E}[\psi_x\psi_x^T]\frac{\partial\gamma_2(x,t)}{\partial t}$$
$$- \epsilon\mathbb{E}[\psi_x\psi_x^T]\frac{\partial\gamma_2(x,t)}{\partial x}\frac{dx}{dt} - \epsilon\gamma_2(x,t)\frac{d\mathbb{E}[\psi_x\psi_x^T]}{dt},$$

$$\frac{db_4}{d\tau} = \epsilon\frac{d\mathbb{E}[\psi_z\psi_z^T]}{dt} - \epsilon\frac{\partial\gamma_3(x,\mathbb{E}[\psi_x\psi_x^T],t)}{\partial t}$$
$$- \epsilon\frac{\partial\gamma_3(x,\mathbb{E}[\psi_x\psi_x^T],t)}{\partial x}\frac{dx}{dt}$$
$$- \epsilon\frac{\partial\gamma_3(x,\mathbb{E}[\psi_x\psi_x^T],t)}{\partial\mathbb{E}[\psi_x\psi_x^T]}\frac{d\mathbb{E}[\psi_x\psi_x^T]}{dt}.$$

Then, substituting for the moment dynamics from equations (9), (B3), (B4) and (B5) we obtain

$$\frac{db_1}{d\tau} = f_z(x,z,t,\epsilon) - \epsilon\frac{\partial\gamma_1(x,t)}{\partial t} - \epsilon\frac{\partial\gamma_1(x,t)}{\partial x}\frac{dx}{dt}, \tag{B24}$$

$$\frac{db_2}{d\tau} = F_x(x,z,t,\epsilon)\mathbb{E}[\psi_x] + F_z(x,z,t,\epsilon)\mathbb{E}[\psi_z]$$
$$- \epsilon\mathbb{E}[\psi_x]\frac{\partial\gamma_2(x,t)}{\partial t} - \epsilon\mathbb{E}[\psi_x]\frac{\partial\gamma_2(x,t)}{\partial x}\frac{dx}{dt}$$
$$- \epsilon\gamma_2(x,t)\frac{d\mathbb{E}[\psi_x]}{dt}, \tag{B25}$$

$$\frac{db_3}{d\tau} = \epsilon\mathbb{E}[\psi_z\psi_x^T]S_x(x,z,t)^T + \epsilon\mathbb{E}[\psi_z\psi_z^T]S_z(x,z,t)^T$$
$$+ F_x(x,z,t,\epsilon)\mathbb{E}[\psi_x\psi_x^T] + F_z(x,z,t,\epsilon)\mathbb{E}[\psi_z\psi_x^T]$$
$$+ \sigma_z(x,z,t,\epsilon)[\sigma_x(x,z,t)\ 0]^T - \epsilon\mathbb{E}[\psi_x\psi_x^T]\frac{\partial\gamma_2(x,t)}{\partial t}$$
$$- \epsilon\mathbb{E}[\psi_x\psi_x^T]\frac{\partial\gamma_2(x,t)}{\partial x}\frac{dx}{dt} - \epsilon\gamma_2(x,t)\frac{d\mathbb{E}[\psi_x\psi_x^T]}{dt}, \tag{B26}$$

$$\frac{db_4}{d\tau} = F_x(x,z,t,\epsilon)\mathbb{E}[\psi_x\psi_z^T] + F_z(x,z,t,\epsilon)\mathbb{E}[\psi_z\psi_z^T]$$
$$+ \mathbb{E}[\psi_z\psi_x^T]F_x(x,z,t,\epsilon)^T + \mathbb{E}[\psi_z\psi_z^T]F_z(x,z,t,\epsilon)^T$$
$$+ \frac{1}{\epsilon}\sigma_z(x,z,t,\epsilon)\sigma_z(x,z,t,\epsilon)^T - \epsilon\frac{\partial\gamma_3(x,\mathbb{E}[\psi_x\psi_x^T],t)}{\partial t}$$
$$- \epsilon\frac{\partial\gamma_3(x,\mathbb{E}[\psi_x\psi_x^T],t)}{\partial x}\frac{dx}{dt}$$

$$-\epsilon\frac{\partial\gamma_3(x,\mathbb{E}[\psi_x\psi_x^T],t)}{\partial\mathbb{E}[\psi_x\psi_x^T]}\frac{d\mathbb{E}[\psi_x\psi_x^T]}{dt}. \tag{B27}$$

where we take $z = b_1 + \gamma_1(x,t)$ and $\mathbb{E}[\psi_z] = b_2 + \gamma_2(x,t)\mathbb{E}[\psi_x]$, $\mathbb{E}[\psi_z\psi_x^T] = b_3 + \gamma_2(x,t)\mathbb{E}[\psi_x\psi_x^T]$ and $\mathbb{E}[\psi_z\psi_z^T] = b_4 + \gamma_3(x,\mathbb{E}[\psi_x\psi_x^T],t)$. From Assumptions 1 - 3 we have that the functions $f_x, f_z, \gamma_1, \gamma_2,$ and γ_3 are continuously differentiable and thus is bounded for a finite time interval $t \in [0, t_1]$. Furthermore due to the linearity of the moment equations we have that $\mathbb{E}[\psi_x]$ and $\mathbb{E}[\psi_x\psi_x^T]$ exist and are bounded for a finite time interval $t \in [0, t_1]$. Thus, setting $\epsilon = 0$ in the equations (B24)–(B27), we obtain the boundary layer system given by

$$\frac{db_1}{d\tau} = f_z(x, b_1 + \gamma_1(x,t), t, 0), \tag{B28}$$

$$\frac{db_2}{d\tau} = F_x(x, b_1 + \gamma_1(x,t), t, 0)\mathbb{E}[\psi_x] \tag{B29}$$
$$+ F_z(x, b_1 + \gamma_1(x,t), t, 0)(b_2 + \gamma_2(x,t)\mathbb{E}[\psi_x]),$$

$$\frac{db_3}{d\tau} = F_x(x, b_1 + \gamma_1(x,t), t, 0)\mathbb{E}[\psi_x\psi_x^T]$$
$$+ F_z(x, b_1 + \gamma_1(x,t), t, 0)(b_3 + \gamma_2(x,t)\mathbb{E}[\psi_x\psi_x^T]), \tag{B30}$$

$$\frac{db_4}{d\tau} = F_x(x, b_1 + \gamma_1(x,t), t, 0)(b_3 + \gamma_2(x,t)\mathbb{E}[\psi_x\psi_x^T])^T$$
$$+ F_z(x, b_1 + \gamma_1(x,t), t, 0)(b_4 + \gamma_3(x, \mathbb{E}[\psi_x\psi_x^T], t))$$
$$+ (b_3 + \gamma_2(x,t)\mathbb{E}[\psi_x\psi_x^T])F_x(x, b_1 + \gamma_1(x,t), t, 0)^T$$
$$+ (b_4 + \gamma_3(x, \mathbb{E}[\psi_x\psi_x^T], t))F_z(x, b_1 + \gamma_1(x,t), t, 0)^T$$
$$+ \Lambda(x, b_1 + \gamma_1(x,t), t),$$
$$\tag{B31}$$

Next, in order to prove that the origin is an exponentially stable equilibrium point of the boundary layer system, we linearize the system (B28)–(B31) about the origin. Towards this end, we first represent the matrix variable b_3 and b_4 in vector form. Let A be an $m \times n$ matrix and let $\text{vec}(A) = [a_{11}, \ldots, a_{m1}, \ldots, a_{1n}, \ldots, a_{nn}]^T$, where a_{ij} are the elements of the matrix A. Then, considering the dynamics of $\text{vec}(b_3)$ and $\text{vec}(b_4)$, we use the Kronecker Product denoted by $\otimes$ to obtain

$$\frac{db_1}{d\tau} = f_z(x, b_1 + \gamma_1(x,t), t, 0),$$
$$\frac{db_2}{d\tau} = F_x(x, b_1 + \gamma_1(x,t), t, 0)\mathbb{E}[\psi_x]$$
$$+ F_z(x, b_1 + \gamma_1(x,t), t, 0)(b_2 + \gamma_2(x,t)\mathbb{E}[\psi_x]),$$

$$\frac{d\text{vec}(b_3)}{d\tau} = (I \otimes F_z(x, b_1 + \gamma_1(x,t), t, 0))\text{vec}(b_3)$$

$$+ \text{vec}(g_2(b_1, x, \psi_x, t)),$$

$$\frac{d\text{vec}(b_4)}{d\tau} = (I \otimes F_z(x, b_1 + \gamma_1(x,t), t, 0)$$

$$+ F_z(x, b_1 + \gamma_1(x,t), t, 0) \otimes I)\text{vec}(b_4)$$

$$+ (I \otimes F_x(x, b_1 + \gamma_1(x,t), t, 0)$$

$$+ F_x(x, b_1 + \gamma_1(x,t), t, 0) \otimes I)\text{vec}(b_3)$$

$$+ \text{vec}(g_3(b_1, x, \psi_x, t)),$$

where

$$g_2(b_1, x, \psi_x, t) = F_x(x, b_1 + \gamma_1(x,t), t, 0)\mathbb{E}[\psi_x \psi_x^T]$$

$$+ F_z(x, b_1 + \gamma_1(x,t), t, 0)\gamma_2(x,t)\mathbb{E}[\psi_x \psi_x^T],$$

$$g_3(b_1, x, \psi_x, t) = F_x(x, b_1 + \gamma_1(x,t), t, 0)\mathbb{E}[\psi_x \psi_x^T]\gamma_2(x,t)^T$$

$$+ F_z(x, b_1 + \gamma_1(x,t), t, 0)\gamma_3(x, \mathbb{E}[\psi_x \psi_x^T], t)$$

$$+ \gamma_2(x,t)\mathbb{E}[\psi_x \psi_x^T]F_x(x, b_1 + \gamma_1(x,t), t, 0)^T$$

$$+ \gamma_3(x, \mathbb{E}[\psi_x \psi_x^T], t)F_z(x, b_1 + \gamma_1(x,t), t, 0)^T$$

$$+ \Lambda(x, b_1 + \gamma_1(x,t), t).$$

Then, considering the state vector $e = [b_1, b_2, \text{vec}(b_3), \text{vec}(b_4)]$ and linearizing about $\hat{e} = 0$, we obtain the dynamics of $\tilde{e} = e - \hat{e}$ in the form

$$\frac{d\tilde{e}}{d\tau} = \begin{bmatrix} J_{11} & 0 & 0 & 0 \\ J_{21} & J_{22} & 0 & 0 \\ J_{31} & 0 & J_{33} & 0 \\ J_{41} & 0 & J_{43} & J_{44} \end{bmatrix} \tilde{e}, \tag{B32}$$

where the diagonal entries are given by $J_{11} = \frac{\partial f_z(x, b_1 + \gamma_1(x,t), t, 0)}{\partial b_1}\big|_{b_1=0}$, $J_{22} = F_z(x, b_1 + \gamma_1(x,t), t, 0)\big|_{b_1=0}$, $J_{33} = (I \otimes F_z(x, b_1 + \gamma_1(x,t), t, 0))\big|_{b_1=0}$ and $J_{44} = (F_z(x, b_1 + \gamma_1(x,t), t, 0) \oplus F_z(x, b_1 + \gamma_1(x,t), t, 0))\big|_{b_1=0}$ where $\oplus$ denotes the Kronecker sum and the lower diagonal entries $J_{21}, J_{31}, J_{41}, J_{43}$ are appropriate functions.

Since the eigenvalues of a block triangular matrix are given by the union of eigenvalues of the diagonal blocks, we consider the eigenvalues of the diagonal entries. We have that

$\frac{\partial f_z(x, b_1 + \gamma_1(x,t), t, 0)}{\partial b_1}\Big|_{b_1=0} = \frac{\partial f_z(x,z,t,0)}{\partial z}\frac{dz}{db_1}\Big|_{z=\gamma_1(x,t)} = \frac{\partial f_z(x,z,t,0)}{\partial z}\Big|_{z=\gamma_1(x,t)}$, which is Hurwitz from As-

sumption 3. Furthermore, we have that $F_z(x, z, t, \epsilon) = \frac{\partial f_z(x,z,t,\epsilon)}{\partial z}$ from the definition of the

full system (8)–(11), and thus, $F_z(x, b_1 + \gamma_1(x,t), t, 0)\big|_{b_1=0} = \frac{\partial f_z(x,z,t,0)}{\partial z}\big|_{z=\gamma_1(x,t)}$ is Hurwitz

under Assumption 3. Thus, the diagonal term J_{22} is Hurwitz. Considering the eigenvalues

of J_{33}, we have that any eigenvalue of a Kronecker product of two matrices are given by

the product of the eigenvalues of the individual matrices[56]. Thus, the eigenvalues of J_{33} are

given by the eigenvalues of $F_z(x, b_1 + \gamma_1(x,t), t, 0)\big|_{b_1=0}$, which is Hurwitz under Assumption

3. Next, we consider the diagonal term J_{44}. From Theorem 13.16 in Laub[56] we have that

any eigenvalue of J_{44} is given by the sum of two eigenvalues of $F_z(x, b_1 + \gamma_1(x,t), t, 0)$. Since

the matrix F_z is Hurwitz under Assumption 3, we have that all eigenvalues of J_{44} have

negative real parts. Therefore, we have that the eigenvalues of the linearized system (B32)

have negative real parts and thus the origin is an exponentially stable equilibrium point

uniformly for x and t.

Next, we prove that the initial conditions of the system

$$\frac{db_1}{d\tau} = f_z(x_0, b_1 + \gamma_1(x_0, 0), 0, 0), \tag{B33}$$

$$\frac{db_2}{d\tau} = F_x(x_0, b_1 + \gamma_1(x_0, 0), 0, 0)\psi_{x_0}$$

$$+ F_z(x_0, b_1 + \gamma_1(x_0, 0), 0, 0)(b_2 + \gamma_2(x_0, 0)\psi_{x_0}), \tag{B34}$$

$$\frac{d\text{vec}(b_3)}{d\tau} = (I \otimes F_z(x_0, b_1 + \gamma_1(x_0, 0), 0, 0))\text{vec}(b_3)$$

$$+ \text{vec}(g_2(b_1, x_0, \psi_{x_0}, 0)), \tag{B35}$$

$$\frac{d\text{vec}(b_4)}{d\tau} = (I \otimes F_z(x_0, b_1 + \gamma_1(x_0, 0), 0, 0))$$

$$+ F_z(x_0, b_1 + \gamma_1(x_0, 0), 0, 0) \otimes I)\text{vec}(b_4)$$

$$+ \text{vec}(g_3(b_1, b_3, x_0, \psi_{x_0}, 0)), \tag{B36}$$

are in the region of attraction of the equilibrium point at the origin. From Assumption 3

we have that the initial condition z_0 is in the region of attraction of the equilibrium point

$z = \gamma_1(x_0, 0)$ of system $\frac{dz}{d\tau} = f_z(x_0, z, 0, 0)$. Thus, it follows that $z_0 - \gamma_1(x_0, 0)$ is in the

region of attraction of equilibrium point of $b_1 = 0$ for system (B33). Furthermore, we have

that the trajectory $b_1(\tau)$ that corresponds to the initial condition $z_0 - \gamma_1(x_0, 0)$ is bounded

and $\lim_{\tau \to \infty} b_1(\tau) = 0$. Using this fact and the linearity of the system (B34)–(B36), we next

prove that any trajectory of the system, starting with any initial condition, converges to

43

zero as $\tau \to \infty$. Towards this end, we first define the vector $r = [\hat{b}_2, \text{vec}(\hat{b}_3), \text{vec}(\hat{b}_4)]$ and write the system (B34)–(B36) in the form

$$
\frac{dr}{d\tau} = \begin{bmatrix} H_{11} & 0 & 0 \\ 0 & H_{22} & 0 \\ 0 & H_{32} & H_{33} \end{bmatrix} r
$$
$$
+ \begin{bmatrix} C_{11}(\tau) & 0 & 0 \\ 0 & C_{22}(\tau) & 0 \\ 0 & C_{32}(\tau) & C_{33}(\tau) \end{bmatrix} r + \begin{bmatrix} 0 \\ D_2(\tau) \\ D_3(\tau) \end{bmatrix},
$$

where

$$
H_{11} = F_z(x_0, \gamma_1(x_0, 0), 0, 0),
$$
$$
H_{22} = I \otimes F_z(x_0, \gamma_1(x_0, 0), 0, 0),
$$
$$
H_{32} = I \otimes F_x(x_0, \gamma_1(x_0, 0), 0, 0)) + F_x(x_0, \gamma_1(x_0, 0), 0, 0) \otimes I,
$$
$$
H_{33} = I \otimes F_z(x_0, \gamma_1(x_0, 0), 0, 0)) + F_z(x_0, \gamma_1(x_0, 0), 0, 0) \otimes I,
$$
$$
C_{11}(\tau) = F_z(x_0, b_1(\tau) - \gamma_1(x_0, 0), 0, 0) - F_z(x_0, \gamma_1(x_0, 0), 0, 0),
$$
$$
C_{22}(\tau) = I \otimes F_z(x_0, b_1(\tau) + \gamma_1(x_0, 0), 0, 0) - I \otimes F_z(x_0, \gamma_1(x_0, 0), 0, 0),
$$
$$
C_{32}(\tau) = I \otimes F_x(x_0, b_1(\tau) + \gamma_1(x_0, 0), 0, 0)) + F_x(x_0, b_1(\tau) + \gamma_1(x_0, 0), 0, 0) \otimes I
$$
$$
\quad - I \otimes F_x(x_0, \gamma_1(x_0, 0), 0, 0)) - F_x(x_0, \gamma_1(x_0, 0), 0, 0) \otimes I,
$$
$$
C_{33}(\tau) = I \otimes F_z(x_0, b_1(\tau) + \gamma_1(x_0, 0), 0, 0)) + F_z(x_0, b_1(\tau) + \gamma_1(x_0, 0), 0, 0) \otimes I
$$
$$
\quad - I \otimes F_z(x_0, \gamma_1(x_0, 0), 0, 0)) - F_z(x_0, \gamma_1(x_0, 0), 0, 0) \otimes I,
$$
$$
D_2(\tau) = \text{vec}(g_2(b_1(\tau), x_0, \psi_{x_0}, 0)),
$$
$$
D_3(\tau) = \text{vec}(g_3(b_1(\tau), x_0, \psi_{x_0}, 0)).
$$

Then, we apply Lemmas 9.4–9.6 from Khalil[4] to show that r tends to zero as $\tau \to \infty$ for any initial condition $r(0)$. From Assumption 3, we have that the matrix H is Hurwitz, and thus the system $\frac{dr}{d\tau} = Hr$ is globally exponentially stable. Therefore, there exists a Lyapunov function $V(r, \tau) = r^T P(\tau) r$ that satisfies (9.3)–(9.5) in Khalil[4]. Furthermore, we have that $\|C(\tau)h + D(\tau)\| \le \|C(\tau)\|\|h\| + \|D(\tau)\|$. We note that $\|D(\tau)\|$ is bounded as the functions $g_2(b_1, x_0, \psi_{x_0}, 0)$ and $g_3(b_1, x_0, \psi_{x_0}, 0)$ are continuous in $b_1(\tau)$ from Assumptions 1 and 2, and $\hat{b}_1(\tau)$ is bounded due to the asymptotic stability of the equilibrium point $\hat{b}_1 = 0$.

Furthermore, since $\lim_{\tau \to \infty} \hat{b}_1(\tau) = 0$, it follows from the definitions of the functions g_2 and g_3 that $\|C(\tau)\| \to 0$ and $\|D(\tau)\| \to 0$. Therefore, from Lemma 9.5, 9.6 and Lemma 9.4 in Khalil[4], we have that $\lim_{\tau \to \infty} r(\tau) = 0$ for all $r(0) \in \mathbb{R}^3$. Thus, it follows that the region of attraction for the system (B34)–(B36) is given by $R_{b1} \times \mathbb{R}^3$ where R_{b1} is the region of attraction of the equilibrium point $z_0 - \gamma_1(x_0, 0)$.

Next, we consider the remaining assumptions of the Tikhonov's theorem. We have that functions f_x, f_z, S_x, S_z, F_x, F_z, $\sigma_x \sigma_x^T$, $\sigma_z [\sigma_x \ 0]^T$ and $\sigma_z \sigma_z^T$ and their first partial derivatives are continuously differentiable from Assumptions 1 and 2. Under Assumption 1 we also have that the $\frac{\partial f_z(x,z,t,0)}{\partial z}$, $\frac{\partial F_x(x,z,t,0)}{\partial z}$, $\frac{\partial F_z(x,z,t,0)}{\partial z}$ have continuous first partial derivatives with respect to their arguments. From Assumptions 1 and 3 the first partial derivatives of $\gamma_1(x,t)$, $\gamma_2(x,t)\mathbb{E}[\psi_x]$, $\gamma_2(x,t)\mathbb{E}[\psi_x \psi_x^T]$, $\gamma_3(x, \mathbb{E}[\psi_x \psi_x^T], t)$ with respect to their arguments are also continuous. Under Assumption 4 there exists a unique, bounded solution for the reduced system (14) for $t \in [0, t_1]$. Furthermore, as the moment dynamics (B9) - (B10) are linear in the variables $\mathbb{E}[\bar{\psi}_x]$, $\mathbb{E}[\bar{\psi}_x \bar{\psi}_x^T]$ there exists a unique, bounded solution to (B9) - (B10) for $t \in [0, t_1]$. Thus, the assumptions of the Tikhonov's theorem are satisfied and applying the theorem to the set of moment equations in (B1)–(B5) and (B9)–(B12) yields the result (19)–(25).

Appendix C

Here, we prove that the moments of the reduced system (14)–(17) can provide a good approximation for the moments of the original variables v and ξ in the system (2)–(3). We only provide a complete proof for the variable v since the proof for $\mathbb{E}[\xi]$ and $\mathbb{E}[\xi^2]$ can be derived in a similar manner. From equation (2)–(3) we have that v and ξ represent the original variables, and from Clam 1 we have that

$$v = T^{-1}[x^T, z^T]^T \text{ and } \xi = T^{-1}[\psi_x^T, \psi_z^T]^T.$$

Then, let $\bar{v}$ and $\bar{\xi}$ represent the species concentrations obtained using the reduced system (14)–(17). Therefore, we have that

$$\bar{v} = T^{-1}[\bar{x}^T, \bar{z}^T]^T \text{ and } \bar{\xi} = T^{-1}[\bar{\psi}_x^T, \bar{\psi}_z^T]^T.$$

Then, computing the error between the moments of the original species concentrations and

the moments obtained using the reduced system, we have

$$\|v - \bar{v}\| = \left\| T^{-1} \begin{bmatrix} x \\ z \end{bmatrix} - T^{-1} \begin{bmatrix} \bar{x} \\ \bar{z} \end{bmatrix} \right\|,$$

$$= \left\| T^{-1} \begin{bmatrix} x - \bar{x} \\ z - \bar{z} \end{bmatrix} \right\|.$$

Using the Cauchy-Schwarz inequality, we have that

$$\|v - \bar{v}\| \leq \|T^{-1}\|_F \left\| \begin{bmatrix} x - \bar{x} \\ z - \bar{z} \end{bmatrix} \right\|,$$

and using the definition of the Euclidean norm we obtain

$$\|v - \bar{v}\| \leq \|T^{-1}\| \sqrt{\|x - \bar{x}\|^2 + \|z - \bar{z}\|^2}. \tag{C1}$$

From Theorem 1, we have that $\|x - \bar{x}\| \leq c_1 \epsilon$ and $\|z - \bar{z}\| \leq c_2 \epsilon$ for appropriate constants c_1 and c_2. Thus, from (C1) it follows that

$$\|v - \bar{v}\| \leq C\epsilon,$$

for an appropriate constant C. Similar result holds for the variables $\mathbb{E}[\xi]$ and $\mathbb{E}[\xi^2]$.

Appendix D

In this section, we present the derivation of the reduced-order moment dynamics of protein M and protein G given in (44)–(48) using the reduced system (38). For ease of analysis, we express the fast variable approximations in the reduced system (38) in terms of the variables $\bar{m}$ and $\bar{g}$ using the coordinate transforms $\bar{w} = \bar{m} - \bar{c}_2$ and $\bar{u} = \bar{g} - \bar{c}_1$. This yields

$$\bar{c}_1 = \frac{\bar{g} p_{T1}}{\bar{g} + k_{d1}}, \tag{D1}$$

$$\bar{c}_2 = \frac{\bar{m} p_{T2}}{\bar{m} + k_{d2}}, \tag{D2}$$

$$\psi_{\bar{c}_1} = \frac{R_1(\bar{g})}{1 + R_1(\bar{g})} \bar{\psi}_u + \sqrt{\frac{\bar{g} R_1(\bar{g})}{1 + R_1(\bar{g})}} N_1(0, 1), \tag{D3}$$

$$\psi_{\bar{c}_2} = \frac{R_2(\bar{m})}{1 + R_2(\bar{m})} \bar{\psi}_w + \sqrt{\frac{\bar{m} R_2(\bar{m})}{1 + R_2(\bar{m})}} N_2(0, 1). \tag{D4}$$

46

where $R_1(\bar{g}) = \frac{p_{T2}/k_{d2}}{(\bar{m}/k_{d2}+1)^2}$ and $R_2(\bar{m}) = \frac{p_{T1}/k_{d1}}{(\bar{g}/k_{d1}+1)^2}$ as defined in the main text.

First considering the deterministic dynamics for protein $\bar{m}$, we have that $\bar{m} = \bar{w} + \bar{c}_2$, which yields

$$\frac{d\bar{m}}{dt} = \frac{d\bar{w}}{dt} + \frac{d\bar{c}_2}{dt} = \frac{d\bar{w}}{dt} + \frac{d\bar{c}_2}{d\bar{m}}\frac{d\bar{m}}{dt}.$$

Then simplifying further and using the equation (D1) and the reduced dynamics for $\bar{w}$ given in (38) we obtain

$$\frac{d\bar{m}}{dt} = \left(\frac{1}{1 + \frac{d\bar{c}_2}{d\bar{m}}}\right)\frac{d\bar{w}}{dt}$$

$$= \left(\frac{1}{1 + R_2(\bar{m})}\right)\left(\beta_1\frac{p_{T1}k_{d1}}{\bar{g} + k_{d1}} - \delta_1\bar{m}\right),$$

in which, we have $\frac{d\bar{c}_2}{d\bar{m}} = \frac{p_{T1}k_{d1}}{(\bar{g}+k_{d1})^2} = R_2(\bar{m})$. Similarly, we have that

$$\frac{d\bar{g}}{dt} = \left(\frac{1}{1 + \frac{d\bar{c}_1}{d\bar{g}}}\right)\frac{d\bar{u}}{dt}$$

$$= \left(\frac{1}{1 + R_1(\bar{g})}\right)\left(\beta_2\frac{\bar{m}p_{T2}}{\bar{m} + k_{d2}} - \delta_2\bar{g}\right).$$

Next, we compute the dynamics for the second moment of proteins M and G. Towards this end we first derive the expressions for these seconds moments in terms of the total protein concentrations $\bar{\psi}_w$ and $\bar{\psi}_u$. Using the fast variable approximations (D3)–(D4) and the fact the normal random variables $N_1(0,1)$ and $N_2(0,1)$ are independent of each other and of $\bar{\psi}_w$ and $\bar{\psi}_u$, we have

$$\mathbb{E}\left[\bar{\psi}_m^2\right] = \mathbb{E}\left[(\bar{\psi}_w - \bar{\psi}_{c2})^2\right]$$

$$= \left(\frac{1}{1 + R_2(\bar{m})}\right)^2\mathbb{E}\left[\bar{\psi}_w^2\right] + \frac{\bar{m}R_2(\bar{m})}{1 + R_2(\bar{m})} \tag{D5}$$

$$\mathbb{E}\left[\bar{\psi}_g^2\right] = \mathbb{E}\left[(\bar{\psi}_u - \bar{\psi}_{c1})^2\right]$$

$$= \left(\frac{1}{1 + R_1(\bar{g})}\right)^2\mathbb{E}\left[\bar{\psi}_u^2\right] + \frac{\bar{g}R_1(\bar{g})}{1 + R_1(\bar{g})} \tag{D6}$$

$$\mathbb{E}\left[\bar{\psi}_m\bar{\psi}_g\right] = \mathbb{E}\left[(\bar{\psi}_w - \bar{\psi}_{c2})(\bar{\psi}_u - \bar{\psi}_{c1})\right]$$

$$= \left(\frac{1}{1 + R_2(\bar{m})}\right)\left(\frac{1}{1 + R_1(\bar{g})}\right)\mathbb{E}\left[\bar{\psi}_w\bar{\psi}_u\right] \tag{D7}$$

Next we consider the dynamics of $\mathbb{E}\left[\bar{\psi}_m^2\right]$. Taking the time derivative of equation (D5) we have

$$\frac{d\mathbb{E}\left[\bar{\psi}_m^2\right]}{dt} = 2\left(\frac{1}{1 + R_2(\bar{m})}\right)\frac{d\left(\frac{1}{1+R_2(\bar{m})}\right)}{d\bar{m}}\frac{d\bar{m}}{dt}\mathbb{E}\left[\bar{\psi}_w^2\right]$$

$$+ \left(\frac{1}{1 + R_2(\bar{m})} \right)^2 \frac{d\mathbb{E}\left[\bar{\psi}_w^2\right]}{dt} + \frac{d\left(\frac{R_2(\bar{m})}{1 + R_2(\bar{m})} \right)}{d\bar{m}} \frac{d\bar{m}}{dt} \bar{m}$$

$$+ \frac{R_2(\bar{m})}{1 + R_2(\bar{m})} \frac{d\bar{m}}{dt},$$

which can be further simplified to

$$\frac{d\mathbb{E}\left[\bar{\psi}_m^2\right]}{dt} = 2 \left(\frac{1}{1 + R_2(\bar{m})} \right) \frac{\frac{dR_2(\bar{m})}{d\bar{m}}}{(1 + R_2(\bar{m}))^2} \frac{d\bar{m}}{dt} \mathbb{E}\left[\bar{\psi}_w^2\right]$$

$$+ \left(\frac{1}{1 + R_2(\bar{m})} \right)^2 \frac{d\mathbb{E}\left[\bar{\psi}_w^2\right]}{dt} - \frac{\frac{dR_2(\bar{m})}{d\bar{m}}}{(1 + R_2(\bar{m}))^2} \frac{d\bar{m}}{dt} \bar{m}$$

$$+ \frac{R_2(\bar{m})}{1 + R_2(\bar{m})} \frac{d\bar{m}}{dt}. \tag{D8}$$

In a similar manner, we derive the dynamics for $\mathbb{E}\left[\bar{\psi}_m^2\right]$ and $\mathbb{E}\left[\bar{\psi}_m^2\right]$, which yields

$$\frac{d\mathbb{E}\left[\bar{\psi}_g^2\right]}{dt} = 2 \left(\frac{1}{1 + R_1(\bar{g})} \right) \frac{\frac{dR_1(\bar{g})}{d\bar{g}}}{(1 + R_1(\bar{g}))^2} \frac{d\bar{g}}{dt} \mathbb{E}\left[\bar{\psi}_u^2\right]$$

$$+ \left(\frac{1}{1 + R_1(\bar{g})} \right)^2 \frac{d\mathbb{E}\left[\bar{\psi}_g^2\right]}{dt} - \frac{\frac{dR_1(\bar{g})}{d\bar{g}}}{(1 + R_1(\bar{g}))^2} \frac{d\bar{g}}{dt} \bar{g}$$

$$+ \frac{R_1(\bar{g})}{1 + R_1(\bar{g})} \frac{d\bar{g}}{dt}, \tag{D9}$$

$$\frac{d\mathbb{E}\left[\bar{\psi}_m \bar{\psi}_g\right]}{dt} = \frac{d\mathbb{E}\left[\bar{\psi}_w \bar{\psi}_u\right]}{dt} \frac{1}{1 + R_2(\bar{m})} \frac{1}{1 + R_1(\bar{g})}$$

$$+ \frac{1}{1 + R_2(\bar{m})} \frac{2R_1(\bar{g})}{(\bar{g} + k_{d1})(1 + R_1(\bar{g}))^2} \mathbb{E}\left[\bar{\psi}_w \bar{\psi}_u\right] \frac{d\bar{g}}{dt} \tag{D10}$$

$$+ \frac{1}{1 + R_1(\bar{g})} \frac{2R_2(\bar{m})}{(\bar{m} + k_{d2})(1 + R_2(\bar{m}))^2} \mathbb{E}\left[\bar{\psi}_w \bar{\psi}_u\right] \frac{d\bar{m}}{dt}. \tag{D11}$$

Next, deriving the moment dynamics for the total protein concentrations $\bar{w}$ and $\bar{u}$, we obtain

$$\frac{d\mathbb{E}\left[\bar{\psi}_w^2\right]}{dt} = \frac{-2\delta_1}{1 + R_2(\bar{m})} \mathbb{E}\left[\psi_w^2\right] - \frac{2\beta_1 R_1(\bar{g})}{1 + R_1(\bar{g})} \mathbb{E}\left[\bar{\psi}_w \psi_u\right]$$

$$+ \beta_1 \frac{p_{T1} k_{d1}}{\bar{g} + k_{d1}} + \delta_1 m, \tag{D12}$$

$$\frac{d\mathbb{E}\left[\bar{\psi}_u^2\right]}{dt} = \frac{-2\delta_2}{1 + R_1(\bar{g})} \mathbb{E}\left[\psi_u^2\right] + \frac{2\beta_2 R_1(\bar{g})}{1 + R_1(\bar{g})} \mathbb{E}\left[\bar{\psi}_w \psi_u\right]$$

$$+ \beta_2 (\frac{\bar{m} p_{T2}}{\bar{m} + k_{d2}}) + \delta_2 g, \tag{D13}$$

$$\frac{d\mathbb{E}\left[\bar{\psi}_w \bar{\psi}_u\right]}{dt} = - \left(\frac{\delta_1}{1 + R_2(\bar{m})} + \frac{\delta_2}{1 + R_1(\bar{g})} \right) \mathbb{E}\left[\psi_w \psi_u\right]$$

$$-\beta_1 \frac{R_1(\bar{g})}{1+R_1(\bar{g})}\mathbb{E}\big[\psi_u^2\big] + \beta_2 \frac{R_2(\bar{m})}{1+R_2(\bar{m})}\mathbb{E}\big[\psi_w^2\big]. \tag{D14}$$

where the deterministic variables $\bar{w}$ and $\bar{u}$ are represented in terms of $\bar{m}$ and $\bar{g}$ using the coordinate transforms $\bar{w} = \bar{m} - \bar{c}_2$ and $\bar{u} = \bar{g} - \bar{c}_1$.

Then, substituting the moment dynamics (D12)–(D14) in the equations (D8)–(D11) and simplifying further using the expressions (D5)–(D7), yields the reduced-order moment dynamics of protein M and protein G given in (44)–(48).

REFERENCES

[1] U. Alon, "Biological networks : Tinkerer as an engineer," Science **301**, 1866–1867 (2003).

[2] G. E. Briggs and J. B. S. Haldane, "A note on the kinetics of enzyme action," Biochemical journal **19**, 338 (1925).

[3] F. Heineken, H. Tsuchiya, and R. Aris, "On the mathematical status of the pseudo-steady state hypothesis of biochemical kinetics," Mathematical Biosciences **1**, 95–113 (1967).

[4] H. K. Khalil, *Nonlinear systems*, Vol. 3 (Prentice hall Upper Saddle River, 2002).

[5] G. A. Pavliotis and A. Stuart, *Multiscale methods: averaging and homogenization* (Springer Science & Business Media, 2008).

[6] A. N. Tikhonov, "Systems of differential equations containing small parameters in the derivatives," Matematicheskii Sbornik **31**, 575 – 586 (1952).

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

[8] B. Choi, G. A. Rempala, and J. K. Kim, "Beyond the michaelis-menten equation: Accurate and efficient estimation of enzyme kinetic parameters," Scientific reports **7**, 17018 (2017).

[9] C. W. Gardiner *et al.*, *Handbook of stochastic methods*, Vol. 4 (Springer Berlin, 1985).

[10] N. G. Van Kampen, "Stochastic processes in physics and chemistry," (1981).

[11] S. Peleš, B. Munsky, and M. Khammash, "Reduction and solution of the chemical master equation using time scale separation and finite state projection," The Journal of chemical physics **125**, 204104 (2006).

[12] E. A. Mastny, E. L. Haseltine, and J. B. Rawlings, "Two classes of quasi-steady-state

model reductions for stochastic kinetics," The Journal of chemical physics **127**, 094106 (2007).

[13]C. A. Gómez-Uribe, G. C. Verghese, and A. R. Tzafriri, "Enhanced identification and exploitation of time scales for model reduction in stochastic chemical kinetics," The Journal of chemical physics **129**, 244112 (2008).

[14]J. Janssen, "The elimination of fast variables in complex chemical reactions. ii. mesoscopic level (reducible case)," Journal of statistical physics **57**, 171–185 (1989).

[15]J. Janssen, "The elimination of fast variables in complex chemical reactions. iii. mesoscopic level (irreducible case)," Journal of Statistical Physics **57**, 187–198 (1989).

[16]T. Shibata, "Reducing the master equations for noisy chemical reactions," The Journal of chemical physics **119**, 6629–6634 (2003).

[17]R. Ghaemi and D. Del Vecchio, "Stochastic analysis of retroactivity in transcriptional networks through singular perturbation," in *American Control Conference (ACC), 2012* (IEEE, 2012) pp. 2731–2736.

[18]Y. Cao, D. T. Gillespie, and L. R. Petzold, "The slow-scale stochastic simulation algorithm," The Journal of chemical physics **122**, 014116 (2005).

[19]C. V. Rao and A. P. Arkin, "Stochastic chemical kinetics and the quasi-steady-state assumption: application to the gillespie algorithm," The Journal of chemical physics **118**, 4999–5010 (2003).

[20]J. K. Kim, K. Josić, and M. R. Bennett, "The validity of quasi-steady-state approximations in discrete stochastic simulations," Biophysical journal **107**, 783–793 (2014).

[21]X. Kan, C. H. Lee, and H. G. Othmer, "A multi-time-scale analysis of chemical reaction networks: Ii. stochastic systems," Journal of mathematical biology **73**, 1081–1129 (2016).

[22]J. K. Kim, K. Josić, and M. R. Bennett, "The relationship between stochastic and deterministic quasi-steady state approximations," BMC systems biology **9**, 87 (2015).

[23]M. J. Lawson, L. Petzold, and A. Hellander, "Accuracy of the michaelis–menten approximation when analysing effects of molecular noise," Journal of The Royal Society Interface **12**, 20150054 (2015).

[24]D. Schnoerr, G. Sanguinetti, and R. Grima, "Approximation and inference methods for stochastic biochemical kinetics?a tutorial review," Journal of Physics A: Mathematical and Theoretical **50**, 093001 (2017).

[25]J. K. Kim and E. D. Sontag, "Reduction of multiscale stochastic biochemical reaction

networks using exact moment derivation," PLOS Computational Biology **13**, e1005571 (2017).

[26] K. R. Sanft, D. T. Gillespie, and L. R. Petzold, "Legitimacy of the stochastic michaelis–menten approximation," IET systems biology **5**, 58–69 (2011).

[27] P. Thomas, A. V. Straube, and R. Grima, "The slow-scale linear noise approximation: an accurate, reduced stochastic description of biochemical networks under timescale separation conditions," BMC systems biology **6**, 39 (2012).

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

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

[30] N. Herath, A. Hamadeh, and D. Del Vecchio, "Model reduction for a class of singularly perturbed differential equations," in *Proc. of American Control Conference (ACC)* (2015) pp. 4404 – 4410.

[31] N. Herath and D. Del Vecchio, "Moment convergence in a class of singularly perturbed stochastic differential equations," in *Australian Control Conference (AUCC)* (2015) pp. 43–47.

[32] N. Herath and D. Del Vecchio, "Model reduction for a class of singularly perturbed stochastic differential equations : Fast variable approximation," in *Proc. of American Control Conference (ACC)* (2016) pp. 3674 – 3679.

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

[34] F. Wu, T. Tian, J. B. Rawlings, and G. Yin, "Approximate method for stochastic chemical kinetics with two-time scales by chemical langevin equations," The Journal of chemical physics **144**, 174112 (2016).

[35] J. Elf and M. Ehrenberg, "Fast evaluation of fluctuations in biochemical networks with the linear noise approximation," Genome research **13**, 2475–2484 (2003).

[36] P. Thomas, R. Grima, and A. V. Straube, "Rigorous elimination of fast stochastic variables from the linear noise approximation using projection operators," Physical Review E **86**, 041110 (2012).

[37] C. D. Pahlajani, P. J. Atzberger, and M. Khammash, "Stochastic reduction method for biological chemical kinetics using time-scale separation," Journal of theoretical biology

272, 96–112 (2011).

[38]A. Sootla and J. Anderson, "Structured projection-based model reduction with application to stochastic biochemical networks," arXiv preprint arXiv:1510.05784 (2015).

[39]D. T. Gillespie, "Stochastic simulation of chemical kinetics," Annu. Rev. Phys. Chem. **58**, 35–55 (2007).

[40]D. T. Gillespie, "Deterministic limit of stochastic chemical kinetics," The Journal of Physical Chemistry B **113**, 1640–1644 (2009).

[41]S. Jayanthi and D. Del Vecchio, "Retroactivity attenuation in bio-molecular systems based on timescale separation," Automatic Control, IEEE Transactions on **56**, 748–761 (2011).

[42]A. Kumar and K. Josić, "Reduced models of networks of coupled enzymatic reactions," Journal of theoretical biology **278**, 87–106 (2011).

[43]D. Barik, M. R. Paul, W. T. Baumann, Y. Cao, and J. J. Tyson, "Stochastic simulation of enzyme-catalyzed reactions with disparate timescales," Biophysical journal **95**, 3563–3574 (2008).

[44]J. A. Borghans, R. J. De Boer, and L. A. Segel, "Extending the quasi-steady state approximation by changing variables," Bulletin of mathematical biology **58**, 43–63 (1996).

[45]A. Ciliberto, F. Capuani, and J. J. Tyson, "Modeling networks of coupled enzymatic reactions using the total quasi-steady state approximation," PLoS computational biology **3**, e45 (2007).

[46]S. MacNamara, A. M. Bersani, K. Burrage, and R. B. Sidje, "Stochastic chemical kinetics and the total quasi-steady-state assumption: application to the stochastic simulation algorithm and chemical master equation," The Journal of chemical physics **129**, 09B605 (2008).

[47]J. E. Marsden and M. J. Hoffman, *Elementary classical analysis* (Macmillan, 1993).

[48]M. R. Bennett, D. Volfson, L. Tsimring, and J. Hasty, "Transient dynamics of genetic regulatory networks," Biophysical journal **92**, 3501–3512 (2007).

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

[50]A. Agarwal, R. Adams, G. C. Castellani, and H. Z. Shouval, "On the precision of quasi steady state assumptions in stochastic dynamics," The Journal of chemical physics **137**, 044105 (2012).

[51]D. Del Vecchio, A. J. Ninfa, and E. D. Sontag, "Modular cell biology: retroactivity and

insulation," Molecular systems biology **4** (2008).

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

[53]K. S. Nilgiriwala, J. Jimnez, P. M. Rivera, and D. Del Vecchio, "Synthetic tunable amplifying buffer circuit in e. coli," ACS synthetic biology **4**, 577–584 (2014).

[54]D. Mishra, P. M. Rivera, A. Lin, D. Del Vecchio, and R. Weiss, "A load driver device for engineering modularity in biological networks," Nature biotechnology **32**, 1268–1275 (2014).

[55]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**, 164109 (2010).

[56]A. J. Laub, *Matrix analysis for scientists and engineers* (Siam, 2005).