Approximation of the Chemical Master Equation using conditional moment closure and time-scale separation

Approximation of the Chemical Master Equation using conditional moment closure and time-scale separation*

Ukjin Kwon[1], Mohammad Naghnaeian[2] and Domitilla Del Vecchio[3]

***Abstract*— To describe the stochastic behavior of biomolecular systems, the Chemical Master Equation (CME) is widely used. The CME gives a complete description of the evolution of a system's probability distribution. However, in general, the CME's dimension is very large or even infinite, so analytical solutions may be difficult to write and analyze. To handle this problem, based on the fact that biomolecular systems are time-scale separable, we approximate the CME with another CME that describes the dynamics of the slow species only. In particular, we assume that the number of each molecular species is bounded, although it may be very large. We thus write Ordinary Differential Equations (ODEs) of the slow-species counts' marginal probability distribution and of the fast-species counts' first n conditional moments. Here, n is an arbitrary (possibly small) number, which can be chosen to compromise between approximation accuracy and the computational burden associated with simulating or analyzing a high dimensional system. Then we apply conditional moment closure and time-scale separation to approximate the first n conditional moments of the fast-species counts as functions of the slow-species counts. By substituting these functions on the right-hand side of the ODEs that describes the marginal probability distribution of the slow-species counts, we can approximate the original CME with a lower dimensional CME. We illustrate the application of this method on an enzymatic and a protein binding reaction.**

I. INTRODUCTION

To analyze and predict the behavior of biomolecular systems, deterministic or stochastic approaches can be used [1]. Deterministic models fail to capture the inherent randomness of biomolecular systems, so stochastic approaches are often necessary. The Chemical Master Equation (CME) gives a complete description of the evolution of a system's probability distribution [2]. However, when the number of molecular counts is large or unbounded, the dimension of the CME is large or countably infinite. As a consequence, analytical or computational solutions are difficult to obtain in general.

When the dimension of the CME is infinite, one can, for example, use the Finite State Projection (FSP) algorithm [3] to truncate the system and find an upper bound to the molecular count of each species, so that the truncated finite dimensional system is arbitrarily close to the original infinite dimensional CME. When the dimension of the CME is finite but very large, [4] [5] [6] approximated the CME with another CME that describes the dynamics of the slow species only, based on singular perturbation theory in [7] and the fact that the biomolecular systems are usually time-scale separable [2] [5]. To achieve this, the stationary conditional probability distributions of the fast-species counts are approximated as functions of the slow-species counts. It was shown that the fast-species counts' conditional distributions converge exponentially fast with their stationary value as time-scale separation becomes more pronounced. This, consequentially, helps with obtaining a CME solely for the slow-species counts. However, the size of these stationary distributions grow exponentially in the number of the fast-species counts. Furthermore, as we will also show, the first few conditional moments of the fast-species counts may be sufficient, as opposed to the conditional distributions, to obtain a CME for the slow species only, which is a good approximation of the original CME. This is the approach that we adopt in this paper.

In this paper, we assume that the number of each molecular count is bounded and consider only first n conditional moments of the fast-species counts. Here, n is an arbitrary number that can be chosen by users from 1 to the bound on the molecular count. For larger n, accuracy increases but the computation burden also increases. Therefore, there is a trade off between accuracy and computation. We quantitatively derive the accuracy of the approximation as a function of n and of the time-scale separation.

Specifically, first, based on the CME, we derive ODEs for the marginal probability distribution of the slow-species counts and for the first n conditional moments of the fast-species counts. In this case, the ODEs are not closed; that is the first n conditional moments depend on the higher-order conditional moments. Therefore, we apply the robust moment closure technique developed in [8] to approximate the higher-order conditional moments as an affine function of the first n conditional moments. Next, we apply singular perturbation theory as in [7] and approximate the first n conditional moments of the fast-species counts as functions of the slow-species counts. Then, for the ODEs of the slow-species counts' marginal probability distribution, we substitute the conditional moments of the fast-species counts as the functions of the slow-species counts, hence obtaining another CME for the slow species only. Finally, we solve a linear program to ensure that the solution of the approximated CME is a proper probability vector. To show the utility of this method, we consider an enzymatic and a protein binding reaction.

*This work was supported by Korea Foundation for Advanced Studies (KFAS) and the Air Force Office of Scientific Research under grant FA9550-14-1-0060

[1]U. Kwon is with the Electrical Engineering and Computer Science Department, Massachusetts Institute of Technology, Cambridge, MA, USA ujkwon@mit.edu

[2]M. Naghnaeian and D. Del Vecchio are with the Mechanical Engineering Department, Massachusetts Institute of Technology, Cambridge, MA, USA mongh@mit.edu, ddv@mit.edu

Manuscript 647 submitted to 2019 American Control Conference.
Received September 24, 2018.

II. PRELIMINARIES

The following notations are used throughout this paper: $\mathbb{R}_{\geq 0}$ and $\mathbb{Z}_{\geq 0}$ are the set of nonnegative real numbers and integers, respectively. For any positive integer n, $\mathbb{R}^n_{\geq 0}$ ($\mathbb{Z}^n_{\geq 0}$) stands for the set of n-dimensional vectors with each entry in $\mathbb{R}_{\geq 0}$ ($\mathbb{Z}_{\geq 0}$). Given an n-dimensional vector $Z = [z_1, z_2, \ldots, z_n]^T$ and a nonnegative integer w, we define $\Psi_w(Z)$ to be the vector composed of entries of the form ${z_1}^{k_1}{z_2}^{k_2}\ldots{z_n}^{k_n}$ where $k_i \in \mathbb{Z}_{\geq 0}$, for $i = 1, 2, \ldots, n$, and $\sum_{i=1}^n k_i = w$. For example, when $Z = [z_1, z_2, z_3]$,

$$\Psi_1(Z) = [z_1, z_2, z_3]^T,$$
$$\Psi_2(Z) = [{z_1}^2, z_1z_2, z_1z_3, {z_2}^2, z_2z_3, {z_3}^2]^T.$$

The l_1 and l_∞ norms of a vector $Z = [z_1, z_2, \ldots, z_n]^T$ are defined as $\|Z\|_1 = \sum_{i=1}^n |z_i|$ and $\|Z\|_\infty = \max_i |z_i|$. We omit the subscript ∞ and simply write $\|Z\|$ for the l_∞ norm. A vector $P \in \mathbb{R}^p_{\geq 0}$ is defined as a probability vector when $\|P\|_1 = 1$. The l_1 to l_∞ induced norm of matrix M is defined as $\|M\|_{l_1-l_\infty} = \max_{i,j} |m_{ij}|$. The l_∞ induced norm of matrix M is defined as

$$\|M\| = \max_i \textstyle\sum_{j=1}^n |m_{ij}|.$$

Given a matrix $M = [m_{ij}] \in \mathbb{R}^{m\times n}$, by $\mathcal{R}[M]_i$ we mean the i^{th} row of M. That is,

$$\mathcal{R}[M]_i = \begin{bmatrix} m_{i1} & m_{i2} & \ldots & m_{in} \end{bmatrix},$$

for $i = 1, 2, \ldots, m$.

The dynamics of chemical reaction networks can be described by Markov processes. Each state of this Markov process represents the accumulated molecule counts of the species. When a chemical reaction fires, a transition from one state to another state occurs and the molecule counts of the species change. To be more specific, suppose there is a reaction network with r number of species and K number of reactions. Let s_i, for $i = 1, 2, \ldots, r$, be the number of each species and $S = [s_1, s_2, \ldots, s_r]^T$. Associated with each reaction $k \in 1, 2, \ldots, K$, there is a corresponding propensity function $a_k(S)$ and a stoichiometry vector γ_k [2]. The propensity function typically assumes that it does not depend on time [2]. For any $q \in \mathbb{Z}^r_{\geq 0}$, we assume that

$$\begin{aligned}\frac{d}{dt}P(S(t) = q) = \sum_{k=1}^{K}[-a_k(q)P(S(t) = q) \\ + a_k(q - \gamma_k)P(S(t) = q - \gamma_k)]\end{aligned} \tag{1}$$

is satisfied. This equation is called the CME [9] [10]. Suppose that the set of reactions can be devided into two subsets, fast and slow reactions. The distinction between them is in their propensity functions. The propensity function of a fast reaction is of order $\frac{1}{\epsilon}$ of that of a slow reaction, where ϵ is a positive number much smaller than 1 quantifying the separation of time scales between the fast and slow reactions. Let K_f and K_s be the number of the fast and slow reactions, respectively, that satisfies $K_f + K_s = K$. Furthermore, suppose that upon firing the fast reactions, the species count of a proper subset of the set of all species changes. Denote this proper subset by $\{Y_j\}_{j=1}^m$. These are referred to as the fast species. The rest of the species are called the slow species and form the set $\{X_i\}_{i=1}^l$, where $l = r - m$. Then the CME can be written as

$$\begin{aligned}&\frac{d}{dt}P(X(t) = x, Y(t) = y) \\ &= \sum_{k=1}^{K_s}[-a^s_k(x; y)P(X(t) = x, Y(t) = y) \\ &+ a^s_k(x - \gamma^s_{x,k}; y - \gamma^s_{y,k})P(X(t) = x - \gamma^s_{x,k}, Y(t) = y - \gamma^s_{y,k})] \\ &+ \sum_{k=1}^{K_f}[-a^f_k(x; y)P(X(t) = x, Y(t) = y) \\ &+ a^f_k(x; y - \gamma^f_{y,k})P(X(t) = x, Y(t) = y - \gamma^f_{y,k})]\end{aligned} \tag{2}$$

for $x \in \mathbb{Z}^l_{\geq 0}$ and $y \in \mathbb{Z}^m_{\geq 0}$, where $a^s_k(x; y)$ and $a^f_k(x; y)$ are propensity functions for the slow and fast reactions, respectively, and $\gamma^s_{x,k}$, $\gamma^s_{y,k}$ and $\gamma^f_{y,k}$ are corresponding stoichiometry vectors [5]. Throughout this paper, we make the following assumptions.

Assumption 2.1: There exist nonnegative integers x^i_{tot} and y^j_{tot} such that

$$0 \leq x_i \leq x^i_{tot},\ 0 \leq y_j \leq y^j_{tot},$$

for $i = 1, 2, \ldots, l$ and for $j = 1, 2, \ldots, m$, where x_i and y_j are number of X_i and Y_j, respectively.

Assumption 2.2 ([10] [11]): The propensity functions are polynomial in S. In addition, the order of each polynomial is less than or equal to 2 .

Assumption 2.3: For the slow reactions, each propensity function can be written as $a^s_k(x; y) = (\theta^{k,s}_0(x) + \theta^{k,s}_1(x)\Psi_1(Y) + \theta^{k,s}_2(x)\Psi_2(Y))$, for $k \in 1, 2, \ldots, K_s$, where $\theta^{k,s}_i(x)$, for $i = 0, 1, 2$ and given x, are matrices with appropriate dimensions.

Assumption 2.4: For the fast reactions, each propensity function can be written as $a^f_k(x; y) = \frac{1}{\epsilon}(\theta^{k,f}_0(x) + \theta^{k,f}_1(x)\Psi_1(Y) + \theta^{k,f}_2(x)\Psi_2(Y))$, for $k \in 1, 2, \ldots, K_f$, where $\theta^{k,s}_i(x)$, for $i = 0, 1, 2$ and given x, are matrices with appropriate dimensions.

Assumption 2.1 states that there exists an upper bound on the number of each species. This is a reasonable assumption in a number of cases. For example, enzymatic reactions do not involve creation and destruction of species, and therefore they are characterized by a bounded total amount of enzymes and substrates. In general, it is still reasonable to assume that the number of fast species y_j are bounded. In fact, in the case of gene regulatory network models, for example, these are often complexes formed by transcription factors with DNA, which is available in a finite amount. In general, the upper bounds may not exist for the slow species. In this case, however, one can use the truncation method given in [3]. Regarding Assumption 2.2, the fact that the propensity functions are polynomial in S is standard and satisfied when the species are well-mixed [10] [11]. The fact that the order of each polynomial is at most two because reactions are either uni-molecular or bi-molecular, which is also a standard assumption since n-molecular reactions with $n > 2$ are considered less probable [2]. Assumption 2.3 and 2.4 are

Manuscript 647 submitted to 2019 American Control Conference.
Received September 24, 2018.

based on Assumption 2.2 that each propensity function is polynomial in X and Y with the order less than or equal to 2. The propensity functions of the fast reactions are order of $\frac{1}{\epsilon}$ of the propensity functions of the slow reactions.

III. Basic Setup

The CME given in (1) with Assumptions 2.1 is a linear system of ODEs with order p, where

$$p = \prod_{i=1}^{l}(x_{tot}^i + 1)\prod_{j=1}^{m}(y_{tot}^j + 1).$$

We define

$$x_{tot} = \prod_{i=1}^{l}(x_{tot}^i + 1), y_{tot} = \prod_{j=1}^{m}(y_{tot}^j + 1).$$

The order of the CME exponentially increases as the number of the fast or slow species increases. Therefore, directly solving the CME is a computationally challenging task. To avoid this computational difficulty and obtain mathematical descriptions suitable for analytical study, one can apply singular perturbation theory [7] to approximate the CME with another CME that describes the dynamics of the slow species only. In this paper, we try to substitute first n conditional moments of the fast-species counts as functions of the slow-species counts to acheive the approximation. To proceed, we have to define some notations. We define

$$\begin{aligned}\Omega_x = \{x \mid x = [x_1, x_2, \ldots, x_l]^T, 0 \le x_i \le x_{tot}^i,\\ \text{for } i = 1, 2, \ldots, l\},\\ \Omega_y = \{y \mid y = [y_1, y_2, \ldots, y_m]^T, 0 \le y_j \le y_{tot}^j,\\ \text{for } j = 1, 2, \ldots, m\},\end{aligned}$$

and X and Y are vectors of random variables taking values in the sets Ω_x and Ω_y, respectively. Let $\{\bar{x}_i\}_{i=1}^{|\Omega_x|}$ be an enumeration of Ω_x. The marginal probability distribution of the slow-species counts is defined as

$$\begin{aligned}P_X(t) = [P(X = [0, 0, \ldots, 0]^T), \ldots, P(X = \bar{x}_i), \ldots,\\ P(X = [x_{tot}^1, \ldots, x_{tot}^l]^T)]^T.\end{aligned}$$

For any $w \in \mathbb{Z}_{\ge 0}$, $1 \le n \le y_{tot}$ and $x \in \Omega_x$, we define

$$\begin{aligned}&\mu_w(x,t)\\ &= E[\Psi_w(Y)|X = x] = \sum_{y\in\mathbb{Z}^m}\Psi_w(y)P(Y = y|X = x),\\ &Y_n(x,t) = [\mu_1(x,t)^T, \mu_2(x,t)^T, \ldots, \mu_n(x,t)^T]^T,\end{aligned} \tag{3}$$

where $\mu_w(x,t)$ and $Y_n(x,t)$ denote fast-species counts' w^{th} and first n conditional moments, respectively. For $i = 1, 2, \ldots, n$, let d_i be a matrix whose multiplication with $Y_n(x,t)$ isolates $\mu_i(x,t)$, i.e.

$$\mu_i(x,t) = d_i Y_n(x,t). \tag{4}$$

Now we can derive ODEs for the marginal probability distribution of the slow-species counts and for first n conditional moments of the fast-species counts as in (5):

Proposition 3.1: For the CME in (2) with Assumptions 2.1 to 2.4, for $1 \le n \le y_{tot}$ and $\bar{x}_i \in \Omega_x$, we can obtain

$$\Sigma_{true}: \begin{cases} \frac{d}{dt}P(X = \bar{x}_i) = \sum_{k=1}^{K_s}(-E[a_k^s(\bar{x}_i; y)|X = \bar{x}_i]P(X = \bar{x}_i)\\ +E[a_k^s(\bar{x}_i - \gamma_{x,k}^s; y)|X = \bar{x}_i - \gamma_{x,k}^s]P(X = \bar{x}_i - \gamma_{x,k}^s))\\ \epsilon\frac{d}{dt}Y_n(x,t) = C(x)Y_n(x,t) + c_1(x)\\ +c_2\mu_{n+1}(x,t) + \epsilon G(t).\end{cases} \tag{5}$$

Proof: ODEs of the slow-species counts' marginal probability distribution are derived in [5]. ODEs of the fast-species counts' conditional probability distribution are derived in [5] as

$$\begin{aligned}&\epsilon\frac{d}{dt}P(Y = y|X = x) = \sum_{k=1}^{K_f}(-\epsilon a_k^f(x; y)P(Y = y|X = x)\\ &+ \epsilon a_k^f(x; y - \gamma_{y,k}^f)P(Y = y - \gamma_{y,k}^f|X = x) + \epsilon G_1(t)\end{aligned} \tag{6}$$

where $G_1(t)$ is bounded. Therefore, from (6), we can derive

$$\begin{aligned}&\epsilon\frac{d}{dt}\mu_w(x,t) = \sum_{y\in\mathbb{Z}^m}\sum_{k=1}^{K_f}[(\Psi_w(y + \gamma_{y,k}^f) - \Psi_w(y))\\ &\epsilon a_k^f(x; y)P(Y = y|X = x)] + \epsilon G_2(t)\\ &= \sum_{y\in\mathbb{Z}^m}\sum_{k=1}^{K_f}[(\Psi_w(y + \gamma_{y,k}^f) - \Psi_w(y))(\theta_0^{k,f}(x)\\ &+ \theta_1^{k,f}(x)\Psi_1(Y) + \theta_2^{k,f}\Psi_2(Y))] + \epsilon G_2(t)\end{aligned} \tag{7}$$

for $1 \le w \le n$. Order of $\Psi_w(y + \gamma_{y,k}) - \Psi_w(y)$ is $w - 1$, so order of the right-hand side of (7) returns at most $(w+1)^{th}$ conditional moments. When we consider w from 1 to n, we can obtain ODEs of the conditional moments of the fast-species counts in Σ_{true}. ■

Given $\bar{x}_i$, conditional expectation of propensity function $a_k^s(\bar{x}_i, y)$ can be expressed as

$$\begin{aligned}&E[a_k^s(\bar{x}_i; y)|X = \bar{x}_i] = \sum_y a_k^s(\bar{x}_i; y)P(Y = y|X = \bar{x}_i)\\ &= \theta_0^{k,s}(\bar{x}_i) + \theta_1^{k,s}(\bar{x}_i)\mu_1(\bar{x}_i, t) + \theta_2^{k,s}(\bar{x}_i)\mu_2(\bar{x}_i, t).\end{aligned} \tag{8}$$

According to [3], we can write ODEs of the slow-species counts' marginal probability distribution in (5) as a single linear expression:

$$\frac{d}{dt}P_X(t) = A(Y_2(x,t))P_X(t), \tag{9}$$

where, for $1 \le i, j \le x_{tot}$,

$$A(Y_2(x,t))_{ij}: \begin{cases} -\sum_{k=1}^{K_s}E[a_k^s(\bar{x}_j; y)|X = \bar{x}_j] \text{ for } i = j\\ E[a_k^s(\bar{x}_j; y)|X = \bar{x}_j] \text{ for all } j \text{ such that}\\ \bar{x}_j = \bar{x}_i - \gamma_{x,k}^s\\ 0 \text{ Otherwise.}\end{cases} \tag{10}$$

In Σ_{true}, when $n = y_{tot}$, the dimension of Σ_{true} is $x_{tot} + y_{tot}$, and it is closed. This is because $\mu_{n+1}(x,t)$ can be represented as an affine function of $Y_n(x,t)$ [5]. However, in general, when $1 \le n < y_{tot}$, the dynamics of the fast-species counts' conditional moments are not closed, because $\mu_{n+1}(x,t)$ is not a function of $Y_n(x,t)$ anymore. Therefore, a robust conditional moment closure method should be

Manuscript 647 submitted to 2019 American Control Conference.
Received September 24, 2018.

applied to approximate $\mu_{n+1}(x,t)$ as a function of $Y_n(x,t)$ to close the dynamics. The next section introduces the robust moment closure technique derived in [8] and we adapt it to the conditional moment case.

IV. Robust Conditional Moment Closure

We are applying Robust Moment Closure (RMC), which was originally developed in [8], to the dynamics of conditional moments. For any $x \in \Omega_x$, we define matrices H_n and V_n such that

$$\mu_{n+1}(x,t) = H_n P_{Y|X}(x,t),\ Y_n(x,t) = V_n P_{Y|X}(x,t),$$

where

$$P_{Y|X}(x,t) = [P(Y=[0,0,\dots,0]^T|X=x),\dots,P(Y=[y_{tot}^1,y_{tot}^2,\dots,y_{tot}^m]^T|X=x)]^T$$

is a conditional probability distribution of the fast-species counts. For example, when $l=m=1$,

$$H_n = \begin{bmatrix} 0 & 1^{n+1} & 2^{n+1} & \dots & (y_{tot}^1)^{n+1} \end{bmatrix}, \tag{11}$$

$$V_n = \begin{bmatrix} 0 & 1 & 2 & \dots & y_{tot}^1 \\ 0 & 1^2 & 2^2 & \dots & (y_{tot}^1)^2 \\ \vdots & \vdots & & & \\ 0 & 1^n & 2^n & \dots & (y_{tot}^1)^n \end{bmatrix}. \tag{12}$$

Our goal is to approximate $\mu_{n+1}(x,t)$ as a function of $Y_n(x,t)$, denoted as

$$\mu_{n+1}(x,t) \approx \phi(Y_n(x,t)),$$

possibly a nonlinear function. According to [8], without a priori information on the probability distribution, the optimal function $\phi(Y_n(x,t))$ that minimizes worst case approximation error between $\mu_{n+1}(x,t)$ and $\phi(Y_n(x,t))$, which can be written as

$$\sup_{P_{Y|X}(x,t)\in\mathbb{P}} \|\mu_{n+1}(x,t) - \phi(Y_n(x,t))\|, \tag{13}$$

is an affine function of $Y_n(x,t)$, which is

$$\phi(Y_n(x,t)) = KY_n(x,t) + K_0.$$

In addition, K and K_0 can be obtained by solving the linear program

$$\begin{aligned} \min_{K_0,K} \quad & \gamma \\ \text{s.t.} \quad & -\gamma 1^T \le \mathcal{R}[H_n - (KV_n + K_0 1^T)]_i \le \gamma 1^T \end{aligned} \tag{14}$$

for $i = 1,2,\dots,m$, where m is the number of rows in H_n. Let the object value of the linear program in (14) be ρ_n, which is a fixed constant that depends on n. Then the approximation error between $\mu_{n+1}(x,t)$ and $\phi(Y_n(x,t))$, which can be written as

$$\|\mu_{n+1}(x,t) - \phi(Y_n(x,t))\| = \big\|H_n P_{Y|X}(x,t) - (KY_n(x,t)+K_0)\big\|,$$

is bounded by ρ_n for all $P_{Y|X}(x,t) \in \mathbb{P}$.

By substituting $\mu_{n+1}(x,t)$ in the right-hand side of Σ_{true} with $KY_n(x,t)+K_0$, we obtain

$$\Sigma_{closed}: \begin{cases} \frac{d}{dt}\tilde{P}_X(t) = A(\tilde{Y}_2^\epsilon(x,t))\tilde{P}_X(t) = \tilde{A}^\epsilon(t)\tilde{P}_X(t) \\ \epsilon\frac{d}{dt}\tilde{Y}_n^\epsilon(x,t) = C(x)\tilde{Y}_n^\epsilon(x,t) + c_1(x) \\ \quad +c_2(K\tilde{Y}_n^\epsilon(x,t)+K_0) + \epsilon G(t). \end{cases} \tag{15}$$

Σ_{closed} is closed and we define $\tilde{\mu}_i^\epsilon(x,t) = d_i\tilde{Y}_n^\epsilon(x,t)$.

Remark 4.1: According to the following Lemma 6.1, the approximation error between $\mu_i(x,t)$ and $\tilde{\mu}_i^\epsilon(x,t)$ is bounded if $C(x)+c_2K$ is a stable matrix, i.e., its eigenvalues have a negative real part. Although the stability of $C(x)+c_2K$ is not guaranteed via (14), we realized that in our examples this matrix is indeed stable. However, to truly enforce the stability, one can augment (14) with a linear matrix inequality and carry out an iterative algorithm. This procedure is in the Appendix. Here we assume that the iterative algorithm is already conducted and $C(x)+c_2K$ is a stable matrix.

V. Time-Scale Separation

We note that (15) is in standard singular perturbation form [7]. As $\epsilon \to 0$, $\tilde{Y}_n^\epsilon(x,t)$ converges exponentially fast to $\tilde{Y}_n^0(x,\infty)$, where $\tilde{Y}_n^0(x,\infty)$ satisfies

$$C(x)\tilde{Y}_n^0(x,\infty) + c_1(x) + c_2(K\tilde{Y}_n^0(x,\infty)+K_0) = 0. \tag{16}$$

This is proved in Lemma 6.2. Let us define $\tilde{\mu}_i^0(x,\infty) = d_i\tilde{Y}_n^0(x,\infty)$. When we substitute $\tilde{Y}_2^\epsilon(x,t)$ with $\tilde{Y}_2^0(x,\infty)$ in the right-hand side of (15), we can obtain

$$\Sigma_{reduced}: \left\{ \ \frac{d}{dt}\bar{P}_X(t) = A(\tilde{Y}_2^0(x,\infty))\bar{P}_X(t) = \bar{A}\bar{P}_X(t), \right. \tag{17}$$

where $\Sigma_{reduced}$ describes the dynamics of the slow species only. In $\Sigma_{reduced}$, $\bar{P}_X(t)$ is a valid probability distribution if and only if $\bar{A}$ is a Metzler matrix [12], which is not guaranteed in general. $\bar{A}$ is a Metzler matrix if and only if $\tilde{a}^{k,s}(x)$, for $k=1,\dots,K_s$, which is defined as

$$\tilde{a}^{k,s}(x) = \theta_0^{k,s}(x) + \theta_1^{k,s}(x)\tilde{\mu}_1^0(x,\infty) + \theta_2^{k,s}(x)\tilde{\mu}_2^0(x,\infty), \tag{18}$$

is non-negative for all $x \in \Omega_x$. We define a linear program

$$\begin{aligned} \min_{h_1(x),h_2(x)} \quad & \big\|h_1(x) - \tilde{\mu}_1^0(x,\infty)\big\| + \big\|h_2(x) - \tilde{\mu}_2^0(x,\infty)\big\| \\ \text{s.t.} \quad & \theta_0^{k,s}(x) + \theta_1^{k,s}(x)h_1(x) + \theta_2^{k,s}(x)h_2(x) \ge 0 \end{aligned} \tag{19}$$

for $k=1,2,\dots,K_s$. Let the object value of (19) be λ^x and the optimal solutions be $h_1(x) = \hat{\mu}_1(x)$ and $h_2(x) = \hat{\mu}_2(x)$. When we substitute $\tilde{\mu}_1^0(x,\infty)$ and $\tilde{\mu}_2^0(x,\infty)$ as $\hat{\mu}_1(x)$ and $\hat{\mu}_2(x)$ in $\Sigma_{reduced}$, we obtain

$$\Sigma_{final}: \left\{ \ \frac{d}{dt}\hat{P}_X(t) = A(\hat{Y}_2(x))\hat{P}_X(t) = \hat{A}\hat{P}_X(t), \right. \tag{20}$$

where $\hat{Y}_2(x) = [\hat{\mu}_1(x)^T, \hat{\mu}_2(x)^T]^T$. In Σ_{final}, $\hat{P}_X(t)$ is a valid probability distribution because $\hat{A}$ is guaranteed to be a Metzler matrix by (19). In addition, $\big\|\hat{\mu}_1(x) - \tilde{\mu}_1^0(x,\infty)\big\|$ and $\big\|\hat{\mu}_2(x) - \tilde{\mu}_2^0(x,\infty)\big\|$ are bounded by λ^x and $\hat{A}$ is a marginally stable matrix with one zero eigenvalue.

Remark 5.1: When $n = y_{tot}$, both ρ_n and λ^x are 0. ρ_n is 0 because $\mu_{n+1}(x,t)$ can be represented as an affine function of $Y_n(x,t)$ in this case. The proof of $\lambda^x = 0$ is in [5].

Manuscript 647 submitted to 2019 American Control Conference.
Received September 24, 2018.

Now we need to quantify the approximation errors.

VI. Error Quantification

A. Conditional Moments of the Fast Species

The following lemmas are proved in the Appendix.

Lemma 6.1: Given $T > t_0 > 0$ and $x \in \Omega_x$, the approximation error between $\mu_i(x,t)$ and $\tilde{\mu}_i^\epsilon(x,t)$ satisfies

$$\sup_{t\in[t_0,T]} \|\mu_i(x,t) - \tilde{\mu}_i^\epsilon(x,t)\| \leq \int_{t_0}^T \left\|d_i \exp\left\{\tfrac{1}{\epsilon}(C(x)+c_2K)(T-\tau)\right\}\right\| d\tau \tfrac{\rho_n}{\epsilon}\|c_2\| = \Delta_{i,\epsilon}^x.$$

Lemma 6.2: Given $T > t_0 > 0$ and $x \in \Omega_x$, the approximation error between $\tilde{\mu}_i^\epsilon(x,t)$ and $\tilde{\mu}_i^0(x,\infty)$ satisfies

$$\sup_{t\in[t_0,T]} \left\|\tilde{\mu}_i^\epsilon(x,t) - \tilde{\mu}_i^0(x,\infty)\right\| \leq \int_{t_0}^T \left\|d_i \exp\left\{\tfrac{1}{\epsilon}(C(x)+c_2K)(T-\tau)\right\}\right\| \|G(\tau)\| d\tau = O(\epsilon).$$

Theorem 6.3: Given $T > t_0 > 0$ and $x \in \Omega_x$, for sufficiently small ϵ, the approximation error between $\mu_i(x,t)$ and $\hat{\mu}_i(x)$ satisfies

$$\sup_{t\in[t_0,T]} \|\mu_i(x,t) - \hat{\mu}_i(x)\| \leq \Delta_{i,\epsilon}^x + \lambda^x + O(\epsilon),$$

for $i = 1, 2$. Furthermore, there exist $\Delta_\epsilon > 0$ and $\epsilon^* > 0$ such that $\sup_{t\in[t_0,T]} \|\mu_i(x,t) - \hat{\mu}_i(x)\| \leq \Delta_\epsilon + O(\epsilon)$ for all x, $\epsilon \in (0, \epsilon^*)$ and $i = 1$ or 2.

Proof: The first inequality of Theorem 6.3 can be directly obtained by combining Lemmas 6.1 and 6.2, result of (19) and triangular inequality. For the second inequality, Δ_ϵ can be obtained by

$$\Delta_\epsilon = \sup_{i\in\{1,2\}, x\in\Omega_x} \Delta_{i,\epsilon}^x + \lambda^x.$$

■

B. Marginal Probability of the Slow Species

We constructed $\hat{P}_X(t)$ such that it is a valid probability distribution, which implies that $\left\|\hat{P}_X(t)\right\|_1 = 1$ for all t. Therefore, $\hat{P}_X(t)(1)$, the first component of $\hat{P}_X(t)$, can be written as a linear combination of other components of $\hat{P}_X(t)$ as

$$\hat{P}_X(t)(1) = 1 - (\hat{P}_X(t)(2) + \ldots + \hat{P}_X(t)(x_{tot})).$$

To remove this linearly-depenent relationship, we define

$$\hat{P}_{X,new}(t) = [\hat{P}_X(t)(2), \hat{P}_X(t)(3), \ldots, \hat{P}_X(t)(x_{tot})]^T,$$

and derive a new equation

$$\frac{d}{dt}\hat{P}_{X,new}(t) = \hat{A}_{new}\hat{P}_{X,new}(t) + \hat{a}_{new}, \tag{21}$$

from Σ_{final}. Here, $\hat{a}_{new}$ is second to x_{tot}^{th} elements of the first column of $\hat{A}$, and for $1 \leq j \leq x_{tot} - 1$, j^{th} column of $\hat{A}_{new}$ is second to x_{tot}^{th} elements of the $(j+1)^{th}$ column of $\hat{A}$ minus $\hat{a}_{new}$. This relationship can be written as

$$\hat{a}_{new} = \hat{A}(2 : x_{tot}, 1),$$
$$\hat{A}_{new}(:, j) = \hat{A}(2 : x_{tot}, j+1) - \hat{a}_{new}.$$

Since we remove the linearly dependent relationship, eigenvalues of $\hat{A}_{new}$ are exactly the same as those of $\hat{A}$ except the zero, so $\hat{A}_{new}$ is a stable matrix. We can repeat the same procedure for Σ_{true} in (9) and derive

$$\frac{d}{dt}P_{X,new}(t) = A_{new}(Y_2(x,t))P_{X,new}(t) + a_{new}(Y_2(x,t)). \tag{22}$$

Now, we regard (21) as the nominal system and (22) as the perturbed system. Then we can rewrite the perturbed system as

$$\frac{d}{dt}P_{X,new}(t) = (\hat{A}_{new} + \Delta_1(t))P_{X,new}(t) + (\hat{a}_{new} + \Delta_2(t)), \tag{23}$$

where

$$\Delta_1(t) = A_{new}(Y_2(x,t)) - \hat{A}_{new},$$
$$\Delta_2(t) = a_{new}(Y_2(x,t)) - \hat{a}_{new}.$$

Lemma 6.4: For sufficiently small ϵ, there are two constants k_1 and k_2 such that

$$\|\Delta_1(t)\|_{l_1-l_\infty} \leq k_1\Delta_\epsilon + O(\epsilon), \|\Delta_2(t)\| \leq k_2\Delta_\epsilon + O(\epsilon).$$

Proof: i^{th} component of $\Delta_2(t)$ is $\theta_1^{k,s}(\bar{x}_1)(\mu_1(\bar{x}_1,t) - \hat{\mu}_1(\bar{x}_1)) + \theta_2^{k,s}(\bar{x}_1)(\mu_2(\bar{x}_1,t) - \hat{\mu}_2(\bar{x}_1))$, which is bounded by $(\left\|\theta_1^{k,s}(\bar{x}_1)\right\| + \left\|\theta_2^{k,s}(\bar{x}_1)\right\|)\Delta_\epsilon + O(\epsilon)$. Therefore,

$$k_2 = \sup_k(\left\|\theta_1^{k,s}(\bar{x}_1)\right\| + \left\|\theta_2^{k,s}(\bar{x}_1)\right\|).$$

With the same procedure, k_1 can be obtained as

$$k_1 = \sup_{k,j}(\left\|\theta_1^{k,s}(\bar{x}_1)\right\| + \left\|\theta_2^{k,s}(\bar{x}_1)\right\| + \left\|\theta_1^{k,s}(\bar{x}_{j+1})\right\| + \left\|\theta_2^{k,s}(\bar{x}_{j+1})\right\|).$$

■

Theorem 6.5: Given $T > t_0 > 0$, the approximation error between $P_{X,new}(t)$ and $\hat{P}_{X,new}(t)$ satisfies

$$\sup_{t\in[t_0,T]} \left\|P_{X,new}(t) - \hat{P}_{X,new}(t)\right\| \leq (k_1+k_2)\int_{t_0}^T \left\|\exp\left\{\hat{A}_{new}(T-\tau)\right\}\right\| d\tau\Delta_\epsilon + O(\epsilon).$$

The proof of Theorem 6.5 is in the Appendix.

Corollary 6.6: As ϵ goes to 0, the right-hand side of the inequality in Theorem 6.5 goes to $k\Delta_0$, where

$$k = (k_1+k_2)\int_{t_0}^T \left\|\exp\left\{\hat{A}_{new}(T-\tau)\right\}\right\| d\tau,$$
$$\Delta_0 = \lim_{\epsilon\to 0}\Delta_\epsilon = \sup_{i\in\{1,2\}, x\in\Omega_x}(\Delta_{i,0}^x + \lambda^x) \text{ , and}$$
$$\Delta_{i,0}^x = \lim_{\epsilon\to 0}\Delta_{i,\epsilon}^x = \int_0^\infty \|d_i \exp\{(C(x)+c_2K)t\}\| dt\rho_n\|c_2\|.$$

Proof: When we substitue $\frac{T-\tau}{\epsilon}$ as t in Lemma 6.1,

$$\lim_{\epsilon\to 0}\Delta_{i,\epsilon}^x = \int_0^\infty \|d_i \exp\{(C(x)+c_2K)t\}\| dt\rho_n\|c_2\|.$$

■

Remark 6.7: When $n = y_{tot}$, as ϵ goes to 0, the right-hand side of the inequality in Theorem 6.5 goes to 0. This is because both ρ_n and λ^x go to 0 by Remark 5.1.

VII. Illustrative Example

In this section, we show the utility of our method with an enzymatic and a protein binding reaction.

A. Enzymatic Reaction

In this example, we consider an enzymatic reaction [2]:

$$E + X \underset{d}{\overset{a}{\rightleftharpoons}} C \xrightarrow{k} E + X^*. \tag{24}$$

In (24), X, E, C, and X^* are the substrate, the enzyme, the binding complex, and the reaction product, respectively. In addition, x, e, c, and x^* are the numbers of corresponding

Manuscript 647 submitted to 2019 American Control Conference.
Received September 24, 2018.

species. We assume that the total numbers of the substrate and enzyme are conserved, which means $x + c + x^* = x_{tot}$ and $e + c = e_{tot}$, for some positive constants x_{tot} and e_{tot}. Therefore, Assumption 2.1 is readily satisfied. In this reaction, ae_{tot} and d are much larger than k, so we can define $\epsilon = \frac{k}{d}$ and let $ae_{tot} = \frac{d}{2}$. When we define $X_1 = X + C$, and consider S = $[x_1, c]^T$, we can verify X_1 is a slow species and C is a fast species based on the following propensity functions and corresponding stoichiometries,

$$a_1^f(x_1; c) = \tfrac{1}{\epsilon}\tfrac{k}{2Ve_{tot}}(e_{tot} - c)(x_1 - c), \gamma_1^f = [0, +1]^T,$$
$$a_2^f(x_1; c) = \tfrac{1}{\epsilon}kc, \gamma_2^f = [0, -1]^T,$$
$$a_1^s(x_1; c) = kc, \gamma_1^s = [-1, -1]^T,$$

where V is the volume. We can derive ODEs of the slow-species counts' marginal probability density function and the fast-species counts' first 2 conditional moments as below:

$$\begin{aligned}
&\frac{d}{dt}P(X_1 = x_1) = -k\mu_1(x_1, t)P(X_1 = x_1) \\
&+ k\mu_1(x_1 + 1, t)P(X_1 = x_1 + 1) \\
&\epsilon\frac{d}{dt}Y_2(x_1, t) = \\
&\begin{bmatrix} -k - \frac{k(e_{tot}+x_1)}{2Ve_{tot}} & \frac{k}{2Ve_{tot}} \\ -\frac{k(e_{tot}+x_1)}{2Ve_{tot}} + \frac{kx_1}{V} + k & \frac{k}{2Ve_{tot}} - \frac{k(e_{tot}+x_1)}{Ve_{tot}} - 2k \end{bmatrix} Y_2(x_1, t) \\
&+ \begin{bmatrix} \frac{kx_1}{2V} \\ \frac{kx_1}{2V} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{k}{Ve_{tot}} \end{bmatrix} \mu_3(x_1, t) + \epsilon G(t) \\
&= C(x_1)Y_2(x_1, t) + c_1(x_1) + c_2\mu_3(x_1, t) + \epsilon G(t).
\end{aligned} \tag{25}$$

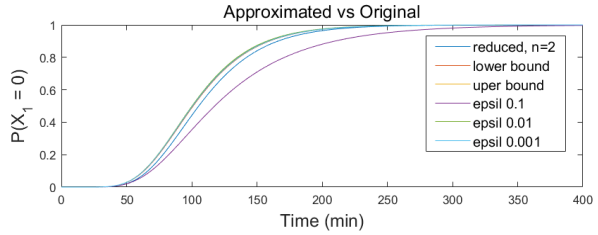
We can check that (25) is in Σ_{true} form. To close the dynamics, we let $e_{tot} = 5[molecules]$ and solve the linear program in (14) and obtain

$$\mu_3(x_1, t) \approx K_{32}\mu_2(x_1, t) + K_{31}\mu_1(x_1, t) + K_{30},$$
where $K_{32} = 7.5$, $K_{31} = -14$, $K_{30} = 3.75$ and $\rho_2 = 3.75$.

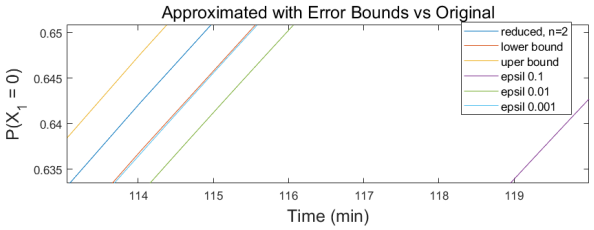
This approximation makes (25) to Σ_{closed} form. Then we let $\epsilon = 0$ and obtain

$$\begin{aligned}
\tilde{\mu}_2^\epsilon(x_1, t) &\approx K_{21}\tilde{\mu}_1^0(x_1, \infty) + K_{20}, \text{where} \\
K_{20} &= \tfrac{e_t x_1 + 2K_{30}}{2(e_t + x_1) - 1 - 2K_{32} + 4e_t V}, \\
K_{21} &= \tfrac{2e_t x_1 - e_t - x_1 + 2K_{31} + 2e_t V}{2(e_t + x_1) - 1 - 2K_{32} + 4e_t V}, \\
\tilde{\mu}_1^\epsilon(x_1, t) &\approx \tfrac{e_t x_1 + K_{20}}{e_t + x_1 - K_{21} + 2e_t V} = \tilde{\mu}_1^0(x_1, \infty).
\end{aligned}$$

When we substitute $\mu_1(x_1, t)$ as $\tilde{\mu}_1^0(x_1, \infty)$ in (25), we can obtain the CME that describes the dynamics of the slow species only, which is in $\Sigma_{reduced}$ form. When we let $k = 0.1[min^{-1}]$, $V = 1[m^3]$, $x_{tot} = 100[molecule]$, we can check that $C(x_1) + c_2K$ is stable and propensity functions defined in (18) are all non-negative for $x_1 = 0, \ldots, x_{tot}$. Therefore, in this case, Σ_{final} is the same as $\Sigma_{reduced}$, which implies that $\lambda^x = 0$. In addition, Δ_0 in Corollary 6.6 can be achieved at $x_1 = 5$ and $i = 1$, $\int_0^\infty \|d_i \exp\{(C(x_1) + c_2K)t\}\|dt = 0.85$ and $\|c_2\| = \frac{1}{50}$. k_1 and k_2 in Corollary 6.6 are 0.2 and 0.1, respectively, and $\left\|\exp\left\{\hat{A}_{new}(T - \tau)\right\}\right\| = 0.1 * \exp\{-0.4(T - \tau)\}$. Based on these values, we can calculate the approximation error bound for the slow-species counts' marginal probability distribution in Corollary 6.6. We can repeat the same procedure for $n = 1$ and $n = 3$ cases.



(a) Comparing $P(X_1 = 0)$ of Σ_{true} with $\epsilon = 0.1, 0.01, 0.001$ and those of Σ_{final} with $n = 1, 2, 3$



(b) Extended view, comparing $P(X_1 = 0)$ of Σ_{true} with $\epsilon = 0.1, 0.01, 0.001$ and those of Σ_{final}, with $n = 2$, with the error bound obtained from Corollary 6.6.

Fig. 1: Comparing P(X_1 = 0), for Σ_{true} and Σ_{final}. For this simulation, $\epsilon = 0.1, 0.01, 0.001$, $n = 1, 2, 3$, $k = 0.1[min^{-1}]$, $V = 1[m^3]$, $x_{tot} = 100[molecule]$, $e_{tot} = 5[molecules]$ are used.

Fig. 1(a) compares $P(X_1 = 0)$ of Σ_{true} with $\epsilon = 0.1, 0.01, 0.001$ and those of Σ_{final} with $n = 1, 2, 3$. The simulation result shows that Σ_{true} and Σ_{final} are almost the same when $\epsilon \leq 0.01$ and $n = 2, 3$. Fig. 1(b) shows as ϵ goes to 0, $P(X_1 = 0)$ of Σ_{true} approaches those of Σ_{final}, with $n = 2$, with the error bound obtained from Corollary 6.6.

B. Protein Binding Reaction

In this example, we consider a protein binding reaction [2]:

$$\emptyset \underset{\delta}{\overset{k}{\rightleftharpoons}} X, X + P \underset{d}{\overset{a}{\rightleftharpoons}} C. \tag{26}$$

In (26), X, P, and C are the protein, the promoter, and the binding complex, respectively. In addition, x, p, and c are the numbers of corresponding species. In this situation, the total numbers of the promoter are conserved, which means $p + c = p_{tot}$ for some positive constant p_{tot}. Therefore, p is a dependant variable and number of complex is bounded by p_{tot}. However, in general, the number of proteins is not bounded, so we need to assume that it is bounded by a positive constant x_{tot} for Assumption 2.1. In this reaction, ap_{tot} and d are much larger than δ and k, so we can define $\epsilon = \frac{\delta}{d}$ and let $k = \delta$ and $ap_{tot} = \frac{d}{2}$. When we define $X_1 = X + C$, and consider $S = [x_1, c]^T$, we can verify X_1 is a slow species, and C is a fast species based on the following propensity functions and correspoinding stoichiometries,

$$a_1^f(x_1; c) = \tfrac{1}{\epsilon}\tfrac{\delta}{2Vp_{tot}}(x_1 - c)(p_{tot} - c), \gamma_1^f = [0, +1]^T,$$
$$a_2^f(x_1; c) = \tfrac{1}{\epsilon}\delta c, \gamma_2^f = [0, -1]^T,$$

$$a_1^s(x_1;c)=\delta, \gamma_1^s=[+1,0]^T,$$
$$a_2^s(x_1;c)=\delta(x_1-c), \gamma_2^s=[-1,0]^T,$$

where V is the volume. We can derive ODEs of the slow-species counts' marginal probability density function and the fast-species counts' first 2 conditional moments as below:

$$\begin{aligned}
&\frac{d}{dt}P(X_1=x_1)=-\delta x_1 P(X_1=x_1)\\
&+\delta(x_1+1)P(X_1=x_1+1)-\delta\mu_1(x_1+1,t)P(X_1=x_1+1)\\
&+\delta\mu_1(x_1,t)P(X_1=x_1)-\delta P(X_1=x_1)+\delta P(X_1=x_1-1)\\
&\epsilon\frac{d}{dt}Y_2(x_1,t)=\\
&\begin{bmatrix} -\delta-\frac{\delta(x_1+p_{tot})}{2Vp_{tot}} & \frac{\delta}{2Vp_{tot}}\\ \delta+\frac{\delta(2\delta x_1 p_{tot}-\delta x_1-\delta p_{tot})}{2VP_{tot}} & -2\delta+\frac{\delta-2\delta x_1-2\delta p_{tot}}{2Vp_{tot}}\end{bmatrix}Y_2(x_1,t)\\
&+\begin{bmatrix}\frac{\delta x_1}{2V}\\ \frac{\delta x_1}{2V}\end{bmatrix}+\begin{bmatrix}0\\ \frac{\delta}{Vp_{tot}}\end{bmatrix}\mu_3(x_1,t)+\epsilon G(t)\\
&=C(x_1)Y_2(x_1,t)+c_1(x_1)+c_2\mu_3(x_1,t)+\epsilon G(t).
\end{aligned}\tag{27}$$

We can check that (27) is in Σ_{true} form. To close the dynamics, we let $p_{tot}=10[molecules]$ and solve the linear program in (14) and obtain

$$\mu_3(x_1,t)\approx K_{32}\mu_2(x_1,t)+K_{31}\mu_1(x_1,t)+K_{30},$$
where $K_{32}=15$, $K_{31}=-56$, $K_{30}=30$ and $\rho_2=30$.

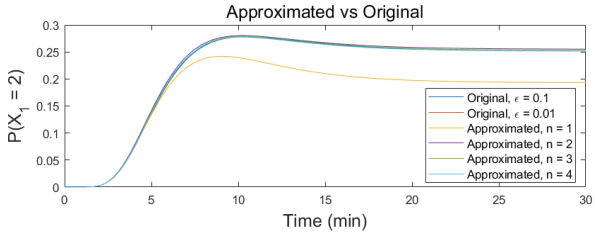
This approximation makes (27) to Σ_{closed} form. Then we let $\epsilon=0$ and obtain

$$\begin{aligned}
\tilde{\mu}_2^\epsilon(x_1,t)&\approx K_{21}\tilde{\mu}_1^0(x_1,\infty)+K_{20}, \text{where}\\
K_{20}&=\tfrac{ax_1p_{tot}+2aK_{30}}{2ap_{tot}+2ax_1-2aK_{32}-a+2dV},\\
K_{21}&=\tfrac{2ax_1p_{tot}-ap_{tot}-ax_1+2aK_{31}+dV}{2ap_{tot}+2ax_1-2aK_{32}-a+2dV},\\
\tilde{\mu}_1^\epsilon(x_1,t)&\approx\tfrac{aK_{20}+ax_1p_{tot}}{ap_{tot}+ax_1-aK_{21}+dV}=\tilde{\mu}_1^0(x_1,\infty).
\end{aligned}$$

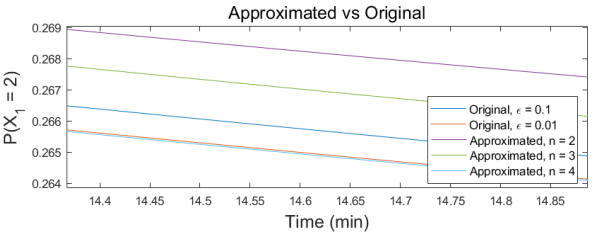
When we substitute $\mu_1(x_1,t)$ as $\tilde{\mu}_1^0(x_1,\infty)$ in (27), we can obtain the CME that describes the dynamics of the slow species only, which is in $\Sigma_{reduced}$ form. When we let $\delta=0.4[min^{-1}]$, $V=1[m^3]$, $x_{tot}=100[molecule]$, we can check that $C(x_1)+c_2K$ is stable and propensity functions defined in (18) are all non-negative for $x_1=0,\ldots,x_{tot}$. Therefore, in this case, Σ_{final} is the same as $\Sigma_{reduced}$, which implies that $\lambda^x=0$. Fig. 2(a) compares $P(X_1=2)$ of Σ_{true} with $\epsilon=0.1, 0.01$ and those of Σ_{final} with $n=1,2,3,4$ and Fig. 2(b) is the extended view of Fig. 2(a). Fig. 2(b) shows that $P(X_1=2)$ of Σ_{true} with $\epsilon=0.01$ and those of Σ_{final} with $n=4$ are almost the same.

VIII. CONCLUSIONS

In this paper, we leveraged time-scale separation, which is intrinsic to many biochemical reaction networks, to develop a method to reduce the order of the CME. We derived a reduced size CME that describes the dynamics of the slow species only. Our approach provides exact error bounds between the true and the reduced systems. This reduction, with quantifiable error bounds, can help the analysis and design of biochemical reaction systems. Our future goal is to apply the FSP algorithm to extend this result to infinite dimensional CME.



(a) Comparing $P(X_1=2)$ of Σ_{true} with $\epsilon=0.1, 0.01$ and those of Σ_{final} with $n=1,2,3,4$



(b) Extended view of the above graph

Fig. 2: Comparing $P(X_1=2)$, for Σ_{true} and Σ_{final}. For this simulation, $\epsilon=0.1,\ 0.01$, $n=1,\ 2,\ 3,\ 4$, $\delta=0.4[min^{-1}]$, $V=1[m^3]$, $x_{tot}=100[molecule]$, $p_{tot}=10[molecules]$ are used.

APPENDIX

A. Iterative algorithm in Remark 4.1

First we solve (14) and let $K=K^*$ be its optimal solution. Then, we will find $K_{(1)}$ such that it is close to K^* but also makes $C(x)+c_2K_{(1)}$ stable. To do this, we will find matrices Z and $P_{(1)}$ such that $K_{(1)}=ZP_{(1)}^{-1}$ where Z and $P_{(1)}$ can be obtained by solving

$$\min_{Z,P_{(1)},\alpha_{(1)}}\|Z-K^*P_{(1)}\|,$$
$$\text{s.t. } C(x)P_{(1)}+c_2Z+(C(x)P_{(1)}+c_2Z)^T\prec-\alpha_{(1)}I,$$
$$P_{(1)}\succ 0,\alpha_{(1)}>0.$$

Then, at each iteration, given $P_{(j)}$ and $\alpha_{(j)}$, we first find $K_{(j+1)}$ by solving

$$\min_{K_0,K_{(j+1)},\gamma}\gamma$$
$$\text{s.t. } -\gamma 1^T\le\mathcal{R}[H_n-(K_{(j+1)}V_n+K_01^T)]_i\le\gamma 1^T,$$
$$[C(x)+c_2K_{(j+1)}]P_{(j)}+P_{(j)}^T[C(x)+c_2K_{(j+1)}]^T\prec-\tfrac{1}{\alpha_{(j)}}I,$$

for i = 1, 2, ..., m. Then, given $K_{(j+1)}$, we find $P_{(j+1)}$ and $\alpha_{(j+1)}$ by solving

$$\min_{P_{(j+1)}}-\alpha_{(j+1)}\ \text{s.t.}$$
$$[C(x)+c_2K_{(j+1)}]P_{(j)}+P_{(j)}^T[C(x)+c_2K_{(j+1)}]^T\prec$$
$$-\alpha_{(j+1)}I,$$
$$P_{(j+1)}\succ 0,\alpha_{(j+1)}>0.$$

We continue until $\|K_{(j+1)}-K_{(j)}\|$ converges.

B. Proof of Lemma 6.1

To quantify the errors, we first define

$$e_1(x,t)=Y_n(x,t)-\tilde{Y}_n^\epsilon(x,t).$$

Manuscript 647 submitted to 2019 American Control Conference.
Received September 24, 2018.

Using (5) and (15), we can derive

$$\epsilon\frac{d}{dt}e_1(x,t) = (C(x)+c_2K)e_1(x,t) + c_2(\mu_{n+1}(x,t) - (KY_n(x,t)+K_0)), e_1(x,t_0) = 0. \tag{28}$$

By solving (28),

$$e_1(x,t) = \int_{t_0}^{t} \exp\left\{\frac{1}{\epsilon}(C(x)+c_2K)(t-\tau)\right\} [\frac{1}{\epsilon}c_2(\mu_{n+1}(x,\tau) - (KY_n(x,\tau)+K_0))]d\tau. \tag{29}$$

Because of Eq (29), we can obtain

$$\begin{aligned}
&\sup_{t\in[t_0,T]}\|\mu_i(x,t)-\tilde{\mu}_i^{\epsilon}(x,t)\| = \sup_{t\in[t_0,T]}\|d_ie_1(x,t)\| \\
&\le \sup_{t\in[t_0,T]}\int_{t_0}^{t}||d_i\exp\left\{\frac{1}{\epsilon}(C(x)+c_2K)(t-\tau)\right\} \\
&[\frac{1}{\epsilon}c_2(\mu_{n+1}(x,\tau)-(KY_n(x,\tau)+K_0))]||d\tau \\
&\le \sup_{t\in[t_0,T]}\int_{t_0}^{t}[||d_i\exp\left\{\frac{1}{\epsilon}(C(x)+c_2K)(t-\tau)\right\}|| \\
&\left\|\frac{1}{\epsilon}c_2(\mu_{n+1}(x,\tau)-(KY_n(x,\tau)+K_0))\right\|]d\tau \\
&\le \int_{t_0}^{T}\left\|d_i\exp\left\{\frac{1}{\epsilon}(C(x)+c_2K)(T-\tau)\right\}\right\|d\tau\frac{\rho_n}{\epsilon}\|c_2\|
\end{aligned} \tag{30}$$

which is the same as Lemma 6.1.

C. Proof of Lemma 6.2

Next we define

$$e_2(x,t) = \tilde{Y}_n^{\epsilon}(x,t) - \tilde{Y}_n^{0}(x,\infty).$$

Using (15) and (17), we can derive

$$\epsilon\frac{d}{dt}e_2(x,t) = (C(x)+c_2K)e_2(x,t)+\epsilon G(t), e_2(x,t_0) = 0. \tag{31}$$

By solving (31), we can derive

$$e_2(x,t) = \int_{t_0}^{t}\exp^{\frac{1}{\epsilon}(C(x)+c_2K)(t-\tau)}G(\tau)d\tau. \tag{32}$$

Because of (32), we can obtain

$$\begin{aligned}
&\sup_{t\in[t_0,T]}\left\|\tilde{\mu}_i^{\epsilon}(x,t)-\tilde{\mu}_i^{0}(x,\infty)\right\| = \sup_{t\in[t_0,T]}\|d_ie_2(x,t)\| \\
&= \sup_{t\in[t_0,T]}\left\|\int_{t_0}^{t}d_i\exp\left\{\frac{1}{\epsilon}(C(x)+c_2K)(t-\tau)\right\}G(\tau)d\tau\right\| \\
&\le \sup_{t\in[t_0,T]}\int_{t_0}^{t}\left\|d_i\exp\left\{\frac{1}{\epsilon}(C(x)+c_2K)(t-\tau)\right\}\right\|\|G(\tau)\|d\tau
\end{aligned} \tag{33}$$

which is the same as Lemma 6.2.

D. Proof of Theorem 6.5

We define

$$e_3(t) = P_{X,new}(t) - \hat{P}_{X,new}(t).$$

Using (21) and (23), we can derive

$$\begin{aligned}
&\frac{d}{dt}e_3(t) = \hat{A}_{new}e_3(t)+\Delta_1(t)P_{X,new}(t)+\Delta_2(t) \\
&= \hat{A}_{new}e_3 + w(t), e_3(t_0) = 0
\end{aligned} \tag{34}$$

By solving (34), we can obtain

$$e_3(t) = \int_{t_0}^{t}\exp\left\{\hat{A}_{new}(t-\tau)\right\}w(\tau)d\tau. \tag{35}$$

Here, norm of $w(t)$ is bounded by

$$\begin{aligned}
&\|w(t)\| = \|\Delta_1(t)P_{X,new}(t)+\Delta_2(t)\| \\
&\le \|\Delta_1(t)P_{X,new}(t)\| + \|\Delta_2(t)\| \\
&\le \|\Delta_1(t)\|_{l_1-l_\infty} + \|\Delta_2(t)\| \le (k_1+k_2)\Delta_\epsilon + O(\epsilon).
\end{aligned} \tag{36}$$

Because of (35) and (36), we can derive

$$\begin{aligned}
&\sup_{t\in[t_0,T]}\|e_3(t)\| = \sup_{t\in[t_0,T]}\left\|\int_{t_0}^{t}\exp\left\{\hat{A}_{new}(t-\tau)\right\}w(\tau)d\tau\right\| \\
&\le \sup_{t\in[t_0,T]}\int_{t_0}^{t}\left\|\exp\left\{\hat{A}_{new}(t-\tau)\right\}w(\tau)\right\|d\tau \\
&\le \sup_{t\in[t_0,T]}\int_{t_0}^{t}\left\|\exp\left\{\hat{A}_{new}(t-\tau)\right\}\right\|\|w(\tau)\|d\tau \\
&\le (k_1+k_2)\sup_{t\in[t_0,T]}\int_{t_0}^{t}\left\|\exp\left\{\hat{A}_{new}(t-\tau)\right\}\right\|d\tau\Delta_\epsilon + O(\epsilon)
\end{aligned} \tag{37}$$

which is the same as Theorem 6.5.

REFERENCES

[1] I. Oppenheim, K. Shuler, and G. Weiss, "Stochastic and deterministic formulation of chemical rate equations," *The Journal of Chemical Physics*, vol. 50, no. 1, pp. 460–466, 1969.

[2] D. Del Vecchio and R. M. Murray, *Biomolecular feedback systems*. Princeton University Press Princeton, NJ, 2015.

[3] B. Munsky and M. Khammash, "The finite state projection algorithm for the solution of the chemical master equation," *The Journal of chemical physics*, vol. 124, no. 4, p. 044104, 2006.

[4] G. G. Yin and Q. Zhang, *Continuous-time Markov chains and applications: A two-time-scale approach*, vol. 37. Springer Science & Business Media, 2012.

[5] 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*, vol. 129, no. 24, p. 244112, 2008.

[6] J. Goutsias, "Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems," *The Journal of chemical physics*, vol. 122, no. 18, p. 184102, 2005.

[7] H. K. Khalil, "Noninear systems," *Prentice-Hall, New Jersey*, vol. 2, no. 5, pp. 5–1, 1996.

[8] M. Naghnaeian and D. Del Vecchio, "Robust moment closure method for the chemical master equation," in *Control Technology and Applications (CCTA), 2017 IEEE Conference on*, pp. 967–972, IEEE, 2017.

[9] N. G. Van Kampen, *Stochastic processes in physics and chemistry*, vol. 1. Elsevier, 1992.

[10] D. T. Gillespie, "A rigorous derivation of the chemical master equation," *Physica A: Statistical Mechanics and its Applications*, vol. 188, no. 1-3, pp. 404–425, 1992.

[11] D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions," *Journal of computational physics*, vol. 22, no. 4, pp. 403–434, 1976.

[12] W. Mitkowski, "Dynamical properties of metzler systems," *Technical Sciences*, vol. 56, no. 4, pp. 309–312, 2008.

Manuscript 647 submitted to 2019 American Control Conference.
Received September 24, 2018.