APPLICATIONS OF
MODELLING AND SIMULATION
http://arqiipubl.com/ams eISSN 2600-8084 VOL 5, 2021, 191-199

Using Virtual Reaction Scheme in the Multiple Time Scales System for Stochastic Modelling

C. F. Khoo

Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka, 76100 Durian Tunggal, Melaka, Malaysia

Corresponding author: chinfoon@utem.edu.my

Submitted 03 September 2021, Revised 11 October 2021, Accepted 17 October 2021.

Abstract: A virtual reaction scheme and the notion of quasi-independent is proposed in this paper to assist in model partitioning for the biochemical reactions with multiple time scales. Subsequently, the singular perturbation method is applied to the stochastic model, in particular the chemical master equation of the multiple time scales system to reduce the model dimension. As a result, a lower dimensional approximation for the chemical master equation is derived through this approach. Therefore, the high dimensional chemical master equation can be solved with a lower computational cost.

Keywords: Multiple time scales system; Singular perturbation method; Stochastic modelling; Virtual reaction scheme.

1. INTRODUCTION

Many biochemical systems found in biological organisms exhibit dynamical stiffness as they have rate constants that vary several orders of magnitudes which result in different dynamics over different time scales. For instance, the enzymatic reactions in intracellular environments are always faster than other protein interactions. Generally, the key idea to deal with such chemical reaction system is the system partitioning based on its kinetics behaviour. Some works regarding with this attempt include the efforts of (1) Haseltine and Rawlings [1] which partitioned the system into slow and fast reaction subsets, then treated the fast system deterministically and the slow system stochastically; (2) Rao and Arkin [2] which partitioned the system into two subsystems with primary and intermediate species respectively, and subsequently simplified the model by eliminating the fast dynamics based on the quasi-steady state assumption (QSSA); (3) Cao et. al. [3] which skipped over the fast reactions by assuming there exists a time-dependent probability function for fast species, and therefore simulated the slow reaction with a slow-scale stochastic simulation algorithm; (4) Goutsias [4] which also partitioned the system into two distinct subsystems of slow and fast reactions. He was then employed the quasi-equilibrium approximation and eliminated the fast reactions kinetics by assuming their distributions are time independent. The are other papers following these ideas such as the later works by Lee and Lui [5], Kan *et al.* [6], Lecca *et al.* [7] and Kim and Sontag [8].

Under the same attempt to reduce the computational complexity of multiple time scales system in this paper, we expand the idea of QSSA, apply the singular perturbation method to the multiple time scales system, that is, find a set of solution which start from the fast and proceeding to the slower time scale in two different regions: the inner (boundary layer) and outer regions.

2. VIRTUAL REACTION SCHEME

Consider a well-stirred biochemical system with N kinds of molecular species $S_i, i = 1, \cdots, N$, interacting via M reaction channels, $R_j, j = 1, \cdots, M$, in the system. From a reaction point of view, we say each reaction involves $n \in [2, N]$ molecular species. Alternatively, from a molecular species point of view, we can also say that each molecular species is involved in $m \in [1, M]$ reactions. For example, the following reaction

$$S_1 \overset{\kappa_1}{\to} S_2 \tag{1}$$

involves two molecular species. Note here κ_1 is the rate constant and hence the reaction rate, α_1, for reaction (1) is $\kappa_1 S_1$ where $[S_i]$ denotes the concentration of species S_i.

If S_1 and S_2 do not interact with other molecular species, then reaction (1) can be represented as

$$S_1 \xrightarrow{\kappa_1} \emptyset,\ \emptyset \xrightarrow{\kappa_1 [S_1]} S_2 \tag{2}$$

where the first reaction represents the decay of species S_1 with reaction rate $\gamma_1 = \kappa_1$ while the second reaction represents the production of S_2 with reaction rate $\gamma_2 = \kappa_1[S_1]$. Note that $\emptyset$ is a "virtual state" to which S_1 decays and from which S_2 is produced.

Obviously, the first type of reaction scheme (scheme (1)) is a more natural way to model the chemical kinetics of the dynamical biochemical processes mathematically. However, the second type of reaction scheme (scheme (2)) is preferable in this paper as we are focusing on the nature of the molecular species instead of the reaction kinetics.

The reaction scheme which is based on the molecular species (such as scheme (2)) is called the "virtual reaction" scheme here. The dynamics of the molecular species in the virtual reaction scheme are said to be "quasi-independent" as their dependencies are hidden in the reaction rate. For instance, the dependency of S_2 on S_1 is not shown directly in the virtual reaction $\emptyset \to S_2$ but on its reaction rate $\gamma_2 = \kappa_1[S_1]$.

To illustrate the notion of virtual reaction scheme in more detail, let's consider the following example:

$$S_1 \underset{\kappa_3}{\overset{\kappa_1}{\rightleftarrows}} S_2 \xrightarrow{\kappa_3} S_3 \tag{3}$$

The virtual reaction scheme of the reactions (3) can be represented as

$$\left.\begin{array}{c} S_1 \to \emptyset\,, \gamma_1 = \kappa_1[S_1]\,,\ \ \phi \to S_1\,,\ \gamma_2 = \kappa_3[S_2]\,, \\ S_2 \to \emptyset\,,\ \gamma_3 = (\kappa_2 + \kappa_3)[S_1]\,,\ \ \phi \to S_2\,,\ \gamma_4 = \kappa_1[S_1]\,, \\ \phi \to S_3\,,\ \gamma_5 = \kappa_2[S_2]\,. \end{array}\right\} \tag{4}$$

We can see that although molecular species S_2 involved in three reactions, the two decay reactions of S_2 have been combined into one as shown in the third reaction of scheme (4). In short, every molecular species in the virtual reaction scheme has at most two virtual reactions only, that is, production and decay. Under such a framework, all the reactions (such as synthesis, dephosphorylation, dissociation) which produce the species S_i are classified in the group of production while all the reactions (for instance degradation, phosphorylation, association) which consume the molecular species S_i in the system are put in the group of decay.

2.1 Stochastic Modelling

Consider a well-stirred system with *N* different kinds of molecular species, $S_i,\ i = 1, \cdots, N$ and *M* reaction channels, $R_j, j = 1, \cdots, M$, in a fixed volume *V*. Let x_i denotes the number of molecules of species S_i. Then the state space of the system is $\Omega \subset \mathbb{N}^N$ where the elements in the state space is $\mathbf{x} = [x_1, \cdots, x_N]^{\mathrm{T}} \in \Omega$. Each reaction is characterised by a non-negative propensity function $\alpha_j(\mathbf{x})$ and a stoichiometric vector $\mathbf{v}_{\mathrm{j}} \in \mathbb{Z}_{\mathrm{N}}$ which satisfies $\mathbf{x} + \mathbf{v}_j \in \Omega$ for all $\mathbf{x} \in \Omega$ such that $\alpha_j(\mathbf{x}) \neq 0$. Usually some mass conservation laws hold in such a system. That is, the sum of the number of molecules of some molecular species is equal to a constant number $y_i \in \mathbb{Z}_+$ where $y_i = \max\{x_i\}$ is the maximum copy number of molecules of a particular species S_i.

In general, the dynamics of the molecular species in this system can be represented stochastically with a chemical master equation (CME) as

$$\frac{dP(\mathbf{x};t)}{dt} = \sum_{j=1}^{M} \alpha_{\mathrm{j}}(\mathbf{x} + \mathbf{v}_j)P(\mathbf{x} + \mathbf{v}_j; t) - \alpha_j(\mathbf{x})P(\mathbf{x}; t) \tag{5}$$

where $P(\mathbf{x}; \mathrm{t})$ is the probability distribution of the states $\mathbf{x}$ at time *t*.

On the other hand, let us consider the virtual reaction scheme of this system. If the system has *N* different kinds of molecular species, there will be η (where $\eta = N, N+1, \cdots, 2N$) virtual reactions in the system. The state space of the system is $\widetilde{\Omega} \subset \mathbb{N}^N$ where the elements in the state space is $\tilde{\mathbf{x}} = [x_1, \cdots, x_N]^{\mathrm{T}} \in \widetilde{\Omega}$. Note that if there are no mass conservation laws applied in the actual reaction scheme, then $|\Omega| = |\widetilde{\Omega}|$, otherwise $|\Omega| < |\widetilde{\Omega}|$. This is because mass conservation laws are not applied in the virtual reaction scheme.

Each virtual reaction is characterised by a propensity function $\gamma_j(\tilde{\mathbf{x}}), j = 1, \cdots, \eta$, and a stoichiometric vector $\tilde{\mathbf{v}}_j \in \mathbb{Z}^N$ which satisfies $\tilde{\mathbf{x}} + \tilde{\mathbf{v}} \in \widetilde{\Omega}$ for all $\tilde{\mathbf{x}} \in \widetilde{\Omega}$ such that $\gamma_j(\tilde{\mathbf{x}}) \neq 0$ in the virtual reaction scheme. More precisely, let $\widetilde{\gamma_j}(\tilde{\mathbf{x}})$ determine the transition rate of the reaction R_j, if reaction R_j is a reaction of production type (which produces species S_{i}), then $\gamma_j(\tilde{\mathbf{x}}) = \widetilde{\gamma_j}(\tilde{\mathbf{x}}) \times \delta_i(\tilde{\mathbf{x}})$ where $\delta_i(\tilde{\mathbf{x}})$ is a Kronecker delta function:

$$\delta_{\mathrm{i}}(\tilde{\mathbf{x}}) = \begin{cases} 0, & if\ x_i \geq y_i \\ & \text{otherwise} \end{cases}. \tag{6}$$

This is reasonable, otherwise the propensity defines a gain of mass for a state where no such gain is expected. That is, in the state $x_i \geq y_i$, we do not expect any production to get $x_i \geq y_i + 1$ which is outside $\widetilde{\Omega}$, hence we set $\gamma_j(\tilde{\mathbf{x}}) = 0$ when $x_i \geq y_i$. On the contrary, if R_j is a reaction of decay type (which consumes species S_i), the Kronecker delta function term is not necessary as $\widetilde{\gamma_j}(\tilde{\mathbf{x}})$ for the decay of species S_i is of the form $\widetilde{\gamma_j}(\tilde{\mathbf{x}}) = \kappa x_i, x_i \in \mathbb{N}$ where κ is the rate constant. Hence when $x_i = 0, \widetilde{\gamma_j}(\tilde{\mathbf{x}}) = 0$. Consequently, the propensity function of any decay reaction R_j is $\gamma_j(\tilde{\mathbf{x}}) = \widetilde{\gamma_j}(\tilde{\mathbf{x}})$.

Under such a framework, the dynamics of the virtual reaction scheme can be represented stochastically as

$$\frac{d\tilde{P}(\tilde{\mathbf{x}};t)}{dt} = \sum_{j=1}^{\eta} \gamma_j\big(\tilde{\mathbf{x}}+\tilde{\mathbf{v}}_j\big)P\big(\tilde{\mathbf{x}}+\tilde{\mathbf{v}}_j;t\big) - \gamma_j(\tilde{\mathbf{x}})P(\tilde{\mathbf{x}};t). \tag{7}$$

Subsequently, we will look at the marginal probability, expected value, variance and covariance of the molecular species to investigate whether a virtual reaction scheme can be used as an approximate model for the actual reaction scheme in any stochastic modelling. Let us take the reactions (1) and (2) as the example again, that is, κ_1 is the rate constant. The CMEs for actual reaction and virtual reaction schemes can be written as follows:

$$\frac{dP(x_1,x_2;t)}{dt} = \kappa_1(x_1+1)P(x_1+1,x_2-1;t) - \kappa_1 x_1 P(x_1,x_2;t), \tag{8}$$

$$\begin{aligned}\frac{d\tilde{P}(x_1,x_2;t)}{dt} &= \kappa_1(x_1+1)\tilde{P}(x_1+1,x_2;t) - \kappa_1 x_1 \tilde{P}(x_1,x_2;t)\\ &+\kappa_1 x_1 \delta_2(x_1,x_2-1)\tilde{P}(x_1,x_2-1;t) - \kappa_1 x_1 \delta_2(x_1,x_2)\tilde{P}(x_1,x_2;t).\end{aligned} \tag{9}$$

The reaction scheme (1) shows that the number of molecules of S_1, x_1, decreases one while number of molecules for S_2, x_2, increases one in each reaction, so the corresponding stoichiometric vector is $[-1, +1]^T$. The propensity function is the product of the rate constant and the number of reactant molecules, so for the reaction scheme (1), $\alpha_j(\mathbf{x}) = \kappa_1 x_1$. Consequently, when reaction fires, $\alpha_j\big(\mathbf{x}+\mathbf{v}_j\big) = \kappa_1(x_1+1)$. Based on the stoichiometric vector, the probability $P\big(\mathbf{x}+\mathbf{v}_j;t\big)$ in Equation (5) becomes $P(x_1+1,x_2-1;t)$. So Equation (5) can be interpreted more specifically as Equation (8) for reaction scheme (1). On the other hand, there are two reactions being considered in reaction scheme (2). In the first equation, x_1 loses one molecule when the reaction fires, while x_2 gains extra one molecule in the second reaction. In general, the dynamics of these virtual reactions can be represented stochastically by Equation (7). We may rewrite Equation (7) as Equation (9) specifically for reaction scheme (2), that is, by considering the first two terms in the right-hand side of the Equation (9) represents the first reaction with a lost of x_1 and the stoichiometric vector is $\tilde{\mathbf{v}}_1 = [-1, 0]^T$, while the next two terms in the right-hand side of the Equation (9) represents the second reaction which shows a gain of x_2 with the stoichiometric vector $\tilde{\mathbf{v}}_2 = [0, +1]^T$.

Based on these equations, the time evolution of the marginal probabilities, expected values, variances and covariances of the molecular species S_1 and S_2 for both reaction schemes are shown in Table 1. Note that expected value can be determined by multiplying x_i to the marginal probablity of x_i and summing over all x_i later. Variance is found by multiplying x_i^2 to the marginal probablity of x_i and summing over all x_i, then minus the square of the expected value of x_i, $\mu_{x_1}^2$. The covariance, on the other hand, can be found by multiplying the marginal probablity with x_1 and x_2, then summing over all x_1 and x_2, subsequently minus the product of expected values of x_1 and x_2. We can see in the Table 1 that if the molecular species is only involved in decay types of reaction, then its marginal distribution, expected value and variance derived from both actual reaction and virtual reaction schemes are exactly same. Otherwise, if the molecular species is involved in production types of reaction, then the virtual reaction scheme is a good approximation to the actual reaction scheme when $\zeta(t)$ is very small, that is, if $\tilde{P}(x_1, y_2; t)$ *i*s approximately zero. Both reaction schemes have different covariances as the relationship between molecular species are different; that is, these species depend on each other in the actual reaction scheme, but have a quasi-independent relationship in the virtual reaction scheme.

Table 1. Comparison of the time evolution of the marginal probability distributions, expected values, variances and covariances of S_1 and S_2 of actual reaction and virtual reaction schemes.

Time evolution of	Actual reaction scheme	Virtual reaction scheme
Marginal probability distribution of S_1	$\frac{dP_{x_1}}{dt} = \sum_{x_2=0}^{y_2} \kappa_1(x_1+1)P(x_1+1,x_2;t) - \kappa_1 x_1 P(x_1,x_2;t)$	$\frac{d\tilde{P}_{x_1}}{dt} = \frac{dP_{x_1}}{dt}$
Marginal probability distribution of S_2	$\frac{dP_{x_2}}{dt} = \sum_{x_1=0}^{y_1} \kappa_1 x_1 P(x_1,x_2-1;t) - \kappa_1 x_1 P(x_1,x_2;t)$	$\frac{d\tilde{P}_{x_2}}{dt} = \frac{dP_{x_2}}{dt} + \zeta$
Expected value of S_1	$\frac{d\mu_{x_1}}{dt} = -\kappa_1\mu_{x_1}$	$\frac{d\tilde{\mu}_{x_1}}{dt} = \frac{d\mu_{x_1}}{dt}$
Expected value of S_2	$\frac{d\mu_{x_2}}{dt} = \kappa_1\mu_{x_1}$	$\frac{d\tilde{\mu}_{x_2}}{dt} = \frac{d\mu_{x_2}}{dt} - \zeta$
Variance of S_1	$\frac{d\sigma_{x_1}^2}{dt} = \kappa_1\mu_{x_1} - 2\kappa_1\sigma_{x_1}^2$	$\frac{d\tilde{\sigma}_{x_1}^2}{dt} = \frac{d\sigma_{x_1}^2}{dt}$

Variance of S_2	$\frac{d\sigma_{x_2}^2}{dt} = \kappa_1\mu_{x_1} + 2\kappa_1\sigma_{x_1x_2}$	$\frac{d\tilde{\sigma}_{x_2}^2}{dt} = \frac{d\sigma_{x_2}^2}{dt} - \zeta(2y_2 + 1 - 2\mu_{x_2})$
Covariance	$\frac{d\sigma_{x_1x_2}}{dt} = \kappa_1(-\mu_{x_1} + \sigma_{x_1}^2 - \sigma_{x_1x_2})$	$\frac{d\tilde{\sigma}_{x_1x_2}}{dt} = \kappa_1(-\mu_{x_1}^2 + \sigma_{x_1}^2 - \sigma_{x_1x_2}) + \zeta\mu_{x_1}$

Note:

$$P_{x_1}(t) = \sum_{x_2=0}^{y_2} P(x_1, x_2; t)\ , \quad \tilde{P}_{x_1}(t) = \sum_{x_2=0}^{y_2} \tilde{P}(x_1, x_2; t)$$

$$P_{x_2}(t) = \sum_{x_1=0}^{y_1} P(x_1, x_2; t)\ , \quad \tilde{P}_{x_2}(t) = \sum_{x_1=0}^{y_1} \tilde{P}(x_1, x_2; t)$$

$$\mu_{x_1}(t) = \sum_{x_1=0}^{y_1}\sum_{x_2=0}^{y_2} x_1P(x_1, x_2; t)\ , \quad \tilde{\mu}_{x_1}(t) = \sum_{x_1=0}^{y_1}\sum_{x_2=0}^{y_2} x_1\tilde{P}(x_1, x_2; t)$$

$$\mu_{x_2}(t) = \sum_{x_1=0}^{y_1}\sum_{x_2=0}^{y_2} x_2P(x_1, x_2; t)\ , \quad \tilde{\mu}_{x_2}(t) = \sum_{x_1=0}^{y_1}\sum_{x_2=0}^{y_2} x_2\tilde{P}(x_1, x_2; t)$$

$$\sigma_{x_1}^2(t) = \sum_{x_1=0}^{y_1}\sum_{x_2=0}^{y_2} x_1^2P(x_1, x_2; t) - \mu_{x_1}^2, \quad \tilde{\sigma}_{x_1}^2(t) = \sum_{x_1=0}^{y_1}\sum_{x_2=0}^{y_2} x_1^2\tilde{P}(x_1, x_2; t) - \tilde{\mu}_{x_1}^2$$

$$\sigma_{x_2}^2(t) = \sum_{x_1=0}^{y_1}\sum_{x_2=0}^{y_2} x_2^2P(x_1, x_2; t) - \mu_{x_2}^2, \quad \tilde{\sigma}_{x_2}^2(t) = \sum_{x_1=0}^{y_1}\sum_{x_2=0}^{y_2} x_2^2\tilde{P}(x_1, x_2; t) - \tilde{\mu}_{x_2}^2$$

$$\sigma_{x_1x_2}(t) = \sum_{x_1=0}^{y_1}\sum_{x_2=0}^{y_2} x_1x_2P(x_1, x_2; t) - \mu_{x_1}\mu_{x_2}\ , \ \tilde{\sigma}_{x_1x_2}(t) = \sum_{x_1=0}^{y_1}\sum_{x_2=0}^{y_2} x_1x_2\tilde{P}(x_1, x_2; t) - \tilde{\mu}_{x_1}\tilde{\mu}_{x_2}$$

$$\zeta(t) = \sum_{x_1=0}^{y_1} \kappa_1x_1\tilde{P}(x_1, y_2; t)$$

3. VIRTUAL REACTION SCHEME OF THE MULTIPLE TIME SCALES SYSTEM

Consider a well-stirred biochemical system. Suppose the chemical kinetics of the molecular species vary several orders of magnitude, so that we can classify the molecular species into two groups, namely the slowly evolving primary species, S_f, and highly reactive intermediate species, S_s. To illustrate the notion of primary species and intermediate species in more detail, consider a reaction scheme where molecular species *A* is converted into molecular species *B* through the formation of another molecular species *C*, that is

$$A \rightarrow C \rightarrow B \tag{10}$$

instead of $A \rightarrow B$ directly. Then the species A and B are called the primary species while C is the intermediate species. Basically, the intermediate species are highly reactive, and their molecular levels always approach a stable value at a rate faster than the other species do, whereas primary species evolve slowly in the system.

In general, a multiple time scales system usually contains reactions of type (10). In this paper, we are interested in a multiple time scales system under the following conditions:

(a) The mass conservation law holds such that nonzero probabilities only occur for $x_f + x_s = x_T$ where x_f and x_s denote the number of molecules of intermediate and primary species respectively and x_T is fixed throughout the system.

(b) Let ϵ be a small positive constant such that $0 < \epsilon \leq 1$ and let $\kappa = \frac{\kappa_d}{\kappa_p}$ where κ_p and κ_d are the production and decay rates respectively of the intermediate species. Then, the values of κ_d are always much larger than κ_p in the system, such that the parameter $\frac{1}{\kappa}$ always less than the small parameter ϵ;

(c) The number of molecules of the primary species is always much smaller than κ and therefore very little amount of intermediate species are formed at any time in the system.

(d) If α_d^* is the maximum of the propensity function $\alpha_d(\mathbf{x})$ for the decay of the primary species, and α_p^* is the maximum of the propensity function $\alpha_p(\mathbf{x})$ for the production of the primary species, then α_p^* is at least less than or equal to α_d^*. Note that $\alpha_d(\mathbf{x})$ and $\alpha_\mathrm{p}(\mathbf{x})$ are the propensities function of actual reaction scheme and $\mathbf{x}$ is the state of the multiple time scales system.

Systems demonstrating multiple time scales always lead to computational difficulties. Anyway, the computational effort required can be reduced if the dynamics of different types of species can be simulated independently of each other in the multiple time scales system. Nonetheless, the occurrence of a reaction in a system may cause the population levels of both primary and intermediate species to change simultaneously. In such a situation, we may not able to simulate the dynamics of the primary species without also simulating dynamics of the intermediate species, and vice versa. With this issue in mind, the virtual reaction scheme which has proposed in the preceding section will be used to replace the actual reaction scheme, hence the dynamics of the primary and intermediate species can be numerically simulated separately. Under the conditions (a)-(d), the virtual reaction scheme is a good approximation to the actual reaction scheme in the multiple time scales system. By assuming that no reaction is shared by any two or more species allows us to simulate reactions related to the primary species without also simulating intermediate species reactions. This enables lower dimensional simulations in the singular perturbation study.

4. SINGULAR PERTURBATION METHOD

Consider the virtual reaction scheme in a multiple time scales system with highly reactive intermediate species S_f and slowly evolving primary species S_s. The state space of the system is $\widetilde{\Omega} \subset \mathbb{N}^N$ where the elements in the state space is $\mathbf{x} = [\mathbf{x_f}, \mathbf{x_s}]^T \in \widetilde{\Omega}$. Let $N = \tilde{n} + \widetilde{m}$, $\mathbf{x_f} = \left[x_{f_1}, x_{f_2}, \cdots, x_{f_{\tilde{n}}}\right]$ where $x_{f_i}, i = 1, \cdots, \tilde{n}$, denotes the number of molecules of i-th intermediate species S_{f_i}, then $\mathbf{x_f} \in \widetilde{\Omega}_1$, $\widetilde{\Omega}_1 \subset \mathbb{N}^{\tilde{n}}$; while $\mathbf{x_s} = \left[x_{s_1}, x_{s_2}, \cdots, x_{s_{\widetilde{m}}}\right]$ where $x_{s\ell}, \ell = 1, \cdots, \widetilde{m}$, denotes the number of molecules of ℓ-th primary species $S_{s\ell}$, hence $\mathbf{x_s} \in \widetilde{\Omega}_2$, $\widetilde{\Omega}_2 \subset \mathrm{N}^{\widetilde{\mathrm{m}}}$. Furthermore, let M_f be the set of reactions $R_{f_j}, j = 1, \cdots, r_1$, which involve the intermediate species, and let $\mathrm{M_s}$ be the set of reaction $\mathrm{R_{s_k}}$, $\mathrm{k} = 1, \cdots, \mathrm{r}_2$, which involve the primary species. Each reaction R_{f_j} is characterised by a propensity function $\lambda_j(x)$ and a stoichiometric vector $\mathbf{v}_j^f$. Likewise, each reaction R_{s_k} is characterised by a propensity function $\beta_k(x)$ and a stoichiometric vector $\mathbf{v}_k^\mathrm{s}$.

The dynamics of this system can be represented stochastically by a compact matrix-vector form of CME as

$$\frac{dP_\epsilon(x;t)}{dt} = \frac{1}{\epsilon}\mathbf{A}_f\, P_\epsilon\big(x_f, x_s; t\big) + \mathbf{A}_s P_\epsilon\big(x_f, x_s; t\big), \quad \mathbf{A}_f, \mathbf{A}_s \in R^{\widetilde{\Omega}\times\widetilde{\Omega}} \tag{11}$$

where matrices $\mathbf{A}_f$ and $\mathbf{A}_s$ contain the propensities representing transition rates of intermediate and primary species respectively. Note that since the multiple time scales system discussed here is a singularly perturbed system, we prefer to use the notation $P_\epsilon(\cdot)$ instead of $P(\cdot)$ in Equation (11) and in the following discussion.

To expand the work in Khoo and Hegland [9], the singular perturbation method is applied instead of QSSA to this multiple time scales system to get two sets of solutions in two different regions (boundary layer and outer region). This is discussed in subsections 4.1 and 4.2. Finally, these solutions are matched to obtain the uniform solution in the subsection 4.3.

4.1 Outer Solution

To find the outer approximation to the solution of (11), let's rewrite Equation (11) as $\epsilon\frac{dP_\epsilon(x;t)}{dt} = \mathbf{A}_f P_\epsilon\big(x_f, x_s; t\big) + \epsilon\mathbf{A}_s P_\epsilon\big(x_f, x_s; t\big)$. Then by setting $\epsilon = 0$, we will get $\mathbf{A}_f \hat{P}\big(x_f, x_s; t\big) = 0$. Here $\hat{P}\big(x_f, x_s; t\big) = \hat{P}\big(x_f, x_s\big)$ as $\hat{P}\big(x_f, x_s; t\big)$ is time invariant. Note that $\hat{P}(\cdot)$ instead of $P_\epsilon(\cdot)$ is used here so that the probability distribution of the unperturbed system can be distinguished from the probability distribution of the original system (11).

For a particular fixed x_s, we have

$$\mathbf{A}_{fs}\hat{P}\big(x_f \big| x_s\big) = 0, \tag{12}$$

hence the conditional probability $\hat{P}\big(x_f \big| x_s\big)$ has a stationary distribution. As a consequence, the dynamics of the primary species are the major concern in the outer approximation. The solution of (12) can be used to determine the conditional expectation; that is, $\hat{E}\big[x_f \big| x_s\big] = \sum_{x_f} x_f \hat{P}\big(x_f \big| x_s\big)$.

On the other hand, by summing the Equation (11) over x_f, we obtain a CME for the marginal probability of x_s:

$$\frac{d\widetilde{P_\epsilon}(x_s;t)}{dt} = \sum_{k\in M_s} \tilde{\gamma}_k(x_s - v_k^s)\widetilde{P_\epsilon}(x_s - v_k^s; t) - \tilde{\gamma}_k(x_s)\widetilde{P_\epsilon}(x_s; t) \tag{13}$$

with the propensity function

$$\tilde{\gamma}_k(x_s) = \sum_{x_f} \beta_k\big(x_f, x_s\big)\, \widetilde{P_\epsilon}\big(x_f \big| x_s; t\big) \approx \beta_k\big(\tilde{E}\big[x_f \big| x_s; t\big], x_s\big). \tag{14}$$

The functional $\tilde{\gamma}_k(x_s)$ in (14) is the conditional expectation of the functional $\beta_k(x_f, x_s)$ where $\tilde{E}[x_f|x_s;t] = \sum_{x_f} x_f \widetilde{P}_\epsilon(x_f\,|x_s;t)$.

Now, use $\hat{E}[x_f|x_s]$ as the approximation of $\tilde{E}[x_f|x_s;t]$, Equation (13) can be approximated by

$$\frac{d\hat{P}(x_s;t)}{dt} = \sum_{k\epsilon M_s} \hat{\gamma}_k(x_s - v_k^s)\hat{P}(x_s - v_k^s;t) - \hat{\gamma}_k(x_s)\hat{P}(x_s;t) \tag{15}$$

with the propensity function

$$\hat{\gamma}_k(x_s) = \sum_{x_f} \beta_k\,(x_f, x_s)\hat{P}(x_f|x_s) \approx \beta_k(\hat{E}[x_f|x_s], x_s). \tag{16}$$

Finally, the marginal probability distribution of the intermediate species, $\widehat{\mathrm{P}}(\mathrm{x_f};\mathrm{t})$, can be computed via the formula of conditional probability, that is,

$$\hat{P}(x_f;t) = \sum_{x_s} \hat{P}(x_f|x_s)\hat{P}(x_s;t) \tag{17}$$

Compared to the dimension of the original system (11), both the homogeneous Equation (12) and the CME (15) are of lower dimension; hence the system (Equations (12) and (15)) in this subsection is known as a reduced or degenerate model.

4.2 Inner Solution

Because the outer solution does not provide a good approximation in the boundary layer (inner region) where $t \in [0, \theta]$, an alternative approximation is used to study the singular limit as $\epsilon \to 0$. We introduce a scaled time variable $\tau = \frac{t}{\epsilon}$ to magnify the boundary layer, thereby disclosing the boundary layer jump in this region.

As a consequence, the CME (11) can be reformulated as

$$\begin{aligned}\frac{d\overline{P}_\epsilon(x;\tau)}{d\tau} &= \sum_{j\in M_f} a_j(x_f - v_j^f, x_s)\overline{P}_\epsilon(x_f - v_j^f, x_s;\tau) - a_j(x_f, x_s)\overline{P}_\epsilon(x;\tau)\\ &\qquad +\epsilon \sum_{k\in M_s} \beta_k(x_f, x_s - v_k^s)\,\overline{P}_\epsilon(x_f, x_s - v_k^s;\tau) - \beta_k(x_f, x_s)\overline{P}_\epsilon(x;\tau)\\ &= \mathbf{A}_f\overline{P}_\epsilon(x;\tau) + \epsilon\mathbf{A}_s\overline{P}_\epsilon(x;\tau)\end{aligned} \tag{18}$$

Summing the Equation (20) over x_f and set $\epsilon = 0$ gives

$$\frac{d\check{P}(x_s;\tau)}{d\tau} = 0\;. \tag{19}$$

That means $x_s(\tau) = x_s(0) = x_{s0}$. Hence, the probability distribution of x_f obeys a lower dimension CME:

$$\frac{d\check{P}(x_f, x_{s0};\tau)}{d\tau} = \sum_{j\in M_f} a_j(x_f - v_j^f, x_{s0})\check{P}(x_f - v_j^f, x_{s0};\tau) - a_j(x_f, x_{s0})\check{P}(x_f, x_{s0};\tau) \tag{20}$$

Here $\check{P}(\cdot)$ is used to replace $\overline{P}(\cdot)$ so that we can distinguish the difference between the probability distributions of perturbed (18) and unperturbed (20) systems.

4.3 Matching and Uniform Approximation

The expected values of complex and substrate in both inner and outer regions are computed in this part, these solutions are conjectured to have a unique common limit at the edge of boundary layer. Consider now $\epsilon \to 0$, $\tau \to \infty$ and $t \to 0$ respectively, the common limit of the inner and outer solutions is defined as

$$\lim_{t\to 0} \hat{E}[\cdot\,;\,t] = \lim_{\tau\to\infty} \check{E}[\cdot\,;\,\tau]. \tag{21}$$

The inner and outer solutions are matched if in the limit of $\epsilon \to 0$, the expected value in inner region as $\tau \to \infty$ is equal to the expected value in outer region as $t \to 0$.

Then, the uniform approximation $E_u(t)$, which well-approximates the actual solution on the entire time domain, is defined as

$$E_u(t) = \hat{E}[\cdot\,;\,t] + \check{E}[\cdot\,;\,t/\epsilon] + E_c \tag{22}$$

where $E_c = \lim_{t\to 0} \hat{E}[\cdot\,;\,t] = \lim_{\tau\to\infty} \check{E}[\cdot\,;\,\tau]$ is the common limit of the outer and inner solutions.

In brief, the steps of the method proposed in this paper are summarised in Figure 1. That is, we first model the reactions of the chemical species by using the virtual reaction model proposed in section 2. Next, by referring to the conditions (a)-(d) which have discussed in section 3, we divide the chemical species to primary species and intermediate species into different

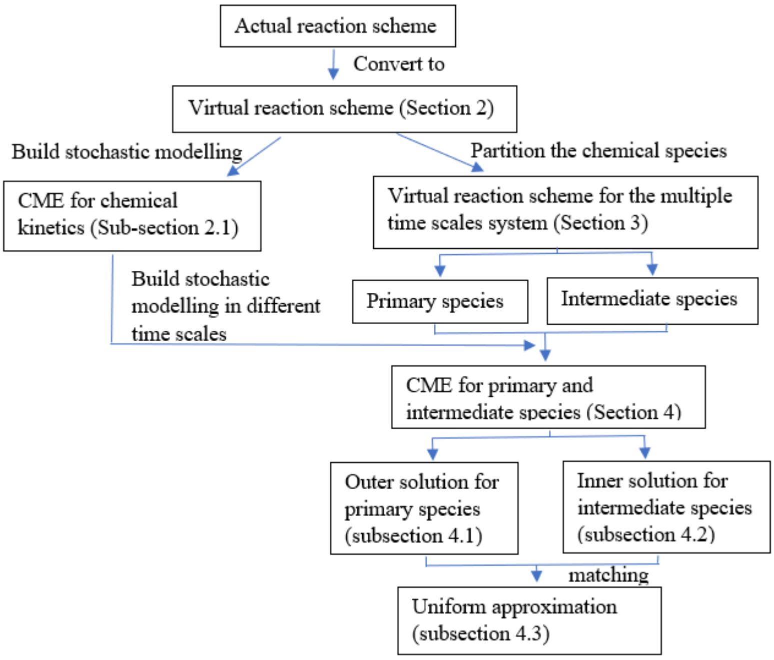


Figure 1. The steps proposed to deal with stochastic modelling of a multiple time scales system

subdomains according to their dynamics. Then by using the singular perturbation method as discussed in section 4, the dynamics of primary and intermediate species are computed separately by using the outer approximation as in subsection 4.1 and inner approximation as in subsection 4.2, finally a uniform approximation is obtained by matching the outer and inner solutions.

5. IMPLEMENTATION

A simple catalytic reaction system will be considered here to illustrate the efficiency of the approach presented in the preceding section.

$$S_1 \xrightarrow{\kappa_1} S_2 \xrightarrow{\kappa_2} S_3 \tag{23}$$

$$S_2 + S_4 \xrightarrow{\kappa_3} S_2 + S_5 \tag{24}$$

with rate constants and initial conditions taken from the paper of Mastny *et al.* [10]: $\kappa_1 = 1, \kappa_2 = 1000, \kappa_3 = 100,\ x_1(0) = 15, x_4(0) = 20, x_2(0) = x_3(0) = x_5(0) = 0$. Furthermore, we choose $\epsilon = 0.005$ for this example. Note that mass conservation laws hold, so that nonzero probabilities only occur for $x_1 + x_2 + x_3 = x_{T1}$ and $x_4 + x_5 = x_{T2}$, where x_{T1} and x_{T2} are positive constants. Hence, we only include x_1, x_2 and x_4 in our modelling.

The reactions (23) and (24) are replaced with the following virtual reaction scheme:

$$\left.\begin{array}{r} S_1 \longrightarrow \emptyset,\ \text{propensity function}\ \beta_1(x_1, x_2, x_4) = \kappa_1 x_1 \\ \emptyset \longrightarrow S_2,\ \text{propensity function}\ \lambda_1(x_1, x_2, x_4) = \kappa_1 x_1 + \kappa_3 x_2 x_4 \\ S_2 \longrightarrow \emptyset,\ \text{propensity function}\ \lambda_2(x_1, x_2, x_4) = \kappa_2 x_2 + \kappa_3 x_2 x_4 \\ \emptyset \longrightarrow S_4,\ \text{propensity function}\ \beta_2(x_1, x_2, x_4) = \kappa_3 x_2 x_4 \end{array}\right\} \tag{25}$$

The dynamics of such system can be represented in the CME as

$$\frac{dP_\epsilon(x_1,x_2,x_4;t)}{dt} = \frac{1}{\epsilon}\mathbf{A}_f P_\epsilon(x_1, x_2, x_4; t) + \mathbf{A}_s P_\epsilon(x_1, x_2, x_4; t) \tag{26}$$

where the reaction's stoichiometric vectors are $\mathbf{v}_1^s = [-1,0,0]^{\mathrm{T}}$, $\mathbf{v}_2^s = [0,0,1]^{\mathrm{T}}$, $\mathbf{v}_1^f = [0,-1,0]^{\mathrm{T}}$ and $\mathbf{v}_2^f = [0,0,1]^{\mathrm{T}}$.

To obtain the outer solution, we set $\epsilon = 0$ and solve

$$\mathbf{A}_{fs}\hat{P}(x_2|[x_1, x_4]) = 0, \tag{27}$$

with propensity functions $a_1(x_1, x_2, x_4) = 0.005x_1 + 0.5x_2x_4$ and $a_2(x_1, x_2, x_4) = 5x_2 + 0.5x_2x_4$ for given x_1 and x_4, as well as

$$\frac{d\hat{P}(x_1,x_4;t)}{dt} = \widehat{\mathbf{A}}_s\hat{P}(x_1, x_4; t) \tag{28}$$

where $\widehat{\mathbf{A}}_s$ contains the propensity functions $\beta_1(x_1, x_2, x_4)$ and $\beta_2(x_1, x_2, x_4)$.

To obtain the inner solution, we set $\tau = \frac{t}{\epsilon}$ and solve

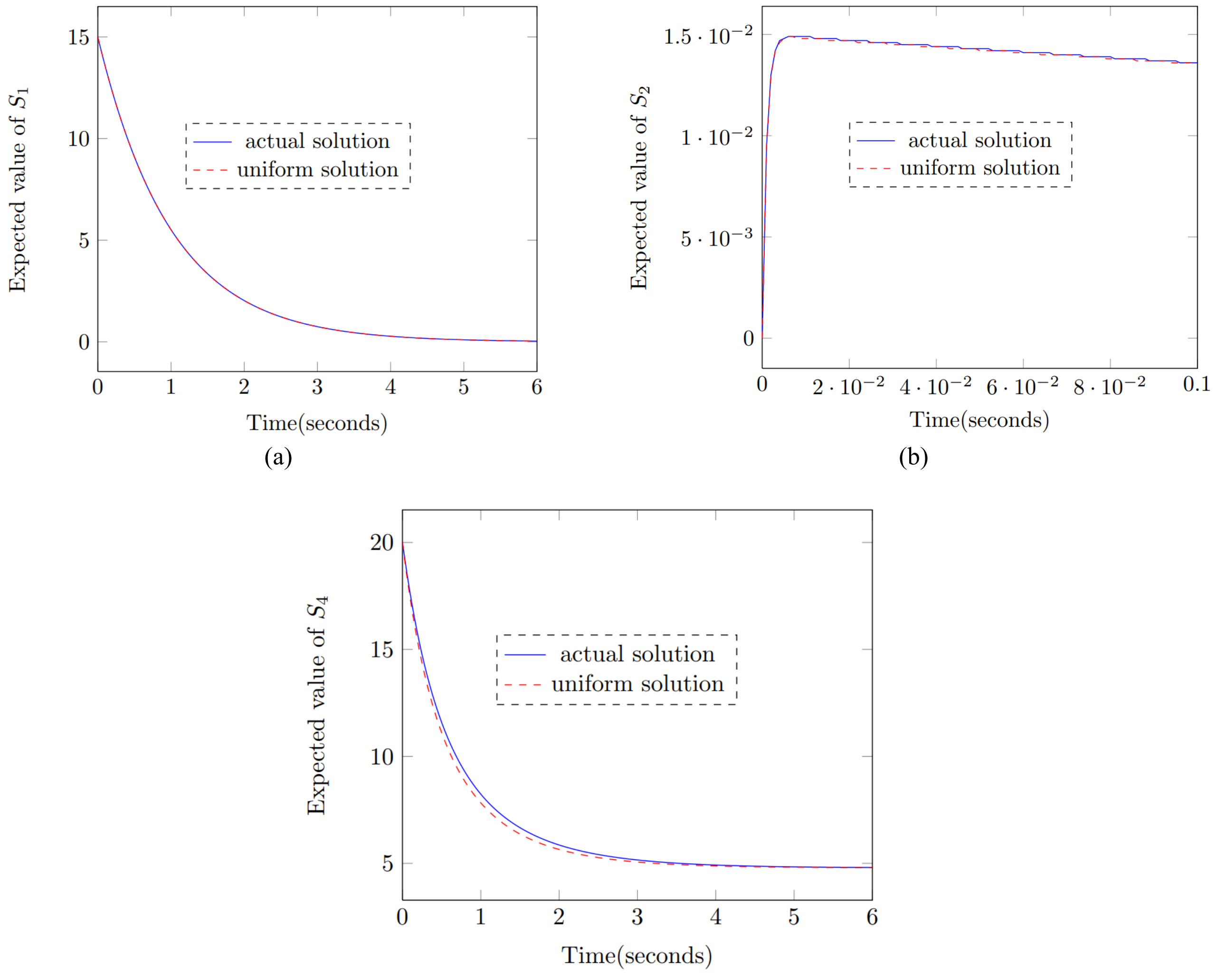
$$\frac{d\check{P}(15,x_2,0;\tau)}{d\tau} = \overline{\mathbf{A}}_f \check{P}(15, x_2, 0; \tau) \tag{29}$$

with propensity functions $\bar{\alpha}_1(15, x_2, 0;) = 0.05x_1 + 0.5x_2x_4$ and $\bar{\alpha}_2(15, x_2, 0;) = 5x_2 + 0.5x_2x_4$. Then the expected values of S_1, S_2 and S_4 are determined for both outer and inner approximation, match them to get the common limit, and determine their uniform approximation.

The results are shown in Figure 2(a), Figure 2(b) and Figure 2(c). The computed expected value of molecular species S_1, S_2 and S_3 obtained via the simulation of reaction scheme (23) and (24) are presented in the blue dashed line. In comparison with the actual solution (blue solid line) which obtained by solving the CME of actual reaction scheme directly, the maximum percent relative errors throughout the time domain are 0.1%, 0.8% and 5.0% respectively. Note here the formula for percent relative error that we used here is

$$\text{percent relative error} = \frac{|\text{actual solution} - \text{uniform approximation}|}{|\text{actual solution}|} \times 100\% \tag{30}$$

As a consequence, we can say the virtual reaction scheme in conjunction with the singular perturbation method has actually provided quite a reasonably good approximation. The most important point is this approach can be done in lower dimension, so it may help in dealing with the high dimensional computational works in chemical kinetics.



(a)

(b)

(c)

Figure 2. The computed expected value of molecular species S_1, S_2 and S_3 obtained via the simulation of reactions (23) and (24) with the parameters $x_1(0) = 15, x_4(0) = 20, x_2(0) = S_3(0) = x_5(0) = 0, \kappa_1 = 1, \kappa_2 = 1000$ and $\kappa_3 = 100$.

6. CONCLUSION

As can be seen from the example discussed in the Section 5, our approach which uses the virtual reaction model and singular perturbation has provided a good approximation to the multiple time scales system. The example regarding catalytic reaction system has been discussed carefully, and the results show the dimension of the problems can be reduced by solving the problem separately in different subdomains. Therefore, we can say such a "divide and conquer" approach enables us to deal with high dimensional stochastic problems. Moreover, this method can actually work in conjunction with some other techniques, such as the aggregation [11], sparse grid method [12] or finite state projection [13] which will then allow us to deal with even higher dimensional stochastic problems.

REFERENCES

[1] E. L. Haseltine and J. B. Rawlings, Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics, *Journal of Chemical Physics*, 117, 2002, 6959-6969.

[2] C. Rao and A. Arkin, Stochastic chemical kinetics and the quasi-steady-state assumption: application to the Gillespie algorithm, *Journal of Chemical Physics*, 118, 2003, 4999-5010.

[3] Y. Cao, D. T. Gillespie and L. R. Petzold, The slow-scale stochastic simulation algorithm, *Journal of Chemical Physics*, 122, 2005, 014116-1-18.

[4] J. Goutsias, Quasi equilibrium approximation of fast reaction kinetics in stochastic biochemical systems, *Journal of Chemical Physics*, 122, 2005, 184102-1-15.

[5] C. H. Lee and R. Lui, A reduction method for multiple time scale stochastic reaction networks, *Journal of Mathematical Chemistry,* 46, 2009**,** 1292–132.

[6] 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, 2016, 1081–1129.

[7] P. Lecca, F. Bagagiolo and M. Scarpa, Hybrid deterministic/stochastic simulation of complex biochemical systems, *Molecular Biosystems*, 3(12), 2017, 2672-2686.

[8] J. K. Kim and E. D. Sontag, Reduction of multiscale stochastic biochemical reaction networks using exact moment derivation, *PLoS Computational Biology,* 13(6), 2017, S1-S7.

[9] C. F. Khoo and M. Hegland, The total quasi-steady state assumption: its justification and its application to the chemical master equation, *ANZIAM Journal (CTAC2008)*, 50, 2008, C429-C443.

[10] A. E. Mastny, E. L. Haseltine and J. L. Rawlings, Two classes of quasi-steady-state model reductions for stochastic kinetics, *Journal of Chemical Physics*, 127, 2007, 094106-1-16.

[11] M. Hegland, C. Burden, L. Santoso, S. MacNamara and H. Booth, A solver for the stochastic master equation applied to gene regulatory network, *Journal of Computational and Applied Mathematics*, 205, 2006, 1-17.

[12] J. Garcke, Sparse grids in a nutshell, in *Sparse Grids and Applications*, J. Garcke and M. Griebel (eds), Berlin: Springer, 88, 2012, 57-80

[13] B. Munsky and M. Khammash, The finite state projection algorithm for the solution of chemical master equation, *Journal of Chemical Physics,* 124, 2006, 044104-1-13.

.