MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **73** (2015) 691-704

ISSN 0340 - 6253

An Analytic Approach to a Stochastic Enzyme Kinetic Model

Hyojung Lee, Chang Hyeong Lee*

Ulsan National Institute of Science and Technology(UNIST)

Department of Mathematical Sciences, School of Natural Science

Ulsan Metropolitan City 689-798, Republic of Korea

chlee@unist.ac.kr

(Received October 29, 2014)

Abstract

In this paper we study a stochastic model of the enzyme-substrate reaction system. We investigate the block structure of the transition rate matrix of the model and find a formula for the exact solution to the chemical master equation of the model. The solution of the model is represented in terms of the eigenvalues and eigenvectors of block matrices whose dimensions are much lower than the original transition rate matrix. The method presented in the paper can reduce greatly computational complexity when the solution of stochastic enzyme-substrate model is sought. We show the accuracy and efficiency of the method by simulating examples.

1 Introduction

Stochastic modeling of chemical reaction systems is used when researches focus on the systems with small number of molecular species. The stochastic models describe the time evolution of the probability of states defined as $\mathbf{n}(t) = (n_1(t), n_2(t), \ldots, n_s(t))^T$ and each $n_i(t)$ denotes the number of molecules of i^{th} species at time t. The time-dependent probability solution is described by the chemical master equation

$$\frac{\partial}{\partial t}p(\mathbf{n}, t) = \sum_{k=1}^{r} a_k(\mathbf{n} - V_k) \cdot p(\mathbf{n} - V_k, t) - \sum_{k} a_k(\mathbf{n}) \cdot p(\mathbf{n}, t), \tag{1}$$

*Corresponding author

where $p(\mathbf{n}, t)$ is the probability of $\mathbf{n}$ at time t, a_k is the propensity function for the k^{th} reaction and V_k is the k^{th} column vector of the stoichiometric matrix V[1]. Moreover, by finding the transition rates between all possible states of $\mathbf{n}$, one can construct a Markov chain whose governing equation is written as the linear system

$$\frac{d\mathbf{p}(t)}{dt} = K\mathbf{p}(t), \tag{2}$$

where K is the matrix of transition rates between all states and $\mathbf{p}$ is the vector of probabilities of the states. One can see that K is a Markov chain generator that satisfies $K_{ij} \geq 0$ for each $i \neq j$ and $K_{jj} = -\sum_{i \neq j} K_{ij}$ for each j[2]. One can find the solution of (2) in the formal form

$$\mathbf{p}(t) = e^{Kt}\mathbf{p}(0). \tag{3}$$

Note that K in (3) is usually a large dimensional matrix in the stochastic models of complex chemical reaction networks and so it is very difficult(or impossible if K is infinite-dimensional) to find the solution of (3) computationally. However, if a reaction network has relatively small number of species and reactions, it may be possible to find the solution of the stochastic model analytically or computationally. One of such models is the enzyme-substrate reaction system which is a fundamental and essential enzyme kinetic model. Enzyme kinetic models are very common in many biochemical systems. Many biochemical reactions can be considered as enzyme kinetic models, including transcription and translation in gene regulatory networks [3, 4], a phase transition in a cell cycle[5] and the Goldbeter-Koshland ultrasensitive switch[6] . The enzyme-substrate model has three different reactions; binding of the enzyme E and the substrate S, unbinding of the enzyme-substrate complex ES and creation of the product P. The mechanism of the model is described by

$$E + S \underset{\longleftarrow}{\longrightarrow} ES \rightarrow E + P. \tag{4}$$

After Michaelis and Menten's pioneering work[7], researches on the deterministic solutions of the enzyme kinetic models have been made by many researchers even recently[8, 9, 10]. For the stochastic model, an analytic solution of (4) was first obtained for the case of a single enzyme molecule by Arányi and Tóth[11]. After their works, many researches have been done for finding the solution of the stochastic model[12, 13, 14]. An analysis of the statistical moments was recently made for a general linear compartment model of chemical kinetic systems including enzyme systems[15].

In this paper we explicitly represent an analytic formula of the solution to the chemical master equation of (4) with general initial conditions by using the block structure of the Markov chain generator K. To the authors' knowledge, the explicit formula of the solution to the stochastic enzyme-substrate model with general initial conditions has not been reported yet.

The outline of the paper is as follows. In Section 2, the analytic expressions of the solution to the master equation of the stochastic enzyme-substrate model are developed. In Section 3, the computational procedure for finding the solution for any given initial condition is presented. Also, we discuss the computational complexity of our method and show the numerical results of examples to illustrate the accuracy and efficiency of our method. Throughout this paper, vectors are boldfaced and matrices are capitalized.

2 Analytic solution of stochastic enzyme-substrate system

In this section we find an analytic solution of (3) for the stochastic enzyme-substrate reaction system

$$E + S \underset{c_{-1}}{\overset{c_1}{\rightleftarrows}} ES \overset{c_2}{\to} E + P, \tag{5}$$

where c_1, c_{-1} and c_2 are probability constants for each reaction. To describe the stochastic model, we denote the number of molecules of E, S, ES, P at time t by $n_1(t), n_2(t), n_3(t)$, $n_4(t)$ and let $\mathbf{n}(t) = (n_1(t), n_2(t), n_3(t), n_4(t))^T$. The stoichiometric matrix is given by

$$V = [\mathbf{v}_1|\mathbf{v}_2|\mathbf{v}_3] = \begin{bmatrix} -1 & 1 & 1 \\ -1 & 1 & 0 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

The stochastic dynamics of this system can be completely described by a Markov chain. For example, if an initial condition is given as $\mathbf{n}(0) = (2, 3, 0, 0)$, then the Markov chain is illustrated as follows;

$$
\begin{array}{ccccccc}
(2,3,0,0) & \underset{c_{-1}}{\overset{6c_1}{\rightleftarrows}} & (1,2,1,0) & \underset{2c_{-1}}{\overset{2c_1}{\rightleftarrows}} & (0,1,2,0) & & \\
 & & \downarrow{\scriptstyle c_2} & & \downarrow{\scriptstyle 2c_2} & & \\
 & & (2,2,0,1) & \underset{c_{-1}}{\overset{4c_1}{\rightleftarrows}} & (1,1,1,1) & \underset{2c_{-1}}{\overset{c_1}{\rightleftarrows}} & (0,0,2,1) \\
 & & & & \downarrow{\scriptstyle c_2} & & \downarrow{\scriptstyle 2c_2} \\
 & & & & (2,1,0,2) & \underset{c_{-1}}{\overset{2c_1}{\rightleftarrows}} & (1,0,1,2) \\
 & & & & & & \downarrow{\scriptstyle c_2} \\
 & & & & & & (2,0,0,3).
\end{array}
\tag{6}
$$

Its governing equation is

$$\frac{d\mathbf{p}}{dt} = K\mathbf{p},$$

where the Markov chain generator K is

$$K = \begin{bmatrix} K_1 & & & \\ K_{21} & K_2 & & \\ & K_{32} & K_3 & \\ & & K_{43} & K_4 \end{bmatrix}$$

and

$$K_1 = \begin{bmatrix} -6c_1 & c_{-1} & 0 \\ 6c_1 & -(2c_1 + c_{-1} + c_2) & 2c_{-1} \\ 0 & 2c_1 & -(2c_{-1} + 2c_2) \end{bmatrix},$$

$$K_2 = \begin{bmatrix} -4c_1 & c_{-1} & 0 \\ 4c_1 & -(c_1 + c_{-1} + c_2) & 2c_{-1} \\ 0 & c_1 & -(2c_{-1} + 2c_2) \end{bmatrix}, \; K_3 = \begin{bmatrix} -2c_1 & c_{-1} \\ 2c_1 & -(c_{-1} + c_2) \end{bmatrix}, \; K_4 = 0,$$

$$K_{21} = \begin{bmatrix} 0 & c_2 & 0 \\ 0 & 0 & 2c_2 \\ 0 & 0 & 0 \end{bmatrix}, \; K_{32} = \begin{bmatrix} 0 & c_2 & 0 \\ 0 & 0 & 2c_2 \end{bmatrix}, \text{ and } \; K_{43} = \begin{bmatrix} 0 & c_2 \end{bmatrix}.$$

For the initial condition $\mathbf{n}(0) = (e_0, s_0, 0, 0)^T$, we define the components D_i as

$$D_i = \{\mathbf{n}(0) + a\mathbf{v}_1 + b\mathbf{v}_2 + (i-1)\mathbf{v}_3 \geq \mathbf{0}, a, b \text{ are positive integers}\},$$

for $i = 1, \ldots, N$ and N is the number of components. Here the inequality $\geq$ means that all entries of the vector are nonnegative. We denote the j^{th} state of D_i by the states $S_j^{(i)}, j = 1, \ldots, m_i$, where m_i is the number of states in the component D_i. For example, the Markov chain (6) has $D_1 = \{S_1^{(1)} = (2,3,0,0), S_2^{(1)} = (1,2,1,0), S_3^{(1)} = (1,2,1,0)\}$, $D_2 = \{S_1^{(2)} =$

$(2,2,0,1), S_2^{(2)} = (1,1,1,1), S_3^{(2)} = (0,0,2,1)\}$, $D_3 = \{S_1^{(3)} = (2,1,0,2), S_2^{(3)} = (1,0,1,2)\}$ and $D_4 = \{S_1^{(4)} = (2,0,0,3)\}$.

One can find that the number of components is $N = s_0 + 1$. Moreover, for the two cases $e_0 \geq s_0$ and $e_0 < s_0$, one can compute the number m_i of states in each component D_i and the number J of the total states; If $e_0 \geq s_0$, then

$$m_i = N + 1 - i, \quad i = 1, 2, \ldots, N$$

and $J = \frac{(s_0+1)(s_0+2)}{2}$. If $e_0 < s_0$, then

$$m_i = \begin{cases} e_0 + 1, & i = 1, \ldots, s_0 - e_0 + 1 \\ s_0 + 2 - i, & i = s_0 - e_0 + 2, \ldots, N \end{cases}$$

and $J = (s_0 - e_0 + 1)(e_0 + 1) + \frac{e_0(e_0+1)}{2}$.

In a generic case, the Markov chain generator for the enzyme-substrate system is

$$K = \begin{bmatrix} K_1 & & & & & & \\ K_{2,1} & K_2 & & & & & \\ & K_{3,2} & K_3 & & & & \\ & & \ddots & \ddots & & & \\ & & & \ddots & K_{n-1} & & \\ & & & & K_{N,N-1} & K_N \end{bmatrix},$$

where $K_{i+1,i}$ is the $m_{i+1} \times m_i$ block matrix that represents transition rates or propensities from the states of component D_i to those of D_{i+1}. Notice that since K is a block lower triangular matrix which consists of diagonal blocks K_i and lower triangular blocks $K_{i+1,i}$, all eigenvalues of K consist of the eigenvalues of the diagonal block $K_i, i = 1, \ldots, N$. Moreover, since each K_i is a real tridiagonal matrix with $[K_i]_{j,j+1} \cdot [K_i]_{j+1,j} > 0$, it has all real and distinct eigenvalue[16] and thus the eigenvectors of each K_i are linearly independent.

Remark 1. *Since each diagonal block K_i is a tridiagonal matrix, one can find the eigenvalues of each K_i by solving the recursive equations presented in [16]. To compute the eigenvectors of K_i, one can use numerical methods such as Inverse iteration[17], Divide and Conquer method[18] and Multiple Relatively Robust Representations algorithm(MR^3)[19].*

Now we let Λ be the diagonal matrix of eigenvalues

$$\Lambda = diag\left(\Lambda^{(1)}, \Lambda^{(2)}, \cdots, \Lambda^{(N)}\right),$$

where $\Lambda^{(i)} = diag\left(\lambda_1^{(i)}, \cdots, \lambda_{m_i}^{(i)}\right),\ i = 1, \ldots, N$ and $\lambda_p^{(i)},\ p = 1, \ldots, m_i$ are eigenvalues of K_i. To find the eigenvectors of K, let us denote V as the matrix of eigenvectors of K

$$V = \left[V^{(1)}|V^{(2)}|\cdots|V^{(N)}\right],$$

where each $V^{(i)},\ i = 1, \ldots, N$ is a submatrix whose columns are the eigenvectors $\mathbf{v}_p^{(i)}$ corresponding to $\lambda_p^{(i)}, p = 1, \ldots, m_i$, denoted by

$$\mathbf{v}_p^{(i)} = \begin{bmatrix} \mathbf{v}_p^{(1,i)} \\ \mathbf{v}_p^{(2,i)} \\ \vdots \\ \mathbf{v}_p^{(N,i)} \end{bmatrix},$$

where each $\mathbf{v}_p^{(j,i)}, j = 1, \ldots, N$ is the $m_j \times 1$ column vector. Now we construct the lower triangular block matrix V; To find the eigenvectors $\mathbf{v}_p^{(1)}, p = 1, \ldots, m_1$, corresponding to the eigenvalue $\lambda_p^{(1)}$ of K_1, we use $K\mathbf{v}_p^{(1)} = \lambda_p^{(1)}\mathbf{v}_p^{(1)}$ and obtain

$$K_1\mathbf{v}_p^{(1,1)} = \lambda_p^{(1)}\mathbf{v}_p^{(1,1)},\quad K_{i,i-1}\mathbf{v}_p^{(i-1,1)} + K_i\mathbf{v}_p^{(i,1)} = \lambda_p^{(1)}\mathbf{v}_p^{(i,1)},\ i = 2, \ldots, N \tag{7}$$

From (7), note that $\mathbf{v}_p^{(1,1)}, p = 1, \ldots, m_1$ are the eigenvectors of K_1 corresponding to $\lambda_p^{(1)}$ and also one finds the recursive equations for $\mathbf{v}_p^{(i,1)}, i = 2, \ldots, N$;

$$\left(K_i - \lambda_p^{(1)} I_{m_i}\right)\mathbf{v}_p^{(i,1)} = -K_{i,i-1}\mathbf{v}_p^{(i-1,1)},\quad i = 2, \ldots, N. \tag{8}$$

To find $\mathbf{v}_p^{(2)}, p = 1, \ldots, m_2$, corresponding to the eigenvalue $\lambda_p^{(2)}$ of K_2, we write

$$K_1\mathbf{v}_p^{(1,2)} = \lambda_p^{(2)}\mathbf{v}_p^{(1,2)},\ K_{i,i-1}\mathbf{v}_p^{(i-1,2)} + K_i\mathbf{v}_p^{(i,2)} = \lambda_p^{(2)}\mathbf{v}_p^{(i,2)},\ i = 2, \ldots, N. \tag{9}$$

If we choose $\mathbf{v}_p^{(1,2)}$ as a zero vector, then $K_{2,1}\mathbf{v}_p^{(1,2)} = \mathbf{0}$ and thus (9) can be rewritten as

$$K_2\mathbf{v}_p^{(2,2)} = \lambda_p^{(2)}\mathbf{v}_p^{(2,2)},\ \left(K_i - \lambda_p^{(2)} I_{m_i}\right)\mathbf{v}_p^{(i,2)} = -K_{i,i-1}\mathbf{v}_p^{(i-1,2)},\ i = 3, \ldots, N. \tag{10}$$

One can see from (10) that $\mathbf{v}_p^{(2,2)}$ is the eigenvector corresponding to the eigenvalue $\lambda_p^{(2)}$ of K_2. If we repeat the similar procedure for the eigenvalues $\lambda_p^{(j)}, j = 3, \ldots, N-1$, we find the eigenvector $\mathbf{v}_p^{(j)}$ of K corresponding to $\lambda_p^{(j)}$ in the form of

$$\mathbf{v}_p^{(j)} = \begin{bmatrix} \mathbf{0}_1 \\ \vdots \\ \mathbf{0}_{j-1} \\ \mathbf{v}_p^{(j,j)} \\ \vdots \\ \mathbf{v}_p^{(N,j)} \end{bmatrix},$$

where $\mathbf{0}_i$, $i = 1, \cdots, j-1$ is the $m_i \times 1$ zero vector, $\mathbf{v}_p^{(j,j)}$ is the eigenvector of K_j corresponding to $\lambda_p^{(j)}$, and each $\mathbf{v}_p^{(i,j)}$, $i = j+1, \cdots, N$, satisfies

$$\left(K_i - \lambda_p^{(j)} I_{m_i}\right) \mathbf{v}_p^{(i,j)} = -K_{i,i-1} \mathbf{v}_p^{(i-1,j)} \tag{11}$$

Since $K_N = 0$, the eigenvector $\mathbf{v}_1^{(N)}$ corresponding to the zero eigenvalue can be chosen as the $m_N \times 1$ vector $(0, \ldots, 0, 1)^T$.

Through the above construction of eigenvectors of K, we have the matrix V as

$$V = \left[V^{(1)}|V^{(2)}|V^{(3)}|\cdots|V^{(N)}\right] = \left[\begin{array}{c|c|c|c|c} V^{(1,1)} & 0 & 0 & \cdots & 0 \\ V^{(2,1)} & V^{(2,2)} & 0 & \cdots & 0 \\ V^{(3,1)} & V^{(3,2)} & V^{(3,3)} & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ V^{(N,1)} & V^{(N,2)} & \cdots & \cdots & V^{(N,N)} \end{array}\right],$$

where each $V^{(i,i)}$ denotes the square matrix whose columns are eigenvectors of the block $K_i, i = 1, \ldots, N$, that is, $V^{(i,i)} = \left[\mathbf{v}_1^{(i,i)}|\cdots|\mathbf{v}_{m_i}^{(i,i)}\right]$, $i = 1, \ldots, N-1$ and $V^{(N,N)} = 1$. Moreover, $V^{(i,j)} = \left[\mathbf{v}_1^{(i,j)}|\cdots|\mathbf{v}_{m_j}^{(i,j)}\right]$ for $i \geq j = 1, \ldots, N-1$. Next we show the independence of eigenvectors of K, which is crucial for finding the exact formula for the solution of (2).

Theorem 2. *The eigenvectors of K are linearly independent.*

Proof. To show the independence of the eigenvectors $\mathbf{v}_p^{(i)}, i = 1, \ldots, N, p = 1, \ldots, m_i$ of K, we prove that if $\sum_{i=1}^{N} \sum_{p=1}^{m_i} a_{i,p} \mathbf{v}_p^{(i)} = \mathbf{0}_J$, we must have $a_{i,p} = 0$ for all $i = 1, \ldots, N, p = 1, \ldots, m_i$. Here $\mathbf{0}_J$ denotes the $J \times 1$ zero vector. Since

$$\sum_{i=1}^{N} \sum_{p=1}^{m_i} a_{i,p} = \left[\begin{array}{c} \sum_{j=1}^{m_1} a_{1,p} \mathbf{v}_p^{(1,1)} \\ \sum_{k=1}^{2} \sum_{p=1}^{m_k} a_{k,p} \mathbf{v}_p^{(2,k)} \\ \sum_{k=1}^{3} \sum_{p=1}^{m_k} a_{k,p} \mathbf{v}_p^{(3,k)} \\ \vdots \\ \sum_{k=1}^{N} \sum_{p=1}^{m_k} a_{k,p} \mathbf{v}_p^{(N,k)} \end{array}\right] = \mathbf{0}_J,$$

$\sum_{p=1}^{m_1} a_{1,p} \mathbf{v}_p^{(1,1)} = \mathbf{0}_{m_1}$, which leads to all $a_{1,p} = 0$ for $p = 1, \ldots, m_1$. Moreover, since

$$\sum_{k=1}^{2} \sum_{p=1}^{m_k} a_{k,p} \mathbf{v}_p^{(2,k)} = \sum_{j=1}^{m_1} a_{1,p} \mathbf{v}_p^{(2,1)} + \sum_{p=1}^{m_2} a_{2,p} \mathbf{v}_p^{(2,2)} = \mathbf{0}_{m_2}$$

and $a_{1,p} = 0$, $a_{2,p} = 0$ for $p = 1, \ldots, m_2$. A similar argument can be applied to show that $a_{i,p} = 0$ for $i = 3, \ldots, N-1$ and $p = 1, \ldots, m_i$. Lastly, since

$$\sum_{k=1}^{N}\sum_{p=1}^{m_k} a_{k,p}\mathbf{v}_p^{(N,k)} = \sum_{k=1}^{N-1}\sum_{p=1}^{m_k} a_{k,p}\mathbf{v}_p^{(N,k)} + \sum_{p=1}^{m_N} a_{N,p}\mathbf{v}_p^{(N,N)} = \mathbf{0}_{m_N}$$

and $a_{1,p} = a_{2,p} = \cdots = a_{N-1,p} = 0$, we must have $a_{N,p}$=0. □

Due to Theorem 2, the matrix V of eigenvectors of K is invertible and the matrix K is diagonalizable. Thus, we can write from (3)

$$P(t) = e^{Kt} = Ve^{\Lambda t}V^{-1}. \tag{12}$$

To compute V^{-1} more conveniently, we define the matrices $C^{(i,j)}, i \geq j = 1, \ldots, N$ as

$$C^{(i,i)} = I_{m_i} \text{ and } C^{(i,j)} = (V^{(i,i)})^{-1}V^{(i,j)} \text{ for } i > j = 1, \ldots, N-1. \tag{13}$$

If $\mathbf{c}_p^{(i,j)}$ and $C_{k,p}^{(i,j)}$ denote the p^{th} column vector of $C^{(i,j)}$ and the $(k,p)^{th}$ entry of $C^{(i,j)}$, respectively, then (13) can be written as

$$\mathbf{v}_p^{(i,j)} = V^{(i,i)}\mathbf{c}_p^{(i,j)} = \sum_{k=1}^{m_i} C_{k,p}^{(i,j)}\mathbf{v}_k^{(i,i)} \text{ for } i > j. \tag{14}$$

Using (11) and (14), we compute $\mathbf{c}_p^{(i,j)}$ by solving

$$\left(K_i - \lambda_p^{(j)}I_{m_i}\right)\left[V^{(i,i)}\right]\left[\mathbf{c}_p^{(i,j)}\right] = -K_{i,i-1}\mathbf{v}_p^{(i-1,j)}. \tag{15}$$

for $i = 2, \ldots, (N-1), j = 1, \ldots, (i-1)$ and $p = 1, \ldots, m_j$, and

$$\mathbf{c}_p^{(N,j)} = \frac{K_{N,N-1}\mathbf{v}_p^{(N-1,j)}}{\lambda_p^{(j)}}, \tag{16}$$

for $j = 1, \ldots, N-1$. After computing (15) and (16), one can find all $C^{(i,j)}, i > j$. Since the inverse of V is also a block lower triangular matrix, one can write

$$V^{-1} = \begin{bmatrix} W^{(1,1)} & & & & & & \\ W^{(2,1)} & W^{(2,2)} & & & & & \\ W^{(3,1)} & W^{(3,2)} & W^{(3,3)} & & & & \\ \vdots & \vdots & \ddots & \ddots & & & \\ W^{(N-1,1)} & W^{(N-1,2)} & & \ddots & W^{(N-1,N-1)} & & \\ W^{(N,1)} & W^{(N,2)} & \ldots & & W^{(N,N-1)} & W^{(N,N)} \end{bmatrix}. \tag{17}$$

Using $VV^{-1} = I$, one can compute the blocks of V^{-1} as

$$W^{(i,j)} = \begin{cases} \left(V^{(i,i)}\right)^{-1}, & \text{if } i = j \\ -C^{(i+1,i)}W^{(i,i)} & \text{if } i = j+1 \\ -\left[\sum_{k=j}^{i-1} C^{(i,k)}W^{(k,j)}\right] & \text{if } i > j+1 \end{cases} \tag{18}$$

where $i, j = 1, 2, \ldots, N$. Thus, we can write

$$P^{(i,j)}(t) = \left(e^{Kt}\right)^{(i,j)} = \begin{cases} V^{(i,i)}e^{\Lambda^{(i)}t}W^{(i,i)} & \textit{for } i = j \\ \sum_{k=j}^{i} V^{(i,k)}e^{\Lambda^{(k)}t}W^{(k,j)} & \textit{for } i > j \end{cases} \tag{19}$$

where $P^{(i,j)}(t)$ denotes the (i,j) block matrix of $P(t) = e^{Kt}$ for $i, j = 1, \ldots, N$. That is,

$$P_{r,q}^{(i,j)}(t) = \sum_{k=j}^{i}\sum_{h=1}^{m_k}\sum_{g=1}^{m_i} C_{g,h}^{(i,k)}V_{r,g}^{(i,i)}W_{h,q}^{(k,j)}e^{\lambda_h^{(k)}t}, \tag{20}$$

where $P_{r,q}^{(i,j)}$, $C_{r,q}^{(i,j)}$, $W_{r,q}^{(i,j)}$ denote the $(r,q)^{th}$ entry of $P^{(i,j)}$, $C^{(i,j)}$ and $W^{(i,j)}$, respectively.

Suppose the initial condition $\mathbf{p}(0)$ is given usually as $\mathbf{p}(0) = (1, 0, 0, \cdots, 0)^T$, which is equivalent to the deterministic condition $\mathbf{n}(0) = (e_0, s_0, 0, 0)^T$. If the k^{th} state is the ℓ^{th} state $S_\ell^{(i)}$ of i^{th} component, i.e., $k = \ell + \sum_{j=1}^{i-1} m_j$, then the exact probability of the k^{th} state is

$$p_k(t) = p_\ell^{(i)}(t) = \sum_{k=j}^{i}\sum_{h=1}^{m_k}\sum_{g=1}^{m_i} C_{g,h}^{(i,k)}V_{\ell,g}^{(i,i)}W_{h,1}^{(k,1)}e^{\lambda_h^{(k)}t}. \tag{21}$$

Moreover, we can find the exact probability for any number of each species at any time as follows; the probability of the number of the product P is

$$P\left(n_4(t) = i-1\right) = \sum_{\ell=1}^{m_i} p_\ell^{(i)}(t) = \sum_{\ell=1}^{m_i}\sum_{k=1}^{i}\sum_{h=1}^{m_k}\sum_{g=1}^{m_i} C_{g,h}^{(i,k)}V_{\ell,g}^{(i,i)}W_{h,1}^{(k,1)}e^{\lambda_h^{(k)}t},$$

for $i = 1, \ldots, s_0 + 1$.

For the substrate S, if $e_0 \geq s_0$, then

$$P\left(n_2(t) = j\right) = \sum_{i=1}^{N-j} p_{\ell_i}(t),$$

where $\ell_i = \sum_{k=1}^{i} m_k - j$ and $j = 0, 1, \cdots, s_0$. If $e_0 < s_0$, then

$$P\left(n_2(t) = j\right) = \begin{cases} \sum_{i=1}^{e_0+1} p_{e_0+2-i+k(s_0-e_0+i-j)}(t) & \text{if } j = 0, 1, \cdots, s_0 - e_0, \\ \sum_{i=1}^{s_0+1-j} p_{s_0+2-i-j+k(i)}(t) & \text{if } j = s_0 - e_0 + 1, \cdots, s_0, \end{cases}$$

where $k(a) = \sum_{j=1}^{a-1} m_j$. Using the conservation quantities $n_1+n_3+n_4 = e_0, n_2+n_3+n_4 = s_0$ and the above result for probabilities for n_2 and n_4, we can also find the probability for n_1 and n_3.

Note that using the similar procedures as discussed in this section, one can find the probability solution of all states from (20) for any initial condition $\mathbf{p}(0)$;

$$p_k(t) = p_\ell^{(i)}(t) = \sum_{s=1}^{i}\sum_{n=1}^{m_s} P_{\ell,n}^{(i,s)}(t)\ p_{n+k(s)}(0)$$
$$= \sum_{s=1}^{i}\sum_{n=1}^{m_s}\sum_{k=s}^{i}\sum_{h=1}^{m_k}\sum_{g=1}^{m_i} C_{g,h}^{(i,k)} V_{\ell,g}^{(i,i)} W_{h,n}^{(k,s)} e^{\lambda_h^{(k)} t}\ p_{n+k(s)}(0),$$

where $k(a) = \sum_{j=1}^{a-1} m_j$.

3 Numerical Computation

In this section, we discuss the numerical efficiency and accuracy of the method presented in the previous section. We first summarize the computation procedure for finding the exact probability solution of the stochastic enzyme-substrate model.

Computational Procedure

Step 0 Construct K and identify the block matrices $K_1, \cdots, K_n$ and $K_{2,1}, \cdots, K_{N,N-1}$.

Step 1. Find the matrix $\Lambda^{(i)}$ of eigenvalues and the matrix $V^{(i,i)}$ of eigenvectors of K_i. (Refer to Remark 1.)

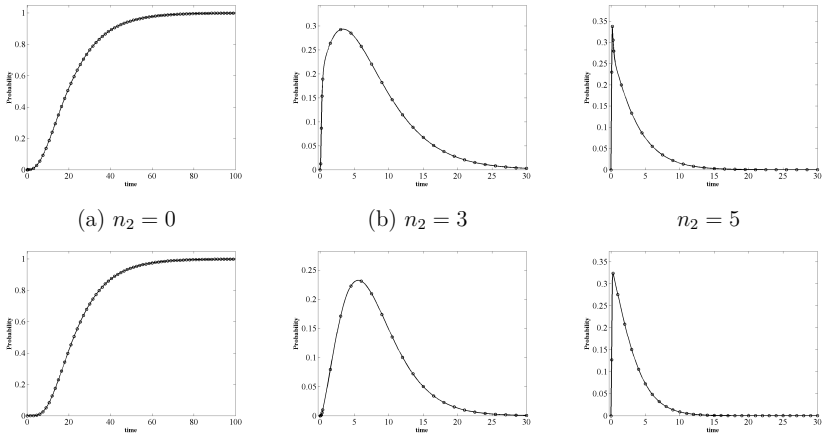
Step 2. Find $C^{(i,j)}$ and $V^{(i,j)}$ by (14), (15) and (16) and then compute $W^{(i,j)}$ by (18).

Step 3. Compute $P^{(i,j)}(t)$ using the explicit formula (19) or (20). If the initial condition $\mathbf{n}(0)$ is given, compute the probability of each state by (21).

Concerning the computational complexity, if $e_0 \geq s_0$, one can see that the matrix exponential solution (12) has $O(J^3) = O(s_0^6)$, because the number of all states is $J = (s_0+1)(s_0+2)/2$, but the block form (19) has $O(s_0^4)$, because the maximal number of states in any component is s_0+1. If $e_0 < s_0$, then the number of all states is $J = (s_0-e_0+1)(e_0+1)+e_0(e_0+1)/2$ and thus the matrix exponential solution (12) has $O(J^3) = O(e_0^3 s_0^3)$, but (19) has $O(s_0^4)$, since the maximal number of states in any component is s_0-1. Moreover, solving (12) is often computationally intractable for most complex chemical systems due

to the large dimension of state space. However, our exact block formula (19) has only $(s_0+1)^2$ as the maximum dimension of block matrices, which means that our method is less constrained from the curse of dimensionality.

We perform numerical experiments for the two cases, $e_0 < s_0$ and $e_0 > s_0$ in Figures 1 and 2. For the case that $e_0 < s_0$, we assume two initial conditions $\mathbf{n}(0) = (8, 10, 0, 0)$(the total number of states is $J = 63$) and $\mathbf{n}(0) = (10, 15, 0, 0)(J = 121)$, respectively. In Figure 1, we compare the time evolution of $P(n_2(t) = i)$ from the original matrix exponential solution (12) and our exact block solution (19). In the computation of the matrix exponential solution, we use **expm** function of MATLAB.



(a) $n_2 = 0$ (b) $n_2 = 3$ $n_2 = 5$

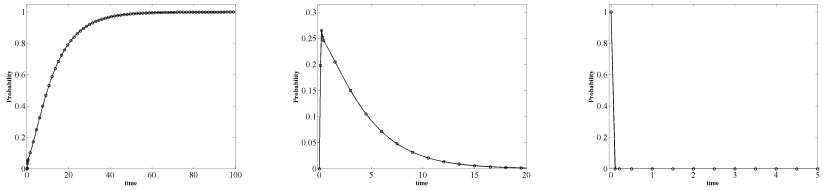
(d) $n_2 = 0$ (e) $n_2 = 4$ (f) $n_2 = 7$

Figure 1: Under the initial condition $\mathbf{n}(0) = (8, 10, 0, 0)$ and $J = 63$(upper three figures) and the condition $\mathbf{n}(0) = (10, 15, 0, 0)$ and $J = 121$(lower three curves), comparison of the time evolution of the probability of $n_2(t)$ between the matrix exponential method(circles) and our method(solid curve).The reaction rate parameters are $c_1 = 1, c_2 = 2, c_3 = 0.1$.

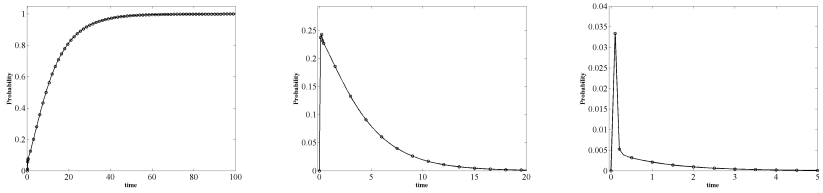
For the case that $e_0 > s_0$, we assume two initial conditions $\mathbf{n}(0) = (30, 20, 0, 0)$ (thus $J = 231$) and $\mathbf{n}(0) = (40, 25, 0, 0)(J = 351)$, respectively. In Figure 2, we compare the time evolution of $P(n_2(t) = i)$ from the original matrix exponential solution (12) and our exact block solution (19).

In Table 1, we compare CPU times taken by computations using the matrix exponential method and our exact method.

One can see from Figures 1, 2 and Table 1 that our method is accurate and more effi-



(a) $n_2 = 0$ (b) $n_2 = 3$ $n_2 = 20$



(d) $n_2 = 0$ (e) $n_2 = 3$ (f) $n_2 = 7$

Figure 2: Under the initial condition $\mathbf{n}(0) = (30, 20, 0, 0)$ and $J = 231$(upper three figures) and the condition $\mathbf{n}(0) = (40, 25, 0, 0)$ and $J = 351$(lower three curves), comparison of the time evolution of the probability of $n_2(t)$ between the matrix exponential method(circles) and our method(solid curve). The reaction rate parameters are $c_1 = 1, c_2 = 2, c_3 = 0.1$.

Cases	$e_0 < s_0$		$e_0 > s_0$	
Number of states J	63	121	231	351
Exact formula	0.79	1.61	3.03	4.99
Matrix exponential	0.96	3.26	17.74	58.37

Table 1: Comparison of CPU time (in seconds) elapsed by the exact formula and the matrix exponential when we compute all the probability solutions of the four models described in Figures 1 and 2 in the time interval $[0, 100]$.

cient than the matrix exponential method. Especially, as the dimension of the system gets large, our method gives much faster computation than the matrix exponential method.

4 Conclusion

In this paper, we proposed an analytic method of the chemical master equation for the stochastic enzyme-substrate model. Using the block structure of the Markov chain generator K of Equation (2), we found a formula for the analytic solution of (2) for given initial initial conditions. Using the computational procedure for finding the solution, we showed the numerical accuracy and efficiency of the method with comparison of the ma-

trix exponential computation by simulating four cases of the enzyme-substrate system. We observed that our method reduces the computational complexity much and as the dimension of the system gets larger, our method shows much better efficiency. We expect that the result presented in this paper will be used for analysis and computation of the stochastic model of the important enzyme-substrate system or any similar systems.

Acknowledgements: This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2010-0024849).

References

[1] D. T. Gillespie, A rigorous derivation of the chemical master equation, *Physica A* **188** (1992) 404–425.

[2] C. H. Lee, R. Lui, A reduction method for multiple time scale stochastic reaction networks, *J. Math. Chem.* **46** (2009) 1292–1321.

[3] M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, Stochastic gene expression in a single cell, *Science* **297** (2002) 1183–1186.

[4] M. Thattai, A. van Oudenaarden, Intrinsic noise in gene regulatory networks, *Proc. Nat. Acad. Sci.* **98** (2001) 8614–8619.

[5] A. Kumar, K. Josić, Reduced models of networks of coupled enzymatic reactions, *J. Theor. Biol.* **278** (2011) 87–106.

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

[7] L. Michaelis, M. L. Menten, Die Kinetik der Invertinwirkung, *Biochem. Z.* **49** (1913) 333–369.

[8] S. Schnell, C. Mendoza, Closed–form solution for time–dependent enzyme kinetics, *J. Theor. Biol.* **187** (1997) 207–212.

[9] M. N. Berberan–Santos, A general treatment of Henri–Michaelis–Menten enzyme kinetics: exact series solution and approximate analytical solutions, *MATCH Commun. Math. Comput. Chem.* **63** (2010) 283–318.

[10] M. Goličnik, 'Die Kinetik der Invertinwirkung' of L. Michaelis and M. L. Menten revisited after 100 years: closed–form solutions of genuine invertase–reaction dynamics, *MATCH Commun. Math. Comput. Chem.* **70** (2013) 63–72.

[11] P. Arányi, J. Tóth, A full stochastic description of the Michaelis–Menten reaction for small systems, *Acta Biochim. Biophys. Acad. Sci. Hung.* **12** (1977) 375–388.

[12] P. J. Staff, A stochastic development of the reversible Michaelis–Menten mechanism, *J. Theor. Biol.* **27** (1970) 221–232.

[13] H. Qian, E.L. Elson, Single–molecule enzymology: stochastic Michaelis–Menten kinetics, *Biophys. Chem.* **101-102** (2002) 565–576.

[14] E. Dóka, G. Lente, Stochastic mapping of the Michaelis–Menten mechanism, *J. Chem. Phys.* **136** (2012) 054111.

[15] R. Varon, J. M. Villalba, F. Garcia–Sevilla, M. Garcia–Moreno, M. Molina–Alarcon, A. Barbero, M. J. Garcia–Meseguer, E. Arribas, General symbolic expressions for statistical moments in any linear compartmental system, *MATCH Commun. Math. Comput. Chem.* **67** (2012) 609–648

[16] C. H. Lee, P. Kim, An analytical approach to solutions of master equations for stochastic nonlinear reactions, *J. Math. Chem.* **50** (2012) 1550–1569.

[17] J. H. Wilkinson, Calculation of the eigenvectors of a symmetric tridiagonal matrix by inverse iteration, *Numer. Math.* **4** (1962) 368–376.

[18] L. Elsner, A. Fasse, E. Langmann, A divide–and–conquer method for the tridiagonal generalized eigenvalue problem, *J. Comput. Appl. Math.* **86** (1997) 141–148.

[19] I. S. Dhillon, B. N. Parlett, Multiple representations to compute orthogonal eigenvectors of symmetric tridiagonal matrices, *Lin. Algebra Appl.* **387** (2004) 1–28.