Published in final edited form as:

Chem Eng Sci. 2011 February 1; 66(3): 268–277. doi:10.1016/j.ces.2010.10.024.

Analytical Derivation of Moment Equations in Stochastic Chemical Kinetics

Vassilios Sotiropoulos and Yiannis N. Kaznessis*

Department of Chemical Engineering and Materials Science, University of Minnesota, 421 Washington Ave. SE, Minneapolis, MN 55455, USA

Abstract

The master probability equation captures the dynamic behavior of a variety of stochastic phenomena that can be modeled as Markov processes. Analytical solutions to the master equation are hard to come by though because they require the enumeration of all possible states and the determination of the transition probabilities between any two states. These two tasks quickly become intractable for all but the simplest of systems. Instead of determining how the probability distribution changes in time, we can express the master probability distribution as a function of its moments, and, we can then write transient equations for the probability distribution moments. In 1949, Moyal defined the derivative, or jump, moments of the master probability distribution. These are measures of the rate of change in the probability distribution moment values, i.e. what the impact is of any given transition between states on the moment values. In this paper we present a general scheme for deriving analytical moment equations for any N-dimensional Markov process as a function of the jump moments. Importantly, we propose a scheme to derive analytical expressions for the jump moments for any N-dimensional Markov process. To better illustrate the concepts, we focus on stochastic chemical kinetics models for which we derive analytical relations for jump moments of arbitrary order. Chemical kinetics models are widely used to capture the dynamic behavior of biological systems. The elements in the jump moment expressions are a function of the stoichiometric matrix and the reaction propensities, i.e the probabilistic reaction rates. We use two toy examples, a linear and a non-linear set of reactions, to demonstrate the applicability and limitations of the scheme. Finally, we provide an estimate on the minimum number of moments necessary to obtain statistical significant data that would uniquely determine the dynamics of the underlying stochastic chemical kinetic system. The first two moments only provide limited information, especially when complex, non-linear dynamics are involved.

Keywords

Multiscale models; Master equation; Model reduction; Statistical thermodynamics; Mathematical modelling; Computational chemistry

© 2010 Elsevier Ltd. All rights reserved.

*Author to whom correspondence should be addressed., sotiropo@cems.umn.edu (Vassilios Sotiropoulos), yiannis@umn.edu (Yiannis N. Kaznessis).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Many stochastic processes in physics and chemistry can be considered to follow the Markov property; the movement of the system in the available state space depends only on the previous state and not on the path that the stochastic process has followed before.

Any stochastic process that obeys the Markov property is a Markov process. The underlying probability distribution, that is the probability of finding the system at any possible state at a certain time, is governed by a single partial differential equation (PDE) called the master equation (ME). If an initial condition is known and given the transitional probabilities, i.e., the probabilities of the system transitioning from any state to any other state, ME uniquely determines the probability distribution at any later time [1, 2].

The solution of ME requires enumerating the states and finding the transition matrix, a task that is tractable only for the simplest of systems. For more complex systems, in lieu of a solution of ME, kinetic Monte Carlo techniques are often used to sample the underlying probability distribution. There is a growing community of researchers that develops computationally efficient and accurate stochastic simulation algorithms [3–14].

Alternatively, instead of sampling the probability distribution all important information for the system's behavior can be had through the moments of the probability distribution [15]. The first moment relates to the mean of the probability distribution, the second to the variance, the third to skew-ness and the fourth to kurtosis. Higher order moments may also contain important information on systems dynamics especially when these systems are complex and highly non-linear.

In this work we start from the master equation and derive a system of ordinary differential equations (ODE) that describe the dynamics of the probability distribution moments. We express the transient moment dynamics in terms of the derivative or jump moments [16]. Jump moments are measures of the rate of change in the probability distribution moment values, i.e. they quantify the impact of any given transition between states on the moment values.

What is new in the current work is a scheme to derive analytical expressions for the jump moments for any N-dimensional Markov process. These expressions result in an infinite linear system of ODEs that describes the moment dynamics of any such process.

We then apply this scheme to stochastic chemical kinetics. Stochastic chemical kinetics have emerged in recent years as an appropriate modeling formalism for biological systems that are away from the thermodynamic limit [17–23].

The idea of using moment equations to predict dynamics of systems in the stochastic chemical kinetics regime has also been considered early on [24]. In recent years, publications have proposed alternative ways to derive the moment equations [25–28]. Our work complements this rich literature, providing useful analytical relations for the jump moments in stochastic chemical kinetics. These relations are then used to provide analytical equations for the probability distribution moments.

The paper is organized as follows. First we briefly present background information on Markov processes and their governing equations. We then define the jump moments of Markov processes and we use this definition to derive the general form of the moment equations. Then we concentrate on the stochastic chemical kinetics regime and we derive analytical relations for the jump moments. We use examples to discuss the applicability of the derived equations. Finally we discuss the drawbacks of any moment based method that

mainly stem from the fact that the resulting ODE system is infinite dimensional. We discuss literature moment-closure schemes and speculate on how the analytical expressions we develop may assist in the development of accurate closing schemes. We also argue that higher order moments, at least up to order six, are necessary for biologically-relevant chemical kinetics systems. This is contrary to the widely held belief that only the mean and variance are important in stochastic chemical kinetics models.

2. Theory

2.1. Markov Processes

Consider a Markov process, $X(t)$, a stochastic process that for any n successive set of times, $t_1 < t_2 < \dots < t_n$, obeys the following property, known as the *Markov property*

$$P_{1|n-1}(X_n, t_n | X_1, t_1; \dots; X_{n-1}, t_{n-1}) = P_{1|1}(X_n, t_n | X_{n-1}, t_{n-1}) \quad (1)$$

where X_k is the state of the system at time t_k . $P_{1|n-1}$ denotes the conditional probability that the outcome at time t_n is X_n provided that the outcomes were $X_1, X_2, \dots, X_{n-1}$ at times $t_1, t_2, \dots, t_{n-1}$, respectively. $P_{1|1}$ denotes the conditional probability that the outcome at time t_n is X_n provided that the outcome at time t_{n-1} was X_{n-1} , irrespective of the previous outcomes. The Markov property states that the transition of the system to state X_n at time t_n depends only on the previous state X_{n-1} at time t_{n-1} and therefore is independent of the history of the process, i.e. $X_1, t_1; \dots; X_{n-2}, t_{n-2}$.

The probability distribution function of any stochastic process with state vector $\underline{X}(t) = (X_1(t), \dots, X_N(t))$ and initial condition $\underline{X}(t=0) = \underline{X}_0$ obeying the Markov property is governed by a differential difference equation in probability known as the *master equation* (ME) [1]

$$\frac{\partial P(\underline{X}, t)}{\partial t} = \int [T(\underline{X}/\underline{X}') P(\underline{X}', t) - T(\underline{X}'/\underline{X}) P(\underline{X}, t)] d\underline{X}' \quad (2)$$

$P(\underline{X}, t)$ is the probability of the system being at state $\underline{X}$ at time t . $T(\underline{X}/\underline{X}')$ is the transition probability per unit time for the system to jump from state $\underline{X}'$ to state $\underline{X}$. In principle, the ME uniquely determines the probability $P(\underline{X}, t)$ of the system being at a state $\underline{X} = \underline{X}(t)$ at time $t > 0$. Note that using integrals in the ME we have made the assumption that variable $\underline{X}(t)$ is continuous.

An alternative description of the Markov process probability distribution is obtained through the *Kramers-Moyal* expansion [1].

$$\frac{\partial P(\underline{X}, t)}{\partial t} = \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \sum_{j_1, \dots, j_m=1}^N \frac{\partial^m}{\partial X_{j_1} \dots \partial X_{j_m}} [\underline{a}_m(\underline{X}) P(\underline{X}, t)] \quad (3)$$

where $\underline{a}_m$ are m -order tensors and were originally called *derivative* or *jump moments* by Moyal [16] and later referred to as propagator moment functions [15]. Jump moments are measures of the rate of change in the probability distribution moment values, i.e. they quantify the impact of any given transition between states on the moment values.

2.2. Moments and Jump Moments

Let us consider an N -dimensional Markov process $\underline{X}(t) = (X_1(t), \dots, X_N(t))$ with probability distribution $P(\underline{X}(t), t)$.

The m -th moment of the stochastic variable X is defined as

$$\langle X_i^m \rangle = \int X_i^m P(\underline{X}, t) d\underline{X} \quad i=1, \dots, N \quad (4)$$

The first order moments of any X_i is usually referred to as the mean while second order moments relate to the variance of the probability distribution through the relation, $var(X_i) = \langle X_i^2 \rangle - \langle X_i \rangle^2$

The joint moments are defined through the following relations

$$\langle X_1^{m_1} X_2^{m_2} \dots X_n^{m_n} \rangle = \int [X_1^{m_1} X_2^{m_2} \dots X_n^{m_n}] P(\underline{X}, t) d\underline{X} \quad (5)$$

$m_1, m_2, \dots, m_n = 1, 2, 3, \dots$

where the sum $m = m_1 + m_2 + \dots + m_n$ is the order of the joint moment. The set of the m -th order moments uniquely determines the underlying probability distribution. The existence of higher order moments depends on how fast $P(\underline{X}, t)$ approaches zero as $\underline{X}$ goes to infinity.

On the other hand, jump moments relate to the transition probability per unit time rather than the probability distribution. Based on Moyal's work jump moments are defined through the following relations [16]

$$a_m^i(\underline{X}) = \int (X'_i - X_i)^m T(\underline{X}'/\underline{X}) d\underline{X}' \quad i=1, \dots, N \quad (6)$$

Index m refers to the order of the moment and index i to the element within the tensor.

Analogously the joint jump moments are

$$a_{m_i+m_j+\dots}^{i,j,\dots}(\underline{X}) = \int [(X'_i - X_i)^{m_i} (X'_j - X_j)^{m_j} \dots] T(\underline{X}'/\underline{X}) d\underline{X}' \quad (7)$$

Jump moments are tensors and the indexes $i, j, \dots$ denote the element of the $m = m_i + m_j + \dots$ order tensor. Jump moments of order 2 and greater are symmetric tensors, i.e. $a_2^{1,2} = a_2^{2,1}$. The same is true for regular moments. In the simple case of an one-dimensional Markov process jump moments are simply scalars.

2.3. Derivation of Moment Equations

Given the definitions for both regular and jump moments we can now start deriving the moment equations, i.e. the system of ODEs that describes moment dynamics.

Starting with the first moment we have

$$\langle X_i \rangle = \int X_i P(\underline{X}, t) d\underline{X} \quad (8)$$

From the defining equation we have the exact identity for the time derivative of the first moment

$$\frac{d\langle X_i \rangle}{dt} = \int X_i \frac{\partial P(\underline{X}, t)}{\partial t} d\underline{X} \quad (9)$$

substituting equation (2) in the last equation we have

$$\frac{d\langle X_i \rangle}{dt} = \int X_i \int [T(\underline{X}/\underline{X}')P(\underline{X}', t) - T(\underline{X}'/\underline{X})P(\underline{X}, t)] d\underline{X}' d\underline{X} = \int \int [X_i T(\underline{X}/\underline{X}')P(\underline{X}', t) - X_i T(\underline{X}'/\underline{X})P(\underline{X}, t)] d\underline{X}' d\underline{X} \quad (10)$$

Noticing that that the integration over $\underline{X}$ and $\underline{X}'$ runs over the same domain we can interchange indexes in the last equation, i.e.

$$X_i T(\underline{X}/\underline{X}')P(\underline{X}', t) = X_i' T(\underline{X}'/\underline{X})P(\underline{X}, t) \quad (11)$$

thus we have

$$\frac{d\langle X_i \rangle}{dt} = \int \int (X_i' - X_i) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \quad (12)$$

Recalling the definition of the jump moment in section 2.2 and in particular that of the first moment

$$a_1^i(\underline{X}) = \int (X_i' - X_i) T(\underline{X}'/\underline{X}) d\underline{X}' \quad (13)$$

the right hand side (RHS) of equation (12) can be recast in the following form

$$\frac{d\langle X_i \rangle}{dt} = \int a_1^i(\underline{X}) P(\underline{X}, t) d\underline{X} \quad (14)$$

where the RHS term denotes the average of $a_1^i(\underline{X})$.

The ODE equation describing the dynamics of the first moment of variable X_i is then simply

$$\frac{d\langle X_i \rangle}{dt} = \langle a_1^i(\underline{X}) \rangle, i=1, \dots, N \quad (15)$$

Following a similar way of thinking and carrying out slightly more complicated algebraic calculations, which we omit in the present section for brevity but we include in Appendix A for completeness, we can write the moment equations for second, third and fourth order moment dynamics.

Using inductive reasoning, from equations (15), (A.5), (A.14) and (A.23) we derive the general moment equation for the m^{th} order moment of an N-dimensional Markov process as a function of joint jump moments

$$\frac{d\langle X_1^{m_1} X_2^{m_2} \dots X_N^{m_N} \rangle}{dt} = \sum_{j_1, j_2, \dots, j_N=0}^{m_1, m_2, \dots, m_N} \binom{m_1}{j_1} \binom{m_2}{j_2} \dots \binom{m_N}{j_N} \times \langle X_1^{m_1-j_1} X_2^{m_2-j_2} \dots X_N^{m_N-j_N} a_{j_1+j_2+\dots+j_N}^{index}(\underline{X}) \rangle \quad (16)$$

with initial conditions

$$\langle X_1^{m_1} X_2^{m_2} \dots X_N^{m_N} \rangle_{t=0} = \langle X_1^{m_1}(t=0) X_2^{m_2}(t=0) \dots X_N^{m_N}(t=0) \rangle$$

where by identity $a_0(\underline{X})=0$, $\binom{m}{j}$ denotes the binomial coefficient and *index* notation is used to symbolize the following operation

$$index=\underbrace{\{1,\dots,1,2,\dots,2,\dots,N,\dots,N\}}_{m_1} - \underbrace{\{1,\dots,1,2,\dots,2,\dots,N,\dots,N\}}_{m_2} - \dots - \underbrace{\{1,\dots,1,2,\dots,2,\dots,N,\dots,N\}}_{m_N} = \underbrace{\{1,\dots,1,2,\dots,2,\dots,N,\dots,N\}}_{m_1-j_1} - \underbrace{\{1,\dots,1,2,\dots,2,\dots,N,\dots,N\}}_{m_2-j_2} - \dots - \underbrace{\{1,\dots,1,2,\dots,2,\dots,N,\dots,N\}}_{m_N-j_N} \quad (17)$$

In other words, the *index* notation refers to the appropriate element of the j^{th} order tensor, $j = j_1 + j_2 + \dots + j_N$.

To our knowledge there have been no other attempts to deduce the general moment equation from jump moments in the case of a N-dimensional Markov process.

In the case of a 1-dimensional Markov process the above equation reduces to the following simpler relation [15]

$$\frac{d\langle X^m \rangle}{dt} = \sum_{j=1}^m \binom{m}{j} \langle X^{m-j} a_j(\underline{X}) \rangle \quad (18)$$

$$\langle X^m \rangle_{t=0} = \langle X^m(t=0) \rangle$$

Alternatively, the multivariate case can be written in efficient matrix form using Kronecker product notation [29]. Kronecker products, ($\otimes$), have interesting implementation properties on computers and thus may improve computational efficiency. We can write

$$\frac{d\langle \underline{X}^{(m)} \rangle}{dt} = \text{vec}(a_m) + \sum_{j=1}^{m-1} \Pi_j^m \langle \underline{X}^{(j)} \otimes \text{vec}(a_{m-j}) \rangle \quad (19)$$

with initial conditions

$$\langle \underline{X}^{(m)} \rangle_{t=0} = \langle \underline{X}^{(m)}(t=0) \rangle$$

where

$$\underline{X}^{(m)} = \underbrace{\underline{X} \otimes \underline{X} \cdots \otimes \underline{X}}_m$$

and where $\text{vec}(\cdot)$ symbolizes the vectorization of any matrix or tensor in general into a column vector, which is obtained by stacking the columns of the matrix or tensor on top of

one another. Finally Π_j^m is the sum of $\binom{m}{j} = \frac{m!}{j!(m-j)!}$ permutation matrices ($N^m \times N^m$), defined as follows

$$\Pi_j^m = \sum_{k=1}^{\binom{m}{j}} \pi_{m,k}$$

where each permutation matrix, $\pi_{m,k}$, is defined as [29]

$$\pi_{m,k} = \sum_{i_1, \dots, i_m=1}^N \underbrace{E_{i_1 i_1}^{N \times N} \otimes E_{i_2 i_2}^{N \times N} \dots \otimes E_{i_m i_m}^{N \times N}}_m$$

where $E_{ij}^{p \times q}$ is a $p \times q$ matrix (elementary matrix) that has a "1" in the (i, j) position and zeros everywhere else and k denotes the different permutation sets for the subscripts of all $E_{ij}^{p \times q}$. The methodology to choose the appropriate permutation sets is detailed in Appendix B. Note that $\Pi_0^m = I_{N^m \times N^m}$, i.e. the $N^m \times N^m$ identity matrix.

The system of equations (16) or its equivalent (19), completely characterizes the moment dynamics for any given N -dimensional Markov process. The only prerequisite is the knowledge of analytical relations for the jump moments. These relations depend on the underlying physics of the problem as it will become evident in the following section. In particular, we derive analytical relations for the components of the jump moments tensors in the case where the underlying Markov process stems from a stochastic chemical kinetics model. The final equations may appear cumbersome but they are very simple to generate with the help of a computer program.

Even though we have made the assumption that variables $\underline{X}(t)$ are continuous, this needs not be the case. The derivation of equation (16) could be carried out by starting from the discrete ME [1]. The main difference in the derivation is that instead of integrals, summation signs over all possible states would be used.

2.4. Jump Moments and Stochastic Chemical Kinetics

Consider a system of N distinct chemical species, X_i ($i = 1, \dots, N$), participating in M chemical reactions in a well-mixed volume V

$$\sum_{i=1}^N r_i^j X_i^{k_j} \quad \sum_{i=1}^N p_i^j X_i, \quad j=1, \dots, M \quad (20)$$

Consider also $\underline{v}_j$ the stoichiometric vector associated with the j^{th} reaction

$$\underline{v}_j = \begin{bmatrix} p_1^j - r_1^j \\ \vdots \\ p_N^j - r_N^j \end{bmatrix} \quad j=1, \dots, M \quad (21)$$

Such systems of reactions are widely used to model biological interactions, such as transcription, translation, degradation, regulation and protein-protein interactions [18, 19, 21, 30, 31]. Dilute and sparse species populations render the traditional continuous-deterministic modeling approach false. Instead, stochastic chemical kinetics models are more appropriate to describe systems of reactions that are far from the thermodynamic limit [32]. Therefore system (20) is frequently considered as a Markov process where a chemical master equation (CME) governs the evolution of the probability distribution.

Under the stochastic chemical kinetics regime reaction rates become reaction propensities, $\alpha_j(\underline{X})$. These are the probabilistic equivalents and are defined as follows [33]

$$\alpha_j(\underline{X}) = k_j c_j(\underline{X}) \quad c_j(\underline{X}) = \prod_{i=1}^N \frac{X_i!}{r_i^j! (X_i - r_i^j)!} \quad (22)$$

Each constant k_j represents the mesoscopic reaction rate of the j^{th} reaction.

In what follows we derive analytical relations for the jump moments for any stochastic chemical kinetics model. Starting with the CME, we can determine the Kramers-Moyal expansion of the probability distribution (cf. eq. (3)). If we truncate the expansion and retain only the first two terms then the resulting equation is the well-known Fokker-Planck equation (FPE) [1]

$$\frac{\partial P(\underline{X}, t)}{\partial t} = - \frac{\partial}{\partial \underline{X}} [\underline{a}_1(\underline{X}) P(\underline{X}, t)] + \frac{1}{2} \frac{\partial}{\partial \underline{X}} \frac{\partial}{\partial \underline{X}} : [\underline{a}_2(\underline{X}) P(\underline{X}, t)] \quad (23)$$

where $\underline{a}_2$ and $\underline{a}_1$ are the first two jump moments tensors or using the terminology most often referred to them the drift vector and diffusion tensor respectively. Instead of solving the Fokker-Planck equations it is usually more convenient to sample the underlying probability distribution generating ensembles of trajectories obtained as solutions of the corresponding chemical Langevin equations (CLE) or systems of CLEs [34].

From the work of Gillespie we know that for systems of chemical reactions under the Markov process regime the corresponding chemical Langevin equation is [34]

$$d\underline{X} = \underline{v}\alpha(\underline{X})dt + \underline{v}\underline{\mathcal{D}}(\sqrt{\alpha(\underline{X})})dW \quad (24)$$

where $\underline{v}$ corresponds to the $N \times M$ stoichiometric matrix, $\alpha(\underline{X})$ corresponds to the $M \times 1$ propensities vector and the notation $\underline{\mathcal{D}}(\sqrt{F})$ denotes the diagonal matrix whose $(i, i)^{th}$ element are the only nonzero elements and their value equals the square root of the i^{th} component of vector F .

Given any FPE we can deduce the corresponding system of CLEs and vice versa. This direct relation between the two allows us to infer through inductive reasoning analytical relations for the jump moment tensors elements. From equation (24) we obtain the following representation for the FPE

$$\frac{\partial P(\underline{X}, t)}{\partial t} = - \frac{\partial}{\partial \underline{X}} [\underline{v}\alpha(\underline{X}) P(\underline{X}, t)] + \frac{1}{2} \frac{\partial}{\partial \underline{X}} \frac{\partial}{\partial \underline{X}} : [\underline{v}\underline{\mathcal{D}}(\alpha(\underline{X}))\underline{v}^T P(\underline{X}, t)] \quad (25)$$

Comparing equations (23) and (25) we retain the following relations for the first two jump moment tensors

$$\begin{aligned} \underline{a}_1(\underline{X}) &= \underline{v}\alpha(\underline{X}) \\ \underline{a}_2(\underline{X}) &= \underline{v}\underline{\mathcal{D}}(\alpha(\underline{X}))\underline{v}^T \end{aligned} \quad (26)$$

or using summation notation each of the tensor elements is defined as follows

$$\begin{aligned} a_1^i(\underline{X}) &= \sum_{k=1}^M v_{ik} \alpha_k(\underline{X}) \\ a_2^{ij}(\underline{X}) &= \sum_{k=1}^M v_{ik} v_{jk} \alpha_k(\underline{X}) \end{aligned} \quad (27)$$

We notice that the difference between the two jump moments relations relies on an additional component of the stoichiometric vector. Following the discussion in the previous section we infer the relations for the third and fourth jump moments, where $a_3(\underline{X})$ and $a_4(\underline{X})$ are $N \times N \times N$ and $N \times N \times N \times N$ tensors respectively.

$$\begin{aligned} a_3^{ijl}(\underline{X}) &= \sum_{k=1}^M v_{ik} v_{jk} v_{lk} \alpha_k(\underline{X}) \\ a_4^{ijlm}(\underline{X}) &= \sum_{k=1}^M v_{ik} v_{lk} v_{jk} v_{mk} \alpha_k(\underline{X}) \end{aligned} \quad (28)$$

In general, for stochastic chemical kinetics models we infer that given the stoichiometric matrix and the reaction propensities vector all jump moments can be defined analytically through the following recursive formula

$$a_m^{\overbrace{ij \dots l}^{m \text{ indices}}}(\underline{X}) = \sum_{k=1}^M \underbrace{v_{ik} v_{jk} \dots v_{lk}}_{m \text{ } v's} \alpha_k(\underline{X}) \quad (29)$$

The formula is relatively simple and intuitive. Substituting the last equation into equation (16) returns analytical relations for the moment equations. A first obvious comment is that the linear or non-linear character of the underlying reaction networks manifests itself through the reaction propensities which in turn render jump moments linear or non-linear functions of the chemical species concentrations. The impact of the non-linear jump moment relations will become evident in the examples following. In general, a similar approach may lead to analytical relations for jump moments for any given Markov process.

In the supplementary material, we present two toy examples, a linear and a non-linear reaction network, to highlight the analytical form of the moment equations for stochastic chemical kinetic systems.

3. Discussion

3.1. Infinite Dimensional Moment Equations

For any given non-linear model, whether it is a chemical kinetics model or an electrical circuit model, jump moments will be non-linear functions of the state vector, hence the system of ODEs describing the moment evolution will be infinite dimensional.

The general form of the ODEs describing the moment dynamics of a stochastic chemical kinetics model up to a desired order can be summarized through the following matrix equation

$$\frac{d\mathbf{Y}}{dt} = \mathbf{A}\mathbf{Y} + \mathbf{A}_h \mathbf{Y}_h + \mathbf{B} \quad (30)$$

where $\mathbf{Y}$ is a vector containing the moment elements of interest and $\mathbf{A}$ is a square matrix with appropriate dimensions. Its elements depend on the kinetic constants and the stoichiometric matrix of the reactions set. Matrix $\mathbf{A}$ may be sparse depending on the connectivity of the system. Vector $\mathbf{B}$ contains the terms that depend on any zeroth order reactions and are always constant as the corresponding reaction rates are independent from the state vector. Finally, the term $\mathbf{A}_h \mathbf{Y}_h$ denotes the dependence of lower order moments on higher order moments. Vector $\mathbf{Y}_h$ contains higher order moments and matrix $\mathbf{A}_h$ depends on the reaction constants and stoichiometric matrix. $\mathbf{A}_h \mathbf{Y}_h$ vanishes only when the underlying system is linear.

3.2. Moment Closure Schemes

As mentioned earlier equation (30) is not solvable when the system of reactions is non-linear, as the dependence on higher moments renders the system infinite. In this section we briefly discuss existing approaches, called moment closure techniques, that attempt to bypass the problem and obtain approximate solutions.

The main idea behind any moment closure scheme is to approximate the infinite dimensional system of ODEs depicted in equation (30) with a finite system of the form

$$\frac{d\mathbf{Y}_t}{dt} = \mathbf{A}\mathbf{Y}_t + \mathbf{f}_h(\mathbf{Y}_t) + \mathbf{B} \quad (31)$$

where $\mathbf{Y}_t$ is a truncated version of the infinite dimensional vector $\mathbf{Y}$ (cf. eq. (30)). The last equation explicitly states that the term $\mathbf{A}_h \mathbf{Y}_h$ in equation (30) should be substituted with a vector containing linear or non-linear functions of $\mathbf{Y}_t$, $\mathbf{f}_h(\mathbf{Y}_t)$, that would closely approximate the effect of the higher order moment terms. In other words, the effect of higher order moment terms is approximated through functions of lower order moment terms. This approximation renders the system finite but also non-linear as all the developed approaches consider $\mathbf{f}$ to be non-linear functions. At the same time the approximation introduces an error in the moment values with its significance depending on how well $\mathbf{f}$ follows $\mathbf{A}_h \mathbf{Y}_h$.

Numerous attempts have been made to develop such closure schemes [25, 27, 35, 36]. A summary of the existing moment closure techniques in stochastic chemical kinetics system can be found in reference [37]. To our knowledge there is no satisfactory solution to the moment closure scheme with broad applicability and the topic is still being actively researched.

3.3. Importance of Higher Moments in Non-linear Chemical Kinetics

In this section we determine the probability distribution of a simple non-linear system of chemical reaction to illustrate how higher than the first two order moments are necessary to capture the dynamic behavior.

Consider the well studied Schlögl model that under specific values for the kinetic rates exhibits bistability, similar to the bistable switch [38] and or the λ -phage infection [31]. Consider the network of reactions shown in Table 1. Kinetic data and initial conditions are taken from reference [39].

The kinetic values depicted in Table 1 render the system bistable. The use of the Schlögl model enables us to determine the minimum number of moments needed to reconstruct the

underlying probability distribution within a reasonable error. Using Hy3S [13] and SynBioSS [14], a suite of multi-scale algorithms, we generate 10^5 independent trajectory trials in the interval $[0,20]$ s. We should note that Hy3S is a publicly available software suite for simulating chemical reaction networks. Hy3S is available with Open Licenses at sourceforge.net.

We use Matlab to compute the underlying probability distribution of species X at three different time points. The shapes and characteristics of the corresponding probability distributions are depicted in Figure 1a. Next we compute the first eight moments of X at the three different time points using the data obtained from the stochastic simulation. Their values are shown in Table 2. Using the numerical values of the moments we can reconstruct the corresponding probabilities using an algorithm based on the maximum entropy principle [40].

In order to determine the minimum number of moments we first reconstruct the distributions using only the first two moments and then we keep increasing the number of moments by two and up to order eight.

In Figures 1b, 1c and 1d the comparison between the actual and reconstructed probabilities are depicted. Reconstructed probabilities using only the first two moments expectedly fail to capture the bimodality of the distribution especially in the cases of $t = 4$ s and 20 s. Using four moments seems to capture the essential features of the distributions, i.e. the two modes, but it fails to weight in the relative peaks of the two modes. Six moments produce adequate results, especially when the separation between the two modes is distinct (cf. Figure 1d), but results are inferior when there is no distinct separation (cf. Figures 1b and 1c). Still results are relatively accurate. Finally, using the first eight moments produces distributions that are almost identical to the actual distribution with minor, if any, deviations.

Overall, it is evident that using the first two moments is not adequate to reconstruct the probability distribution. Our analysis shows that use of at least the first six, or in some cases eight, moments is needed for adequate reconstruction of the probability distribution.

4. Conclusions

A new derivation of the moment equations for any N-dimensional Markov process using the definition of jump moments was presented. The applicability of the scheme is general and leads to analytical relations given that the functional form of the jump moments is known.

Focusing on stochastic chemical kinetics models, we derived analytical relations for the elements of any jump moment tensor, demonstrating the ease of equation setup. The elements of the analytical relations are a function of the stoichiometric matrix and the reaction propensities, i.e the probabilistic reaction rates.

Using a simple example of non-linear kinetics we illustrated how higher than second order moments are important to accurately reconstruct the probability distribution for biologically-relevant chemical kinetics systems. Certainly, then, more effort is warranted on the development of moment closure schemes.

We should stress that there is no general approach to choose the number of moments necessary to accurately capture the probability distribution. This choice remains system dependent and is one driven by empiricism.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a grant from the National Institutes of Health (American Recovery and Reinvestment Act grant GM086865), and grants from the National Science Foundation (CBET-0425882 and CBET-0644792). Computational support from the Minnesota Supercomputing Institute (MSI) is gratefully acknowledged. This work was also supported by the National Computational Science Alliance under TG-MCA04N033.

Appendix A. Derivation of Second Through Fourth Order Moment Equations

In this section of the Appendix we derive the second through fourth order moment equations. For the second order moment equations, the starting point is the defining equation (5) from which the following identity is derived for the time derivative

$$\frac{d\langle X_i X_j \rangle}{dt} = \int X_i X_j \frac{\partial P(\underline{X}, t)}{\partial t} d\underline{X}, j=1, \dots, N \quad (\text{A.1})$$

Analogously to the derivation of the first moment equations, we substitute equation (2) in the last equation and by interchanging notation using the same arguments as previously (cf. eq. (11)). Then

$$\frac{d\langle X_i X_j \rangle}{dt} = \int \int (X'_i X'_j - X_i X_j) T(\underline{X}' / \underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \quad (\text{A.2})$$

Through algebraic manipulations we can rearrange the terms in the last equation. In particular we use the following identity

$$(X'_i X'_j - X_i X_j) = (X'_i - X_i)(X'_j - X_j) + X_j(X'_i - X_i) + X_i(X'_j - X_j) \quad (\text{A.3})$$

Substituting the term in the RHS of equation (A.2) yields

$$\begin{aligned} \frac{d\langle X_i X_j \rangle}{dt} &= \int \int (X'_i - X_i)(X'_j - X_j) T(\underline{X}' / \underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\ &\quad + \int \int X_j (X'_i - X_i) T(\underline{X}' / \underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\ &\quad + \int \int X_i (X'_j - X_j) T(\underline{X}' / \underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \end{aligned} \quad (\text{A.4})$$

Invoking the definition of the joint jump moments (cf. eq. (7)) and also that of averages the RHS becomes

$$\frac{d\langle X_i X_j \rangle}{dt} = \langle a_2^{ij}(\underline{X}) \rangle + \langle X_i a_1^j(\underline{X}) \rangle + \langle X_j a_1^i(\underline{X}) \rangle, i, j = 1, \dots, N \quad (\text{A.5})$$

where $\underline{a}_2(\underline{X})$ is a second order tensor, whereas $\underline{a}_1(\underline{X})$ is a first order tensor. In the trivial case where $i=j$ the last equation simplifies to

$$\frac{d\langle X_i X_i \rangle}{dt} = \langle a_2^{ii}(\underline{X}) \rangle + 2\langle X_i a_1^i(\underline{X}) \rangle, i = 1, \dots, N \quad (\text{A.6})$$

In the derivation of the third moment equations we start again with the defining moment equation (cf. section 2.2)

$$\langle X_i X_j X_l \rangle = \int X_i X_j X_l P(\underline{X}, t) d\underline{X} \quad (\text{A.7})$$

and the time derivative is defined as follows

$$\frac{d\langle X_j X_i X_l \rangle}{dt} = \int X_i X_j X_l \frac{\partial P(\underline{X}, t)}{\partial t} d\underline{X} \quad (\text{A.8})$$

substituting equation (2) in the last equation we have

$$\begin{aligned} \frac{d\langle X_j X_i X_l \rangle}{dt} &= \int X_j X_i X_l \int [T(\underline{X}/\underline{X}') P(\underline{X}', t) - T(\underline{X}'/\underline{X}) P(\underline{X}, t)] d\underline{X}' d\underline{X} \\ &= \int \int [X_j X_i X_l T(\underline{X}/\underline{X}') P(\underline{X}', t) - X_j X_i X_l T(\underline{X}'/\underline{X}) P(\underline{X}, t)] d\underline{X}' d\underline{X} \end{aligned} \quad (\text{A.9})$$

Noticing that the integration over $\underline{X}$ and $\underline{X}'$ runs over the same domain we can interchange indexes in the last equation, i.e.

$$X_j X_i X_l T(\underline{X}/\underline{X}') P(\underline{X}', t) = X_j' X_i' X_l' T(\underline{X}'/\underline{X}) P(\underline{X}, t) \quad (\text{A.10})$$

thus we have

$$\frac{d\langle X_j X_i X_l \rangle}{dt} = \int \int (X_j' X_i' X_l' - X_j X_i X_l) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \quad (\text{A.11})$$

Skipping through some tedious calculations we use the following relation to transform the RHS of equation (A.11)

$$\begin{aligned}
& (X'_j X'_i X'_l \\
& \quad - X_i X_j X_l) \\
& = (X'_i - X_i)(X'_j \\
& \quad - X_j)(X'_l \\
& \quad - X_l) + X_i(X'_j \\
& \quad - X_j)(X'_l \\
& \quad - X_l) + X_j(X'_i - X_i)(X'_l \\
& \quad - X_l) + X_l(X'_i \\
& \quad - X_i)(X'_j \\
& \quad - X_j) + X_i X_j(X'_l \\
& \quad - X_l) + X_i X_l(X'_j \\
& \quad - X_j) + X_j X_l(X'_i - X_i)
\end{aligned} \quad (\text{A.12})$$

substituting yields

$$\begin{aligned}
& \frac{d\langle X_i X_j X_l \rangle}{dt} \\
& = \int \int (X'_i \\
& \quad - X_i)(X'_j \\
& \quad - X_j)(X'_l \\
& \quad - X_l) T(\underline{X}' / \underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} + \int \int X_i(X'_j \\
& \quad - X_j)(X'_l \\
& \quad - X_l) T(\underline{X}' / \underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} + \int \int X_j(X'_i \\
& \quad - X_i)(X'_l \\
& \quad - X_l) T(\underline{X}' / \underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} + \int \int X_l(X'_i \\
& \quad - X_i)(X'_j \\
& \quad - X_j) T(\underline{X}' / \underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} + \int \int X_i X_j(X'_l \\
& \quad - X_l) T(\underline{X}' / \underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} + \int \int X_i X_l(X'_j \\
& \quad - X_j) T(\underline{X}' / \underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} + \int \int X_j X_l(X'_i \\
& \quad - X_i) T(\underline{X}' / \underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X}
\end{aligned} \quad (\text{A.13})$$

Using the definitions of the joint jump moments (cf. eq. (7)) and that of the averages the RHS eventually becomes

$$\frac{d\langle X_i X_j X_l \rangle}{dt} = \langle a_3^{ijl}(\underline{X}) \rangle + \langle X_i a_2^{jl}(\underline{X}) \rangle + \langle X_j a_2^{li}(\underline{X}) \rangle + \langle X_l a_2^{ij}(\underline{X}) \rangle + \langle X_i X_j a_1^l(\underline{X}) \rangle + \langle X_i X_l a_1^j(\underline{X}) \rangle + \langle X_l X_i a_1^i(\underline{X}) \rangle \quad (\text{A.14})$$

$$i, j, l = 1, \dots, N$$

where $\underline{g}_3(\underline{X})$ is a third order tensor. In the trivial case where $i = j = l$ the last equation simplifies to

$$\frac{d\langle X_i^3 \rangle}{dt} = \langle a_3^{iii}(\underline{X}) \rangle + 3\langle X_i a_2^{ii}(\underline{X}) \rangle + 3\langle X_i^2 a_1^i(\underline{X}) \rangle \quad i = 1, \dots, N \quad (\text{A.15})$$

For the fourth order moments the starting point is again the moment defining equation (cf. section 2.2)

$$\langle X_i X_j X_l X_m \rangle = \int X_i X_j X_l X_m P(\underline{X}, t) d\underline{X} \quad (\text{A.16})$$

and the time derivative is defined as follows

$$\frac{d\langle X_j X_i X_l X_m \rangle}{dt} = \int X_j X_i X_l X_m \frac{\partial P(\underline{X}, t)}{\partial t} d\underline{X} \quad (\text{A.17})$$

substituting equation (2) in the last equation we have

$$\begin{aligned} \frac{d\langle X_j X_i X_l X_m \rangle}{dt} &= \int X_j X_i X_l X_m \int [T(\underline{X}/\underline{X}')P(\underline{X}', t) - T(\underline{X}'/\underline{X})P(\underline{X}, t)] d\underline{X}' d\underline{X} \\ &= \int \int X_j X_i X_l X_m [T(\underline{X}/\underline{X}')P(\underline{X}', t) - T(\underline{X}'/\underline{X})P(\underline{X}, t)] d\underline{X}' d\underline{X} \end{aligned} \quad (\text{A.18})$$

Similarly to the third order moments case the integration in equation (A.18) over $\underline{X}$ and $\underline{X}'$ runs over the same domain thus we can interchange indexes, i.e.

$$X_j X_i X_l X_m T(\underline{X}/\underline{X}')P(\underline{X}', t) = X_j' X_i' X_l' X_m' T(\underline{X}'/\underline{X})P(\underline{X}, t) \quad (\text{A.19})$$

thus we have

$$\frac{d\langle X_j X_i X_l \rangle}{dt} = \int \int (X_j' X_i' X_l' - X_j X_i X_l) T(\underline{X}'/\underline{X})P(\underline{X}, t) d\underline{X}' d\underline{X} \quad (\text{A.20})$$

Skipping through some tedious calculations we use the following relation to transform the RHS of equation (A.11)

$$\begin{aligned}
& (X'_j X'_l X'_m) \\
& \quad - X_i X_j X_l X_m) \\
& = (X'_i - X_i)(X'_j \\
& \quad - X_j)(X'_l \\
& \quad - X_l)(X'_m \\
& \quad - X_m) + X_i(X'_j \\
& \quad - X_j)(X'_l \\
& \quad - X_l)(X'_m \\
& \quad - X_m) + X_j(X'_i \\
& \quad - X_i)(X'_l \\
& \quad - X_l)(X'_m \\
& \quad - X_m) + (X_l(X'_i \\
& \quad - X_i)(X'_j \\
& \quad - X_j)(X'_m \\
& \quad - X_m) + X_m(X'_i \\
& \quad - X_i)(X'_j \\
& \quad - X_j)(X'_l \\
& \quad - X_l) + X_i X_j(X'_l \\
& \quad - X_l)(X'_m \\
& \quad - X_m) + X_i X_l(X'_j \\
& \quad - X_j)(X'_m \\
& \quad - X_m) + X_i X_m(X'_j \\
& \quad - X_j)(X'_l \\
& \quad - X_l) + X_j X_l(X'_i \\
& \quad - X_i)(X'_m \\
& \quad - X_m) + X_j X_m(X'_i \\
& \quad - X_i)(X'_l \\
& \quad - X_l) + X_l X_m(X'_i \\
& \quad - X_i)(X'_j \\
& \quad - X_j) + X_i X_j X_l(X'_m \\
& \quad - X_m) + X_i X_j X_m(X'_l \\
& \quad - X_l) + X_i X_l X_m(X'_j \\
& \quad - X_j) + X_j X_l X_m(X'_i \\
& \quad - X_i)
\end{aligned} \tag{A.21}$$

substituting yields

$$\begin{aligned}
& \frac{d\langle X_i X_j X_l \rangle}{dt} \int \int (X'_i - X_i)(X'_j - X_j)(X'_l - X_l)(X'_m - X_m) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_i (X'_j - X_j)(X'_l - X_l)(X'_m - X_m) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_j (X'_i - X_i)(X'_l - X_l)(X'_m - X_m) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_l (X'_i - X_i)(X'_j - X_j)(X'_m - X_m) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_m (X'_i - X_i)(X'_j - X_j)(X'_l - X_l) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_i X_j (X'_l - X_l)(X'_m - X_m) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_i X_l (X'_j - X_j)(X'_m - X_m) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_i X_m (X'_j - X_j)(X'_l - X_l) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_j X_l (X'_i - X_i)(X'_m - X_m) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_j X_m (X'_i - X_i)(X'_l - X_l) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_l X_m (X'_i - X_i)(X'_j - X_j) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_i X_j X_l (X'_m - X_m) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_i X_j X_m (X'_l - X_l) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_i X_l X_m (X'_j - X_j) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X} \\
& + \int \int X_j X_l X_m (X'_i - X_i) T(\underline{X}'/\underline{X}) P(\underline{X}, t) d\underline{X}' d\underline{X}
\end{aligned} \tag{A.22}$$

Using the definitions of the joint jump moments (cf. eq. (7)) and that of the averages the RHS eventually becomes

$$\begin{aligned}
& \frac{d\langle X_i X_j X_l X_m \rangle}{dt} \\
&= \langle a_4^{ijlm}(\underline{X}) \rangle \\
&+ \langle X_i a_3^{ilm}(\underline{X}) \rangle \\
&+ \langle X_j a_3^{lmi}(\underline{X}) \rangle \\
&+ \langle X_l a_3^{mij}(\underline{X}) \rangle \\
&+ \langle X_m a_3^{ijl}(\underline{X}) \rangle \\
&+ \langle X_i X_j a_2^{lm}(\underline{X}) \rangle \\
&+ \langle X_j X_l a_2^{mi}(\underline{X}) \rangle \quad (A.23) \\
&+ \langle X_l X_m a_2^{ij}(\underline{X}) \rangle \\
&+ \langle X_m X_i a_2^{jl}(\underline{X}) \rangle \\
&+ \langle X_j X_m a_2^{li}(\underline{X}) \rangle \\
&+ \langle X_l X_i a_2^{jm}(\underline{X}) \rangle \\
&+ \langle X_i X_j X_l a_1^m(\underline{X}) \rangle \\
&+ \langle X_j X_l X_m a_1^i(\underline{X}) \rangle \\
&+ \langle X_l X_m X_i a_1^j(\underline{X}) \rangle \\
&+ \langle X_m X_i X_j a_1^l(\underline{X}) \rangle
\end{aligned}$$

$$i, j, l, m = 1, \dots, N$$

where $\underline{a}_4(\underline{X})$ is a fourth order tensor. In the trivial case where $i = j = l = m$ the last equation simplifies to

$$\frac{d\langle X_i^4 \rangle}{dt} = \langle a_4^{iiii}(\underline{X}) \rangle + 4\langle X_i a_3^{iii}(\underline{X}) \rangle + 6\langle X_i^2 a_2^{ii}(\underline{X}) \rangle + 4\langle X_i^3 a_1^i(\underline{X}) \rangle, i = 1, \dots, N \quad (A.24)$$

Appendix B. Appropriate permutation sets for Π_j^m

Recall that Π_j^m is defined as follows

$$\Pi_j^m = \sum_{k=1}^{\binom{m}{j}} \pi_{m,k} \quad (B.1)$$

where each permutation matrix, $\pi_{m,k}$, is defined as

$$\pi_{m,k} = \sum_{i_1, \dots, i_m=1}^N \underbrace{E_{i_1 i_1}^{N \times N} \otimes E_{i_2 i_2}^{N \times N} \dots \otimes E_{i_m i_m}^{N \times N}}_m \quad (B.2)$$

where $E_{ij}^{p \times q}$ is a $p \times q$ matrix (elementary matrix) that has a "1" in the (i, j) position and zeros everywhere else and k denotes the different permutation sets for the subscripts of all $E_{ij}^{p \times q}$

The number of the necessary permutation matrices to construct $\underline{\Pi}_j^m$ depends on both m and j , as the binomial coefficient denotes in equation (B.1). The choice of the appropriate $\binom{m}{j}$ permutation matrices in turn depends on the terms that $\underline{\Pi}_j^m$ multiplies, i.e. $\langle \underline{X}^{(j)} \otimes \text{vec}(a_{m-j}) \rangle$. These terms are column vectors ($N^m \times 1$) and each of the N^m elements can be depicted as $\langle X_{i_1}, X_{i_2} \dots X_{i_j} a_{m-j}^{j_{j+1}, \dots, i_m} \rangle$.

$$\langle \underline{X}^{(j)} \otimes \text{vec}(a_{m-j}) \rangle = \begin{bmatrix} \vdots \\ \langle X_{i_1} X_{i_2} \dots X_{i_j} a_{m-j}^{j_{j+1}, \dots, i_m} \rangle \\ \vdots \end{bmatrix} \quad (\text{B.3})$$

The idea behind choosing the appropriate permutation matrices, $\underline{\pi}_{m,k}$, is simple and straight forward. Through equation (B.2), the k permutation matrices are identified by keeping the first subscript of the elementary matrices constant and changing the second subscript

according to the order in which subscript indices appear in the $\langle X_{i_1, i_2, \dots, i_j} a_{m-j}^{j_{j+1}, \dots, i_m} \rangle$ terms. Note that we are only interested in unique subscript permutations sets.

The methodology presented above will become evident in the following examples. We consider a system with $N=2$ and we construct the $\underline{\Pi}_j^m$ matrices for the cases $m=2$ through $m=4$. • $m=2$

According to equation (19) we have

$$\frac{d\langle \underline{X}^{(2)} \rangle}{dt} = \text{vec}(a_2) + \underline{\Pi}_1^2 \langle \underline{X} \otimes \text{vec}(a_1) \rangle$$

The column vector $\langle \underline{X} \otimes \text{vec}(a_1) \rangle$ has entries of the form $\langle X_{i_1} a_1^{i_2} \rangle$ or $\langle X_{i_2} a_1^{i_1} \rangle$. Note that these two permutation of indices i_1 and i_2 are the only ones producing unique terms. Therefore the corresponding permutation matrices, $\underline{\pi}_{m,k}$, are

$$\begin{aligned} \langle X_{i_1} a_1^{i_2} \rangle: \quad \underline{\pi}_{2,1} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_1}^{2 \times 2} \otimes E_{i_2 i_2}^{2 \times 2} \\ \langle X_{i_2} a_1^{i_1} \rangle: \quad \underline{\pi}_{2,2} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_2}^{2 \times 2} \otimes E_{i_2 i_1}^{2 \times 2} \end{aligned}$$

The sum of the above two terms yields the desired $\underline{\Pi}_1^2$ matrix. In the general case where we have N state variables then

$$\underline{\Pi}_1^2 = \sum_{i_1, i_2=1}^N E_{i_1 i_1}^{N \times N} \otimes E_{i_2 i_2}^{N \times N} + \sum_{i_1, i_2=1}^N E_{i_1 i_2}^{N \times N} \otimes E_{i_2 i_1}^{N \times N} \quad (\text{B.4})$$

• $m = 3$

According to equation (19) we have

$$\frac{d\langle \underline{X}^{(3)} \rangle}{dt} = \text{vec}(a_3) + \underline{\Pi}_1^3 \langle \underline{X} \otimes \text{vec}(a_2) \rangle + \underline{\Pi}_2^3 \langle \underline{X}^{(2)} \otimes \text{vec}(a_1) \rangle$$

Starting with $\underline{\Pi}_1^3$, the column vector $\langle \underline{X} \otimes \text{vec}(a_2) \rangle$ has entries of the form $\langle X_{i_1} a_2^{i_2 i_3} \rangle$ or $\langle X_{i_2} a_2^{i_3 i_1} \rangle$ or $\langle X_{i_3} a_2^{i_1 i_2} \rangle$. Note that these three permutation of indices i_1 , i_2 and i_3 are the only ones producing unique terms. For instance, the term $\langle X_{i_1} a_2^{i_3 i_2} \rangle$ is equivalent to $\langle X_{i_1} a_2^{i_2 i_3} \rangle$ and therefore is not unique. Then the corresponding permutation matrices, $\underline{\pi}_{m,k}$, are

$$\begin{aligned} \langle X_{i_1} a_2^{i_2 i_3} \rangle: \quad \underline{\pi}_{3,1} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_1}^{2 \times 2} \otimes E_{i_2 i_2}^{2 \times 2} \otimes E_{i_3 i_3}^{2 \times 2} \\ \langle X_{i_2} a_2^{i_3 i_1} \rangle: \quad \underline{\pi}_{3,2} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_2}^{2 \times 2} \otimes E_{i_2 i_3}^{2 \times 2} \otimes E_{i_3 i_1}^{2 \times 2} \\ \langle X_{i_3} a_2^{i_1 i_2} \rangle: \quad \underline{\pi}_{3,3} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_3}^{2 \times 2} \otimes E_{i_2 i_1}^{2 \times 2} \otimes E_{i_3 i_2}^{2 \times 2} \end{aligned}$$

The sum of the above $\binom{3}{1}=3$ terms yields the desired $\underline{\Pi}_1^3$ matrix. In the general case where we have N state variables then

$$\underline{\Pi}_1^3 = \sum_{i_1, i_2=1}^N E_{i_1 i_1}^{N \times N} \otimes E_{i_2 i_2}^{N \times N} \otimes E_{i_3 i_3}^{N \times N} + \sum_{i_1, i_2=1}^N E_{i_1 i_2}^{N \times N} \otimes E_{i_2 i_3}^{N \times N} \otimes E_{i_3 i_1}^{N \times N} + \sum_{i_1, i_2=1}^N E_{i_1 i_3}^{N \times N} \otimes E_{i_2 i_1}^{N \times N} \otimes E_{i_3 i_2}^{N \times N} \quad (\text{B.5})$$

Continuing with $\underline{\Pi}_2^3$, the column vector $\langle \underline{X}^{(2)} \otimes \text{vec}(a_1) \rangle$ has entries of the form $\langle X_{i_1} X_{i_2} a_1^{i_3} \rangle$ or $\langle X_{i_2} X_{i_3} a_1^{i_1} \rangle$ or $\langle X_{i_3} X_{i_1} a_1^{i_2} \rangle$. Note that these three permutation of indices i_1 , i_2 and i_3 are the only ones producing unique terms. For instance, the term $\langle X_{i_1} X_{i_3} a_1^{i_2} \rangle$ is equivalent to $\langle X_{i_3} X_{i_1} a_1^{i_2} \rangle$ and therefore is not unique. Then the corresponding permutation matrices, $\underline{\pi}_{m,k}$, are

$$\begin{aligned} \langle X_{i_1} X_{i_2} a_1^{i_3} \rangle: \quad \underline{\pi}_{3,1} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_1}^{2 \times 2} \otimes E_{i_2 i_2}^{2 \times 2} \otimes E_{i_3 i_3}^{2 \times 2} \\ \langle X_{i_2} X_{i_3} a_1^{i_1} \rangle: \quad \underline{\pi}_{3,2} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_2}^{2 \times 2} \otimes E_{i_2 i_3}^{2 \times 2} \otimes E_{i_3 i_1}^{2 \times 2} \\ \langle X_{i_3} X_{i_1} a_1^{i_2} \rangle: \quad \underline{\pi}_{3,3} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_3}^{2 \times 2} \otimes E_{i_2 i_1}^{2 \times 2} \otimes E_{i_3 i_2}^{2 \times 2} \end{aligned}$$

The sum of the above $\binom{3}{2}=3$ terms yields the desired Π_2^3 matrix, which if you notice is equal to Π_1^3 . This also implies that in the general case for N state variables

$$\Pi_2^3 = \Pi_1^3 \quad (\text{B.6})$$

• m = 4

According to equation (19) we have

$$\frac{d\langle X^{(4)} \rangle}{dt} = \text{vec}(a_4) + \Pi_1^4 \langle X \otimes \text{vec}(a_3) \rangle + \Pi_2^4 \langle X^{(2)} \otimes \text{vec}(a_2) \rangle + \Pi_3^4 \langle X^{(3)} \otimes \text{vec}(a_1) \rangle$$

Starting with Π_1^4 , the column vector $\langle X \otimes \text{vec}(a_3) \rangle$ has unique entries of the form $\langle X_{i_1} a_3^{i_2 i_3 i_4} \rangle$ or $\langle X_{i_2} a_3^{i_3 i_4 i_1} \rangle$ or $\langle X_{i_3} a_3^{i_4 i_1 i_2} \rangle$ or $\langle X_{i_4} a_3^{i_1 i_2 i_3} \rangle$. Then the corresponding permutation matrices, $\pi_{m,k}$ are

$$\begin{aligned} \langle X_{i_1} a_3^{i_2 i_3 i_4} \rangle : \pi_{\underline{4},1} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_1}^{2 \times 2} \otimes E_{i_2 i_2}^{2 \times 2} \otimes E_{i_3 i_3}^{2 \times 2} \otimes E_{i_4 i_4}^{2 \times 2} \\ \langle X_{i_2} a_3^{i_3 i_4 i_1} \rangle : \pi_{\underline{4},2} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_2}^{2 \times 2} \otimes E_{i_2 i_3}^{2 \times 2} \otimes E_{i_3 i_4}^{2 \times 2} \otimes E_{i_4 i_1}^{2 \times 2} \\ \langle X_{i_3} a_3^{i_4 i_1 i_2} \rangle : \pi_{\underline{4},3} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_3}^{2 \times 2} \otimes E_{i_2 i_4}^{2 \times 2} \otimes E_{i_3 i_1}^{2 \times 2} \otimes E_{i_4 i_2}^{2 \times 2} \\ \langle X_{i_4} a_3^{i_1 i_2 i_3} \rangle : \pi_{\underline{4},4} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_4}^{2 \times 2} \otimes E_{i_2 i_1}^{2 \times 2} \otimes E_{i_3 i_2}^{2 \times 2} \otimes E_{i_4 i_3}^{2 \times 2} \end{aligned}$$

The sum of the above $\binom{4}{1}=4$ terms yields the desired Π_1^4 matrix. In the general case where we have N state variables then

$$\begin{aligned} \Pi_1^4 &= \sum_{i_1, i_2=1}^N E_{i_1 i_1}^{N \times N} \otimes E_{i_2 i_2}^{N \times N} \otimes E_{i_3 i_3}^{N \times N} \otimes E_{i_4 i_4}^{N \times N} \\ &+ \sum_{i_1, i_2=1}^N E_{i_1 i_2}^{N \times N} \otimes E_{i_2 i_3}^{N \times N} \otimes E_{i_3 i_4}^{N \times N} \otimes E_{i_4 i_1}^{N \times N} \\ &+ \sum_{i_1, i_2=1}^N E_{i_1 i_3}^{N \times N} \otimes E_{i_2 i_4}^{N \times N} \otimes E_{i_3 i_1}^{N \times N} \otimes E_{i_4 i_2}^{N \times N} \\ &+ \sum_{i_1, i_2=1}^N E_{i_1 i_4}^{N \times N} \otimes E_{i_2 i_1}^{N \times N} \otimes E_{i_3 i_2}^{N \times N} \otimes E_{i_4 i_3}^{N \times N} \end{aligned} \quad (\text{B.7})$$

Using similar arguments we can also conclude that the above permutation matrices are those that apply in the case of Π_3^4 . Therefore

$$\underline{\Pi}_3^4 = \underline{\Pi}_1^4 \quad (\text{B.8})$$

In the case of $\underline{\Pi}_2^4$ the corresponding column vector is $\langle \underline{X}^{(2)} \otimes \text{vec}(a_2) \rangle$ and the unique entries of column are of the form

$$\langle X_{i_1} X_{i_2} a_2^{i_3 i_4} \rangle \text{ or } \langle X_{i_2} X_{i_3} a_2^{i_4 i_1} \rangle \text{ or } \langle X_{i_3} X_{i_4} a_2^{i_1 i_2} \rangle \text{ or } \langle X_{i_4} X_{i_1} a_2^{i_2 i_3} \rangle \text{ or } \langle X_{i_2} X_{i_4} a_2^{i_3 i_1} \rangle \text{ or } \langle X_{i_3} X_{i_1} a_2^{i_2 i_4} \rangle.$$

Note that there are no other unique sequences i_1, i_2, i_3 and i_4 that would create unique column entries. On the other hand, for example we could have substituted

$\langle X_{i_2} X_{i_4} a_2^{i_3 i_1} \rangle$ with $\langle X_{i_2} X_{i_4} a_2^{i_3 i_1} \rangle$, but we could not use both as they are dependent. This dependence is even more evident if we recall that the tensors a_m are symmetric. Then the corresponding permutation matrices, $\underline{\pi}_{m,k}$, are

$$\begin{aligned} \langle X_{i_1} X_{i_2} a_2^{i_3 i_4} \rangle : \underline{\pi}_{4,1} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_1}^{2 \times 2} \otimes E_{i_2 i_2}^{2 \times 2} \otimes E_{i_3 i_3}^{2 \times 2} \otimes E_{i_4 i_4}^{2 \times 2} \\ \langle X_{i_2} X_{i_3} a_2^{i_4 i_1} \rangle : \underline{\pi}_{4,2} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_2}^{2 \times 2} \otimes E_{i_2 i_3}^{2 \times 2} \otimes E_{i_3 i_4}^{2 \times 2} \otimes E_{i_4 i_1}^{2 \times 2} \\ \langle X_{i_3} X_{i_4} a_2^{i_1 i_2} \rangle : \underline{\pi}_{4,3} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_3}^{2 \times 2} \otimes E_{i_2 i_4}^{2 \times 2} \otimes E_{i_3 i_1}^{2 \times 2} \otimes E_{i_4 i_2}^{2 \times 2} \\ \langle X_{i_4} X_{i_1} a_2^{i_2 i_3} \rangle : \underline{\pi}_{4,4} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_4}^{2 \times 2} \otimes E_{i_2 i_1}^{2 \times 2} \otimes E_{i_3 i_2}^{2 \times 2} \otimes E_{i_4 i_3}^{2 \times 2} \\ \langle X_{i_2} X_{i_4} a_2^{i_3 i_1} \rangle : \underline{\pi}_{4,5} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_2}^{2 \times 2} \otimes E_{i_2 i_4}^{2 \times 2} \otimes E_{i_3 i_3}^{2 \times 2} \otimes E_{i_4 i_1}^{2 \times 2} \\ \langle X_{i_3} X_{i_1} a_2^{i_2 i_4} \rangle : \underline{\pi}_{4,6} &= \sum_{i_1, i_2=1}^2 E_{i_1 i_3}^{2 \times 2} \otimes E_{i_2 i_1}^{2 \times 2} \otimes E_{i_3 i_2}^{2 \times 2} \otimes E_{i_4 i_4}^{2 \times 2} \end{aligned}$$

The sum of the above $\binom{4}{2}=6$ terms yields the desired $\underline{\Pi}_2^4$ matrix. In the general case where we have N state variables then

$$\underline{\Pi}_2^4 = \underline{\Pi}_1^4 + \sum_{i_1, i_2=1}^N E_{i_1 i_2}^{N \times N} \otimes E_{i_2 i_4}^{N \times N} \otimes E_{i_3 i_3}^{N \times N} \otimes E_{i_4 i_1}^{N \times N} + \sum_{i_1, i_2=1}^2 E_{i_1 i_3}^{N \times N} \otimes E_{i_2 i_1}^{N \times N} \otimes E_{i_3 i_2}^{N \times N} \otimes E_{i_4 i_4}^{N \times N} \quad (\text{B.9})$$

The same methodology can be applied to construct any given $\underline{\Pi}_j^m$ matrix. Though as m increases the terms in the sum of equation (B.1) increase which implies that $\underline{\pi}_{m,k}$ increase too. The idea is to find all the permutations sets of $i_1, i_2, \dots, i_m$ that uniquely describe the elements of the column vector, $\langle \underline{X}^{(j)} \otimes \text{vec}(a_{m-j}) \rangle$ and then substitute them in equation (B.2).

References

1. van Kampen, N. Stochastic processes in physics and chemistry. North Holland: 1992.
2. Gardiner, C. Handbook of Stochastic Methods for Physics, Chemistry and Natural Sciences, Synergetics. 3rd edition. Berlin: Springer; 2004.
3. Salis H, Kaznessis Y. Accurate hybrid stochastic simulation of a system of coupled chemical or biochemical reactions. J Chem Phys. 2005; 122:54103. [PubMed: 15740306]

4. Salis H, Kaznessis YN. An equation-free probabilistic steady-state approximation: dynamic application to the stochastic simulation of biochemical reaction networks. *J Chem Phys.* 2005; 123:214106. [PubMed: 16356038]
5. Gibson MA, Bruck J. Efficient exact stochastic simulation of chemical systems with many species and many channels. *Journal of Physical Chemistry A.* 2000; 104:1876.
6. Cao Y, Gillespie DT, Petzold LR. Efficient step size selection for the tau-leaping simulation method. *J Chem Phys.* 2006; 124:44109.
7. Rathinam M, Petzold LR, Yang C, Gillespie DT. Stiffness in stochastic chemically reacting systems: the implicit tau-leaping method. *Journal of Chemical Physics.* 2003; 119:12784.
8. Cao Y, Gillespie DT, Petzold LR. The slow-scale stochastic simulation algorithm. *The Journal of Chemical Physics.* 2005; 122 014116.
9. Samant A, Vlachos DG. Overcoming stiffness in stochastic simulation stemming from partial equilibrium: a multiscale monte carlo algorithm. *J Chem Phys.* 2005; 123:144114. [PubMed: 16238381]
10. Munsky B, Khammash M. The finite state projection algorithm for the solution of the chemical master equation. *J Chem Phys.* 2006; 124:44104.
11. Sotiropoulos V, Kaznessis YN. An adaptive time step scheme for a system of stochastic differential equations with multiple multiplicative noise: Chemical langevin equation, a proof of concept. *The Journal of Chemical Physics.* 2008; 128 014103.
12. Sotiropoulos V, Contou-Carrere M, Daoutidis P, Kaznessis YN. Model reduction of multiscale chemical langevin equations: A numerical case study. *Computational Biology and Bioinformatics, IEEE/ACM Transactions on.* 2009; 6:470–482.
13. Salis H, Sotiropoulos V, Kaznessis YN. Multiscale hy3s: Hybrid stochastic simulation for supercomputers. *BMC Bioinformatics.* 2006; 7:93. [PubMed: 16504125]
14. Hill AD, Tomshine JR, Weeding EMB, Sotiropoulos V, Kaznessis YN. Synbioss: the synthetic biology modeling suite. *Bioinformatics.* 2008; 24:2551–2553. [PubMed: 18757873]
15. Gillespie, D. *Markov processes: An introduction for physical scientists.* Academic Press; 1992.
16. Moyal J. Stochastic processes and statistical physics. *Journal of the Royal Statistical Society. Series B (Methodological).* 1949; 11:150–210.
17. Kaznessis YN. Multi-scale models for gene network engineering. *Chemical Engineering Science.* 2006; 61:940.
18. Sotiropoulos V, Kaznessis YN. Synthetic tetracycline-inducible regulatory networks: computer-aided design of dynamic phenotypes. *BMC Systems Biology.* 2007; 1:7. [PubMed: 17408514]
19. Kaznessis YN. Models for synthetic biology. *BMC Systems Biology.* 2007; 1:47. [PubMed: 17986347]
20. Tomshine J, Kaznessis YN. Optimization of a stochastically simulated gene network model via simulated annealing. *Biophysical journal.* 2006; 91:3196–3205. [PubMed: 16920827]
21. Tuttle L, Salis H, Tomshine J, Kaznessis Y. Model-driven designs of an oscillating gene network. *Biophysical journal.* 2005; 89:3873–3883. [PubMed: 16183880]
22. Salis H, Kaznessis YN. Computer-aided design of modular protein devices: Boolean and gene activation. *Phys Biol.* 2006; 3:295–310. [PubMed: 17200605]
23. Ramalingam K, Tomshine J, Maynard J, Kaznessis Y. Forward engineering of synthetic bio-logical and gates. *Biochemical Engineering Journal.* 2009; 47:38–47.
24. McQuarrie DA. Stochastic approach to chemical kinetics. *Journal of Applied Probability.* 1967; 4:413.
25. Hespanha JP, Singh A. Stochastic models for chemically reacting systems using polynomial stochastic hybrid systems. *International Journal of Robust and Nonlinear Control.* 2005; 15:669–689.
26. Goutsias J. Classical versus stochastic kinetics modeling of biochemical reaction systems. *Biophys. J.* 2007; 92:2350–2365. [PubMed: 17218456]
27. Gillespie CS. Moment-closure approximations for mass-action models. *Systems Biology, IET.* 2009; 3:52.

28. Lee CH, Kim K, Kim P. A moment closure method for stochastic reaction networks. *The Journal of Chemical Physics*. 2009; 130:134107. [PubMed: 19355717]
29. Brewer J. Kronecker products and matrix calculus in system theory. *IEEE Transactions on Circuits and Systems*. 1978; 25:772–781.
30. Wolf DM, Arkin AP. Fifteen minutes of fim: control of type 1 pili expression in e. coli. *Omics*. 2002; 6:91–114. [PubMed: 11881836]
31. St-Pierre F, Endy D. Determination of cell fate selection during phage lambda infection. *Proceedings of the National Academy of Sciences*. 2008; 105:20705.
32. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. *Science*. 2002; 297:1183–1186. [PubMed: 12183631]
33. Gillespie DT. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem*. 1977; 81:2340–2361.
34. Gillespie DT. The chemical langevin equation. *Journal of Chemical Physics*. 2000; 113:297.
35. Thakur AK, Rescigno A, DeLisi C. Stochastic theory of second-order chemical reactions. *The Journal of Physical Chemistry*. 1978; 82:552–558.
36. Krishnarajah I, Cook A, Marion G, Gibson G. Novel moment closure approximations in stochastic epidemics. *Bulletin of mathematical biology*. 2005; 67:855–873. [PubMed: 15893556]
37. Singh, A.; Hespanha, J. , . Moment closure techniques for stochastic models in population biology; *American Control Conference*; p. 6
38. Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in escherichia coli. *Nature*. 2000; 403:339–342. [PubMed: 10659857]
39. Gunawan R, Cao Y, Petzold L, Doyle r, J F. Sensitivity analysis of discrete stochastic systems. *Biophys J*. 2005; 88:2530–2540. [PubMed: 15695639]
40. Mohammad-Djafari A. A matlab program to calculate the maximum entropy distributions. 2001

Figure 1.

Actual and reconstructed probability distributions for the Schlögl model, depicted in Table 1, using the moment data in Table 2 (a) Probability distribution of species X at times $t = 2$ s, 4 s and 20 s. (b) Comparison between the actual and reconstructed probability distributions using different moment sets at time $t = 2$ s (c) Comparison between the actual and reconstructed probability distributions using different moment sets at time $t = 4$ s (d) Comparison between the actual and reconstructed probability distributions using different moment sets at time $t = 20$ s.

Table 1

Reactions and Parameters for the Schlögl Model

Set of Reactions ^a	Mesosopic Reaction Rates ^b	Initial Values ^c
$A + 2X \xrightarrow{k_1} 3X$	$k_1 = 3 \times 10^{-7}$	$[X]_0 = 247$
$3X \xrightarrow{k_2} 2X + A$	$k_2 = 10^{-4} \times 10^2$	
$B \xrightarrow{k_3} X$	$k_3 = 10^{-3}$	$[A] = 10^5$
$X \xrightarrow{k_4} B$	$k_4 = 3.5 \times 10^2$	$[B] = 2 \times 10^5$

^a A and B are buffer species.^b for 1st order reactions the units are s⁻¹, for 2nd order reactions the units are molecules⁻¹s⁻¹ and for 3rd order reactions the units are molecules⁻²s⁻¹.^c number of molecules

Table 2

First Eight Moment Values for X in the Schlögl Model

Moments	Value at t = 2 s	Value at t = 4 s	Value at t = 20 s
$\langle X \rangle$	268.9871	295.6759	301.6574
$\langle X^2 \rangle$	9.2903×10^4	1.3364×10^5	1.4733×10^5
$\langle X^3 \rangle$	3.7966×10^7	7.0001×10^7	8.1413×10^7
$\langle X^4 \rangle$	1.7179×10^{10}	3.8403×10^{10}	4.6203×10^{10}
$\langle X^5 \rangle$	8.2743×10^{12}	2.1488×10^{13}	2.6470×10^{13}
$\langle X^6 \rangle$	4.1528×10^{15}	1.2170×10^{16}	1.5262×10^{16}
$\langle X^7 \rangle$	2.1465×10^{18}	6.9578×10^{18}	8.8511×10^{18}
$\langle X^8 \rangle$	1.1347×10^{21}	4.0100×10^{21}	5.1617×10^{21}