Published in final edited form as:

Methods Mol Biol. 2015 ; 1244: 179–191. doi:10.1007/978-1-4939-1878-2_9.

Chemical Master Equation Closure for Computer-Aided Synthetic Biology

Patrick Smadbeck and Yiannis N. Kaznessis*

Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455, USA

SUMMARY

With inexpensive DNA synthesis technologies, we can now construct biological systems by quickly piecing together DNA sequences. Synthetic biology is the promising discipline that focuses on the construction of these new biological systems. Synthetic biology is an engineering discipline, and as such, it can benefit from mathematical modeling. This chapter focuses on mathematical models of biological systems. These models take the form of chemical reaction networks. The importance of stochasticity is discussed and methods to simulate stochastic reaction networks are reviewed. A closure scheme solution is also presented for the master equation of chemical reaction networks. The master equation is a complete model of randomly evolving molecular populations. Because of its ambitious character, the master equation remained unsolved for all but the simplest of molecular interaction networks for over seventy years. With the first complete solution of chemical master equations, a wide range of experimental observations of biomolecular interactions may be mathematically conceptualized. We anticipate that models based on the closure scheme described herein may assist in rationally designing synthetic biological systems.

Keywords

Synthetic Biology; Computer-Aided Design; Multiscale Models; Chemical Master Equation; Closure Schemes

1. INTRODUCTION

Numerous synthetic gene circuits have been created in the past decade, including bistable switches, oscillators, inducible activators and logic gates [1-6]. Applications abound from biofuel synthesis to antibiotic technologies [7-10]. Designing synthetic gene regulatory networks now takes advantage of an ever-expanding toolbox of molecular components and of developed technologies for inexpensively manipulating DNA sequences. Although recently developed designs of regulatable gene networks are ingenious, there are still numerous limitations in rationally engineering synthetic biological systems. This is especially true if targeted phenotypes are nonlinear, such as the ones observed in bistable or oscillatory gene networks.

* Corresponding author's, yiannis@umn.edu.

Multiscale mathematical tools have been developed in an attempt to rationalize synthetic biology [11-18]. Multiscale models that expand on traditional mathematics developed by scientists and engineers are necessary to model kinetic and thermodynamic processes. Biological phenotypes can be reduced to networks of biomolecular interactions, with the time and length scales of these often spanning several orders of magnitude. Although the principles of thermodynamics, kinetics and transport phenomena apply to biological systems, these systems differ from traditional, industrial-scale chemical systems in an important, fundamental way: they are occasionally far from the thermodynamic limit. This theoretical limit is attained when the number of molecules of molecular species in the system increases toward infinity. However, the fact that biomolecular systems can be very far from the thermodynamic limit, with reactants/products numbering only very small numbers of molecules in the system, hinders the use of continuous-deterministic models. Indeed, using ordinary differential equations for simulating the reaction kinetics of these systems can be distinctly false, especially if experimentally observed nonlinearities are to be captured by the models. The need arises then for stochastic models that account for inherent, thermal noise, which is manifest as phenotypic distributions at the population growth/interaction levels.

This assessment is not new. The importance of modeling formalisms appropriate for systems away from the thermodynamic limit was recognized more than 50 years ago by McQuarrie, Moyal and Oppenheim [19-23], among others. These physical chemists developed the chemical master equation (CME) that follows the time changes of the probability the state is at any point in the available state space.

Without the ability to solve the CME, scientists turned to approximations. In 1976, Daniel Gillespie developed a computer algorithm that could sample the master probability distribution with numerical simulations of networks of reactions [24, 25]. Although Gillespie's methods were not widely recognized for almost 20 years, his algorithms found fertile ground for development in efforts to model biological systems. Nowadays, a community of scientists and engineers is continually working to improve the computational efficiency and accuracy of algorithms that simulate chemical reacting systems [26-35].

In the following section, we introduce the chemical master equation and discuss the challenges faced when developing solutions. We then summarize algorithms based on Gillespie's stochastic simulation algorithm (SSA). We return to the master equation and present the recently developed zero-information closure scheme. Finally we present a simple example of a bistable reaction networks and the solution of the master equation. We conclude by speculating on the impact a closure scheme may have on computer-aided synthetic biology.

2. METHODS FOR MULTISCALE MODELS OF BIOMOLECULAR SYSTEMS

2.1 Chemical Master Equation

The chemical master equation (CME) is the mathematical foundation for modeling stochastic chemical reactions [19-23]. The common form is equivalent to the more general Chapman-Kolmogorov equation as applied to Markov processes. In particular, it will be a

Markov chain with a discrete set of possible states, or “state space”, occurring in continuous time. Here, states refer to numbers of molecules present in the system. In general, for time t , this will be a vector $\underline{X}(t) = [X_1 \dots X_N]$, where X_i is number of molecules of the i -th chemical species with $i = 1, \dots, N$.

Transitions between states of the Markov chain occur when a chemical reaction occurs. Reactions in biological systems may include covalent reactions, bindings, conformational changes, transcriptional elongation events, etc.

The general form of the CME is:

$$\frac{\partial P(\underline{X};t)}{\partial t} = \sum_{\underline{X}'} \left[T(\underline{X}|\underline{X}') P(\underline{X}';t) - T(\underline{X}'|\underline{X}) P(\underline{X};t) \right] \quad (1)$$

Where $P(\underline{X};t)$ is the probability of being in state $\underline{X}$ at time t , and $T(\underline{X}|\underline{X}')$ is the transition probability of going from state $\underline{X}'$ to state $\underline{X}$ per unit time. The CME describes the dynamics of stochastic systems exactly, but has been until recently, for all but the simplest systems, mathematically intractable [19].

The reason analytical solutions to the chemical master equation remained elusive becomes clear when the master equation is recast in equivalent terms of probability moments – the probability distribution average, the variance, and so on:

$$\frac{\partial \underline{\mu}}{\partial t} = A \underline{\mu} + A' \underline{\mu}' \quad (2)$$

where $\underline{\mu}$ is the vector of moments up to order M and A is the matrix describing the linear portion of the moment equations. On the right, $\underline{\mu}'$ is the vector of higher-order moments, and the corresponding matrix A' . Generating the matrices in Equation 2 can be performed either analytically [36, 37] or numerically [38]. For linear systems with only 0th or 1st-order reactions, A' is empty. For other systems, A' is not empty and the set of ODEs becomes infinite, and thus intractable.

Consequently, the Gillespie stochastic simulation algorithm (SSA) and the Chemical Langevin Equation (CLE) formalism were developed to approximate the dynamic solution to the CME.

2.2 Stochastic-discrete and stochastic-continuous algorithms

The SSA utilizes Monte Carlo sampling to circumvent the mathematical difficulties inherent to stochastic simulation. Gillespie proved that, with an assumption of a well-mixed volume, chemical reactions can be defined as exponentially distributed random events. For example, given a reaction j with propensity $a_j(\underline{X})$, which is analogous to the reaction macroscopic rate and that depends on the current state $\underline{X}$, the time to the next reaction is determined using a uniform random number (URN) as:

$$\tau_j = -\frac{\ln URN}{a_j(\underline{X})} \quad (3)$$

Using this characterization, individual trajectories through time can be generated for any arbitrary chemical reaction network. The exact probability distribution can be rebuilt from an ensemble of individual trajectories. This method was extended and made efficient by Gibson and Bruck [26] and is best described as a stochastic-discrete-space algorithm.

In principle, the SSA produces exact distributions at the limit of infinite simulated trajectories and thus can be made arbitrarily accurate. The primary drawback is that producing sufficient numbers of trajectories can be computationally taxing. For systems in which many reactions occur in short periods of time the computational load is substantial.

The common situation in which the Gillespie algorithm becomes prohibitively expensive is when the state space can be accurately described as a continuum. In such cases the number of reactions occurring in a short period of time can become effectively infinite. The chemical Langevin equation is then an alternative to the SSA that maintains accuracy but avoids the computational costs of a discrete-space algorithm. A Langevin equation is a stochastic differential equation (SDE) that adds a random variable to a deterministic system to produce fluctuations. For stochastic chemical systems the CLE is commonly written as:

$$d\underline{X} = \sum_j \left[a_j(\underline{X}) \underline{v}_j dt + \sqrt{a_j(\underline{X})} \underline{v}_j dW_j \right] \quad (4)$$

where $a_j(\underline{X})$ is the propensity for reaction j as a function of the state $\underline{X}$, $\underline{v}_j$ is the stoichiometric vector for reaction j , and dW_j is a Wiener process, a continuous-time stochastic process, also called Brownian motion. The CLE is a stochastic-continuous-space algorithm.

The CLE is a good approximation of the CME when the state space is roughly continuous. The primary drawback is accuracy. When the continuous state space approximation does not hold, accuracy is not guaranteed.

2.3 Hybrid Algorithms

The drawbacks of SSA algorithms are especially important when considering biological systems. First, many biological models can be described as stiff in that there are multiple time scales involved. For biological systems, using the SSA can be time consuming, but using the CLE can be inaccurate. Second, in biological systems many reactions change in the course of a simulation. In particular, consider an oscillatory system in which components will alternate from nearly zero to a large number and back relatively quickly. Thus, any solution must be able to handle the determination of fast and slow reaction sets dynamically. The key to fast and accurate biological simulation is in selectively utilizing the appropriate algorithms.

Algorithms that utilize both the SSA and CLE are called hybrid stochastic algorithms. The reaction set is split into two regimes, fast reactions and slow reactions, and continuous-space stochastic algorithms are utilized, when possible, for the fast reactions. For stiff systems a hybrid algorithm must be able to handle disparate time scales efficiently and to dynamically characterize reactions as slow and fast. The Hybrid Stochastic Simulator for Supercomputers (Hy3S) is such an algorithm, designed specifically for biological model simulation on supercomputers [13].

The first aspect of Hy3S to be discussed is dynamic partitioning of a reaction network into a set of fast reactions and slow reactions, such that different algorithms can be applied to each set. Initially all reactions are defined as slow reactions and re-defined as fast if two conditions hold. First, a reaction is fast if many reaction events occur in a small increment of time. This condition is defined by the parameter λ and mathematically defined as:

$$a_j(\underline{X}) \Delta t \geq \lambda \gg 1 \quad (5)$$

Second, for fast reactions the effect of the reaction on the reactants and products must be sufficiently small. This condition is described by a parameter ε , and mathematically defined as:

$$X_i(t) > \varepsilon \cdot |v_{i,j}| \quad (6)$$

where i is a reactant or product of reaction j . The common values for λ and ε are 10 molecules/s and 100 molecules, respectively [13].

The reactions classified as fast are then simulated using the Chemical Langevin Equation, while the slow reactions utilize the SSA. Stochastic differential equations like the CLE can be integrated using different algorithms depending on the desired accuracy. In the case of Hy3S two methods in particular are used, the Euler-Maruyama method and the Milstein method [13].

In order to maintain accuracy in both regimes a method for determining when a slow reaction, and which slow reaction, occurs is necessary. In the Hy3S suite this is done by tracking zero-crossing events for each slow reaction. The variable tracked is denoted by R_j where j is one of the slow reactions in a system. The governing equation is then:

$$\frac{dR_j}{dt} \Big|_t = a_j(\underline{X}) \Big|_t \quad (7)$$

with an initial condition of $R_j|_{t_0} = \ln URN_j$. Zero-crossing events occur when $R_j > 0$ at the end of a time step.

Numerous methods have been developed for reducing run times for complex stochastic system, such as quasi-steady state approximations [29], and dynamic integration stepping techniques [30].

In all, there are eight hybrid algorithms in the Hy3S suite depending on which integration method is used (Euler-Maruyama or Milstein), the time-step (Fixed or adaptive), and whether the algorithm is parallelized (serial or MPI). Along with these algorithms, a standard SSA (serial and MPI) is included making the total number of algorithms ten.

2.4 Zero-information closure of the master chemical equation

Recently, we presented a zero information (ZI) closure scheme for the chemical master equation of reaction networks that may include nonlinear reaction kinetics. This section explains the theoretical underpinnings of the zero-information moment closure method. The section starts with an outline of ZI-closure and then provides the algorithm for determining the maximum entropy distribution. It concludes with a Schlögl model example to clarify the calculation of several important equations [39, 40]. This section expands upon a previously described ODE solving scheme and steady-state determination method [39].

A closure scheme is used to determine the key relationship between higher and lower-order moments necessary to solve the moment equations in Equation (2):

$$\underline{\mu}' = F(\underline{\mu}) \quad (8)$$

In most closure schemes $F(\underline{\mu})$ is an analytical expression related to a well-known characteristic equation. In the case of ZI-closure the closure scheme refers to the maximum-entropy distribution or most-likely distribution.

For a single component system with a probability distribution $p(x)$, the information entropy is defined as:

$$H = - \sum_{x=0}^{\infty} p(x) \ln p(x) \quad (9)$$

For an unconstrained system the resulting maximum-entropy distribution is simply uniform. However, values of the lower-order moments, $\underline{\mu}$, act as constraints on the system. The result is best solved using a Lagrange multiplier method (here assuming a simple component with M known lower order moments):

$$\begin{aligned} \Lambda &= H - \lambda_0 g_0 - \lambda_1 g_1 - \dots - \lambda_M g_M \\ g_0 &= \sum_{x=0}^{\infty} p(x) - 1 \\ g_1 &= \sum_{x=0}^{\infty} x p(x) - \langle x \rangle \\ &\vdots \\ g_M &= \sum_{x=0}^{\infty} x^M p(x) - \langle x^M \rangle \end{aligned} \quad (10)$$

Taking Equation 10, the maximum is found by differentiating by $p(x)$ and setting the result to zero:

$$\frac{\partial \Lambda}{\partial p(x)} = -\ln p(x) - 1 - \lambda_0 - \lambda_1 x - \dots - \lambda_M x^M = 0 \quad (11)$$

or, trivially:

$$p_H(x) = \exp(-1 - \lambda_0 - \lambda_1 x - \dots - \lambda_M x^M) \quad (12)$$

An analytical expression for the maximum entropy distribution is determined with the same number of parameters as the number of known lower-order moments.

With Equation (9) it should be obvious how Equation (2) is satisfied. Using the lower-order moments, $\underline{\mu}$, the maximum entropy distribution is found by determining the Lagrange parameters, $\underline{\lambda}$. The maximum-entropy moments, denoted with a subscript H, can be determined from Equation (9). Take, for example, the determination of $\langle x^m \rangle_H$ in a single component system:

$$\langle x^m \rangle_H = \sum_{x=0}^{\infty} x^m p_H(x) \quad (13)$$

Again, this method is extended to multi-component systems. Zero-information closure uses maximum-entropy moments as approximations for higher-order moments in simulation:

$$\underline{\mu}' = \underline{\mu}'_H \quad (14)$$

Note that the order of closure, M, used throughout the section is applied without drawing any conclusions about accuracy. At present the order of closure is chosen in a trial-and-error fashion. In the future, error analysis may be used to determine the optimal closure order using ZI-Closure.

The method used within this study for the determination of the Lagrange parameters, $\underline{\lambda}$, given a set of known lower-order moments, $\underline{\mu}$, is a simple Newton-Raphson optimization scheme, although more sophisticated algorithms may be explored.

2.5 ZI-closure example

To conclude this section, results for zero-information closure will be presented along with an example of a novel analytical method now available for network analysis. The Schlögl model was chosen for two main reasons. First, it may exhibit two main peaks in its probability distribution (bimodality). It can thus require up to 12 probability distribution moments to accurately reproduce the underlying distribution, and demonstrates some of the necessary versatility of the ZI-closure scheme. Second, the network characteristics are very sensitive to the chosen kinetic parameters, and it is thus ideal for demonstrating some of the novel stochastic analytical methods available using moment closure. See Table 1 for the network description.

As a dual peaked distribution, the Schlögl model requires up to twelve lower-order moments in order to reproduce its complex function, which is impossible with analytical closure schemes. Figure 1 demonstrates how accuracy improves as the closure-order is increased. Several maximum entropy optimizations are executed using an increasing closure order (from 2nd to 10th-order closure).

With moment closure it is now possible to obtain steady-states quickly using appropriate closure schemes. In particular, by setting the left side of Equation 2 to zero, a steady-state distribution is immediately available. Figure 2 shows an example of the profile obtained by solving for steady-states across a range of key values. The parameter range (k_4 from 2 to 5 s⁻¹) was chosen such that the left and right sides of the graph have a single peaked distribution, whereas the blue shaded region in the middle is where the steady-state distribution has two peaks. This is explicitly shown in Figure 3 as well, where three example parameter values show how the Schlögl model flows from normally distributed to bimodal and back again.

These examples show two important features of ZI-Closure. First, there are 300 steady-state values obtained using ZI-Closure in Figure 2. While this number of simulations would be impractical for SSA simulations (indeed, hundreds of hours of simulation time were used to obtain the 31 SSA points for comparison), it takes a matter of minutes with ZI-closure scheme. With ZI-closure, steady-state optimization can be almost immediate for some small systems. Second, the line itself is, for all practical purposes, a sensitivity analysis for parameter k_4 for the Schlögl system. Sensitivity analysis is an important tool in chemical reaction design because it can identify the parameters that most influence a specific behavior in a system. In this case it is clear that in the bimodal region the distribution mean is very sensitive to the value of k_4 . This type of analysis could have otherwise taken hours of computation time with a direct SSA simulation.

The Schlögl model is a particularly interesting network to investigate because the underlying distribution can take on a range of functions, from a simple normal distribution to a complex bimodal distribution that takes a dozen moments to describe. ZI-Closure can obtain steady-state distributions immediately, and subsequently certain analytical techniques can be performed. A sensitivity analysis for k_4 shows that the steady-state optimization can be obtained quickly relative to the SSA, and also that the identification of important network parameters can be determined immediately. This is just the start of what we think is possible using moment closure techniques. The vast toolbox of deterministic analytical techniques, like eigenvalue and nonlinear analysis, may also become available for stochastic systems by applying ZI-closure to small chemical reaction networks.

3. CONCLUSIONS - FUTURE DIRECTIONS

In order to make stochastic modeling more accessible, we have developed the Synthetic Biology Software Suite (SynBioSS), a software suite that automates the steps for building models of and conducting numerical simulations for synthetic biological systems [41-43]. There are three components in SynBioSS: *Designer*, *Wiki* and *Simulator*.

With SynBioSS Designer, gene network models are created automatically after the user enters molecular components and their relationships. Every reaction in the model has a corresponding kinetic rate that describes the rate of association of its reactant molecules and the formation or destruction of any covalent bonds or stable non-covalent interactions. SynBioSS Wiki has been specifically created to store and recall just this sort of kinetic data. Simulating gene regulatory networks is feasible with the third component of SynBioSS, the Desktop Simulator. SynBioSS DS can be downloaded as an installation executable for Windows. At the heart of SynBioSS DS has been Hy3S.

We are planning to implement the ZI-closure scheme within SynBioSS. We anticipate that an augmented SynBioSS will aid scientists in modeling gene regulatory networks in a way fit for analysis and design. In particular, the calculation of steady states with ZI-closure affords a parameter sensitivity analysis and stability analysis for stochastic reaction networks. Modeling ranges of kinetic and thermodynamic parameters may rationalize the design of complex, nonlinear networks of reactions. This in turn may assist in identifying components and interactions for synthetic biological systems.

Admittedly, the field of computational synthetic biology is still nascent. A long distance separates theory from practical implementation, but with advances in modeling techniques, from Gillespie's SSA, to hybrid simulation algorithms, to CME closure schemes, this distance is shortened.

ACKNOWLEDGEMENTS

This work was supported by a grant from the National Institutes of Health (American Recovery and Reinvestment Act grant GM086865) and a grant from the National Science Foundation (CBET-0644792) with computational support from the Minnesota Supercomputing Institute (MSI). Support from the University of Minnesota Digital Technology Center and the University of Minnesota Biotechnology Institute is also acknowledged.

REFERENCES

1. Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. *Nature*. 2000; 403:335–338. [PubMed: 10659856]
2. Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*. 2000; 403:339–342. [PubMed: 10659857]
3. Andrianantoandro E, Basu S, Karig DK, Weiss R. Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol*. 2006; 2:2006.0028. [PubMed: 16738572]
4. Volzing K, Borrero J, Sadowsky MJ, Kaznessis YN. Antimicrobial Peptides Targeting Gram-negative Pathogens, Produced and Delivered by Lactic Acid Bacteria. *ACS Synth Biol*. 2013 2013 Jul 10. [Epub ahead of print] PubMed PMID: 23808914.
5. Ramalingam K, Maynard J, Kaznessis YN. Forward engineering of synthetic biological AND gates. *Biochemical Engineering Journal*. 2009; 47:38–47.
6. Alon U. Biological networks: the tinkerer as an engineer. *Science*. 2003; 301:1866–1867. [PubMed: 14512615]
7. Endy D. Foundations for engineering biology. *Nature*. 2005; 438:449–453. [PubMed: 16306983]
8. Volzing K, Biliouris K, Kaznessis YN. proTeOn and proTeOff, New Protein Devices That Inducibly Activate Bacterial Gene Expression. *ACS Chem Biol*. 2011; 6(10):1107–16. [PubMed: 21819083]
9. Kaern M, Blake WJ, Collins JJ. The engineering of gene regulatory networks. *Annu Rev Biomed Eng*. 2003; 5:179–206. [PubMed: 14527313]
10. Keasling J. The promise of synthetic biology. *Bridge Natl Acad Eng*. 2005; 35:18–21.

11. Salis H, Kaznessis YK. Accurate Hybrid Stochastic Simulation of a System of Coupled Chemical or Biochemical Reactions. *J. Chem. Phys.* 2005; 122:1–13.
12. Haseltine EL, Rawlings JB. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J Chem Phys.* 2002; 117:6959–6969.
13. Salis H, Kaznessis YN. Numerical simulation of stochastic gene circuits. *Comp Chem Eng.* 2005; 29:577–588.
14. Cao Y, Li H, Petzold L. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *J Chem Phys.* 2004; 121:4059–4067. [PubMed: 15332951]
15. Chatterjee A, Mayawala K, Edwards JS, Vlachos DG. Time accelerated monte carlo simulations of biological networks using the binomial {tau}-leap method. *Bioinformatics.* 2005; 21:2136–2137. [PubMed: 15699024]
16. Tian T, Burrage K. Binomial leap methods for simulating stochastic chemical kinetics. *J Chem Phys.* 2004; 121:10356–10364. [PubMed: 15549913]
17. E. W. Liu D, Vanden-Eijnden E. Nested stochastic simulation algorithm for chemical kinetic systems with disparate rates. *J Chem Phys.* 2005; 123:194107. [PubMed: 16321076]
18. Munsky B, Khammash M. The finite state projection algorithm for the solution of the chemical master equation. *J Chem Phys.* 2006; 124:044104. [PubMed: 16460146]
19. McQuarrie DA. Stochastic approach to chemical kinetics. *J Appl Probab.* 1967; 4:413–478.
20. Moyal JE. Stochastic Processes and Statistical Physics. *Journal of the Royal Statistical Society Series B (Methodological).* 1949; 11:150–210.
21. Oppenheim I, Shuler KE. Master equations and markov processes. *Phys Rev.* 1965; 138:1007–1011.
22. Oppenheim I, Shuler KE, Weiss GH. Stochastic theory of multistate relaxation processes. *Adv Mol Relax Processes.* 1967; 1:13–68.
23. Oppenheim, I.; Shuler, KE.; Weiss, GH. Stochastic processes in chemical physics: The master equation. The MIT Press; Cambridge, MA: 1977.
24. Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J Comp Phys.* 1976; 22:403–434.
25. Gillespie DT. Exact Stochastic Simulation of Coupled Chemical Reactions. *J Phys Chem.* 1977; 81:2340–2361.
26. Gibson MA, Bruck J. Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels. *J Phys Chem.* 2000; 104:1876–1889.
27. Cao Y, Gillespie DT, Petzold LR. Avoiding negative populations in explicit poisson tau-leaping. *J Chem Phys.* 2005; 123:054104. [PubMed: 16108628]
28. Chatterjee A, Vlachos DG. Temporal acceleration of spatially distributed kinetic monte carlo simulations. *J Comput Phys.* 2006; 211:596–615.
29. 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]
30. Sotiropoulos V, Kaznessis YN. An adaptive time step scheme for a system of SDEs with multiple multiplicative noise. Chemical Langevin equation, a proof of concept. *J Chem Phys.* 2008; 128:014103. [PubMed: 18190181]
31. Kaznessis Y. Multi-Scale Models for Gene Network Engineering. *Chem Eng Sci.* 2006; 61:940–953.
32. Kaznessis Y. Models for Synthetic Biology. *BMC Systems Biology.* 2007; 1:47. [PubMed: 17986347]
33. Harris LA, Clancy PA. A partitioned leaping” approach for multiscale modeling of chemical reaction dynamics. *J Chem Phys.* 2006; 125:144107. [PubMed: 17042579]
34. Tuttle L, Salis H, Tomshine J, Kaznessis YN. Model-Driven Design Principles of Gene Networks: the Oscillator. *Biophys J.* 2005; 89:3873–83. [PubMed: 16183880]
35. Tomshine J, Kaznessis YN. Optimization of a stochastically simulated gene network model via simulated annealing. *Biophys J.* 2006; 91:3196–3205. [PubMed: 16920827]

36. Gillespie CS. Moment closure approximations for mass-action models. *IET systems biology*. 2009; 3:52–58. [PubMed: 19154084]
37. Sotiropoulos V, Kaznessis YN. Analytical Derivation of Moment Equations in Stochastic Chemical Kinetics. *Chem Eng Sci*. 2011; 66:268–277. [PubMed: 21949443]
38. Smadbeck P, Kaznessis YN. Efficient Moment Matrix Generation for Arbitrary Chemical Networks. *Chem Eng Sci*. 2012; 84:612–618. [PubMed: 23175571]
39. Smadbeck P, Kaznessis YN. A closure scheme for chemical master equations. *Proc Natl Acad Sci U S A*. 2013; 110(35):14261–5. doi: 10.1073/pnas.1306481110. [PubMed: 23940327]
40. Schlögl F. Chemical reaction models for non-equilibrium phase transition. *Z. Physik*. 1972; 253:147–161.
41. Salis H, Sotiropoulos V, Kaznessis YN. Multiscale Hy3S: Hybrid Stochastic Simulations for Supercomputers. *BMC Bioinformatics*. 2006; 7:93. 2006. [PubMed: 16504125]
42. Hill A, Tomshine J, Wedding E, Sotiropoulos V, Kaznessis YK. SynBioSS: the Synthetic Biology Modeling Suite. *Bioinformatics*. 2008; 24:2551–2553. [PubMed: 18757873]
43. Weeding E, Houle J, Kaznessis YN. SynBioSS designer: a web-based tool for the automated generation of kinetic models for synthetic biological constructs. *Brief Bioinform*. 2010; 11(4): 394–402. [PubMed: 20639523]

Figure 1. Maximum Entropy Demonstration

The parameters are $k_1 = 0.15$ (molec. $^{-1}$ s $^{-1}$), $k_2 = 0.0015$ (molec. $^{-2}$ s $^{-1}$), $k_3 = 20$ molec. $^{-1}$ s $^{-1}$, and $k_4 = 3.5$ s $^{-1}$. The stochastic simulation algorithm results are an exact reproduction of the chemical master equation solution (blue, solid line). As more moments are added to the maximum entropy optimization program (from 2 to 10 moments), the optimization results become more representative of the exact solution distribution.

Figure 2. Schlögl Sensitivity Analysis

The parameters are $k_1 = 0.15 \text{ molec.}^{-1}\text{s}^{-1}$, $k_2 = 0.0015 \text{ molec.}^{-2}\text{s}^{-1}$, and $k_3 = 20 \text{ molec.}^{-1}\text{s}^{-1}$. The range for k_4 was chosen such that at low values (left side) the distribution has a single peak, for middle range values (blue region) the distribution has two peaks, and at high values (right side) it returns to a single peaked distribution.

Figure 3. Schlögl Distribution

The parameters are $k_1 = 0.15 \text{ molec.}^{-1}\text{s}^{-1}$, $k_2 = 0.0015 \text{ molec.}^{-2}\text{s}^{-1}$, and $k_3 = 20 \text{ molec.}^{-1}\text{s}^{-1}$.

Three k_4 values were chosen: $k_4 = 2 \text{ s}^{-1}$ (red), $k_4 = 3.5 \text{ s}^{-1}$ (green), and $k_4 = 5 \text{ s}^{-1}$ (blue).

The distribution goes from a single peak to two peaks and back. The SSA results are represented as squares, while the ZI-Closure results (12th order closure) are lines, demonstrating the versatility of the method.

Table 1

Schlögl model network description

Model	Schlögl
Reactions	$\begin{array}{l} 2X + A \xrightarrow{k_1} 3X \\ 3X \xrightarrow{k_2} 2X + A \\ B \xrightarrow{k_3} X \\ X \xrightarrow{k_4} B \end{array}$
Degrees of freedom	X
Initial condition	$X_0 = 25$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript