

IMT School for Advanced Studies, Lucca

Lucca, Italy

**The Effect of Compression and Expansion on
Stochastic Reaction Networks**

PhD Program in Institutions, Markets and Technologies
Curriculum in Computer Science and Systems Engineering
(CSSE)

XXXI Cycle

By

Tabea Waizmann

2021

The dissertation of Tabea Waizmann is approved.

Program Coordinator: Prof. Rocco De Nicola, IMT Lucca

Supervisor: Prof. Mirco Tribastone, IMT Lucca

The dissertation of Tabea Waizmann has been reviewed by:

Dr. Catia Trubiani, Gran Sasso Science Institute

Prof. Michele Loreti, University of Camerino

IMT School for Advanced Studies, Lucca

2021

To everyone who believed in me.

Contents

List of Figures	ix
List of Tables	xi
Acknowledgements	xiii
Vita and Publications	xv
Abstract	xviii
1 Introduction	1
2 Background	6
2.1 Multisets	6
2.2 Reaction Networks	7
2.3 Ordinary Lumpability	11
2.4 Layered Queuing Networks	12
2.5 PEPA	16
3 Coarse graining mass-action stochastic reaction networks by species equivalence	19
3.1 Species Equivalence	20
3.1.1 Species equivalence as a generalization of Markov chain ordinary lumpability	27
3.1.2 Characterization of SE for mass-action networks	27
3.1.3 Computation of the maximal SE and reduced network	28

3.2	Applications	31
3.2.1	Computational systems biology	31
3.2.2	Epidemic processes in networks	33
3.3	Discussion	36
4	DiffLQN: Differential Equation Analysis of Layered Queuing Networks	37
4.1	DiffLQN	38
4.1.1	Architecture	38
4.1.2	Capabilities	38
4.1.3	Syntax	39
4.2	Case Study: Client-Server dynamics	41
4.3	Discussion	45
5	The Finite State Expansion Method for Stochastic Reaction Networks	47
5.1	The Finite State Expansion Method	48
5.1.1	Detailed example	50
5.1.2	Stochastic equivalence	55
5.2	Applications	61
5.2.1	Genetic Feedback Switch	62
5.2.2	Gene regulatory system with inhibition feedback loop	66
5.2.3	Genetic Toggle Switch	68
5.2.4	Queuing systems	72
5.2.5	Schlögl	75
5.3	Discussion	79
6	Conclusion	81
	References	83

List of Figures

1	Graphical representation of an LQN	14
2	Illustration of SE on a simple network with four species . .	29
3	Simple multisite phosphorylation process (SH09)	31
4	Example of SE reduction of SIS dynamics on a coarse-grained network	34
5	Example input file for the network of Fig. 1.	43
6	Conceptualisation of FSE as a hybrid method bridging the gap between ME and DRE	48
7	Stochastic simulation traces of the Schlögl model, illustrating the bimodality of the system.	50
8	Expansion of the Schlögl network.	51
9	Graphical representation of the CTMC for the Schlögl model before and after finite state expansion	52
10	Deterministic Rate Equations before and after application of FSE.	53
11	FSE results for the Schlögl model compared to the original DRE solution and simulation average	54
12	FSE results for the genetic feedback switch in scheme (5.6) with two different sets of observation bounds, compared to the DRE solution and simulation average	62
13	Comparison of FSE results to other state-of-the-art methods for the gene feedback switch model from Eq. 5.6	63

14	The longer-term behavior of FSP on the gene feedback switch model, using the same settings as in Fig. 13.	64
15	Examination of the influence of observation bound choice on the accuracy of FSE solutions for the genetic feedback switch model (Eq. 5.6)	65
16	Comparison of FSE results to other state-of-the-art methods for the inhibition feedback loop model (Eq. 5.7)	67
17	Examination of the influence of observation bound choice on the accuracy of FSE solutions for the inhibition feedback loop model (Eq. 5.7)	68
18	FSE results for the genetic toggle switch model (5.8) with two different sets of observation bounds, compared to the DRE solution and simulation average	69
19	Comparison of FSE results to other state-of-the-art methods for the genetic toggle switch model (Eq. 5.8)	70
20	Examination of the influence of observation bound choice on the accuracy of FSE solutions for the genetic toggle switch model (Eq. 5.8)	71
21	Finite state expansion applied to the queueing network model (Eq. 5.9) for varying parameter sets	73
22	Examination of the influence of observation bound choice and arrival rate λ on FSE accuracy in the queueing network model (Eq. 5.9)	74
23	Comparison of FSE results to other state-of-the-art methods for the Schlögl model with parameter set (5.10)	77
24	Examination of the influence of observation bound choice on FSE accuracy in the Schlögl model with parameter set (5.10)	77
25	FSE accuracy and buffer species population in the Schlögl model with parameter set (5.11) using different observation bounds	78
26	Influence of the distribution of elements between the auxiliary and buffer species in the initial state in the Schlögl model with parameter set (5.11)	78

List of Tables

2	Coarse graining of SIS models on real-world networks. . .	35
3	Numerical results of DiffLQN on the example LQN from Fig. 1. The errors are measured as percentage relative errors from the simulation estimate.	46

List of Abbreviations

A	
AMVA	Approximate Mean Value Analysis
C	
CMC	Conditional Moment Closure
CTMC	Continuous-time Markov Chain
D	
DRE	Deterministic Rate Equation
E	
EMRE	Effective Mesoscopic Rate Equation
F	
FSE	Finite State Expansion
FSP	Finite State Projection
L	
LQN	Layered Queuing Network
M	
MC	Markov Chain
MCA	Moment Closure Approximation
ME	Master Equation
O	
ODE	Ordinary Differential Equation
P	
PEPA	Performance Evaluation Process Algebra
R	
RN	Reaction Network
S	
SCV	Squared Coefficient of Variation
SE	Species Equivalence
SIS	Susceptible-Infected-Susceptible network

Acknowledgements

First of all, I want to thank Mirco Tribastone for supporting me throughout the process of researching and writing this thesis. He was the one who motivated me to pursue a PhD in the first place after I completed my Master's thesis and it is no exaggeration to say that without him, this thesis would not exist today. His encouragement gave me hope when progress was slow and difficult and his constructive criticism always inspired me towards further improvement.

I am now very glad I took the leap and moved to Lucca despite my initial doubts. The three years I spent living at IMT have been an experience I wouldn't want to miss for the world and I found more of a second home in Lucca than I ever expected. In this context, I also have to mention my friends from Camarilla Italia, who made my time in Italy what it was and helped me learn more of the language than I ever would have managed otherwise.

Next, I want to thank Andrea Vandin, with whom I worked during my Erasmus stay in Denmark. He not only welcomed me to DTU for two very productive and enjoyable months, he also showed me around Copenhagen, shared the secret of the best coffee machine on campus and introduced me to some great board games.

Thanks also go to my parents, whose unwavering moral support was much appreciated.

This thesis is composed from material based on the three papers I co-authored. These papers, their other authors and my contributions therein are as follows:

Chapter 3 is based on (CPVT⁺21). The project involved as coauthors Isabel Cristina Pérez-Verona, Luca Cardelli, Mirco Tribastone, Max Tschaikowski, and Andrea Vandin. I made two main contributions to this paper:

- I expanded the definition of Species Equivalence and its proof of lumpability from just binary reaction networks to the general case of reaction networks with an arbitrary number of reactants
- I devised the first proof of characterization for binary reaction networks. This was however later superseded by the general proof presented in this thesis, which was a collaborative effort with main contributions from Max Tschaikowski.

Chapter 4 is based on (WT16). This paper is a continuation of the work I started with my Master's thesis, and was written in collaboration with Mirco Tribastone. I want to thank him again for proofreading, editing, and always being available when I was working with the code of PEPAtO.

Chapter 5 presents my main research project and is based on (WBVT21), for which I collaborated with Mirco Tribastone, Andrea Vandin and Luca Bortolussi. While I am the one responsible for developing the theories and programming the implementations, my coauthors' help especially with writing, but also with advice and asking the occasional important question was invaluable. Special thanks go to Mirco Tribastone for letting me develop an abandoned idea of his, and to Andrea Vandin for helping with integrating FSE into his program ERODE.

Vita

- September 6, 1989** Born, Fürstenfeldbruck (Bavaria), Germany
- 2008** Abitur
Final mark: 1.6
Graf Rasso Gymnasium, Fürstenfeldbruck, Germany
- 2012** Bachelor in Computer Science
Final mark: 1.59
LMU, Munich, Germany
- 2014** Master in Computer Science
Final mark: 1.29
LMU, Munich, Germany
- 2015** PhD Student in Southampton
University of Southampton
Southampton, England
- 2015** PhD Student in Lucca
IMT School for Advanced Studies Lucca
Lucca, Italy

Publications

1. Tabea Waizmann, Mirco Tribastone “DiffLQN: Differential Equation Analysis of Layered Queuing Networks,” in *Companion Publication for ACM/SPEC on International Conference on Performance Engineering 2016*, pages 63–68, March 2016.
2. Luca Cardelli, Isabel Cristina Pérez-Verona, Mirco Tribastone, Max Tsch aikowski, Andrea Vandin, Tabea Waizmann “Exact maximal reduction of stochastic reaction networks by species lumping,” in *Bioinformatics*, February 2021.
3. Tabea Waizmann, Luca Bortolussi, Andrea Vandin, Mirco Tribastone “Improved estimations of stochastic chemical kinetics by finite-state expansion,” in *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 477(2251):20200964, July 2021.

Presentations

1. T. Waizmann, "DiffLQN: Differential Equation Analysis of Layered Queuing Networks," at *2nd Workshop on Challenges in Performance Methods for Software Development (WOSP-C 2016)*, Delft, Netherlands, 2016.
2. T. Waizmann, "The Finite State Expansion Method for Stochastic Reaction Networks," at *7th International Workshop on Hybrid Systems Biology (HSB 2020)*, Online, 2020.

Abstract

Markov chains are a fundamental model to study systems with stochastic behavior. However, their state space is often of an unmanageable size, making the use of approximations and simplifications necessary for analytic solutions. This thesis considers reaction networks as a well-known representation for Markov chains describing interactions between species populations. It presents several methods using model transformations to aid with the effective analysis of such systems.

Species equivalence is a reduction technique that lifts the concept (and related algorithms) of Markov chain lumpability from lumping of states to directly lumping species in a reaction network. This allows the simplification of a reaction network without first examining its state space.

The tool DiffLQN implements a method for the analysis of large-scale stochastic models for the performance evaluation of software systems using an approach based on deterministic rate equations, by means of a compact system of ordinary differential equations that approximate only mean estimates for stochastic reaction networks.

Deterministic rate equations are generally accurate for networks with large populations, but may incur errors when elements are only present in low copy numbers. This thesis presents finite state expansion, which aims to solve that problem. It does so by converting a given reaction network into an expanded one with additional species and reactions such that the overall stochastic behavior is preserved. The resulting rate equations, however, may enjoy increased accuracy. Several tests on example models show that finite state expansion proves competitive with other state-of-the-art methods.

Chapter 1

Introduction

Reaction networks (RNs) are a fundamental model across many scientific disciplines including chemistry, biochemistry, epidemiology, and computer science (GJ13). An RN describes species populations that interact stochastically through a set of reaction channels according to dynamics governed by a continuous-time Markov chain through the well-known master equation (ME). This is a system of linear ordinary differential equations (ODEs) whose solution gives the time course of the probability that the process is in a specific configuration, represented as a state vector of nonnegative integers counting the population levels for each species (Van07).

It is widely understood that analytical solutions of the ME are unaccessible in practice (Gil77). Numerical methods are hindered by the problem that the state space grows combinatorially with the population of species under consideration. Moreover, if the model represents an open system then the number of states grows unboundedly, which requires some form of truncation to enable numerical treatment of the ME (MK06).

One possible solution is to circumvent ME analysis through simulating sample paths by means of Gillespie's stochastic simulation algorithm or variants thereof (Gil07). However, stochastic simulations are computationally expensive. In addition, forgoing an analytical description of

the model may preclude other important studies such as stability, perturbation analysis, bifurcation, and parameter inference (SSG17).

Deterministic rate equations (DREs, also known as mean field equations) provide a macroscopic dynamical view of the network by means of a system of ODEs where one equation is associated with each species, instead of a single discrete configuration as in the ME microscopic description. The DRE solution can be interpreted as the mean population level of each species as a function of time. This is exact if each propensity function associated with a reaction channel is linear, corresponding, for instance, to monomolecular chemical reaction networks (Gil77). With nonlinear propensity functions, theoretical results of asymptotic convergence rigorously establish that, under mild conditions, the DREs do give the true expectations in the thermodynamic limit when the species populations go to infinity (Kur72).

In general, however, nonlinear propensities lead to DREs that provide only an approximation to the true dynamical behavior. For example, cell regulation can be modeled with low-abundance species (in the order of a few units) to describe the behavior of genes (Gup95). Processes such as activation and deactivation that vary with time as a result of various bio-molecular processes (Pau05) may introduce significant variability in gene expression (ELSS02), caused by inherent stochasticity in the bio-molecular processes involved (SES02). Such forms of noise are not accounted for in the macroscopic DRE representation, which may lead to large approximation errors.

This thesis examines the effects of *compressing* and *expanding* reaction networks for the purpose of obtaining accurate and manageable estimations of their stochastic behavior. *Compression* is here meant in two distinct forms.

The first is by means of lumping, in its technical sense of computing projections of the original state space into a lower-dimensional one that retains the relevant stochastic behavior of the model (OM98). Markov chain lumping is a well-established technique (KS76). It characterizes a partition of the state space such that it is possible to derive a reduced Markov chain over partition blocks (Buc94). Complementary methods

such as finite state projection (MK06) and sliding window abstraction (HMW09) truncate the state space by discarding states which accumulate low probability mass. However, all these techniques work at the Markov chain level, requiring the computationally expensive enumeration of the state space.

To make the lumping process more efficient, we present an exact aggregation method for reaction networks with mass-action kinetics which hinges on an equivalence relation between the species. This results in a structure-preserving coarse-grained network where the dynamics of each macro-species is stochastically equivalent to the sum of the original species in each equivalence class. By how much a model can be compressed using this method depends entirely on the network's structure, as is typical for lumping. An algorithm based on partition refinement (PT87) can be used to compute the largest equivalence, thus the coarsest aggregation, without enumerating the state space. We demonstrate the effectiveness and scalability of our technique, as well as the physical interpretability of resulting reductions, in several models of signaling pathways in systems biology and epidemic processes on complex networks.

Reducing the size of a model via lumping can be very helpful in decreasing analysis complexity. If the modeling goal is to obtain coarser estimates such as expectations of the underlying stochastic process, a full stochastic analysis may however still be excessively expensive and the advantages of analytical methods remain relevant.

The second way of compressing stochastic reaction networks investigated in this thesis is the use of DRE models for the performance evaluation of computer systems.

Layered Queuing Networks (LQN) are a popular model in software performance engineering because they support, as first-class citizens, frequently used high-level mechanisms such as synchronous and asynchronous communication, layered services (i.e., entities that work as clients as well as servers), and fork/join synchronization (FAOW⁺09).

In addition to stochastic simulation, LQNs are traditionally solved analytically using approximate mean value analysis (AMVA), an efficient

technique that provides estimates of steady-state performance indices such as throughput, utilization, and response time. The approximation comes from two main sources: first, a recursive algorithm depending on the number of jobs in the network is replaced by a fixed-point iteration independent on them, e.g., (CN82); second, heuristics are used to capture non-product-form behavior in the AMVA framework, e.g., (RS95).

Although the approximation has been empirically shown to be satisfactory, no guarantees can be provided (PKT90). To partially mitigate this problem, more recently an alternative interpretation of LQNs as a type of RN has been proposed, making DRE solutions for LQNs viable (Tri10; Tri11; Tri13). This has been motivated by (unrelated) results on DRE approximations for process algebra models (TGH12; TDGH12). In this view, every basic activity in the LQN can be seen as an RN species, associating it with an equation of the DRE which in turn provides an estimate of the average number of jobs executing that activity. The tendency of DRE estimates to get more reliable as the multiplicity of jobs and servers rises is in contrast to the behavior of AMVA, which may even increase errors with larger multiplicities (Tri13). A useful side product of this approach is that it also readily gives indices of performance for the transient behavior. Indeed, while AMVA only considers the steady state, this is computed by running a transient analysis for long enough until convergence is numerically detected.

To make practical use of these theories, the tool DiffLQN was developed. It analyses LQNs using DREs, in order to estimate average performance indices such as throughputs, utilizations, and response times of software and hardware devices. Being based on DREs, the complexity of computing the solution is independent of the concurrency levels in the model (i.e., thread multiplicities and processing units) and the estimates are theoretically guaranteed to be asymptotically correct for large enough concurrency levels.

There are several approaches available in the literature to improve the accuracy of DRE estimates whilst avoiding stochastic simulation, including moment-closure approximations (Kue16), correction terms to van Kampen's well-known system size expansion (Gri10; TMG12), and hy-

brid techniques (HWKT14); ref. (SSG17) offers an up-to-date review. However, these methods are applicable under certain assumptions such as smoothness of the propensity functions (GUV07; Gri10; AKS13), mass-action kinetics (LKK09; SK11; Gil09; SH11; SK13), specific structure of the RN, e.g., to describe gene regulatory systems (TPG14), and species that can be partitioned into low- and high-abundance classes (Jah11; MLSH12; HWKT14).

As a new option, we propose finite state expansion, an analytical method that mediates between the microscopic and the macroscopic interpretations of a reaction network by coupling the dynamics of a chosen subset of discrete state configurations with whole-population species that buffer the probability mass that falls off the tracked state space. Finite state expansion is an algorithmic translation of a source network into a target one where each tracked state configuration is modeled as a further auxiliary species. The translation produces stochastically equivalent dynamics, but the DREs of the expanded network, with more variables, can be interpreted as a correction to the original ones. We demonstrate the effectiveness of finite state expansion to effectively account for stochasticity using models from biochemistry and computer science which challenge state-of-the-art techniques due to the presence of intrinsic noise, multi-scale population dynamics, and multi-stability.

Chapter 2

Background

Stochastic systems are relevant in a variety of contexts from computer science to biochemistry and beyond, and, accordingly, many different models have been devised to describe and analyse them. This chapter provides the background and definitions for the model types relevant to our research.

A common factor between the model types described here is that the underlying mathematical mechanic they describe is that of a Markov Chain. Throughout this thesis, the focus is on analysing a given model, not on judging its accuracy towards the real-world process it describes. As such, no explicit time units are used; time units are always relative to the model's transition rates.

2.1 Multisets

Much of the syntax of reaction networks is based on multisets, so we begin by fixing the notation that will be used throughout this document to describe them. Multiset operations are particularly relevant for the proofs given in chapters 3 and 5.

Let \mathcal{S} be a set. Then, $\mathbb{N}^{\mathcal{S}}$ describes the set of all multisets over the elements of \mathcal{S} and $\mathbb{R}^{\mathcal{S}}$ is the set of all real-valued multisets over the elements of \mathcal{S} . For a given multiset $\sigma \in \mathbb{R}^{\mathcal{S}}$, we denote by σ_S or $\sigma(S)$ the

multiplicity of the element $S \in \mathcal{S}$. Given a set of elements $H \subseteq \mathcal{S}$, and a multiset $\rho \in \mathbb{R}^{\mathcal{S}}$, we use $\rho(H)$ to denote the cumulative multiplicity of all the elements in H in ρ .

We use S to denote either the element S or the singleton $\{S\}$.

Let $\sigma, \mu \in \mathbb{R}^{\mathcal{S}}$ be two multisets. We define the following operations:

- **Intersection:** $\sigma \cap \mu$ is the multiset such that $(\sigma \cap \mu)_S = \min(\sigma_S, \mu_S)$ for all $S \in \mathcal{S}$.
- **Addition:** $\sigma + \mu$ is the multiset such that $(\sigma + \mu)_S = \sigma_S + \mu_S$ for all $S \in \mathcal{S}$.
- **Subtraction:** $\sigma - \mu$ is the multiset such that $(\sigma - \mu)_S = \sigma_S - \mu_S$ for all $S \in \mathcal{S}$. This can lead to some species being associated with negative numbers in the resulting multiset.
- **Limited Subtraction:** $\sigma \ominus \mu$ is the multiset such that $(\sigma \ominus \mu)_S = \max(0, \sigma_S - \mu_S)$ for all $S \in \mathcal{S}$.
- **Projection:** Given $\mathcal{P} \subseteq \mathcal{S}$, we denote by $\sigma|_{\mathcal{P}} \in \mathbb{R}^{\mathcal{P}}$ the projection of σ such that $(\sigma|_{\mathcal{P}})_P = \sigma_P$ for all $P \in \mathcal{P}$.
- **Transformation:** Given a set $\mathcal{P} \subseteq \mathcal{S}$, and a function $f : \mathcal{S} \rightarrow \mathcal{P}$, we denote by σ^f the multiset in $\mathbb{R}^{\mathcal{P}}$ such that $\sigma_P^f = \sum_{f(S)=P} \sigma_S$ for all $\sigma \in \mathbb{R}^{\mathcal{S}}, P \in \mathcal{P}$.

The binary operations of intersection, addition, subtraction, and limited subtraction are defined also in the case of multisets of different sets. That is, if $\sigma \in \mathbb{R}^{\mathcal{S}_1}$ and $\mu \in \mathbb{R}^{\mathcal{S}_2}$, a binary operation treats them as multisets of $\mathbb{R}^{\mathcal{S}_1 \cup \mathcal{S}_2}$.

2.2 Reaction Networks

We can now formally define reaction networks to provide a basis for chapters 3 and 5. The accompanying concepts of an underlying Markov Chain, the Master Equation and Deterministic Rate Equations are also relevant to chapter 4.

Definition 1 (Reaction Network). A reaction network (RN) is a pair $(\mathcal{S}, \mathcal{R})$, where \mathcal{S} is a set of species and \mathcal{R} is a set of reactions. A reaction is a triple denoted by $\rho \xrightarrow{f} \pi$ where $\rho \in \mathbb{N}^{\mathcal{S}}$ is the multiset of reactants, $\pi \in \mathbb{N}^{\mathcal{S}}$ is the multiset of products, and f is the propensity function, $f : \mathbb{R}^{\mathcal{S}} \rightarrow \mathbb{R}_0^+$.

Propensities are arbitrary functions that map a configuration of the RN to a non-negative real which gives the parameter of the exponential distribution of the firing time of that reaction. From a multiset $\sigma \in \mathbb{N}^{\mathcal{S}}$ a reaction $\rho \xrightarrow{f} \pi$ induces a transition with propensity equal to $f(\sigma)$ to the multiset $\sigma + \pi - \rho$.

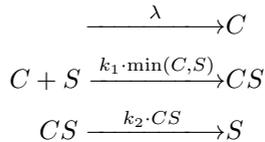
Throughout the remainder of this document we will consider well-defined RNs where each propensity function evaluates to zero for all multisets that do not have the minimum population counts described by the reactants. This guarantees that the underlying Markov chain does not reach states with negative population counts.

Definition 2. We call an RN well-defined if every reaction $\rho \xrightarrow{f} \pi$ is such that $f(\sigma) = 0$ if $\rho \not\subseteq \sigma$.

We further use $\rho(\mathcal{R})$ and $\pi(\mathcal{R})$ to denote the set of reactants and products in the reactions \mathcal{R} of an RN, by making explicit the dependence on the reactions:

$$\rho(\mathcal{R}) = \{\rho \mid (\rho \xrightarrow{\alpha} \pi) \in \mathcal{R}\} \quad \text{and} \quad \pi(\mathcal{R}) = \{\pi \mid (\rho \xrightarrow{\alpha} \pi) \in \mathcal{R}\}.$$

Example 1. The RN $(\mathcal{S}^q, \mathcal{R}^q)$ with $\mathcal{S}^q = \{S, C, CS\}$ represents a simple client-server interaction as can typically be found in queuing networks (ACK10):



New clients C arrive in the system with a constant arrival rate λ and have to wait to be assigned to an idle server S . Once assigned, the server has to work on the task for a time (CS) and then returns to the idle state while the client leaves the system. In a system state $3C + 2S + CS$, a client is assigned to a server with rate $k_1 \cdot \min(3, 2) = 2k_1$ and the third server will become available again for new clients with rate k_2 .

A popular class of RNs are those with mass-action kinetics, i.e., where the propensity with which a given reaction occurs is proportional to the product of the population levels of the species involved. This is a ubiquitous model of opportunistic interactions that has applications in several disciplines including chemistry (Gil77), computer science (ZNKT07), ecology (BM12), epidemiology (WTSB17), and systems biology (VMO15).

To simplify notation, the arrow of a reaction in a mass-action RN is typically marked with its *kinetic parameter* instead of the function. The propensity function of a mass-action reaction $\rho \xrightarrow{\alpha} \pi$ is

$$f_{\alpha}(\sigma) = \alpha \prod_{S \in \mathcal{S}} \binom{\sigma_S}{\rho_S}.$$

We recall the well-known Markov chain interpretation of RNs, see, e.g., (Gil77).

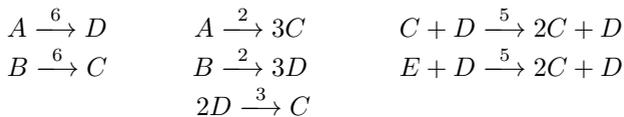
The complete state space of the Markov chain is the countable infinite set $\mathbb{N}^{\mathcal{S}}$. This is at the basis of the aforementioned combinatorial explosion of the state space. Starting from a given initial state $\hat{\sigma}$, the state space of interest can be restricted to those states generated by exhaustively applying the reactions to discover new states, until no further is found. The presence of a reaction such as $S_1 \xrightarrow{f} S_1 + S_1$ or the client arrival reaction in Example 1 may give rise to infinite state spaces in any case (FH74).

Since there may be several reactions connecting two CTMC states $\sigma, \theta \in \mathbb{N}^{\mathcal{S}}$ (e.g., reactions $A + B \xrightarrow{\alpha_1} B + C$ and $A \xrightarrow{\alpha_2} C$ contribute both to the transition from state $A + B$ into state $C + B$), the following auxiliary notion will be needed.

Definition 3 (Outgoing Reactions). *Let $(\mathcal{S}, \mathcal{R})$ be an RN. The multiset of outgoing transitions $out(\sigma)$ from state $\sigma \in \mathbb{N}^{\mathcal{S}}$ is obtained as*

$$out(\sigma) = \left\{ \left[\sigma \xrightarrow{\lambda} \sigma - \rho + \pi \mid (\rho \xrightarrow{f} \pi) \in \mathcal{R}, \lambda = f(\sigma) \right] \right\}$$

Example 2. *Consider the simple mass-action RN $(\mathcal{S}^e, \mathcal{R}^e)$ with $\mathcal{S}^e = \{A, B, C, D, E\}$:*



With an initial population $\sigma^{0e} = 2A + C + D$, we have

$$\begin{aligned} \text{out}(\sigma^{0e}) &= \{ \sigma^{0e} \xrightarrow{6 \cdot 2} A + C + 2D, \sigma^{0e} \xrightarrow{2 \cdot 2} A + 4C + D, \\ &\quad \sigma^{0e} \xrightarrow{5} 2A + 2C + D \}. \end{aligned}$$

The three outgoing reactions are due to the first, the second and the third reaction of \mathcal{R}^e , respectively.

Now we can define the Markov chain of an RN. Its transitions are obtained by collapsing the outgoing reactions between any two of its states.

Definition 4 (CTMC). Let $(\mathcal{S}, \mathcal{R})$ be an RN. The CTMC of $(\mathcal{S}, \mathcal{R})$ is denoted by $MC(\mathcal{S}, \mathcal{R})$ and has $\mathbb{N}^{\mathcal{S}}$ as state space. Instead, the transition rate from σ to θ , where $\sigma, \theta \in \mathbb{N}^{\mathcal{S}}$ are arbitrary, is defined as

$$q(\sigma, \theta) = \begin{cases} \sum_{(\sigma \xrightarrow{\lambda} \theta) \in \text{out}(\sigma)} \lambda & \text{if } \theta \neq \sigma \\ - \sum_{\theta' \in \mathbb{N}^{\mathcal{S}} \text{ s.t. } \theta' \neq \sigma} q(\sigma, \theta') & \text{if } \theta = \sigma \end{cases}$$

For any set of states $\mathcal{M} \subseteq \mathbb{N}^{\mathcal{S}}$, we define $q[\sigma, \mathcal{M}] = \sum_{\theta \in \mathcal{M}} q(\sigma, \theta)$.

Together, all values $q(\sigma, \theta)$ form the generator matrix Q .

Remark 1. Please note that the infinite sum $q(\sigma, \sigma)$ will have at most finitely many non-zero summands because \mathcal{R} is finite. Likewise, thanks to the fact that \mathcal{R} is finite, each state has finitely many incoming transitions.

The dynamical evolution of the CTMC is described by the master equation (ME). Each component of its solution, $p_{\sigma}(t)$, is the probability of being in state σ at time t (Van07), starting from an initial condition where the probability mass is concentrated at the initial state $\hat{\sigma}$, i.e., $p_{\hat{\sigma}}(0) = 1$.

Definition 5 (ME). Let $(\mathcal{S}, \mathcal{R})$ be an RN. Its master equation is:

$$\forall \sigma \in \mathbb{N}^{\mathcal{S}} : \quad \frac{dp_{\sigma}(t)}{dt} = \sum_{(\rho \xrightarrow{f} \pi) \in \mathcal{R}} -f(\sigma)p_{\sigma}(t) + f(\sigma + \rho - \pi)p_{\sigma + \rho - \pi}(t).$$

Using the CTMC's generator matrix, it may also be written as $\dot{p} = p^T Q$.

The Markov chain and ME describe the ground truth of a RN's behavior. Meanwhile, deterministic rate equations (DRE) can give an approximation with drastically fewer equations. Like the master equation, the DRE is a system of ordinary differential equations (ODE). Instead of one equation for each possible state of the Markov chain, however, a DRE solution needs only one equation per species, keeping track of that species' approximate mean population level.

Definition 6 (DRE). *Let $(\mathcal{S}, \mathcal{R})$ be an RN. Its deterministic rate equation is:*

$$\forall X \in \mathcal{S}, \sigma \in \mathbb{R}^{\mathcal{S}} : \quad \frac{d\sigma_X(t)}{dt} = \sum_{(\rho \xrightarrow{f} \pi) \in \mathcal{R}} f(\sigma(t))(\pi_X - \rho_X)$$

This approximation of the mean completely disregards variance, covariance, and other higher-order moments. In effect, the DRE is a first-order moment closure in which all moments of higher orders are set to zero (AKS13; Eng06; SSG15). Notably, building on a fundamental result by Kurtz on fluid limits of Markov population processes (Kur70), the DRE approximation has an asymptotic guarantee of exactness when the multiplicity of all species is large enough.

2.3 Ordinary Lumpability

Given a generator matrix Q of a Markov Chain, ordinary lumpability is an equivalence relation that partitions the state space into N blocks P_1, \dots, P_N , such that all states in a block have equal aggregate rate toward any block (KS76; Buc94), that is,

$$\sum_{\vartheta \in P_j} q(\sigma, \vartheta) = \sum_{\vartheta \in P_j} q(\sigma', \vartheta) \quad \text{for any } P_i, P_j \text{ with } \sigma, \sigma' \in P_i.$$

These aggregate rates form the generator matrix of the lumped Markov chain, where each macro-state corresponds to a partition block. Ordinary lumpability preserves the stochastic behavior in that the probability of being in one block in the lumped Markov chain is equal to the sum of the probabilities of being in states of that block in the original Markov chain (Buc94).

We now recast the notion of ordinary lumpability to our notation.

Definition 7 (Ordinary Lumpability, (Buc94)). *Let $(\mathcal{S}, \mathcal{R})$ be an RN and \mathbb{H} a partition of $\mathbb{N}^{\mathcal{S}}$ whose blocks are finite (that is, $|\mathcal{M}| < \infty$ for any $\mathcal{M} \in \mathbb{H}$). Then, $MC(\mathcal{S}, \mathcal{R})$ is ordinarily lumpable with respect to \mathbb{H} when for any two states $\sigma_1, \sigma_2 \in \mathbb{N}^{\mathcal{S}}$ in the same block of \mathbb{H} we have $q[\sigma_1, \mathcal{M}] = q[\sigma_2, \mathcal{M}]$ for all $\mathcal{M} \in \mathbb{H}$.*

It is worth remarking that the original result of ordinary lumpability (cf. Definition 7, ref. (Buc94)) applies to finite Markov chains. However, using concepts from functional analysis and the theory of linear ODEs on Banach spaces, this statement can be extended, under certain assumptions, to CTMCs with countably infinite state spaces (RT03). A sufficient condition for the theory to apply is to assume that a) each state of the CTMC has finitely many incoming and outgoing transitions and that b) the state space of the CTMC is partitioned in blocks of finite size.

The concept of Lumpability will be used mainly in chapter 3 and to a lesser degree in chapter 5 of this thesis.

2.4 Layered Queuing Networks

LQNs are the model type analysed by the tool DiffLQN, which will be presented in chapter 4. While this thesis gives a brief overview of the essential LQN syntax supported by DiffLQN, illustrated on a graphical example (Fig. 1), we refer to (FAOW⁺09) for more details. While LQNs can be described by an elaborate graphical syntax that makes them easier to understand for a human observer, they are typically defined in a text format for computer input. For a comprehensive description of this textual syntax beyond the scope of DiffLQN, we refer to (FMW⁺13).

The text format defines elements grouped in blocks by their type. Each block starts with a capital letter defining the respective element type and ends with a -1 . In the following, unless explicitly declared otherwise, brackets are not part of the syntax but only represent a grouping of elements. $\langle X \rangle$ designates an element X with multiple possible values, $(X)^*$ an element X that can be repeated an arbitrary number of times, and $(X)?$ an optional element X . $(a | b)$ is a choice between a and b .

The basic computational resource is the *processor*, drawn as an oval. The textual syntax for the processor block is:

```

1 P <number-of-defined-processors>
2 (p <Processor> f (m <mult>?)*)
3 -1

```

It is to note that the `<number-of-defined-processors>` parameter is usually ignored by solvers and often given as 0 regardless of the number of defined processors. The PEPA translation defined in (Tri13) does not distinguish between scheduling disciplines, so for the purposes of this thesis we can assume First Come First Served scheduling, designated by the letter `f` for processors. `<Processor>` defines a name for each processor and the `<mult>` parameter finally is an integer determining its number of identical copies.

On the processor, *tasks* (large parallelograms), for instance software services, are deployed. In text form, the task block is designated by the letter `T`.

```

1 T <number-of-defined-tasks>
2 (t <Task> (r|n) <entry>* -1 <Processor> (m <mult>?)*)
3 -1

```

Tasks are split into *reference tasks* (modelling clients/jobs) that actively send requests but do not receive them (keyword `r`) and *non-reference tasks* (modelling servers) that only start working when they receive a request (keyword `n`). In either case, a list of entries belonging to the task is given; each entry specified here needs to also be defined in the entry definition block. Like in the processor block, `<number-of-defined-tasks>` can be set to 0 and ignored.

A task consists of different *entries* (smaller parallelograms) that represent distinct kinds of services. An entry can be a basic *activity*, the atomic unit of operation in LQN if it is not further specified. Otherwise, it points to a diagram of basic activities (rectangles) which are performed in sequence (linked by an arrow), through probabilistic choice via *decision/merge* nodes ('+' operator), or by means of fork/join synchronization ('&' operator).

Activities can call entries synchronously (closed arrowhead) or asynchronously (open arrowhead, not shown in example). Asynchronous

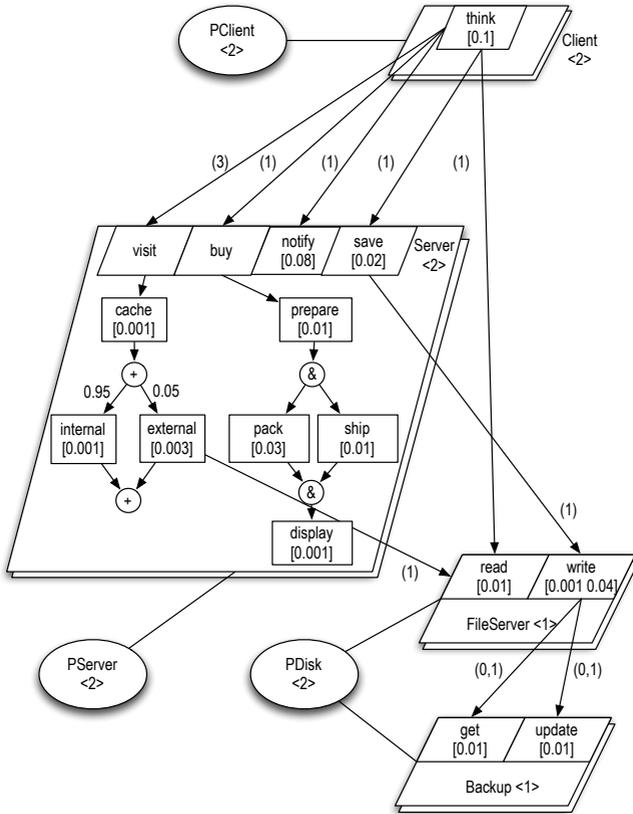


Figure 1: Graphical representation of an LQN, taken from (Tri13). Refer to Fig. 5 in chapter 4 to compare with the textual definition.

requests just start the requested entry while continuing the current activity, without waiting for a reply. In synchronous requests, the activity is stopped until a reply arrives. An activity can send any number of requests to the same entry; the number of requests is written next to the call arrow in brackets. When called, an activity consumes time on the processor where the task is deployed. Mean time demands are shown below the activity name within square brackets. When two time demands are listed, as in `write`, second-phases are modeled. The second time demand

happens after the activity has returned control to the caller, effectively executing asynchronously.

In the textual definition, the entry definition block is marked with an E and contains a line for each entry and each request that is not sent from an activity in an activity diagram.

```
1 E <number-of-defined-entries>
2 (s <entry> <time-demand>* -1)*
3 ((y|z) <sendingEntry> <receivingEntry> <request-number>* -1)*
4 (A <entry> <activity>)*
5 -1
```

The syntax given in line 2 of the entry block defines a new entry without an activity diagram. If multiple time demands are given, this defines multiple phases.

The syntax in line 3 defines a request. A request defined with y is synchronous, one defined with z asynchronous. Either way, the parameter <request-number> gives the number of requests sent in each phase of the sending entry. Accordingly the repetitions of <request-number> need to match the number of phases defined for the entry. If an entry contains an activity diagram, this is defined by the syntax in line 4. Like in the previous blocks, the parameter <number-of-defined-entries> is typically unused and can be set to 0.

For each task containing activity diagrams, a separate activity block starting with an A is defined:

```
1 A <Task>
2 (s <activity> <time-demand>)*
3 ((y|z) <activity> <entry> <request-number>)*
4 :
5 (<activity> -> <activity>)*
6 (<activity> -> <activity> (& <activity>)*)*
7 (<activity> (& <activity>)* -> <activity>)*
8 (<activity> -> (<prob>)<activity> (+ (<prob>)<activity>)*)*
9 <activity>[<entry>]*
10 -1
```

First, time demands and requests are defined. The only difference from the syntax of the entry block in this part is that activities cannot have phases.

After the colon, the diagram structure is defined. Line 5 defines the simple case of one activity following another, line 6 defines a fork, line 7 a

join and line 8 a decision node. Merge nodes do not need to be modelled explicitly.

The parameters (`<prob>`) are the only place where brackets are actually part of the syntax. Here, the probabilities of choosing each branch of the decision node are given as numbers between 0 and 1. Of course the probabilities should sum up to 1 for each node. When an activity diagram has reached its end, the syntax from line 9 is used to formalize that and return to the idle state of the given entry while sending a reply if the diagram was activated by a synchronous request. The final line of the diagram description is not ended with a semicolon.

2.5 PEPA

Translating an LQN to the process algebra PEPA (Hil96) makes it possible to leverage existing algorithms for DRE analysis (GTD⁺14; Dug06).

PEPA describes processes as chains of actions with associated rates. A process $(a, r).P$ will first execute action a and then continue to behave as process P . The time demand of a is exponentially distributed with rate r .

Constants can be used to define cyclic behavior: The process $P \stackrel{\text{def}}{=} (a, r_1).(b, r_2).P$ alternately executes actions a and b in an endless loop.

Nondeterministic choice is modelled with the $+$ operator: A process $(a_1, r_1).P_1 + (a_2, r_2).P_2$ can proceed to either process P_1 or P_2 , depending of which action is executed. The probability of choosing either path is proportional to the actions' rates, so the probability of choosing a_1 is $r_1/(r_1 + r_2)$.

Concurrent processes can be linked by the cooperation operator \bowtie , synchronising them over a set of actions L . In the process $P_1 \bowtie_{\{a_1, a_2\}} P_2$, the actions a_1 and a_2 have to be executed by P_1 and P_2 simultaneously, while other actions can be performed independently. When an action is performed by two processes in cooperation, the effective rate of that action is the minimum of the two rates associated to the action in the

cooperating processes.

In the translation of an LQN model, its modularity is preserved, as each element can be translated individually. Requests and replies, as well as the action of accessing a processor to execute an activity, are implied by the arrows and connections in an LQN and happen instantly. They are modelled as explicit actions in the PEPA translation, with extremely fast rates so as to approximate the instantaneous event of the original. These almost-instantaneous actions are then used to synchronize the technically parallel processes of the translated LQN elements. To translate a decision node, the rates of the request actions are chosen specifically to reflect the required probability distribution.

As an example, the processor $PDisk$ from Figure 1 translates as follows:

$$\begin{aligned}
 PDisk_1 &\stackrel{\text{def}}{=} (get_{PDisk}, v).PDisk_2 \\
 PDisk_2 &\stackrel{\text{def}}{=} (read, 1/0.01).PDisk_1 \\
 &\quad + (write', 1/0.001).PDisk_1 \\
 &\quad + (write'', 3/0.04).PDisk_1 \\
 &\quad + (get, 1/0.01)PDisk_1 \\
 &\quad + (update, 1/0.01).PDisk_1
 \end{aligned}$$

The action get_{PDisk} is shared by all entries running on this processor, so any of them can synchronise with $PDisk_1$ to claim the processor. The variable v stands in for the mentioned extremely fast rate, since the assignment of a processor is not technically supposed to take any time. In its active state $PDisk_2$, the processor can then execute any of its associated activities before returning to the idle state. The processor does not need to be locked into a specific activity in state $PDisk_2$, as the available processes for synchronisation are only those that have already previously synchronised on the action get_{PDisk} . The rate of the action $write''$ is proportional to executing a third of the original activity, since it is split into three slices to accommodate sending two requests during its runtime.

The corresponding DRE has one equation for each PEPA constant in the model. The processor $PDisk$ in our example is accordingly associ-

ated with two equations of the DRE.

More details on the complete translation process are provided in (Tri13), which served as a basis for the development of DiffLQN.

Chapter 3

Coarse graining mass-action stochastic reaction networks by species equivalence

In this chapter, we present an aggregation method for mass-action RNs that rests on an equivalence relation over the species. It can be identified by inspecting the set of reactions only and induces a lumpable partition on the underlying Markov chain, collapsing states that preserve the same overall population counts across equivalent species. Such *species equivalence* (SE) enables physical intelligibility of the reduction because it yields a reduced mass-action network defined for macro-species, each representing a distinct equivalence class.

From a mathematical viewpoint, SE can be seen as a lifting of lumpability from Markov chains to stochastic mass-action reaction networks. Indeed, our definition of SE precisely collapses to lumpability when the reaction network encodes the generator matrix of a Markov chain.

Importantly, two fundamental features of lumpability, with key practical implications, also carry over to SE. The first feature is the existence of a maximal SE, i.e., the equivalence that leads the coarsest aggrega-

tion of the reaction network. The second feature is an algorithm for the computation of such maximal SE. We build on the fundamental results by Paige and Tarjan (PT87), which have been applied to several other problems in computer science beyond lumpability (DHS03; VF10), such as the aggregation of polynomial differential equations (CTTV17a). Our algorithm computes the largest SE that refines a given initial partition of species by iteratively splitting partition blocks until a fixed point. This algorithm runs in polynomial time and space complexity with respect to the number of species and reactions in the network. As with all partition-refinement algorithms, an important advantage arising from the freedom in choosing the initial partition to be refined is that one can fully observe the dynamics of species of interest. This can be simply obtained by isolating any such observable as a singleton block in the initial partition.

The earlier approach to species lumping for stochastic mass-action networks from ref. (CTTV17b) supported only elementary networks (where reactions involve at most two distinct individuals as reactants), motivated by their relevance as a model of independent or pairwise interactions, such as unimolecular or bimolecular reactions which have a rigorous physical derivation from first principles (Gil77). SE is a generalization of the method in ref. (CTTV17b) in two fundamental ways. First, it is applicable to arbitrary higher-order reactions. Second, we prove that SE is the coarsest possible aggregation that yields a Markov chain lumping according to an equivalence over species, effectively leading to coarser aggregations in benchmark models.

3.1 Species Equivalence

Here we lift ordinary lumpability to reaction networks as an equivalence between species instead of Markov chain states. For the rest of this chapter, the notation for the number of elements of a species X in a multiset σ will be restricted to the variant $\sigma(X)$, leaving indexes usable to differentiate between multisets.

Our definition of reaction rate intuitively plays the role of the transition rate in the generator matrix.

Definition 8 (Reaction rate). Let $(\mathcal{S}, \mathcal{R})$ be an RN, and $\rho, \pi \in \mathbb{N}^{\mathcal{S}}$. The reaction rate from ρ to π is defined as

$$\mathbf{rr}(\rho, \pi) = \begin{cases} \sum_{(\rho \xrightarrow{\alpha} \pi) \in \mathcal{R}} \alpha & , \text{ if } \rho \neq \pi \\ - \sum_{\pi' \in \pi(\mathcal{R}), \rho \neq \pi'} \mathbf{rr}(\rho, \pi') & , \text{ if } \rho = \pi \end{cases}$$

For any $\mathcal{M} \subseteq \mathbb{N}^{\mathcal{S}}$, we may use $\mathbf{rr}[\rho, \mathcal{M}]$ for $\sum_{\pi \in \mathcal{M}} \mathbf{rr}(\rho, \pi)$.

Species equivalence (SE) induces an equivalence over Markov chain states (identified by the partition blocks of an ordinarily lumpable partition) starting from equivalences between species in an RN. We do so by providing the notion of *multiset lifting*. Given an RN $(\mathcal{S}, \mathcal{R})$ and an equivalence relation \mathcal{R} over \mathcal{S} , the lifting of \mathcal{R} relates multisets with same number of \mathcal{R} -equivalent species. In this way, a partition \mathcal{H} of species of a network naturally induces one over multisets of species, by relating those that have same number of species of each equivalence class. Thus we say that two multisets of species σ, σ' are equal up to the partition \mathcal{H} if and only if

$$\sum_{S_i \in H} \sigma_i = \sum_{S_i \in H} \sigma'_i, \quad \text{for all blocks of species } H \in \mathcal{H}.$$

Definition 9 (Multiset lifting). Let $(\mathcal{S}, \mathcal{R})$ be an RN, $\mathcal{R} \subseteq \mathcal{S} \times \mathcal{S}$ be an equivalence relation over \mathcal{S} , and $\mathcal{H} = \mathcal{S}/\mathcal{R}$ be the partition induced by \mathcal{R} over \mathcal{S} . We define the multiset lifting of \mathcal{R} on $\mathbb{N}^{\mathcal{S}}$, denoted by $\mathcal{R}^\uparrow \subseteq \mathbb{N}^{\mathcal{S}} \times \mathbb{N}^{\mathcal{S}}$, as

$$\mathcal{R}^\uparrow \triangleq \{(\sigma_1, \sigma_2) \mid \sigma_1, \sigma_2 \in \mathbb{N}^{\mathcal{S}} \wedge \forall H \in \mathcal{H} : \sigma_1(H) = \sigma_2(H)\}$$

Let $\mathcal{H}^\uparrow = \mathbb{N}^{\mathcal{S}}/\mathcal{R}^\uparrow$ be the partition induced by \mathcal{R}^\uparrow over $\mathbb{N}^{\mathcal{S}}$. For any block $\mathcal{M} \in \mathcal{H}^\uparrow$ we use $\mathcal{M}(H)$ to denote $\sigma(H)$, with σ being any element of \mathcal{M} .

Example 3. Continuing to use the model given in Example 2, consider the equivalence relation \mathcal{R}_m over S_e inducing $\mathcal{H}_m = \{\{A\}, \{B\}, \{C, E\}, \{D\}\}$. Examples of multisets related by \mathcal{R}_m^\uparrow are C and E , $2C$ and $2E$, and $C + E$ and $2E$, while $(A + C, B + C) \notin \mathcal{R}_m^\uparrow$.

Remark 2. A multiset lifting \mathcal{R}^\uparrow of \mathcal{R} induces an equivalence relation over $\mathbb{N}^{\mathcal{S}}$ with equivalence classes of finite size. Hence, $\mathcal{H}^\uparrow = \mathbb{N}^{\mathcal{S}}/\mathcal{R}^\uparrow$ is a potential candidate for an ordinary lumpability, see Remark 1 and the background given in chapter 2.3.

We now can define SE.

Definition 10 (Species equivalence). Let $(\mathcal{S}, \mathcal{R})$ be an RN, \mathcal{R} an equivalence relation over \mathcal{S} , and $\mathcal{H}^\uparrow = \mathbb{N}^{\mathcal{S}} / \mathcal{R}^\uparrow$. We say that \mathcal{R} is a species equivalence (SE) for $(\mathcal{S}, \mathcal{R})$ if and only if

$$\mathbf{rr}[X + \rho, \mathcal{M}] = \mathbf{rr}[Y + \rho, \mathcal{M}], \quad \forall (X, Y) \in \mathcal{R}, \rho \in \mathbb{N}^{\mathcal{S}}, \text{ and } \mathcal{M} \in \mathcal{H}^\uparrow.$$

Remark 3. Note that the definition of species equivalence in Definition 10 could be equivalently given in finitary terms, considering only all ρ in $\{\rho' \mid X + \rho' \in \rho(\mathcal{R}) \vee Y + \rho' \in \rho(\mathcal{R})\}$, and all \mathcal{M} containing at least one product in $\pi(\mathcal{R})$. All ρ and π multisets not appearing in such sets do not give any contribution to the summations in the definition. However, the definition given here simplifies the forthcoming related technical developments.

Example 4. Looking again at the model from Example 2, we infer that \mathcal{R}_m is an SE since $\mathbf{rr}(C + D, 2C + D) = \mathbf{rr}(E + D, 2C + D)$.

The following shows that there exists the largest SE of an RN, i.e., the coarsest partition of species that satisfies Definition 10. We show this by proving that the transitive closure of a union of species equivalences is a species equivalence.

Proposition 1. Let $(\mathcal{S}, \mathcal{R})$ be an RN, I a set of indices, and \mathcal{R}_i an SE for $(\mathcal{S}, \mathcal{R})$, for all $i \in I$. The transitive closure of their union $\mathcal{R} = (\bigcup_{i \in I} \mathcal{R}_i)^*$ is an SE for $(\mathcal{S}, \mathcal{R})$.

Proof. We first note that \mathcal{R} is an equivalence relation over \mathcal{S} , as it is the transitive closure of the union of equivalence relations over \mathcal{S} . For $i \in I$, let \mathcal{H}_i denote the partition induced over \mathcal{S} by \mathcal{R}_i , and \mathcal{H} the one induced by \mathcal{R} . For any $i \in I$, any block $H^i \in \mathcal{H}_i$ is contained in a block $H \in \mathcal{H}$, implying that any $H \in \mathcal{H}$ is the union of blocks of \mathcal{H}_i . For $(X_1, X_2) \in \mathcal{R}$, we have that $(X_1, X_2) \in (\bigcup_{i \in I} \mathcal{R}_i)^n$, for some $n > 0$. We now show that \mathcal{R} is an SE by induction over n . Let \mathcal{R}^n be $(\bigcup_{i \in I} \mathcal{R}_i)^n$, and $\rho \in \mathbb{N}^{\mathcal{S}}$.

Base case ($n = 1$): $(X_1, X_2) \in \mathcal{R}^1$ implies that $(X_1, X_2) \in \mathcal{R}_i$, for some $i \in I$. In order to prove that the condition on \mathbf{rr} required by SE in Definition 10 holds, we use that for any $H \in \mathcal{H}$ and any $i \in I$ we have that there exists some set of indices J^i such that $H = \bigcup_{j \in J^i} H_j^i$, with H_j^i a block of \mathcal{H}_i ; hence, $\mathbf{rr}[X_1 + \rho, H] = \sum_{j \in J^i} \mathbf{rr}[X_1 + \rho, H_j^i]$.

Inductive step: we assume that the condition on \mathbf{rr} required by SE holds for $\mathcal{R}^m, \forall m < n$. If $(X_1, X_2) \in \mathcal{R}^n$, then there exists an $X_3 \in \mathcal{S}$ such that $(X_1, X_3) \in \mathcal{R}_i$ for some $i \in I$, and $(X_3, X_2) \in \mathcal{R}^{n-1}$. Then, the claim follows from a similar argument as in the base case and the induction hypothesis. \square

Before stating our first major result, Theorem 1, we provide three propositions used in the proof of the theorem.

Proposition 2. *Let \mathcal{S} be a set and \mathcal{R} an equivalence relation on \mathcal{S} . Let $\mathcal{H} = \mathcal{S}/\mathcal{R}$ and $\mathcal{H}^\uparrow = \mathcal{R}^\uparrow/\mathbb{N}^\mathcal{S}$. For all $\sigma, \sigma', \pi, \pi', \rho, \rho' \in \mathbb{N}^\mathcal{S}$, we have*

- $(\sigma \cup \pi, \sigma \cup \pi') \in \mathcal{R}^\uparrow$ if and only if $(\pi, \pi') \in \mathcal{R}^\uparrow$,
- if $(\sigma, \sigma') \in \mathcal{R}^\uparrow$, then $(\sigma \cup \pi, \sigma' \cup \pi') \in \mathcal{R}^\uparrow$ if and only if $(\pi, \pi') \in \mathcal{R}^\uparrow$,
- $(\sigma - \rho, \sigma - \rho') \in \mathcal{R}^\uparrow$ if and only if $\rho \subseteq \sigma \supseteq \rho'$ and $(\rho, \rho') \in \mathcal{R}^\uparrow$,
- if $(\sigma, \sigma') \in \mathcal{R}^\uparrow$, then $(\sigma - \rho, \sigma' - \rho') \in \mathcal{R}^\uparrow$ if and only if $\rho \subseteq \sigma, \rho' \subseteq \sigma'$ and $(\rho, \rho') \in \mathcal{R}^\uparrow$.

Moreover, for any $\mathcal{M}, \tilde{\mathcal{M}} \in \mathcal{H}^\uparrow$, if it is possible to obtain multisets in $\tilde{\mathcal{M}}$ by adding species to those in \mathcal{M} , i.e., if $\mathcal{M}(H) \leq \tilde{\mathcal{M}}(H)$ for all $H \in \mathcal{H}$, then there exists exactly one $\hat{\mathcal{M}} \in \mathcal{H}^\uparrow$ such that $\tilde{\mathcal{M}} = \{\sigma + \hat{\sigma} \mid \sigma \in \mathcal{M}, \hat{\sigma} \in \hat{\mathcal{M}}\}$. That is, we obtain $\tilde{\mathcal{M}}$ by pairwise merging the multisets in \mathcal{M} with those in $\hat{\mathcal{M}}$. In addition, for all $\tilde{\mathcal{M}} \in \mathcal{H}^\uparrow$ we have that there exists exactly one equivalence class in \mathcal{H}^\uparrow denoted by $\bar{\mathcal{M}}_{\mathcal{M} \rightarrow \tilde{\mathcal{M}}}$ such that

$$\bar{\mathcal{M}} = \{\sigma - \rho + \pi \mid \sigma \in \mathcal{M}, \rho \in \tilde{\mathcal{M}} \text{ s.t. } \rho \subseteq \sigma, \pi \in \bar{\mathcal{M}}_{\mathcal{M} \rightarrow \tilde{\mathcal{M}}}\}$$

Proof. Straightforward. \square

Proposition 3. *Let $(\mathcal{S}, \mathcal{R})$ be an RN, \mathcal{R} an equivalence relation over \mathcal{S} , and $\mathcal{H}^\uparrow = \mathbb{N}^\mathcal{S}/\mathcal{R}^\uparrow$. Further, let $\mathcal{M}, \tilde{\mathcal{M}} \in \mathcal{H}^\uparrow$ such that $\mathcal{M} \neq \tilde{\mathcal{M}}$ and $\sigma \in \mathcal{M}$. Then, for any $\sigma \in \mathbb{N}^\mathcal{S}$, it holds that*

$$q[\sigma, \tilde{\mathcal{M}}] = \sum_{\tilde{\mathcal{M}} \in \mathcal{H}^\uparrow} \sum_{\substack{\rho \in \tilde{\mathcal{M}} \\ \text{s.t. } \rho \subseteq \sigma}} \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} \cdot \mathbf{rr}[\rho, \bar{\mathcal{M}}_{\mathcal{M} \rightarrow \tilde{\mathcal{M}}}] .$$

Proof. By Definitions 3 and 4 we have

$$\begin{aligned}
q[\sigma, \tilde{\mathcal{M}}] &= \sum_{\theta \in \tilde{\mathcal{M}}} q(\sigma, \theta) \\
&= \sum_{\theta \in \tilde{\mathcal{M}}} \sum_{(\sigma \xrightarrow{\lambda} \theta) \in \text{out}(\sigma)} \lambda \\
&= \sum_{\theta \in \tilde{\mathcal{M}}} \sum_{\substack{(\rho \xrightarrow{r} \pi) \in \mathcal{R} \\ \sigma - \rho + \pi = \theta}} \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} \cdot r \\
&= \sum_{\substack{(\rho \xrightarrow{r} \pi) \in \mathcal{R} \\ \sigma - \rho + \pi \in \tilde{\mathcal{M}}} } \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} \cdot r \\
&= \sum_{\tilde{\mathcal{M}} \in \mathcal{H}^\uparrow} \sum_{\rho \in \tilde{\mathcal{M}}} \sum_{\substack{(\rho \xrightarrow{r} \pi) \in \mathcal{R} \\ \sigma - \rho + \pi \in \tilde{\mathcal{M}}} } \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} \cdot r \\
&= \sum_{\tilde{\mathcal{M}} \in \mathcal{H}^\uparrow} \sum_{\rho \in \tilde{\mathcal{M}}} \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} \sum_{\substack{(\rho \xrightarrow{r} \pi) \in \mathcal{R} \\ \sigma - \rho + \pi \in \tilde{\mathcal{M}}} } r \\
&= \sum_{\tilde{\mathcal{M}} \in \mathcal{H}^\uparrow} \sum_{\substack{\rho \in \tilde{\mathcal{M}} \\ \text{s.t. } \rho \subseteq \sigma}} \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} \sum_{\substack{(\rho \xrightarrow{r} \pi) \in \mathcal{R} \\ \sigma - \rho + \pi \in \tilde{\mathcal{M}}} } r \tag{3.1} \\
&= \sum_{\tilde{\mathcal{M}} \in \mathcal{H}^\uparrow} \sum_{\substack{\rho \in \tilde{\mathcal{M}} \\ \text{s.t. } \rho \subseteq \sigma}} \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} \sum_{\pi \in \tilde{\mathcal{M}}_{\mathcal{M} \rightarrow \tilde{\mathcal{M}}}} \sum_{(\rho \xrightarrow{r} \pi) \in \mathcal{R}} r \tag{3.2} \\
&= \sum_{\tilde{\mathcal{M}} \in \mathcal{H}^\uparrow} \sum_{\substack{\rho \in \tilde{\mathcal{M}} \\ \text{s.t. } \rho \subseteq \sigma}} \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} \cdot \mathbf{rr}[\rho, \tilde{\mathcal{M}}_{\mathcal{M} \rightarrow \tilde{\mathcal{M}}}]
\end{aligned}$$

In particular, we have rewritten Equation [3.1] in Equation [3.2] by using Proposition 2. The proof is thus complete. \square

Proposition 4. Let \mathcal{S} be a set and \mathcal{R} an equivalence relation on \mathcal{S} . Let $\mathcal{H} = \mathcal{S}/\mathcal{R}$ and $\mathcal{H}^\uparrow = \mathbb{N}^{\mathcal{S}}/\mathcal{R}^\uparrow$. For any $\mathcal{M}, \mathcal{M}' \in \mathcal{H}^\uparrow$, for any $\sigma \in \mathcal{M}$, it holds that

$$\sum_{\rho \in \mathcal{M}'} \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} = \prod_{H \in \mathcal{H}} \binom{\mathcal{M}(H)}{\mathcal{M}'(H)} \tag{3.3}$$

Proof. We can rewrite the left-hand side of Equation [3.3] as follows:

$$\begin{aligned} \sum_{\rho \in \mathcal{M}'} \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} &= \sum_{\rho \in \mathcal{M}'} \prod_{H \in \mathcal{H}} \prod_{\substack{X \in \rho \\ \text{s.t. } X \in H}} \binom{\sigma(X)}{\rho(X)} \\ &= \prod_{H \in \mathcal{H}} \sum_{\rho_H \in (\mathcal{M}')_{|H}} \prod_{X \in H} \binom{\sigma(X)}{\rho_H(X)} \end{aligned} \quad (3.4)$$

In Equation [3.4], we use $(\mathcal{M}')_{|H}$ to denote the set obtained by projecting each $\rho \in \mathcal{M}'$ to H . Taking elements ρ_H from the set $(\mathcal{M}')_{|H}$ instead of taking each ρ in \mathcal{M}' and projecting it to H makes sure that each distinct multiset ρ_H is only counted once.

Considering that $\sum_{X \in H} \rho_H(X) = \rho_H(H) = (\mathcal{M}')_{|H}(H) = \mathcal{M}'(H)$ is constant for each $H \in \mathcal{H}$, just as $\sum_{X \in H} \sigma(X) = \sigma(H) = \mathcal{M}(H)$ for all $\sigma \in \mathcal{M}$, Vandermonde's identity can be applied:

$$\begin{aligned} \prod_{H \in \mathcal{H}} \sum_{\rho_H \in (\mathcal{M}')_{|H}} \prod_{X \in H} \binom{\sigma(X)}{\rho_H(X)} &= \prod_{H \in \mathcal{H}} \binom{\sum_{X \in H} \sigma(X)}{\sum_{X \in H} \rho_H(X)} \\ &= \prod_{H \in \mathcal{H}} \binom{\sigma(H)}{\rho_H(H)} \\ &= \prod_{H \in \mathcal{H}} \binom{\mathcal{M}(H)}{\mathcal{M}'(H)} \end{aligned}$$

□

We are now ready to state our first major result.

Theorem 1. *Let \mathcal{R} be an SE for the RN $(\mathcal{S}, \mathcal{R})$. Then, the multiset lifting \mathcal{R}^\uparrow induces the ordinarily lumpable partition \mathcal{H}^\uparrow on $MC(\mathcal{S}, \mathcal{R})$.*

Proof. We have to prove that for any $\mathcal{M}, \tilde{\mathcal{M}} \in \mathcal{H}^\uparrow$ and $\sigma, \sigma' \in \mathcal{M}$, we have that

$$q[\sigma, \tilde{\mathcal{M}}] = q[\sigma', \tilde{\mathcal{M}}].$$

We distinguish between the cases $\mathcal{M} \neq \tilde{\mathcal{M}}$ and $\mathcal{M} = \tilde{\mathcal{M}}$.

Case $\mathcal{M} \neq \tilde{\mathcal{M}}$. By Proposition 3, we have

$$q[\sigma, \tilde{\mathcal{M}}] = \sum_{\mathcal{M} \in \mathcal{H}^\uparrow} \sum_{\substack{\rho \in \mathcal{M} \\ \text{s.t. } \rho \subseteq \sigma}} \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} \cdot \mathbf{rr}[\rho, \tilde{\mathcal{M}}_{\mathcal{M} \rightarrow \tilde{\mathcal{M}}}]$$

which, using $\rho^{\bar{\mathcal{M}}}$ to denote any element of $\bar{\mathcal{M}}$, by Definition 10 can be rewritten as

$$\begin{aligned}
& \sum_{\bar{\mathcal{M}} \in \mathcal{H}^\uparrow} \sum_{\substack{\rho \in \bar{\mathcal{M}} \\ \text{s.t. } \rho \subseteq \sigma}} \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} \cdot \mathbf{rr}[\rho, \bar{\mathcal{M}}_{\mathcal{M} \rightarrow \bar{\mathcal{M}}}] = \\
& \sum_{\bar{\mathcal{M}} \in \mathcal{H}^\uparrow} \mathbf{rr}[\rho^{\bar{\mathcal{M}}}, \bar{\mathcal{M}}_{\mathcal{M} \rightarrow \bar{\mathcal{M}}}] \sum_{\substack{\rho \in \bar{\mathcal{M}} \\ \text{s.t. } \rho \subseteq \sigma}} \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} = \\
& \sum_{\bar{\mathcal{M}} \in \mathcal{H}^\uparrow} \mathbf{rr}[\rho^{\bar{\mathcal{M}}}, \bar{\mathcal{M}}_{\mathcal{M} \rightarrow \bar{\mathcal{M}}}] \sum_{\rho \in \bar{\mathcal{M}}} \prod_{X \in \rho} \binom{\sigma(X)}{\rho(X)} \quad (\text{by Proposition 4}) = \\
& \sum_{\bar{\mathcal{M}} \in \mathcal{H}^\uparrow} \mathbf{rr}[\rho^{\bar{\mathcal{M}}}, \bar{\mathcal{M}}_{\mathcal{M} \rightarrow \bar{\mathcal{M}}}] \prod_{H \in \mathcal{H}} \binom{\mathcal{M}(H)}{\bar{\mathcal{M}}(H)}
\end{aligned}$$

The same holds for σ' , obtaining

$$q[\sigma', \bar{\mathcal{M}}] = \sum_{\bar{\mathcal{M}} \in \mathcal{H}^\uparrow} \mathbf{rr}[\rho^{\bar{\mathcal{M}}}, \bar{\mathcal{M}}_{\mathcal{M} \rightarrow \bar{\mathcal{M}}}] \prod_{H \in \mathcal{H}} \binom{\mathcal{M}(H)}{\bar{\mathcal{M}}(H)}$$

Given that σ and σ' belong to the same block \mathcal{M} , the considered $\bar{\mathcal{M}}_{\mathcal{M} \rightarrow \bar{\mathcal{M}}}$ are the same for σ and σ' . This allows us to close the case, as, for any $\bar{\mathcal{M}} \in \mathcal{H}^\uparrow$, $\mathbf{rr}[\rho^{\bar{\mathcal{M}}}, \bar{\mathcal{M}}_{\mathcal{M} \rightarrow \bar{\mathcal{M}}}] \cdot \prod_{H \in \mathcal{H}} \binom{\mathcal{M}(H)}{\bar{\mathcal{M}}(H)}$ does not depend on σ or σ' .

Case $\mathcal{M} = \bar{\mathcal{M}}$. We have to prove that $q[\sigma, \mathcal{M}] = q[\sigma', \mathcal{M}]$, which, given that $\sigma, \sigma' \in \mathcal{M}$, can be rewritten as

$$q[\sigma, \mathcal{M} \setminus \{\sigma\}] + q(\sigma, \sigma) = q[\sigma', \mathcal{M} \setminus \{\sigma'\}] + q(\sigma', \sigma'). \quad (3.5)$$

From Definition 4 we have that $q(\sigma, \sigma) = -q[\sigma, \mathbb{N}^{\mathcal{S}} \setminus \{\sigma\}]$. If we partition $\mathbb{N}^{\mathcal{M}}$ according to \mathcal{H}^\uparrow , we obtain

$$q(\sigma, \sigma) = -q[\sigma, \mathcal{M} \setminus \{\sigma\}] - \sum_{\widehat{\mathcal{M}} \in \mathcal{H}^\uparrow \text{ s.t. } \mathcal{M} \neq \widehat{\mathcal{M}}} q[\sigma, \widehat{\mathcal{M}}]$$

The same holds for σ' , allowing us to rewrite Equation [3.5] as

$$- \sum_{\widehat{\mathcal{M}} \in \mathcal{H}^\uparrow \text{ s.t. } \mathcal{M} \neq \widehat{\mathcal{M}}} q[\sigma, \widehat{\mathcal{M}}] = - \sum_{\widehat{\mathcal{M}} \in \mathcal{H}^\uparrow \text{ s.t. } \mathcal{M} \neq \widehat{\mathcal{M}}} q[\sigma', \widehat{\mathcal{M}}] \quad (3.6)$$

Finally, we close the proof noticing that Equation [3.6] follows from the case $\mathcal{M} \neq \bar{\mathcal{M}}$. Indeed, we have shown that for every $\widehat{\mathcal{M}} \neq \mathcal{M}$, we have $q[\sigma, \widehat{\mathcal{M}}] = q[\sigma', \widehat{\mathcal{M}}]$. \square

3.1.1 Species equivalence as a generalization of Markov chain ordinary lumpability

SE was previously described as a lifting of ordinary lumpability of CTMCs to RNs. As this section will show, the notion of SE in fact collapses to that of ordinary lumpability when each state of a CTMC is encoded as a species of an RN, and each transition of a CTMC is encoded as a reaction with unary reactants and products corresponding to the source and target states of the transition, respectively.

Lemma 2. *Let \mathcal{C} be a CTMC with state space $\{1, \dots, n\}$ and let $q_{i,j}$ denote the transition rate from state i into state j . The RN-encoding of \mathcal{C} is given by*

- $\mathcal{S}_{\mathcal{C}} = \{X_i \mid 1 \leq i \leq n\}$
- $\mathcal{R}_{\mathcal{C}} = \{X_i \xrightarrow{q_{i,j}} X_j \mid q_{i,j} > 0\}$

Let \mathcal{R} be an equivalence relation on $\{1, \dots, n\}$, and

$\text{ReacSet}_{\mathcal{C}} = \{(X_i, X_j) \mid (i, j) \in \mathcal{R}\}$. Then \mathcal{C} is ordinarily lumpable with respect to \mathcal{R} if and only if

$\text{ReacSet}_{\mathcal{C}}$ is an SE of $(\mathcal{S}_{\mathcal{C}}, \mathcal{R}_{\mathcal{C}})$.

Proof. Given that we only have unary reactants, the only ρ to be considered in Definition 10 is \emptyset . Moreover, thanks to the fact that products are unary as well, it holds that $\text{rr}[X_i, \mathcal{M}] = \text{rr}[X_i, \mathcal{M}_{|1}]$ for all $1 \leq i \leq n$ and $\mathcal{M} \in (\mathbb{N}^{\mathcal{S}_{\mathcal{C}}})/\mathcal{R}_{\mathcal{C}}^{\uparrow}$, where $\mathcal{M}_{|1} = \{\sigma \in \mathcal{M} \mid |\sigma| = 1\}$. The claim then follows by noting that $q_{i,j} = \text{rr}(X_i, X_j)$ for all $1 \leq i, j \leq n$. \square

This result also applies to infinite state spaces, while lumping algorithms can be used only for finite ones (DHS03; VF10).

3.1.2 Characterization of SE for mass-action networks

This section provides our second major result, stating that SE is also a necessary condition for ordinary CTMC lumpability, for partitions obtained via multi-set liftings of equivalences on species. Before doing so, we have to introduce an auxiliary notion and statement.

Definition 11. *Let $(\mathcal{S}, \mathcal{R})$ be an RN. For $m \geq 1$ and $\bowtie \in \{=, \leq\}$, let $R_{\bowtie m} := \{(\rho \xrightarrow{\alpha} \pi) \in \mathcal{R} \mid |\rho| \bowtie m\}$. With this, define $\text{out}_{\bowtie m}(\sigma)$ using $R_{\bowtie m}$ and let $q_{\bowtie m}(\sigma, \theta)$ refer to the transition rate from σ into θ in $MC(\mathcal{S}, R_{\bowtie m})$.*

The following auxiliary statement will be needed.

Lemma 3. *Let $(\mathcal{S}, \mathcal{R})$ be an RN, \mathcal{R} an equivalence relation over \mathcal{S} and $\mathcal{H}^\uparrow = \mathbb{N}^{\mathcal{S}} / \mathcal{R}^\uparrow$. Then, if \mathcal{R} is an SE of $(\mathcal{S}, R_{\leq m})$ for some $m \geq 1$, it holds that \mathcal{H}^\uparrow is an ordinary lumpable partition of $MC(\mathcal{S}, R_{\leq m})$.*

Proof. Follows by applying Theorem 1 to $R_{\leq m}$ instead of \mathcal{R} . \square

Armed with Lemma 3, we can show our second main result in its final form.

Theorem 4. *Let $(\mathcal{S}, \mathcal{R})$ be an RN, \mathcal{R} an equivalence relation over \mathcal{S} and $\mathcal{H}^\uparrow = \mathbb{N}^{\mathcal{S}} / \mathcal{R}^\uparrow$. If \mathcal{H}^\uparrow is an ordinary lumpable partition of $MC(\mathcal{S}, \mathcal{R})$, then \mathcal{R} is an SE.*

Proof. We next prove

$$\forall \rho \in \mathbb{N}^{\mathcal{S}}, \forall \rho' \in \mathbb{N}^{\mathcal{S}}, \forall \tilde{\mathcal{M}} \in \mathcal{H}^\uparrow \text{ it holds that} \\ (\rho, \rho') \in \mathcal{R}^\uparrow \text{ implies } \mathbf{rr}[\rho, \tilde{\mathcal{M}}] = \mathbf{rr}[\rho', \tilde{\mathcal{M}}].$$

To this end, we proceed by induction on $|\rho| = m$:

- $|\rho| = 1$: Then $\mathbf{rr}[\rho, \tilde{\mathcal{M}}] = q[\rho, \tilde{\mathcal{M}}] = q[\rho', \tilde{\mathcal{M}}] = \mathbf{rr}[\rho', \tilde{\mathcal{M}}]$, where the second identity follows from the assumption.
- $|\rho| = m + 1$: By observing that $q[\rho, \tilde{\mathcal{M}}] = q_{=m+1}[\rho, \tilde{\mathcal{M}}] + q_{\leq m}[\rho, \tilde{\mathcal{M}}]$, the induction hypothesis and Lemma 3 ensure that $q_{\leq m}[\rho, \tilde{\mathcal{M}}] = q_{\leq m}[\rho', \tilde{\mathcal{M}}]$. Thanks to the fact that $q[\rho', \tilde{\mathcal{M}}] = q_{=m+1}[\rho', \tilde{\mathcal{M}}] + q_{\leq m}[\rho', \tilde{\mathcal{M}}]$, the assumption ensures that $q_{=m+1}[\rho, \tilde{\mathcal{M}}] = q_{=m+1}[\rho', \tilde{\mathcal{M}}]$. This, in turn, yields

$$\mathbf{rr}[\rho, \tilde{\mathcal{M}}] = q_{=m+1}[\rho, \tilde{\mathcal{M}}] = q_{=m+1}[\rho', \tilde{\mathcal{M}}] = \mathbf{rr}[\rho', \tilde{\mathcal{M}}].$$

\square

3.1.3 Computation of the maximal SE and reduced network

Starting from an initial partition of species, we use a partition refinement algorithm to compute the largest SE that is a refinement of the initial partition. Therefore, the maximal SE is a special case that can be computed

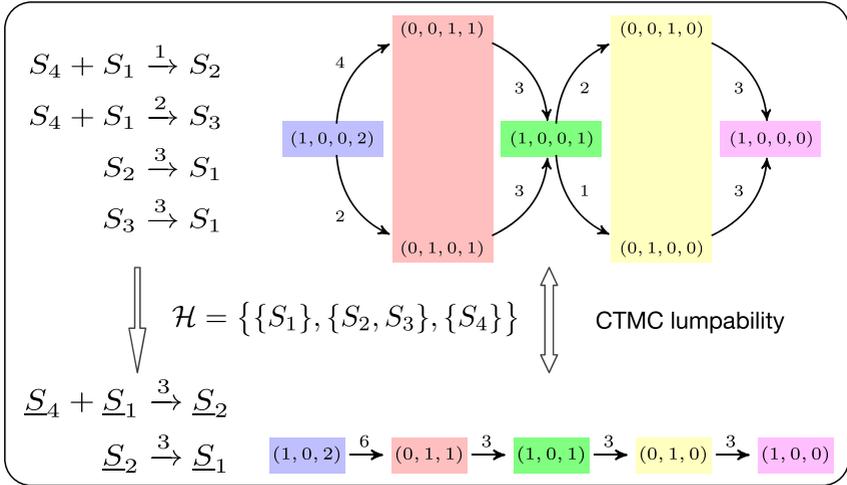


Figure 2: Illustration of SE on a simple network with species S_1, \dots, S_4 . The top-right diagram shows the graphical representation of the underlying Markov chain starting from the initial condition $\hat{\sigma} = (1, 0, 0, 2)$. Each node is a state and each arc is labeled with the transition rate according to mass-action kinetics. The colored boxes represent five blocks of an ordinary lumpable partition of the Markov chain (here it suffices to check that the outgoing transitions are equal for states in blocks of size two). The partition of species denoted by \mathcal{H} can be shown to be an SE, hence states that are equal up to the sum of the second and third coordinate form a lumpable partition block. This SE gives rise to a reduced network by choosing the representatives S_1, S_2 and S_4 for each block (underlined in the figure to distinguish them from original species names). The reduced network has fewer reactions due to the fact that reactions in the original network are merged into a single one after renaming. The bottom-right diagram shows the underlying Markov chain of the reduced network starting from the matching initial state $(1, 0, 2)$. The Markov chain of the reduced network essentially corresponds to the lumped Markov chain of the original network (as indicated by the matching colors of the nodes).

by initializing the algorithm with the partition with the trivial singleton block containing all species. The algorithm maintains a reference to the current candidate SE partition and a set of splitters, i.e., blocks of species against which the candidate partition is to be checked. Both structures are initialized with the input partition.

A fixed-point iteration splits a block of the current candidate SE partition whenever it falsifies the condition in Def. 10 with respect to a splitter. If no such block is found then the algorithm terminates and the current candidate partition is proven to be the largest SE refinement of the initial partition. Else, the falsifying block is split into sub-blocks that have equal values for the quantities in the equation of Def. 10. Each as-refined sub-block is a potential splitter for a further iteration of the algorithm; importantly, however, the largest may be ignored (PT87), and thus is not added to the set of splitters to be analyzed in the subsequent iterations. In (CPVT⁺21), we prove that this algorithm has $O(pm)$ space and $O(n^2m^3p(p + \log m))$ time complexity, where p is the largest number of different species appearing in the reactants or products of every reaction. A more detailed description of the algorithm can also be found there.

Given an SE \mathcal{H} , it is possible to construct a reduced network using an algorithm similarly to ref. (CTTV15), where it was developed for network reductions with deterministic reaction-rate interpretation. Briefly, the reduced network is obtained by applying the following four steps: (i) choose a representative species for each equivalence class of \mathcal{H} ; (ii) discard all reactions whose reactants have species that are not representatives; (iii) replace the species in the products of the remaining reactions with their representatives; (iv) reduce the set of reactions by merging all those that have same reactants and products by summing their kinetic parameters.

Each representative in the reduced network can be interpreted as a macro-species that tracks the sum of the populations of the distinct species in the original network that belong to the same SE equivalence class. Therefore, for any given initial condition $\hat{\sigma}$ of the original network, it is possible to directly generate its lumped Markov chain from the re-

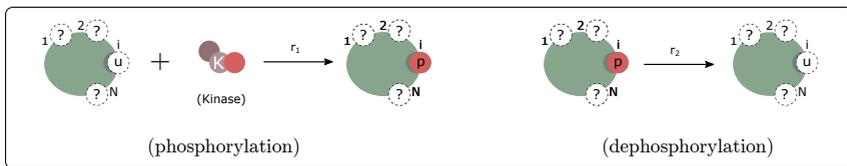


Figure 3: Simple multisite phosphorylation process (SH09). A protein (green circle) has N sites which can be in two states: phosphorylated (p) or unphosphorylated (u). Phosphorylation occurs according to a random mechanism with a kinase (brown circle) which can target the protein at each site i , independently from other sites (which can be in any state, as indicated by the '?' symbol). The protein exhibits the same affinity with the kinase at all sites through the kinetic parameter r_1 . Dephosphorylation is modeled as a spontaneous reaction at each site with kinetic parameter r_2 .

duced network by fixing a matching initial condition up to sums of populations (Fig. 2).

The algorithms for species equivalence have been implemented in the software tool ERODE (CTVT17), available at <https://sysma.imtlucca.it/tools/erode/>.

3.2 Applications

We now present some examples to show the real-world applicability of SE. All analyzed models are available as ERODE specifications at <http://bit.ly/ERODE-BENCHMARKS-SE>.

3.2.1 Computational systems biology

Mechanistic models of signaling pathways are prone to a rapid growth in the number of species and reactions because of the combinatorial effects due to the distinct configurations in which a molecular complex can be found (SH09). A prototypical situation is multisite phosphorylation, a fundamental process in eukaryotic cells that is responsible for various mechanisms such as the regulation of switch-like behavior (Gun05; TG09). For example, let us consider a protein with N sites

that can be phosphorylated according to a random mechanism (Fig. 3). This requires 2^N distinct molecular species in the network to track the state of each individual site (SH09). To simplify the mathematical model it is often assumed that the kinetic parameters are equal at all phosphorylation sites (SFE11). With this, the maximal SE provides a stochastically equivalent reduction using N equivalence classes, each aggregating the behavior of all distinct protein configurations that have the same number of phosphorylated sites, independently of their identity.

The assumption of equal kinetic parameters is not necessary to achieve aggregation with SE. This can be shown on a model from ref. (PKUMK10) of the interactions between calcium (Ca^{2+}), calmodulin (CaM), and the Ca^{2+} -CaM dependent protein kinase II (CaMKII), which play a fundamental role in the mechanism of synaptic plasticity (LSC02). An overall reduction from 156 species and 480 reactions to 76 species and 264 reactions was achieved here. Notably, important quantities to observe in this model are the amounts of free and bound CaM (LYR12), both recoverable from the reduced network.

Further, SE can also aggregate species that exhibit contrasting functionality, such as in signal transduction switches realized by GTP- and GDP-bound forms of GTPases. An example for this is the mechanistic model of the spindle position checkpoint (SPOC)(CLG⁺12), a mechanism in the budding yeast responsible for detecting the correct alignment of the nucleus between between mother and daughter cells (LB03).

Here, the original network with 24 species and 71 reactions is reduced to 16 species and 36 reactions, from which one may recover observables of interest such as the total amount of active Bfa1 (CLG⁺12).

Both models were obtained from the BioModels database (LDR⁺10) and analyzed using the original kinetic parameters: the CaM-CaMKII model is identified as MODEL1001150000, the SPOC model is identified as BIOMD0000000699.

3.2.2 Epidemic processes in networks

Models of epidemic processes are well established since the celebrated work by Mermack and McKendrick (KM27). They have received considerable attention from several disciplines including biology, computer science, mathematics, physics, and sociology due to the generality with which, in addition to the diffusion of pathogens, the epidemic analogy can be applied to a variety of phenomena such as the spreading of rumor (MT73), opinion (CFL09), as well as computer viruses (WGHB09). The availability of large datasets in a range of socio-technical systems has prompted the study of epidemic processes on complex networks that consider the heterogeneity of real-world processes, which is neglected in simpler variants that assume a well-mixed, uniform environment (PSCVMV15).

Here we show that SE is an effective aggregation method for epidemic processes on complex networks. In (CPVT⁺21), we also include a direct comparison to the orbit partition lumping algorithm by Simon *et al.* (STK11). As an example, we study the well-known susceptible-infected-susceptible (SIS) model, where each node in the network in the susceptible state can be infected with a rate proportional to the number of infected neighbors, and recover from the infection according to an independent Poissonian process. Let $A = (a_{ij})$, with $A \in \mathbb{R}^{N \times N}$, define the adjacency matrix of a graph with N nodes representing the network topology, with $a_{ij} > 0$ denoting the presence of a possibly weighted edge between node i and j .

The SIS epidemic process can be described by the network

$$S_i + I_j \xrightarrow{a_{ij}\lambda} I_i + I_j, \quad I_i \xrightarrow{\gamma} S_i, \quad 1 \leq i, j \leq N, \quad j \neq i, \quad (3.7)$$

where the first reaction models infections by neighbors and the second reaction is the spontaneous recovery, with parameters λ and μ respectively. In a similar fashion, different variants of the process, such as SIR, SIRS, and SEIR (PSCVMV15), can be described. Any physically meaningful initial condition $\hat{\sigma}$ for this network must be such that each node i is initially in infected ($\hat{\sigma}_{S_i} = 0, \hat{\sigma}_{I_i} = 1$) or susceptible ($\hat{\sigma}_{S_i} = 1, \hat{\sigma}_{I_i} = 0$).

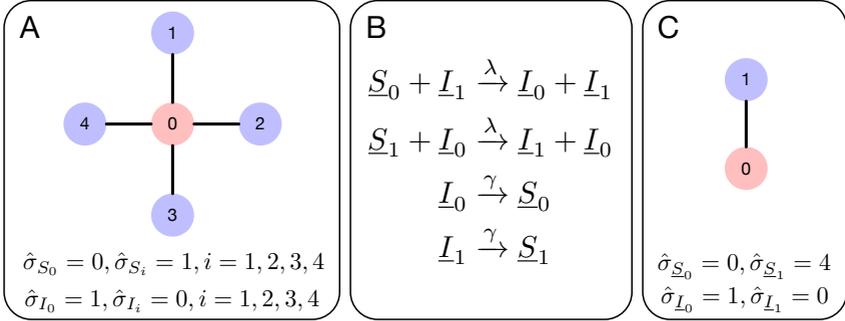


Figure 4: Example of SE reduction of SIS dynamics on a coarse-grained network. (A) Star network over which an SIS process evolves according to Eq. 3.7, starting from an initial condition where the infection starts at node 0. (B) Reduced network (species representatives are underlined in the figure for clarity) according to the largest SE refinement of the initial partition with blocks $\mathcal{S} = \{S_0, S_1, S_2, S_3, S_4\}$ and $\mathcal{I} = \{I_0, I_1, I_2, I_3, I_4\}$. This SE has blocks $\{S_0\}$, $\{I_0\}$, $\{S_1, S_2, S_3, S_4\}$ and $\{I_1, I_2, I_3, I_4\}$. (C) The SE partition induces a partition on the graph with blocks $\{0\}$ and $\{1, 2, 3, 4\}$. The reduced network corresponds to the description of the SIS dynamics on the quotient graph. The lumpability relation holds for an initial condition of the reduced network that is consistent with the initial condition of the original network up to SE.

This setting makes stochastic models of epidemics spreading on complex networks difficult to study exactly because the state of each individual node is tracked explicitly (WTSB17), leading to a state space size with 2^N distinct configurations (STK11). SE provides an ordinary lumpability of the underlying Markov chain, without ever generating it, on the network of Eq. 3.7, which has exponentially smaller size because it has $2N$ species and $E + N$ reactions, where E is the number of nonzero entries in the adjacency matrix of the graph.

For the SIS model, the maximal SE is the trivial partition where all the species are in a single block. This is an invariant property stating that the total population of individuals in the system is constant (STK11). Thus, we consider non-degenerate reductions using initial partitions with two blocks, hereafter denoted by $\mathcal{S} = \{S_i \mid 1 \leq i \leq N\}$ and $\mathcal{I} = \{I_i \mid 1 \leq i \leq N\}$, that separate species associated with nodes in the susceptible state

Table 2: Coarse graining of SIS models on real-world networks.

<i>Network</i>	<i>Original size</i>		<i>Reduced size (SE)</i>	
	<i>N</i>	<i>E</i>	<i>N</i>	<i>E</i>
tntp-ChicagoRegional (ECLB83)	1 467	2 596	635 (43%)	932 (36%)
ego-facebook (ML12)	2 888	5 962	35 (1%)	104 (2%)
as20000102 (LKF07)	6 474	27 790	3 885 (60%)	19 437 (70%)
arenas-pgp (BPSDGA04)	10 680	48 632	8 673 (81%)	44 074 (91%)
web-webbase-2001 (BCSV04)	16 062	51 186	5 253 (33%)	24 232 (47%)
as-caida20071105 (LKF07)	26 475	106 762	13 393 (51%)	69 184 (65%)
ia-email-EU (LKF07)	32 430	108 794	6 262 (19%)	53 228 (49%)
topology (ZLMZ05)	34 761	215 440	19 246 (55%)	168 782 (78%)
douban (ZL09)	154 908	654 324	59 524 (38%)	462 128 (71%)

from those in the infected state, respectively.

On a simple star graph (Fig. 4), an inspection of the obtained SE equivalence classes reveals that each refinement of the initial block \mathcal{S} matches a refinement of block \mathcal{I} for the same subset of nodes of the graph. Such an SE naturally induces a partitioning of the graph, and the reduction can be understood as an SIS dynamics on the quotient graph where each macro-node subsumes a partition block of nodes induced by SE. A similar observation can be made from the analysis of the SE reductions of SIS evolving on several real-world benchmark networks (Table 2).

The networks were produced from the network specification files available in public repositories: tntp-ChicagoRegional, ego-facebook, as20000102, arenas-pgp, as-caida20071105, topology, and douban were taken from the Koblenz Network Collection (Kun13); networks web-webbase-2001 and ia-email-EU were taken from the Network Data Repository (RA15).

Since the reduced model is an epidemic process, it is still amenable to a wide range of analysis techniques developed for such models (PSCVMV15; WTSB17). These include mean-field and pair approximation (VM11; CVM12; MF13), whose computational cost for the generation and solution of the resulting nonlinear differential equations may

benefit from the availability of a stochastically equivalent reduced model.

3.3 Discussion

Species Equivalence can efficiently reduce the size of a lumpable network. It evades the problem of state space explosion by directly working with the species and reactions of the network instead of the potentially intractable Markov Chain state space. Since species equivalence gives rise to a network where the reactions preserve the structure (up to a renaming of the species into equivalence classes), the reduction maintains a physical interpretation in terms of coarse-grained interactions between populations of macro-species.

Another consequence of the availability of a reduced network is that our method is orthogonal to any of the analysis techniques developed for stochastic reaction networks. Numerical simulations may run faster because they traverse fewer reactions at each time step (CTTV17b); when feasible, one can generate the underlying Markov chain to be further analyzed or reduced (VF10; MK06; HMW09); the reduced network can be subjected to complementary coarse-graining techniques concerned with time-scale separation (SHN09; GUVT08; KK13; CW16). More generally, since the reduced network preserves the stochastic dynamics in the sense specified above, it can be used as the basis for various forms of approximate analysis such as linear noise or moment closure approximation (SSG17), where the complexity of the resulting system of equations grows rapidly with the network size.

Chapter 4

DiffLQN: Differential Equation Analysis of Layered Queuing Networks

This chapter presents DiffLQN, the first software tool that supports DRE analysis of LQNs. We envisage two use cases:

- i) For end users, it represents an alternative to Carleton University's `lqns` (Car14), the state-of-the-art tool for LQNs
- ii) For researchers, it is a testbed to push forward further efficient analysis techniques for LQNs.

To facilitate i), we designed DiffLQN having in mind compatibility with `lqns`, supporting a text-based syntax that presents only minor deviations (detailed later). To facilitate ii) we make available the source code to welcome further integration, extensions, and optimizations.

DiffLQN is 100% Java. A single-JAR executable is available for download at <https://sysma.imtlucca.it/tools/difflqn/>.

4.1 DiffLQN

4.1.1 Architecture

Currently, the front-end of DiffLQN is a command line that accepts a text-based representation of an LQN. The parser is automatically generated from Eclipse's Xtext framework. In (Tri13), the algorithmic derivation of the DRE was mediated by a translation of an LQN into a model written in the process algebra PEPA (Hil96). We exploit this fact, by converting the abstract syntax tree of an LQN into a PEPA model in order to leverage the tool support for this process algebra: in particular, we use `PEPATO`, the API of the PEPA Eclipse Plugin (TDG09) in order to generate, analyze, and manipulate the DRE of the LQN model.

In essence, DiffLQN tracks the correspondence from LQN elements to process-algebra models, and back for the propagation of the analysis results. In this way, the intermediate translation step into PEPA is hidden to the end user.

4.1.2 Capabilities

DiffLQN provides the following LQN performance indices:

- *Throughput*, at different levels of granularity: it provides the average number of activities, entries, or tasks completed per unit time at steady state.
- *Utilization* for processors and tasks, giving the average number of busy entities at steady state. Utilization estimates are also provided per single entry and activity, giving their contribution to the utilization of the processor on which they are deployed.
- *Response time*, the average response time at steady state for the execution of entries or tasks.

These can be computed by numerical integration of the DRE or by stochastic simulation, with user-tuneable settings that will be described below. Because of the presence of very fast rates in the LQN encoding

presented in (Tri13), we extended PEPAtO with a stiff ODE solver, which was not originally included in the API. Our contribution is adapted from the JVode component of the BioUML workbench (Ins15). Stochastic simulation is done using Gillespie’s algorithm (Gil77), directly leveraging the implementation available in the PEPA Eclipse plug-in.

4.1.3 Syntax

Model specification To favour compatibility, DiffLQN accepts a slight variant of the text-based input format for lqns. We refer to (Car14) for a complete documentation on the grammar. Here we highlight three main differences. First, names have to be unique across tasks and entries, regardless of the capitalization. Unlike lqns, reusing a task name for an entry running inside it is not allowed. Second, phases of an entry are designated by appending digits to its name, so having e.g. one entry called Name and another element called Name1 will lead to errors if the entry Name has more than one phase. Third, labels for calls can only be integers, and they are always interpreted as the number of calls every time the caller is accessed. This is in contrast to lqns where also a probability distribution over the outgoing calls of an activity may be specified.

Some LQN features, e.g. loops and tasks with infinite multiplicity, are not currently supported. The input file template that is available on the website explains all supported keywords.

DiffLQN accepts LQN models containing unsupported features as valid syntax, but the solver either emits a warning, or explains the problem in an error message.

Solver settings ODE analysis is performed by solving an initial value problem numerically until convergence to steady state is detected (or if a threshold time horizon is reached, in which case a warning is issued if convergence has not been reached). The convergence criteria are based on *absolute* and *relative tolerances*. The former considers the Euclidean norm of the derivatives of the solution at the current time point, and absolute convergence is reached when this value is below a given threshold

(formally, the norm must be equal to zero in the steady state); the latter compares the norm of the difference between the solutions at successive time points. By default, the analysis terminates successfully when both the absolute and the relative convergence criteria are met.

Stochastic simulation is performed using the method of batch means (Ste09): roughly speaking, a single simulation run is performed and statistics are collected across different non-overlapping parts of the run (the batches) which are assumed to be long enough that the system has reached steady state.

Solver settings for DiffLQN are backward compatible with lqns since every line must start with '#!', which is treated by lqns as a comment. Below we list the settings that are currently supported.

- `v` specifies the value for a fast rate v that approximates the behavior of certain operations, such as forks and joins, that are assumed to be instantaneous in LQNs. This is the only mandatory setting.
- `solver [ode | sim]` specifies whether to use DRE analysis or stochastic simulation.
- `stoptime` specifies the maximum time horizon for the numerical ODE integration or the length of an initial transient simulation run that is removed before batch statistics are collected.
- `solver_abs_tol` and `solver_rel_tol` are typical absolute and relative tolerances for the ODE numerical integration (AP88).
- `steady_abs_tol` and `steady_rel_tol` specify the tolerances for ODE steady-state detection, as discussed above.
- `[absolute | relative] steady state` is a flag for using only one of the two criteria of steady-state convergence.
- `batch_length_factor` specifies the length of a batch, relative to the initial transient defined with `stoptime`.
- `confidence_level` together with `confidence_percent_error` specify the usual termination criteria for stochastic simulation.

Output settings By default, DiffLQN computes all possible performance measures discussed in Sect. 4.1.2. Optionally the user can explicitly choose which measures to track. This can be speed up the computation, especially for large networks analyzed using simulation (TDGH12). This is done in a block with lines starting (in order) with keywords `throughput`, `utilisation`, and `response time`, followed by a list of desired elements for the respective performance index.

Exporting options By default, analysis results are outputted to the screen in a human-readable format. However, the LQN model as well as the results can be exported in different formats. Each export command is specified in a new line with the `export` keyword, followed by the type of export requested (and an optional file path). Available export types are:

- `pepa`: Export of the PEPA encoding of the input LQN, in a format that is compatible with the PEPA Eclipse plug-in.
- `matlab`: A function file in Matlab-compatible form which can be used in conjunction with Matlab’s ODE solvers (e.g., the stiff solver `ode15s`)
- `csv`: Results are saved to a comma-separated values file.

Fig. 5 shows the input file for the network of Fig. 1. The `G`-block in lines 1–7 is ignored by DiffLQN because it provides parameter settings specific to `lqns`. Lines 11–61 contain the various model definition blocks as explained in chapter 2. Finally, the options specific to DiffLQN are in lines 64–69.

4.2 Case Study: Client-Server dynamics

As a case study we evaluate DiffLQN on the running example. For this, we consider a comparison between the DRE results and the simulation results. The latter are taken to be the “true” values of the performance indices, following the successful validation against the simulation results

```
1 G
2 "Example-LQN"
3 0.0001
4 500
5 1
6 0.5
7 -1
8
9 # processor definition block
10 P 0
11 p PClient f m 2
12 p PServer f m 2
13 p PDisk f m 2
14 -1
15
16 # task definition block
17 T 0
18 t Client r think -1 PClient m 2
19 t Server n visit buy notify save -1 PServer m 2
20 t FileServer n read write -1 PDisk
21 t Backup n get update -1 PDisk
22 -1
23
24 # entry definition block
25 E 0
26 s think 0.1 -1
27 y think visit 3 -1
28 y think save 1 -1
29 y think notify 1 -1
30 y think read 1 -1
31 y think buy 1 -1
32 A visit cache
33 A buy prepare
34 s save 0.02 -1
35 y save write 1 -1
36 s notify 0.08 -1
37 s read 0.01 -1
38 s write 0.001 0.04 -1
39 y write get 0 1 -1
40 y write update 0 1 -1
41 s get 0.01 -1
42 s update 0.01 -1
43 -1
```

```
44 # activity definition block for task Server
45 A Server
46 s prepare 0.01
47 s pack 0.03
48 s ship 0.01
49 s display 0.001
50 s cache 0.001
51 s internal 0.001
52 s external 0.003
53 y external read 1
54 :
55 prepare -> pack & ship;
56 pack & ship -> display;
57 cache -> (0.95)internal + (0.05)external;
58 internal[visit];
59 external[visit];
60 display[buy]
61 -1
62
63 # DiffLQN settings block
64 #! v 1.0e5
65 #! solver sim
66 #! confidence_level 0.98
67 #! confidence_percent_error 2.0
68 #! stoptime 1000.0
69 #! export csv
```

Figure 5: Example input file for the network of Fig. 1.

of lqns performed in (Tri13). In particular, the settings in Fig. 5 indicate that the simulations were set to stop when the 98% confidence levels were within 2% of the estimated averages.

To show the advantages in using DRE analysis for larger multiplicities we consider two scenarios: the first scenario uses the parameters as shown in Fig. 5 (we denote this by the label $x1$); the second scenario uses the same service demands, but all multiplicities for processors and tasks are increased by a factor 10 (label $x10$). (For convenience, both scenarios are available for download as separate input files.)

The numerical results are presented in Tab. 3. However, to reduce clutter only a selection of all performance estimates are presented. In particular, we removed repeated throughput estimates that were equal to those already found in the table. (This can happen when certain activities are performed sequentially, for instance `prepare` and `display` have the same steady-state throughput). The first column gives the type of the measure as a triple consisting of a metric — throughput (Th), utilization (Ut), processor utilization (PU), or response time (RT) — kind of LQN entity, and LQN entity name. The other columns show the performance estimates from DRE analysis and simulation in both scenarios, together with the percentage relative errors.

Overall, we can make the following main observations:

- Despite the low multiplicities of processors and tasks in scenario $x1$, the DRE estimates enjoy good accuracy in most cases.
- The highest error in scenario $x1$, 66.46%, occurs for a response-time metric (entry `save`). This confirms that response times can be challenging to approximate, because the errors of the basic metrics from which their are computed through Little’s law can propagate (TDGH12).
- Scaling up multiplicities in scenario $x10$ shows a considerable improvement on the accuracy, despite the fact that the model has populations of entities in the order of tens, which is significantly away from a limiting regime with infinitely many entities, where the DRE estimate is asymptotically exact.

- In scenario $x10$, the largest error is roughly halved. Although still large, it is possible to notice that the trend of that response-time metric is followed fairly well. Indeed, we remark that this is an instance where the percentage relative error may not be very informative because it tends to penalize small variations in metrics that have small “true” values to start with (EL88).

4.3 Discussion

DiffLQN is a tool that supports differential-equation analysis for layered queuing networks. In its current version, DiffLQN already shows the potential of DRE approximation for LQNs. Considering the DRE-typical trait of increasing accuracy for systems with higher multiplicity, it especially presents an interesting complement to AMVA with its contrasting tendency to be more reliable for smaller systems (Tri13). There is however room to further increase the usability and applicability of DiffLQN by adding additional features. Being based on Eclipse’s Xtext framework, a natural evolution would be to provide a graphical user interface as an Eclipse plug-in, with the possibility of drawing LQNs in addition to specifying them textually. To enhance the capability of conducting large experiments such as what-if scenarios or capacity-planning studies, the syntax can be augmented with parametric variables that can be instantiated (and the resulting model evaluated) over user-defined ranges. The numerical analysis of DREs gives the time-course evolution of the queue-length process at each station, from which the steady-state LQN metrics are derived. Making these traces available to the user could allow them to obtain performance indices of the transient regime on an LQN as well.

<i>Metric/Kind/Name</i>	<i>Scenario x1</i>			<i>Scenario x10</i>		
	<i>DRE</i>	<i>Sim.</i>	<i>Error</i>	<i>DRE</i>	<i>Sim.</i>	<i>Error</i>
Th / act / prepare	7.681	6.251	22.88	76.815	72.519	5.92
Th / act / cache	23.044	18.680	23.36	230.444	217.812	5.80
Th / act / internal	21.892	17.754	23.31	218.922	206.831	5.85
Th / act / external	1.152	0.928	24.11	11.522	10.845	6.25
Th / entry / read	8.834	7.141	23.70	88.337	83.370	5.96
Th / entry / get	7.681	6.237	23.17	76.815	72.746	5.59
Th / entry / update	7.681	6.258	22.75	76.815	72.585	5.83
Th / entry / visit	23.044	18.705	23.20	230.444	217.561	5.92
Th / task / Server	46.089	37.396	23.24	460.889	435.145	5.92
Th / task / FileServer	16.515	13.373	23.50	165.152	155.830	5.98
Th / task / Backup	15.363	12.494	22.96	153.630	145.331	5.71
Ut / proc / PClient	0.768	0.624	23.12	7.681	7.251	5.94
Ut / proc / PServer	1.208	0.980	23.26	12.083	11.397	6.02
Ut / proc / PDisk	0.557	0.452	23.25	5.569	5.258	5.92
Ut / task / Client	2.000	2.000	0.00	19.999	19.999	0.00
Ut / task / Server	1.142	1.264	9.65	11.424	11.994	4.75
Ut / task / FileServer	0.578	0.610	5.32	5.776	6.003	3.78
Ut / task / Backup	0.154	0.125	22.96	1.539	1.456	5.71
PU / act / pack	0.230	0.187	22.95	2.304	2.173	6.06
PU / act / get	0.077	0.062	23.17	0.768	0.727	5.59
RT / entry / think	0.260	0.321	18.78	0.260	0.276	5.61
RT / entry / save	0.021	0.063	66.46	0.021	0.034	37.58
RT / entry / notify	0.080	0.080	0.00	0.080	0.080	0.00
RT / entry / read	0.010	0.010	0.01	0.010	0.010	0.01
RT / entry / write	0.061	0.061	0.08	0.061	0.061	0.09
RT / entry / get	0.010	0.010	0.00	0.010	0.010	0.00
RT / entry / update	0.010	0.010	0.00	0.010	0.010	0.00
RT / entry / visit	0.002	0.005	59.95	0.002	0.003	34.92
RT / entry / buy	0.041	0.044	5.85	0.041	0.042	1.18
RT / task / Client	0.260	0.321	18.78	0.260	0.276	5.61
RT / task / Server	0.025	0.034	26.70	0.025	0.028	10.07
RT / task / FileServer	0.034	0.034	0.21	0.034	0.034	0.06
RT / task / Backup	0.010	0.010	0.000	0.010	0.010	0.00
Average percentage errors			18.38			6.17

Table 3: Numerical results of DiffLQN on the example LQN from Fig. 1. The errors are measured as percentage relative errors from the simulation estimate.

Chapter 5

The Finite State Expansion Method for Stochastic Reaction Networks

Here we present finite state expansion, a method which offers a principled way to interpolate between the ME and the DRE, resting on a systematic transformation of an RN with arbitrary propensity functions into a new RN augmented with additional species and reactions. The key insight is to explicitly track a finite subset of the discrete microscopic state space, treating each such state as a new individual species of the network and coupling its dynamics with those of the original species. Roughly speaking, the role of the original set of species is to buffer the probability mass that falls out of the state space that is tracked. In this respect, finite state expansion can be seen as a mass-preserving variation to the well-known finite state projection method (MK06), which truncates the state space.

Our transformation is proved to be stochastically exact when the RN is analyzed with the ME. The DRE resulting from finite state expansion, instead, can be interpreted as a master equation projected on the subset of the state space that is explicitly tracked, coupled with the macroscopic population-based description provided through the origi-

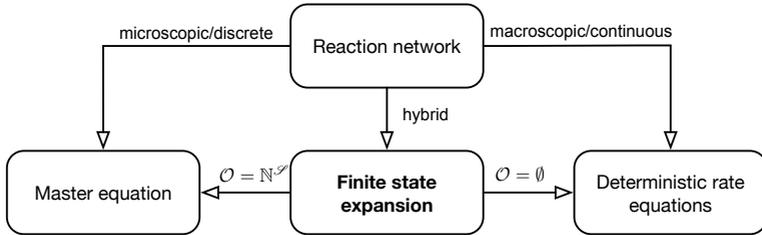


Figure 6: The microscopic dynamics of an RN with species \mathcal{S} is given by the master equation, which tracks the probability distribution over all possible discrete configurations of the species populations. The deterministic rate equations approximate the average population levels. Finite state expansion is a stochastically equivalent translation into a target RN where discrete configurations levels belonging to a modeler-provided set \mathcal{O} are represented as auxiliary species, defining a hybrid analytical model that mixes discrete and whole-population variables. Finite state expansion collapses to the master equation when all configurations are tracked, and to the deterministic rate equations when no configuration is tracked.

nal species. Through examples drawn from biochemistry and computer science we show that the DREs of the expanded RN improve the accuracy of the mean estimates, with a quality of the approximation that increases with larger sets of observed state configurations. This proposes finite state expansion as an automatic technique to trade off precision and computational cost in the analysis of RNs, without resorting to simulations.

5.1 The Finite State Expansion Method

As input, finite state expansion (FSE) takes a user-defined bound $\bar{O} \in \mathbb{N}^{\mathcal{S}}$ that defines the tracked subset \mathcal{O} of the state space as all state configurations that are (component-wise) smaller, i.e., $\mathcal{O} = \{o \in \mathbb{N}^{\mathcal{S}} \mid o \leq \bar{O}\}$. Each tracked configuration $o \in \mathcal{O}$ is associated with an auxiliary species denoted by $\llbracket o \rrbracket$, resulting in an expanded set of species that we denote by $\mathcal{S}_{\bar{O}} = \mathcal{S} \cup \{\llbracket o \rrbracket \mid o \leq \bar{O}\}$. The expanded set of reactions $\mathcal{R}_{\bar{O}}$ is built by replacing each reaction $\rho \xrightarrow{f} \pi$ in the original RN with a set of reac-

tions that couple the behavior of each tracked configuration $\llbracket o \rrbracket$ with the original species:

$$\llbracket o \rrbracket + \eta \xrightarrow{f_o} \llbracket o' \rrbracket + \psi, \quad \text{for } o \in \mathcal{O}, \quad (5.1)$$

Intuitively, for each original reaction, Equation [5.1] considers its behavior with respect to each tracked configuration $\llbracket o \rrbracket$. Any expanded reaction maintains the same overall counts of reactants and products as the originating reaction, with a target observed configuration $\llbracket o' \rrbracket$ that results from the addition of products and removal of reactants within the upper bound \bar{O} . The vectors η and ψ refer to the original species of \mathcal{S} , which act as buffer for the configurations that are not explicitly tracked. Finally, the propensity function f_o is derived from that of the original reaction f , accounting for the fact that the observed state $\llbracket o \rrbracket$ encodes additional population counts, as given by o .

Definition 12 (Finite State Expansion). *Let $(\mathcal{S}, \mathcal{R})$ be an RN, $\bar{O} \in \mathbb{N}^{\mathcal{S}}$, and $\mathcal{O} = \{\sigma \in \mathbb{N}^{\mathcal{S}} \mid \sigma \subseteq \bar{O}\}$. Define the set of auxiliary species as $\llbracket \mathcal{O} \rrbracket = \{\llbracket o \rrbracket \mid o \in \mathcal{O}\}$. The RN obtained by finite state expansion (FSE) is given by the pair $(\mathcal{S}_{\bar{O}}, \mathcal{R}_{\bar{O}})$ where*

$$\mathcal{S}_{\bar{O}} = \mathcal{S} \cup \llbracket \mathcal{O} \rrbracket \quad \text{and} \quad \mathcal{R}_{\bar{O}} = \bigcup_{\rho \xrightarrow{f} \pi \in \mathcal{R}} \mathcal{R}_{\bar{O}} \left[\rho \xrightarrow{f} \pi \right],$$

with

$$\begin{aligned} \mathcal{R}_{\bar{O}} \left[\rho \xrightarrow{f} \pi \right] := & \left\{ \llbracket o \rrbracket + \eta \xrightarrow{f_o} \llbracket o' \rrbracket + \psi \mid o \in \mathcal{O}, \eta = \rho \ominus o, \right. \\ & o' = \bar{O} \cap (o \ominus \rho + \pi), \psi = o \ominus \rho + \pi \ominus \bar{O}, \\ & \left. f_o : \mathbb{R}^{\mathcal{S}_{\bar{O}}} \rightarrow \mathbb{R}_0^+, f_o(z) = z_{\llbracket o \rrbracket} \cdot f(o + z_{\setminus \mathcal{S}}) \right\}. \end{aligned}$$

We say that $\mathcal{R}_{\bar{O}} \left[\rho \xrightarrow{f} \pi \right]$ is the set of reactions *generated from the source reaction* $\rho \xrightarrow{f} \pi \in \mathcal{R}$. Similarly, the RN $(\mathcal{S}_{\bar{O}}, \mathcal{R}_{\bar{O}})$ is said to be *generated from its source RN* $(\mathcal{S}, \mathcal{R})$.

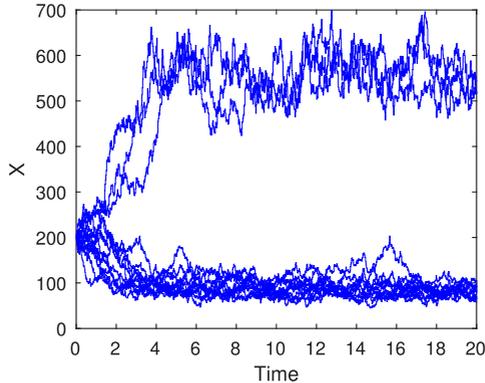


Figure 7: Stochastic simulation traces of the Schlögl model, illustrating the bimodality of the system.

5.1.1 Detailed example

As an example, consider the effect of applying FSE to the Schlögl system (Sch72):



The model describes an autocatalytic process for species X (reactions R1–R4), evolving according to the law of mass action.

Stochastic simulations (Fig. 7) show the well-known bimodality of the steady-state probability distribution of species X under an appropriate choice of the kinetic parameters, here set as

$$k_1 = 0.03 \quad k_2 = 0.0001 \quad k_3 = 200 \quad k_4 = 3.5,$$

taken from ref. (LCPG08). For a given parameter \bar{O}_X , finite state expansion rewrites the original reactions by introducing auxiliary

$$\llbracket n \rrbracket \xrightarrow{f_1(n)} \llbracket n \rrbracket + X, \quad n = \overline{O}_X \quad (\text{R1.1})$$

$$\begin{aligned} \llbracket n \rrbracket &\xrightarrow{f_1(n)} \llbracket n + 1 \rrbracket, \\ f_1(n) &= \llbracket n \rrbracket k_1 (X + n)(X + n - 1)/2 \end{aligned} \quad 0 \leq n < \overline{O}_X \quad (\text{R1.2})$$

$$\llbracket n \rrbracket \xrightarrow{f_2(n)} \llbracket n - 1 \rrbracket, \quad 3 \leq n \leq \overline{O}_X \quad (\text{R2.1})$$

$$\llbracket n \rrbracket \xrightarrow{f_2(n)} \llbracket n - 1 \rrbracket, \quad 0 < n < 3 \quad (\text{R2.2})$$

$$\begin{aligned} X + \llbracket n \rrbracket &\xrightarrow{f_2(n)} \llbracket n \rrbracket, \\ f_2(n) &= \llbracket n \rrbracket k_2 (X + n)(X + n - 1)(X + n - 2)/6 \end{aligned} \quad n = 0 \quad (\text{R2.3})$$

$$\llbracket n \rrbracket \xrightarrow{f_3(n)} \llbracket n \rrbracket + X, \quad n = \overline{O}_X \quad (\text{R3.1})$$

$$\begin{aligned} \llbracket n \rrbracket &\xrightarrow{f_3(n)} \llbracket n + 1 \rrbracket, \\ f_3(n) &= \llbracket n \rrbracket k_3 \end{aligned} \quad 0 \leq n < \overline{O}_X \quad (\text{R3.2})$$

$$\llbracket n \rrbracket \xrightarrow{f_4(n)} \llbracket n - 1 \rrbracket, \quad 0 < n \leq \overline{O}_X \quad (\text{R4.1})$$

$$\begin{aligned} X + \llbracket n \rrbracket &\xrightarrow{f_4(n)} \llbracket n \rrbracket, \\ f_4(n) &= \llbracket n \rrbracket k_4 (X + n) \end{aligned} \quad n = 0 \quad (\text{R4.2})$$

Figure 8: Expansion of the Schlögl network.

species $\llbracket 0 \rrbracket, \llbracket 1 \rrbracket, \dots, \llbracket \overline{O}_X \rrbracket$, which explicitly track discrete population levels (Fig. 8).

The original species X acts as buffer which collects populations levels that are not explicitly tracked. For example, reaction R1.1 derives from reaction R1 when the autocatalytic formation of a new molecule occurs when the system tracks the discrete state $\llbracket \overline{O}_X \rrbracket$, thus requiring to increase the buffer species X by one element. Similarly, the effect of the degradation reaction R4 is to remove one molecule from the buffer when the system tracks the empty discrete state $\llbracket 0 \rrbracket$ (reaction R4.2). The cou-

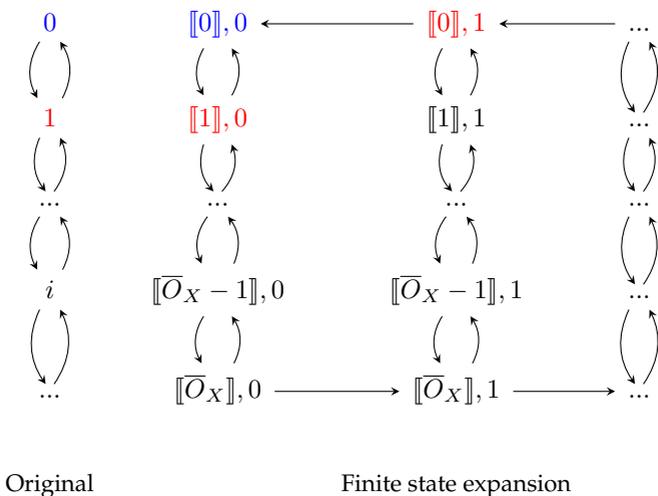


Figure 9: Graphical representation of the Continuous-time Markov chains for the Schlögl model before and after finite state expansion.

pling between the tracked states and X also occurs within the modified propensity functions. For example, even when the system tracks a discrete configuration which does not require buffering (R1.2) the propensity function $f_1(n)$ of the reaction effectively considers an overall kinetics of mass-action type, since the factor $k_1(X+n)(X+n-1)/2$ models the total rate due to number of possible collisions between pairs of $X+n$ indistinguishable molecules. Intuitively, the factor $\llbracket n \rrbracket$ conditions these events to the system tracking n discrete molecules.

The original CTMC is a birth-death Markov process counting the population of X molecules in each state. The state in the CTMC arising from finite state expansion consists of the pair tracked discrete configuration/population level of the buffer species (Fig. 9). The expansion is proven to be stochastically correct in the sense that the probability across all pairs that have the same overall population is preserved (as exemplified by matching colors of the states).

The DRE of the original Schlögl model is a single differential equation. The DRE of the finite state expansion however has $\bar{O}_X + 1$ vari-

$$\begin{aligned} \frac{dX}{dt} &= k_1 X^2/2 - k_2 X^3/6 + k_3 - k_4 X \\ &\text{Original} \\ \frac{d\llbracket n \rrbracket}{dt} &= -\mathbf{I}_{n < \bar{O}_X} \{f_1(n) + f_3(n)\} \\ &\quad - \mathbf{I}_{n > 0} \{f_2(n) + f_4(n)\} \\ &\quad + \mathbf{I}_{1 \leq n \leq \bar{O}_X} f_1(n-1) \\ &\quad + \mathbf{I}_{0 \leq n \leq \bar{O}_X - 1} f_2(n+1) \\ &\quad + \mathbf{I}_{1 \leq n \leq \bar{O}_X} f_3(n-1) \\ &\quad + \mathbf{I}_{0 \leq n \leq \bar{O}_X - 1} f_4(n+1) \\ \frac{dX}{dt} &= f_1(\bar{O}_X) + f_3(\bar{O}_X) - f_2(0) - f_4(0) \\ &\text{Finite state expansion} \end{aligned}$$

Figure 10: Deterministic Rate Equations before and after application of FSE.

ables; each variable $\llbracket n \rrbracket$ can be interpreted as the probability of the system tracking that discrete configuration (Fig. 10). Hence, an estimate of the total mean population is given by the solution $X(t) + \sum_n n \llbracket n \rrbracket(t)$.

The DRE features two equilibrium points owing to the strong (cubic) nonlinearity in the ODEs (BQ10), deterministically converging only to one (VQ09), here at ca. 85.50 (blue line in Fig. 11). The noticeable discrepancy with respect to the true mean (dotted line, computed as the average of 10^4 simulations) has been observed for a long time (ZR91). Finite state expansion achieves excellent agreement with an upper bound $\bar{O}_X = 650$.

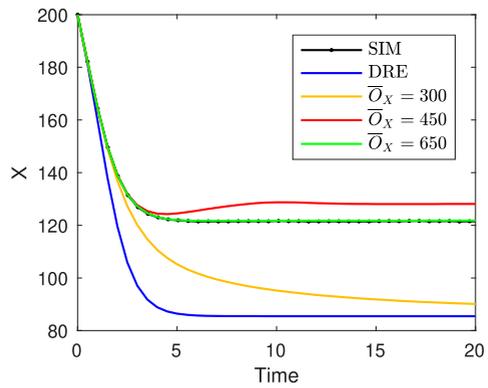


Figure 11: Effects of deterministic approximation on the Schlögl model after applying FSE with different observation bounds, compared to the original DRE solution and the simulation average across 100 000 repetitions

5.1.2 Stochastic equivalence

The previously described translation preserves the stochastic properties of the RN in the sense of ordinary lumpability of Markov chains (Buc94), which guarantees that the solution of the ME is correctly preserved by the expansion. In other words, the ME solution for a state σ in the original RN will exactly correspond to the sum of the ME solutions for all states in the expanded RN that track the same overall population levels. Furthermore, when the RN is fully expanded, i.e., when $\mathcal{O} = \mathbb{N}^{\mathcal{S}}$, we recover the original ME; by construction, instead, when $\mathcal{O} = \emptyset$ then the source and target networks coincide.

We now prove these statements, starting with the preservation of the overall population changes in the expansion.

Lemma 5. *Let $(\mathcal{S}, \mathcal{R})$ be an RN and $(\mathcal{S}_{\overline{\mathcal{O}}}, \mathcal{R}_{\overline{\mathcal{O}}})$ be an expanded RN according to Definition 12. Then, for all $[[o]] + \eta \xrightarrow{f_o} [[o']] + \psi \in \mathcal{R}_{\overline{\mathcal{O}}}[\rho \xrightarrow{f} \pi]$, it holds that:*

1. $(o + \eta) \ominus (o' + \psi) = \rho \ominus \pi$;
2. $(o' + \psi) \ominus (o + \eta) = \pi \ominus \rho$;
3. $\sigma \ominus (o + \eta) + (o' + \psi) = \sigma \ominus \rho + \pi$, for all $\sigma \in \mathbb{N}^{\mathcal{S}}$ such that $(o + \eta) \subseteq \sigma$.

Proof. For case (1):

$$\begin{aligned}
 (o + \eta) \ominus (o' + \psi) &= (o + (\rho \ominus o)) \ominus ((\overline{\mathcal{O}} \cap (o \ominus \rho + \pi) + ((o \ominus \rho + \pi) \ominus \overline{\mathcal{O}})) \\
 &= (o + (\rho \ominus o)) \ominus (o \ominus \rho + \pi) \\
 &= (\rho + (o \ominus \rho)) \ominus ((o \ominus \rho) + \pi) \\
 &= \rho \ominus \pi.
 \end{aligned}$$

For case (2):

$$\begin{aligned}
 (o' + \psi) \ominus (o + \eta) &= ((\overline{\mathcal{O}} \cap (o \ominus \rho + \pi) + ((o \ominus \rho + \pi) \ominus \overline{\mathcal{O}})) \ominus (o + (\rho \ominus o)) \\
 &= (o \ominus \rho + \pi) \ominus (o + (\rho \ominus o)) \\
 &= ((o \ominus \rho) + \pi) \ominus (\rho + (o \ominus \rho)) \\
 &= \pi \ominus \rho.
 \end{aligned}$$

For case (3): We prove this with the following multiset operations:

$$\begin{aligned}
\sigma \ominus \rho + \pi &= \sigma \ominus (\rho \ominus \pi) + (\pi \ominus \rho) \\
&= \sigma \ominus ((o + \eta) \ominus (o' + \psi)) + (o' + \psi) \ominus (o + \eta) \\
&= \sigma \ominus ((o + \eta) \ominus ((o + \eta) \cap (o' + \psi))) + (o' + \psi) \\
&\quad \ominus ((o + \eta) \cap (o' + \psi)) \tag{5.2}
\end{aligned}$$

$$\begin{aligned}
&= \sigma \ominus (o + \eta) + ((o + \eta) \cap (o' + \psi)) + (o' + \psi) \\
&\quad \ominus ((o + \eta) \cap (o' + \psi)) \tag{5.3}
\end{aligned}$$

$$\begin{aligned}
&= \sigma \ominus (o + \eta) + (o' + \psi) + ((o + \eta) \cap (o' + \psi)) \\
&\quad \ominus ((o + \eta) \cap (o' + \psi)) \\
&= \sigma \ominus (o + \eta) + (o' + \psi),
\end{aligned}$$

where Eq. 5.3 follows from Eq. 5.2 because of the relations:

$$(o + \eta) \subseteq \sigma \quad \text{and} \quad (o + \eta) \supseteq ((o + \eta) \cap (o' + \psi)) \subseteq (o' + \psi).$$

□

The correspondence between the original RN and the expanded one is proven in the sense of ordinary lumpability (Buc94). In particular, we prove that the probability of a state in the original RN follows the same law as the sum of the probabilities of all states in the expanded RN which track the same discrete state configuration.

Theorem 6. *Let the RN $(\mathcal{S}, \mathcal{R})$ be the source of the expanded RN $(\mathcal{S}_{\overline{O}}, \mathcal{R}_{\overline{O}})$. Then it holds that*

$$\sum_{o+\xi=\sigma} \hat{p}_{\llbracket o \rrbracket + \xi}(0) = p_{\sigma}(0) \implies \sum_{o+\xi=\sigma} \hat{p}_{\llbracket o \rrbracket + \xi}(t) = p_{\sigma}(t), \quad \text{for all } t.$$

Proof. We prove the following equivalence for the derivatives of the solutions of the respective master equations

$$\sum_{o+\xi=\sigma} \frac{d\hat{p}_{\llbracket o \rrbracket + \xi}}{dt} = \frac{dp_{\sigma}}{dt} \quad \text{for all } \sigma \in \mathbb{N}^{\mathcal{S}},$$

from which the statement holds under the assumption of consistent initial conditions. Using Def. 5, the equation on the next page concludes the proof.

$$\begin{aligned}
\sum_{o+\xi=\sigma} \frac{d\hat{p}_{[o]+\xi}}{dt} &= \sum_{o+\xi=\sigma} \sum_{([\epsilon]+\eta) \xrightarrow{f_\epsilon} ([o']+\psi) \in \mathcal{R}_{\overline{\mathcal{O}}}} \left(f_\epsilon([o] + \xi + [\epsilon] + \eta - [o'] - \psi) \hat{p}_{[o]+\xi+[\epsilon]+\eta-[o']-\psi} - f_\epsilon([o] + \xi) \cdot \hat{p}_{[o]+\xi} \right) \\
&= \sum_{o+\xi=\sigma} \left(\sum_{([\epsilon]+\eta) \xrightarrow{f_\epsilon} ([o]+\psi) \in \mathcal{R}_{\overline{\mathcal{O}}}} f_\epsilon([\epsilon] + \xi + \eta - \psi) \cdot \hat{p}_{[\epsilon]+\xi+\eta-\psi} - \sum_{([o]+\eta) \xrightarrow{f_o} ([o']+\psi) \in \mathcal{R}_{\overline{\mathcal{O}}}} f_o([o] + \xi) \cdot \hat{p}_{[o]+\xi} \right) \\
&= \sum_{([\epsilon]+\eta) \xrightarrow{f_\epsilon} ([o]+\psi) \in \mathcal{R}_{\overline{\mathcal{O}}}} f_\epsilon([\epsilon] + \xi + \eta - \psi) \cdot \hat{p}_{[\epsilon]+\xi+\eta-\psi} - \sum_{([o]+\eta) \xrightarrow{f_o} ([o']+\psi) \in \mathcal{R}_{\overline{\mathcal{O}}}} f_o([o] + \xi) \cdot \hat{p}_{[o]+\xi} \\
&= \sum_{([\epsilon]+\eta) \xrightarrow{f_\epsilon} ([o]+\psi) \in \mathcal{R}_{\overline{\mathcal{O}}}} f(\epsilon + \xi + \eta - \psi) \cdot \underbrace{([\epsilon] + \xi + \eta - \psi)_{[\epsilon]}}_{=1} \cdot \hat{p}_{[\epsilon]+\xi+\eta-\psi} - \sum_{([o]+\eta) \xrightarrow{f_o} ([o']+\psi) \in \mathcal{R}_{\overline{\mathcal{O}}}} f(o + \xi) \cdot \underbrace{([o] + \xi)_{[o]}}_{=1} \cdot \hat{p}_{[o]+\xi} \\
&= \sum_{\substack{[\epsilon]+\eta \xrightarrow{f_\epsilon} ([o]+\psi) \in \mathcal{R}_{\overline{\mathcal{O}}} \\ \epsilon+\xi+\eta-\psi=\sigma-(o+\psi)+(\epsilon+\eta)}} f(\epsilon + \xi + \eta - \psi) \cdot \hat{p}_{[\epsilon]+\xi+\eta-\psi} - \sum_{\substack{([o]+\eta) \xrightarrow{f_o} ([o']+\psi) \in \mathcal{R}_{\overline{\mathcal{O}}} \\ o+\xi=\sigma}} f(\sigma) \cdot \hat{p}_{[o]+\xi} \\
&= \sum_{\substack{\rho \xrightarrow{f} \pi \in \mathcal{R} \\ \epsilon+\xi+\eta-\psi=\sigma-\pi+\rho}} f(\sigma - \pi + \rho) \cdot p_{\sigma-\pi+\rho} - \sum_{\substack{\rho \xrightarrow{f} \pi \in \mathcal{R} \\ o+\xi=\sigma}} f(\sigma) \cdot p_\sigma \\
&= \sum_{\rho \xrightarrow{f} \pi \in \mathcal{R}} f(\sigma - \pi + \rho) \cdot p_{\sigma-\pi+\rho} - \sum_{\rho \xrightarrow{f} \pi \in \mathcal{R}} f(\sigma) \cdot p_\sigma \\
&= \sum_{\rho \xrightarrow{f} \pi \in \mathcal{R}} \left(f(\sigma - \pi + \rho) \cdot p_{\sigma-\pi+\rho} - f(\sigma) \cdot p_\sigma \right) = \frac{dp_\sigma}{dt} \quad \square
\end{aligned}$$

Although the stochastic behavior of the source RN and any expansion are equivalent in this specific sense, their respective DREs may not. The target RN has $|\mathcal{O}| + |\mathcal{S}|$ variables, and the corresponding DRE can be seen as an interpolation of the ME and the original DRE.

Now, we show that when $\mathcal{O} = \mathbb{N}^{\mathcal{S}}$, corresponding to an infinite FSE bound \bar{O} , the DRE solution of an expanded RN corresponds to the ME.

To this end, we need to first establish two contributing features of FSE.

We begin with the conservation of the well-defined property.

Proposition 5. *The expansion of a well-defined RN is well-defined.*

Proof. Let $(\mathcal{S}, \mathcal{R})$ be a well-defined RN, $\rho \xrightarrow{f} \pi$ one of its reactions, $[[o]] + \eta \xrightarrow{f_o} [[o']] + \psi \in \mathcal{R}_{\bar{O}} \left[\rho \xrightarrow{f} \pi \right]$, and $z \in \mathbb{N}^{\mathcal{S}\bar{O}}$.

If $([[o]] + \eta) \not\subseteq z$, this is either because z does not contain the auxiliary species $[[o]]$, or because it lacks some element of η . Assuming $[[o]] \notin z$, the case is simple: The rate function of the expanded RN is by definition zero, keeping with the requirement for it being well-defined. If $\eta \not\subseteq z$, it is necessary to use the definition of η as $\rho \ominus o$ in the expansion algorithm. Next, we can reduce z to $z_{|\mathcal{S}}$, because $\rho, o \in \mathbb{N}^{\mathcal{S}}$.

$$\eta \not\subseteq z \Leftrightarrow (\rho \ominus o) \not\subseteq z \Leftrightarrow (\rho \ominus o) \not\subseteq z_{|\mathcal{S}}$$

Now, it is possible to draw conclusions about the rate function:

$$\begin{aligned} (\rho \ominus o) \not\subseteq z_{|\mathcal{S}} &\Rightarrow \rho \not\subseteq (o + z_{|\mathcal{S}}) && \text{with } (\mathcal{S}, \mathcal{R}) \text{ well-defined} \\ &\Rightarrow f(o + z_{|\mathcal{S}}) = 0 \\ &\Rightarrow f_o(z) = 0 \end{aligned}$$

The proof is thereby closed. □

The following lemma states that if the FSE bound \bar{O} is high enough to contain the products of a reaction, no elements are added to the buffer pool.

Lemma 7. *Let $(\mathcal{S}_{\bar{O}}, \mathcal{R}_{\bar{O}})$ be an expanded RN generated from the source RN $(\mathcal{S}, \mathcal{R})$. For all reactions $[[o]] + \xi \xrightarrow{f_o} [[o']] + \psi \in \mathcal{R}_{\bar{O}}$ that are generated from a reaction $\rho \xrightarrow{f} \pi \in \mathcal{R}$ it holds that if $o \ominus \rho + \pi \subseteq \bar{O}$ then $\psi = \emptyset$.*

Now we are ready to prove the following theorem.

Theorem 8. Consider a well-defined RN $(\mathcal{S}, \mathcal{R})$ and let $(\mathcal{S}_{\overline{\mathcal{O}}}, \mathcal{R}_{\overline{\mathcal{O}}})$ be its expansion so that $\mathcal{O} = \mathbb{N}^{\mathcal{S}}$. Let $X(t)$ be the DRE solution of the expanded RN and $p(t)$ the ME solution of the original RN at time t . Then it holds that:

$$\begin{aligned} X_{|\mathcal{S}}(0) &= \emptyset \text{ and } X_{\llbracket o \rrbracket}(0) = p_o(0) \quad \forall o \in \mathbb{N}^{\mathcal{S}} \\ \implies X_{|\mathcal{S}}(t) &= \emptyset \text{ and } X_{\llbracket o \rrbracket}(t) = p_o(t) \quad \forall o \in \mathbb{N}^{\mathcal{S}}, \quad \text{for all time points } t. \end{aligned}$$

Proof. Similarly to Theorem 6, this statement holds if

$$\begin{aligned} X_{|\mathcal{S}} &= \emptyset \text{ and } X_{\llbracket o \rrbracket} = p_o \quad \forall o \in \mathbb{N}^{\mathcal{S}} \\ \implies \frac{dX_S}{dt} &= 0 \quad \forall S \in \mathcal{S} \text{ and } \frac{dX_{\llbracket o \rrbracket}}{dt} = \frac{dp_o}{dt} \quad \forall o \in \mathbb{N}^{\mathcal{S}}. \end{aligned}$$

Using Definition 6, the DRE for the expanded RN can be written as follows:

$$\frac{dX_{S'}}{dt} = \sum_{\rho \xrightarrow{f_o} \pi \in \mathcal{R}_{\overline{\mathcal{O}}}} (\pi_{S'} - \rho_{S'}) \cdot f_o(X), \quad \text{for all } S' \in \mathcal{S}_{\overline{\mathcal{O}}}. \quad (5.4)$$

We now distinguish between two cases concerning the DRE for the original and expanded species, respectively.

- Case $S' \equiv S \in \mathcal{S}$.

The condition $\mathcal{O} = \mathbb{N}^{\mathcal{S}}$ implies that $\sigma \subseteq \overline{\mathcal{O}} \quad \forall \sigma \in \mathbb{N}^{\mathcal{S}}$. It follows from Lemma 7 that $\pi_S = 0$ in Eq. 5.4 for all reactions $\rho \xrightarrow{f_o} \pi \in \mathcal{R}_{\overline{\mathcal{O}}}$. Accordingly, we note that $\frac{dX_S}{dt} \neq 0$ only if both $\rho_S \neq 0$ and $f_o(X) \neq 0$ for a $\rho \xrightarrow{f_o} \pi \in \mathcal{R}_{\overline{\mathcal{O}}}$. Let $\rho \xrightarrow{f_o} \pi \in \mathcal{R}_{\overline{\mathcal{O}}}$ be a reaction generated from $\rho' \xrightarrow{f} \pi' \in \mathcal{R}$ so that $\rho_S \neq 0$ and $f_o(X) \neq 0$ for an X with $X_{|\mathcal{S}} = \emptyset$. By Definition 12, for a given auxiliary species $\llbracket o \rrbracket$ with $o \in \mathcal{O}$, the propensity function is in the form $f_o(X) = X_{\llbracket o \rrbracket} \cdot f(o + X_{|\mathcal{S}})$. Since $X_{|\mathcal{S}} = \emptyset$, this reduces to $f_o(X) = X_{\llbracket o \rrbracket} \cdot f(o)$. Since the RN $(\mathcal{S}, \mathcal{R})$ is well-defined, $f(o) > 0$ implies that $\rho' \subseteq o$. In this case, ρ must be in the form $\rho = \llbracket o \rrbracket + \emptyset$, i.e. $\rho_S = 0 \quad \forall S \in \mathcal{S}$, closing this case by contradiction.

- Case $S' \equiv \llbracket o \rrbracket \in \mathcal{S}_{\overline{\sigma}} \setminus \mathcal{S}$.

Then, from Definition 12, the DRE can be written as:

$$\begin{aligned} \frac{dX_{\llbracket o \rrbracket}}{dt} &= \sum_{(\llbracket \epsilon \rrbracket + \rho) \xrightarrow{f_\epsilon} (\llbracket o \rrbracket + \xi) \in \mathcal{R}_{\overline{\sigma}}} f_\epsilon(X) - \sum_{(\llbracket o \rrbracket + \xi) \xrightarrow{f_o} (\llbracket o' \rrbracket + \pi) \in \mathcal{R}_{\overline{\sigma}}} f_o(X) \\ &= \sum_{(\llbracket \epsilon \rrbracket + \rho) \xrightarrow{f_\epsilon} (\llbracket o \rrbracket + \xi) \in \mathcal{R}_{\overline{\sigma}}} X_{\llbracket \epsilon \rrbracket} \cdot f(\epsilon + X_{|\mathcal{S}}) - \sum_{(\llbracket o \rrbracket + \xi) \xrightarrow{f_o} (\llbracket o' \rrbracket + \pi) \in \mathcal{R}_{\overline{\sigma}}} X_{\llbracket o \rrbracket} \cdot f(o + X_{|\mathcal{S}}). \end{aligned}$$

As $X_{|\mathcal{S}} = \emptyset$ and the expanded RN is well-defined by Theorem 5, this simplifies to:

$$\begin{aligned} \frac{dX_{\llbracket o \rrbracket}}{dt} &= \sum_{(\llbracket \epsilon \rrbracket + \theta) \xrightarrow{f_\epsilon} (\llbracket o \rrbracket + \theta) \in \mathcal{R}_{\overline{\sigma}}} X_{\llbracket \epsilon \rrbracket} \cdot f(\epsilon + \emptyset) - \sum_{(\llbracket o \rrbracket + \theta) \xrightarrow{f_o} (\llbracket o' \rrbracket + \theta) \in \mathcal{R}_{\overline{\sigma}}} X_{\llbracket o \rrbracket} \cdot f(o + \emptyset) \\ &= \sum_{(\llbracket \epsilon \rrbracket + \theta) \xrightarrow{f_\epsilon} (\llbracket o \rrbracket + \theta) \in \mathcal{R}_{\overline{\sigma}}} X_{\llbracket \epsilon \rrbracket} \cdot f(\epsilon) - \sum_{(\llbracket o \rrbracket + \theta) \xrightarrow{f_o} (\llbracket o' \rrbracket + \theta) \in \mathcal{R}_{\overline{\sigma}}} X_{\llbracket o \rrbracket} \cdot f(o). \end{aligned}$$

Using Definition 12, it follows that:

$$\begin{aligned} \frac{dX_{\llbracket o \rrbracket}}{dt} &= \sum_{(\llbracket \epsilon \rrbracket + \theta) \xrightarrow{f_\epsilon} (\llbracket o \rrbracket + \theta) \in \mathcal{R}_{\overline{\sigma}}} X_{\llbracket \epsilon \rrbracket} \cdot f(\epsilon) - \sum_{(\llbracket o \rrbracket + \theta) \xrightarrow{f_o} (\llbracket o' \rrbracket + \theta) \in \mathcal{R}_{\overline{\sigma}}} X_{\llbracket o \rrbracket} \cdot f(o) \\ &= \sum_{\substack{\rho \xrightarrow{f} \pi \in \mathcal{R} \\ o = \epsilon \ominus \rho + \pi}} X_{\llbracket \epsilon \rrbracket} \cdot f(\epsilon) - \sum_{\substack{\rho \xrightarrow{f} \pi \in \mathcal{R} \\ o' = o \ominus \rho + \pi}} X_{\llbracket o \rrbracket} \cdot f(o) \\ &= \sum_{\substack{\rho \xrightarrow{f} \pi \in \mathcal{R} \\ \epsilon = o + \rho - \pi}} X_{\llbracket \epsilon \rrbracket} \cdot f(\epsilon) - \sum_{\rho \xrightarrow{f} \pi \in \mathcal{R}} X_{\llbracket o \rrbracket} \cdot f(o) \\ &= \sum_{\rho \xrightarrow{f} \pi \in \mathcal{R}} \left(f(o - \pi + \rho) \cdot X_{\llbracket o - \pi + \rho \rrbracket} - f(o) \cdot X_{\llbracket o \rrbracket} \right) \\ &\quad \text{with } X_{\llbracket \sigma \rrbracket} = p_\sigma \quad \forall \sigma \in \mathbb{N}^{\mathcal{S}} \\ &= \sum_{\rho \xrightarrow{f} \pi \in \mathcal{R}} \left(f(o - \pi + \rho) \cdot p_{o - \pi + \rho} - f(o) \cdot p_o \right) \\ &= \frac{dp_o}{dt}. \end{aligned} \quad \square$$

5.2 Applications

Through various examples discussed next, we show how FSE can achieve increased accuracy of the mean estimates with modest expansions.

Throughout all the figures' captions, the following abbreviations are used:

- **CMC**: method of conditional moments (HWKT14), a method that uses the discrete representation of selected species to improve the moment-based approximation of the others;
- **DRE**: deterministic rate equations;
- **EMRE**: effective mesoscopic rate equation (Gri10), a method based on linear noise approximation with some additional corrective terms;
- **FSE**: finite state expansion;
- **FSP**: finite state projection (MK06), a method that restricts the state space by replacing low-probability states with a single sink state;
- **MCA**: moment-closure approximation, in particular second-order low-dispersion moment closure (KFR⁺16; SSG15), in which variance and covariance are the highest observed moments and all higher-order central moments are set to zero (in all models considered in this paper, computing approximations with higher-order moments did not improve the accuracy);
- **SIM**: stochastic simulation using Gillespie's direct method (Gil77).

Several of the following figures report analyses of the sensitivity of the accuracy of FSE with respect to the choice of the observation bound. For this, we consider the notion of percentage error of the FSE estimate against the average value obtained by simulation (SIM) as follows:

$$\text{Error \%} = \left| \frac{\text{FSE} - \text{SIM}}{\text{SIM}} \right| \cdot 100. \quad (5.5)$$

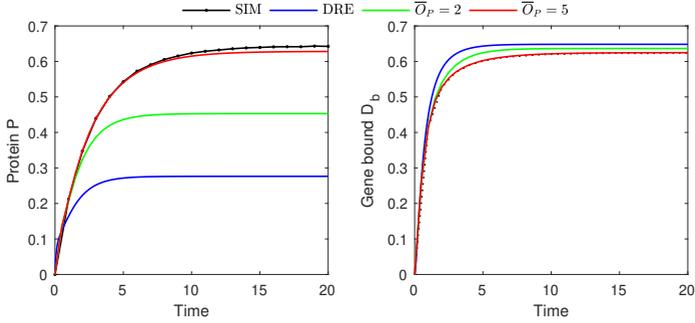
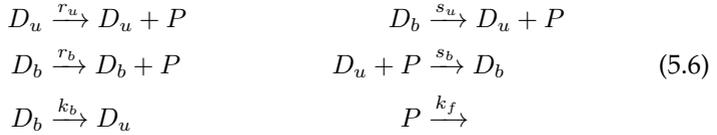


Figure 12: Numerical simulations of the genetic feedback switch in scheme (5.6) comparing stochastic simulation, DRE and finite state expansions for fixed $\overline{O}_{D_u} = \overline{O}_{D_b} = 1$ and different upper bounds \overline{O}_P . The resulting DRE from finite state expansion has $2 \cdot \overline{O}_P + 2$ equations.

5.2.1 Genetic Feedback Switch

We first consider a chemical reaction network for a genetic feedback switch taken from refs. (HSI⁺05; GSN12):



Species D_u and D_b represent the state of a single gene when its promoter region is unbound (respectively, bound) to a protein P . The reaction propensities obey the law of mass action through the kinetic parameters r_u , s_u , r_b , s_b , k_b , and k_f . This is a basic model for negative autoregulation, a well-known motif appearing in more than 40% of the known transcription factors in *E.coli* (SOMMA02).

The values for these kinetic parameters were set as follows in the examples shown here:

$$r_u = 1.0 \quad r_b = 0.5 \quad k_f = 0.1 \quad k_b = 1.0 \quad s_b = 10.0 \quad s_u = 0.5$$

The initial state is the configuration $(P, D_u, D_b) = (0, 1, 0)$, so a natural choice of upper bounds for the gene species is $\overline{O}_{D_u} = \overline{O}_{D_b} = 1$,

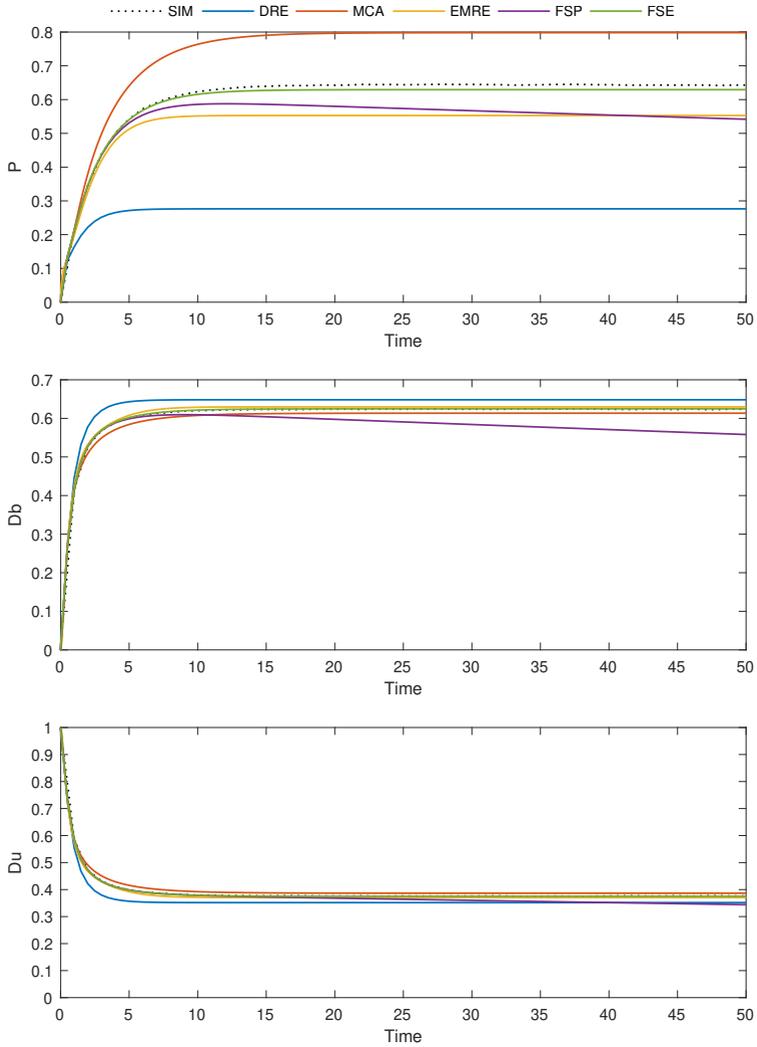


Figure 13: Numerical evaluation of the gene feedback switch model from Eq. 5.6. The average across $1E+6$ repetitions of stochastic simulation is compared against DRE, MCA, EMRE, FSP, and FSE. For FSE and FSP the following bounds were used: $\overline{O}_{Db} = \overline{O}_{Du} = 1$ and $\overline{O}_P = 5$.

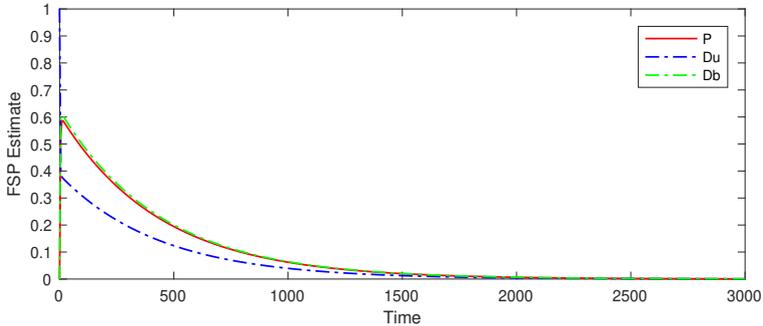


Figure 14: The longer-term behavior of FSP on the gene feedback switch model, using the same settings as in Fig. 13.

by which the DRE arising from finite state expansion can be interpreted as the solution of the conditional expectation of the protein population based on the gene state. Small values of \overline{O}_P may achieve a significant correction of the protein levels as well as of the marginal probability distribution of the gene state (Fig. 12).

As seen in Fig. 13, the largest difference in the results of different approximation methods occurs in the estimation of species P . All methods except FSP can be seen to reach their steady state within the first 20 time points and successfully approximate the populations of D_u and D_b very closely. At a bound of $\overline{O}_P = 5$, FSE predicts the population of P better than the two moment-based methods. The behavior of FSP can be observed better over a longer time period (Fig. 14): With the same observed state space as FSE, FSP causes the probability mass to absorb into the sink state, leading to estimates for the average populations that vanish for long time horizons.

In this model, we can observe a monotonic decrease of the percentage error of FSP as defined in Equation 5.5 with increasing observation bound (Fig. 15).

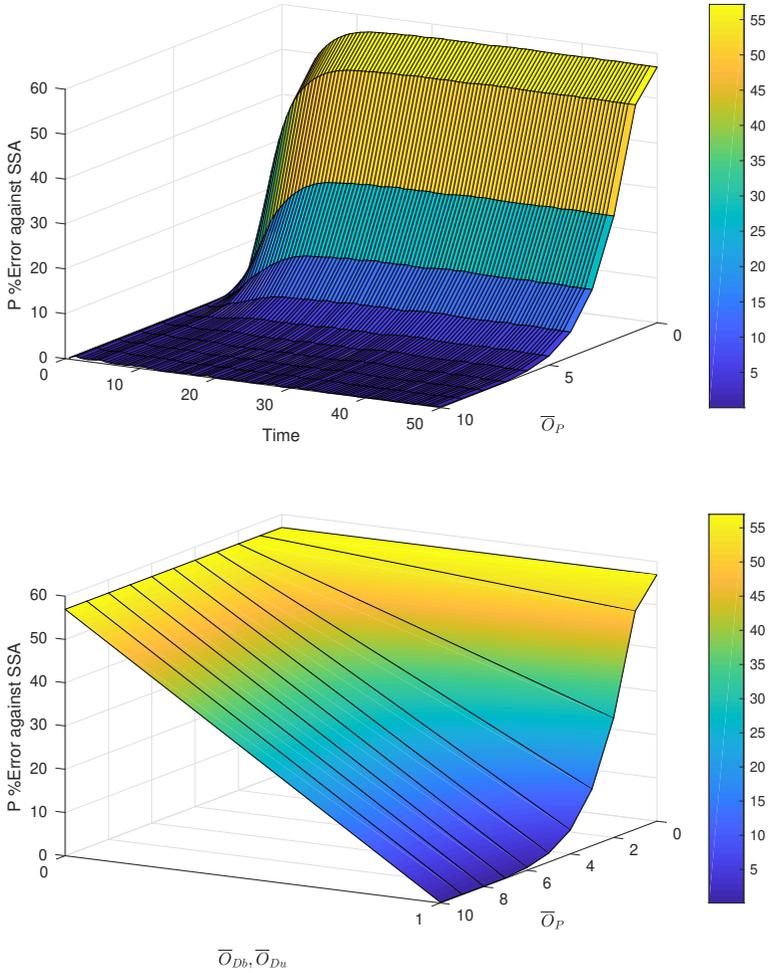
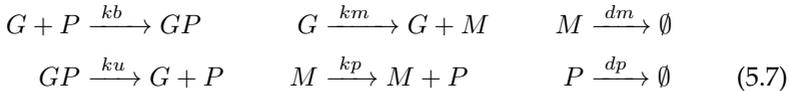


Figure 15: (Top) Monotonic behavior of the time-dependent accuracy of FSE against stochastic simulation (1 000 000 repetitions) of the genetic feedback switch model (Eq. 5.6) for increasing observation bounds of \bar{O}_P , fixing $\bar{O}_{D_b} = \bar{O}_{D_u} = 1$. (Bottom) Error behavior in the steady state (estimated at time point $t = 50$) shows a significant impact of explicitly tracking the discrete states D_b, D_u of the gene.

5.2.2 Gene regulatory system with inhibition feedback loop

In a second reaction network for a gene regulatory system with inhibition feedback loop taken from ref. (WS17), the production of the protein is mediated through a buffer species M , representing the mRNA in the biological interpretation.



As before, reaction propensities obey the law of mass action. Kinetic parameters were chosen as follows:

$$\begin{array}{lll}
 kb = 0.6 & km = 1.0 & dm = 0.0001 \\
 ku = 0.001 & kp = 0.005 & dp = 0.005
 \end{array}$$

Although this model looks very similar to the previous one at first glance, its dynamic behavior is significantly more complex: Whereas the previously described genetic feedback switch balances out with some stochastic variation around its mean, this version exhibits some multimodality.

Looking at Fig. 16, the dynamics of species M and P are predicted by both FSE and EMRE with very high accuracy, however EMRE is slightly less accurate on species G . MCA is not far off the mark either, but cannot deliver truly accurate results for either of the species. FSP has the same issue of probability mass absorbing into the sink state as previously, when run with bounds that are sufficient for FSE. Finally, while this model looks like it could be a prime candidate for CMC, this is not actually the case. In the most successful attempt, species G , GP and P were observed stochastically using an upper bound of 4 for species P . This was the largest bound for which CMC returned results at all; with larger bounds the CMC implementation did not return results within 6h. Additionally, CMC could not provide results past time point 73.073.

While the error still goes against zero for large enough observation bounds across the board, it is in this case not generally decreasing mono-

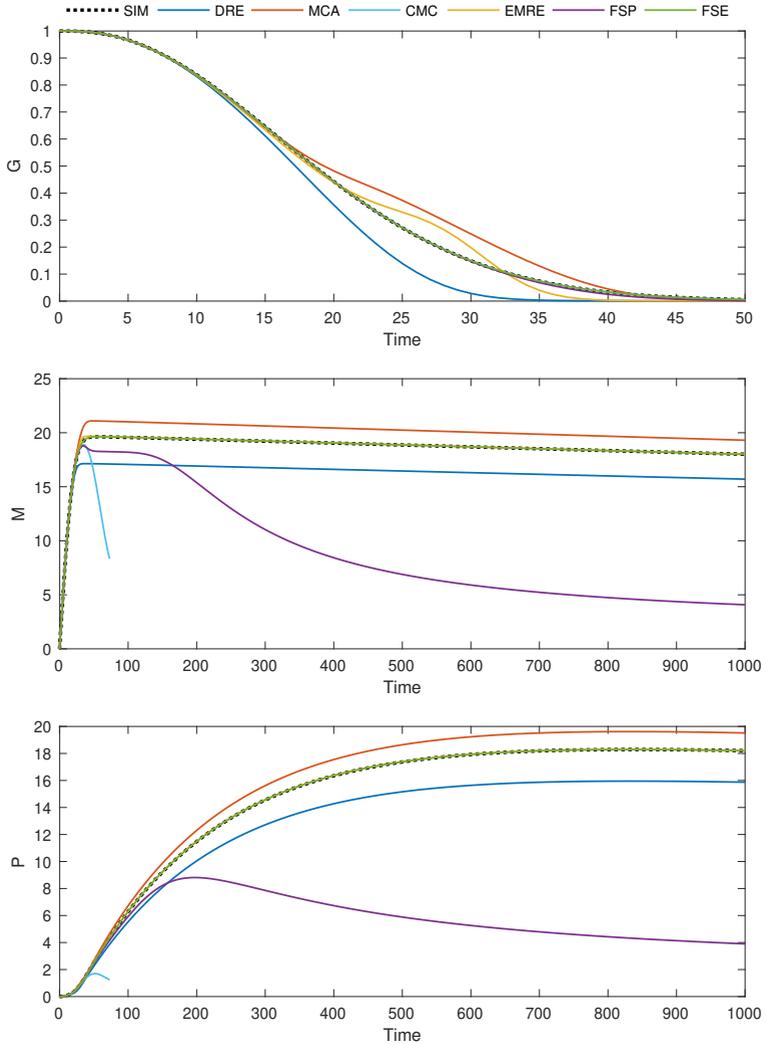


Figure 16: Numerical evaluation of the inhibition feedback loop model from Eq. 5.7. Observation bounds for both FSP and FSE were set as follows: $\overline{O}_G = \overline{O}_{GP} = 1$, $\overline{O}_M = 40$ and $\overline{O}_P = 20$. For CMC, species G , GP and P were observed stochastically using an upper bound of 4 for species P . The top plot considers only the first 50 time units because the dynamics of the G species is faster than in the other subplots. Average from stochastic simulation was computed over $2E+6$ independent repetitions.

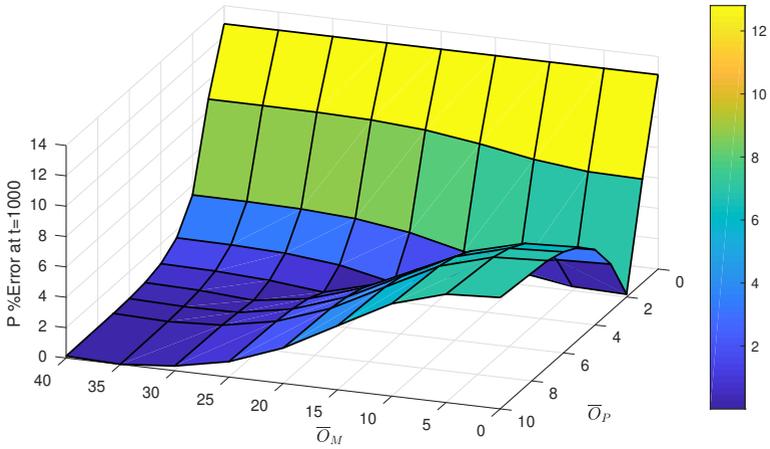


Figure 17: Non-monotonic behavior of the accuracy of FSE against stochastic simulation (20 000 repetitions) of the inhibition feedback loop model from Eq. 5.7 for increasing observation bounds \bar{O}_M and \bar{O}_P at representative time point $t = 1000$. The bounds \bar{O}_G and \bar{O}_{GP} were fixed at 1.

tonically (Fig. 17). A monotonic improvement can only be observed when raising one observation bound while the other is already sufficiently high.

5.2.3 Genetic Toggle Switch

The toggle switch network is a fundamental regulatory system of two mutually repressing genes (GCC00). Models of toggle-switch networks are mathematically challenging because of multimodality (TB06; TPG14) and stochastic noise due to the species such as mRNA present in low molecular abundances (KEBC05). Here we study the reaction scheme analyzed in ref. (HGK15), consisting of a mass-action version of the RN

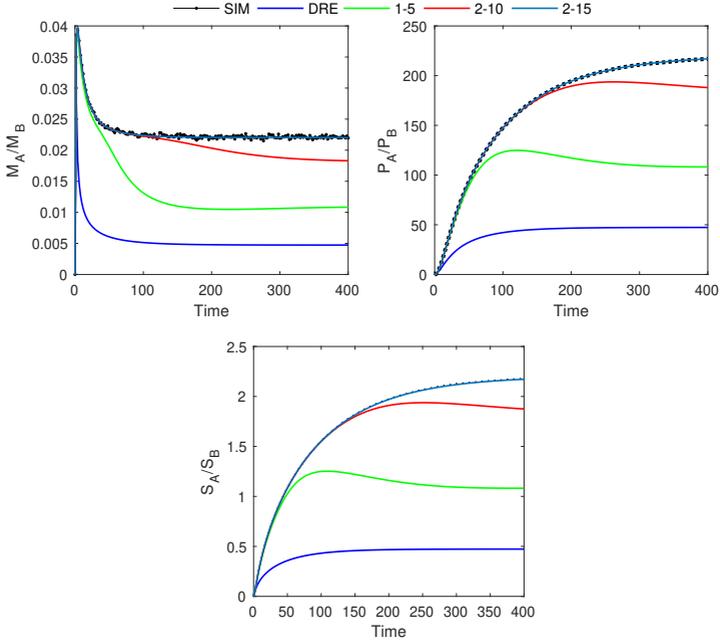
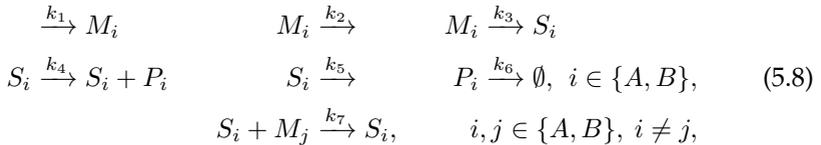


Figure 18: Numerical simulations of the genetic toggle switch in scheme (5.8) comparing stochastic simulation, DRE and finite state expansions fixing $\bar{O}_{P_A} = \bar{O}_{P_B} = 0$ while using different upper bounds \bar{O}_M and \bar{O}_S ($\bar{O}_M - \bar{O}_S$ in short) for the number of copies of M_A/M_B and S_A/S_B (as indicated in the legend), respectively. Initial condition was the zero state.

presented in ref. (GCC00):



where M_i and S_i denote the precursor mRNA and the mRNA for target protein P_i . The last two reactions model mutual inhibition by means of a precursor of one protein repressing the mRNA of the other.

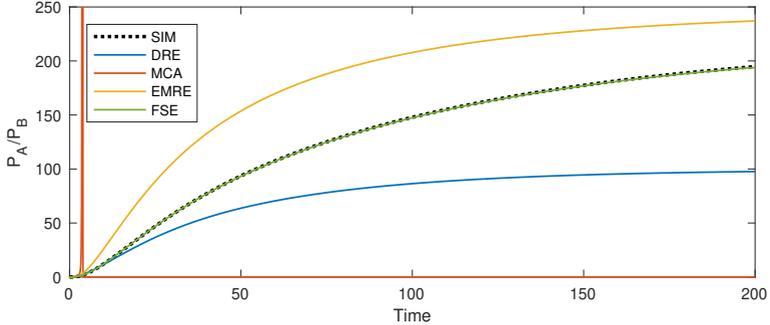


Figure 19: Comparison of stochastic simulation (500 000 repetitions), DRE, MCA, EMRE, and FSE for species P_A/P_B of the genetic toggle switch model (Eq. 5.8). FSE is run with upper bounds $\bar{O}_P = 0$, $\bar{O}_M = 2$ and $\bar{O}_S = 10$. The MCA estimates population levels approaching 75 000 (out of scale in this plot to improve readability) before dropping to zero, while the EMRE approximation overestimates the average dynamics.

Kinetic parameters were chosen as follows:

$$\begin{aligned}
 k_1 &= 0.05 & k_2 &= 0.1 & k_3 &= 1.0 & k_4 &= 10.0 \\
 k_5 &= 0.01 & k_6 &= 0.1 & k_7 &= 20.0
 \end{aligned}$$

When protein production is controlled by low populations of precursor mRNA, the stochastic fluctuations are not adequately approximated with DRE. By explicitly observing few copies of mRNA (up to tens) our method provides precise estimates of the time courses of the mean populations (Fig. 18). To preserve the symmetry, it makes sense to choose identical observation bounds for the pairs of analogous species M_A/M_B and S_A/S_B . For species P_A and P_B , being only involved in a simple decay reaction, the DRE approximation is sufficient and FSE observation bounds can be set to zero. The ODE system size for the tested choices of upper bounds is equal to $(\bar{O}_M+1)^2 \cdot (\bar{O}_S+1)^2 + 6$. The resulting equations, of size at most 2310 in our tests, can be analyzed effectively, as opposed to time-consuming stochastic simulations using hybrid approaches such as those reported in ref. (HGK15).

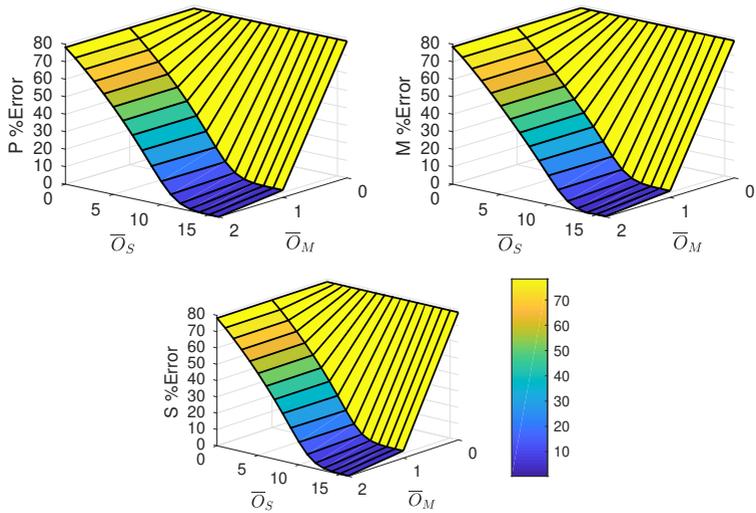


Figure 20: Monotonic behavior of the accuracy of FSE against stochastic simulation (500 000 repetitions) of the genetic toggle switch model (Eq. 5.8) for increasing observation bounds of \bar{O}_M and \bar{O}_S at representative time point $t = 400$, fixing $\bar{O}_P = 0$.

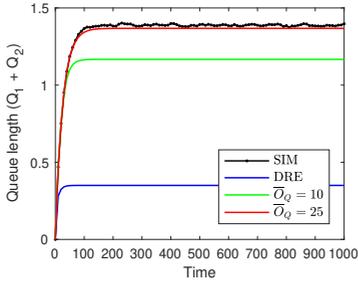
5.2.4 Queuing systems

Queuing systems are an established model of resource contention for the quantitative analysis of computer and communication systems (Kle75; BGdMT05). They can be represented as an RN that tracks the population of clients that are waiting for service at a queuing center. There has been considerable interest in the development of DRE-type approximations for queuing networks (BHL13), which has also stimulated a line of research into the development of methods to correct the mean estimates (Gas17).

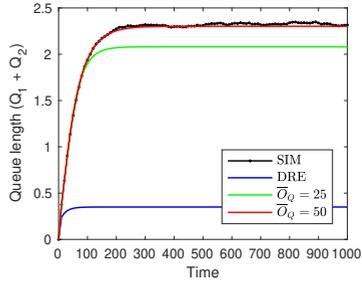
Here we consider a simple queuing system for an exogenous Poissonian arrival of clients (with rate λ) at a queuing station that serves jobs according to a two-stage Coxian distribution, which has three degrees of freedom hereafter identified by parameters p , μ_1 , and μ_2 . This allows us testing the finite state expansion method against service times characterized by different variance, following the observation that this is a crucial statistics affecting the accuracy of the DRE approximation, e.g., ref. (Gri10). In particular, from ref. (BGdMT05) we have that a Coxian service-time distribution with mean E , variance V , and squared coefficient of variation $V/E^2 \geq 1/2$ can be obtained by setting $p = E^2/2V$, $\mu_1 = 2/E$ and $\mu_2 = E/V$. The RN for the queuing system is

$$\begin{aligned}
 & \xrightarrow{\lambda} Q_1 \\
 Q_1 + S_1 & \xrightarrow{(1-p)\mu_1 \min(Q_1, S_1)} S_1 \\
 Q_1 + S_1 & \xrightarrow{p\mu_1 \min(Q_1, S_1)} Q_2 + S_2 \\
 Q_2 + S_2 & \xrightarrow{\mu_2 \min(Q_2, S_2)} S_1
 \end{aligned} \tag{5.9}$$

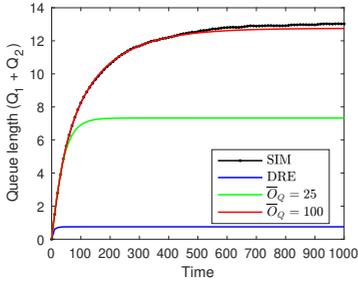
where Q_1 tracks the population of jobs that are either waiting for service or are in the first stage of the Coxian-distributed service; Q_2 tracks the population of jobs in the second stage of service; S_1 and S_2 indicate the population of servers that are in the first or second stage, respectively. Initializing the system with k elements in class S_1 (and zero elements otherwise) models a queuing system with k parallel independent servers. From Equation [5.9], it can be seen that any state of the RN is such that



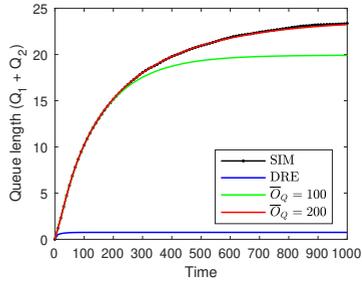
(a) $SCV = 10, \lambda = 0.35$



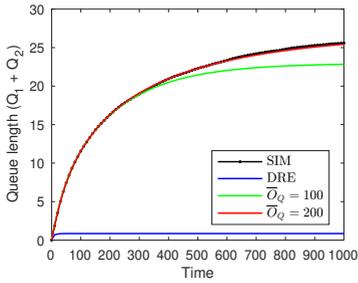
(b) $SCV = 20, \lambda = 0.35$



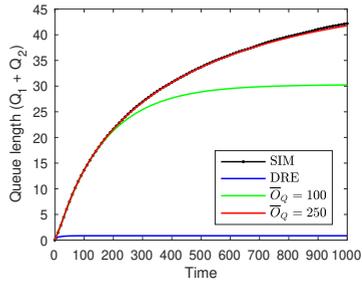
(c) $SCV = 10, \lambda = 0.75$



(d) $SCV = 20, \lambda = 0.75$



(e) $SCV = 10, \lambda = 0.85$



(f) $SCV = 20, \lambda = 0.85$

Figure 21: Finite state expansion applied to the queueing network in Equation [5.9] for varying exogenous arrival rates λ and service-time distributions with the same mean 1.0 and squared coefficients of variation (SCV) equal to 10.0 and 20.0, with an initial state containing only a single server S_1 .

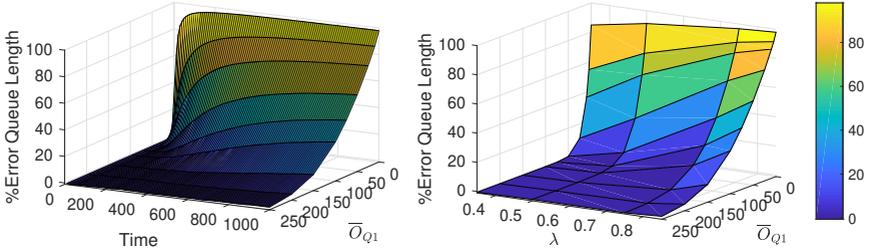


Figure 22: (Left) Monotonic behavior of the time-dependent accuracy of FSE against stochastic simulation (200 000 repetitions) of the queuing network model with squared coefficient of variation $SCV = 20$ and arrival rate $\lambda = 0.85$ for increasing observation bounds of \bar{O}_{Q_1} , fixing $\bar{O}_{Q_2} = \bar{O}_{S_1} = \bar{O}_{S_2} = 1$. The y -axis plots the queue length, given by the total populations of species Q_1 and Q_2 . (Right) Error behavior in the steady state of the queue length (estimated at time point $t = 1000$) as a function of λ and \bar{O}_{Q_1} .

the sum of the populations of S_1 and S_2 is equal to k , while Q_1 may grow unboundedly.

The average queue length is the expected population of jobs that are in the queuing system, given by the sum of the populations of species Q_1 and Q_2 . In the model of Equation [5.9] such metric is affected by the squared coefficient of variation of the service time distribution as well as the rate of arrivals of jobs λ . DREs provide rather inaccurate estimates of the average queue lengths for service time distributions with large squared coefficient of variations. By contrast, finite state expansion may significantly improve the approximation.

Figure 21 shows the effect of finite state expansion in this case for varying exogenous arrival rates λ and service-time distributions with the same mean 1.0 and squared coefficients of variation (SCV) equal to 10.0 and 20.0, with an initial state containing only a single server S_1 . The corresponding parameter values are:

$$\begin{array}{llll}
 SCV = 10 : & p = 1.3123 & \mu_1 = 0.0244 & \mu_2 = 0.9942 \\
 SCV = 20 : & p = 12.3869 & \mu_1 = 0.1696 & \mu_2 = 0.8441
 \end{array}$$

As predicted, the DRE provides a poor approximation of the average queue length of the system. Furthermore, confirming an early result for

this class of queueing networks (TT13), the DRE estimate in equilibrium is insensitive to the SCV, whereas the ground-truth average queue length computed by simulation ($2 \cdot 10^5$ individual traces) grows proportionally with the variance of the service-time distribution. In all cases, finite state expansion can provide very accurate corrections of the mean estimate by observing configurations with an upper bound \bar{O}_{Q_1} on the number of Q_1 jobs and $\bar{O}_S = 1$ on species S_1, S_2 and Q_2 . The size of the ODE system of the expanded RN is $(\bar{O}_{Q_1} + 1) \cdot (\bar{O}_S + 1)$.

5.2.5 Schlögl

The conversion of the Schlögl model is already described in detail in Section 5.1.1, but examining the behavior of its expansions more closely helps with the direct comparison to the other models presented here. For convenience of reference, the model definition is repeated:



To start with, parameters are chosen as previously:

$$k_1 = 0.03 \quad k_2 = 0.0001 \quad k_3 = 200 \quad k_4 = 3.5 \quad (5.10)$$

Fig. 23 shows the trouble MCA and EMRE have with this model: MCA and DRE converge towards the lower equilibrium point, while EMRE pushes the mean prediction towards the higher stable point. Only FSE provides high accuracy at all time points with an observation bound of $\bar{O}_X = 650$. FSP could not be applied to a model with ternary reactions.

In the sensitivity analysis in Fig. 24, it can however be seen that FSE is affected by the complexities of this model as well: despite only containing one species and without interactions between several observation bounds, the error behavior with increasing observation bound is non-monotonic.

Finally, we examine the model with a second set of parameter values:

$$k_1 = 0.03 \quad k_2 = 0.0004 \quad k_3 = 2620 \quad k_4 = 11. \quad (5.11)$$

With these parameters, the DRE has two stable equilibrium points at ca. 543 and 940. With the given initial conditions $X(0) = 700$ the DRE converges to the lower equilibrium point. More importantly, the raised lower attractor basin means that typically a population level of at least several hundred X is maintained at all time points when running stochastic simulations.

Fig. 25 shows high accuracy of the mean predictions from $\bar{O}_X = 600$ upwards (with respect to stochastic simulation over 100 000 repetitions) when significant mass is accumulated in the buffer species. For $\bar{O}_X = 1300$ most of the probability mass is kept in the discrete state space, as shown by the low population values of the buffer species X , providing excellent agreement with stochastic simulations.

The inaccuracies in the initial phase occur for all observation levels that have a significant population of the buffer species X . Moreover, looking at the development of the buffer species population, it becomes obvious that the inaccuracies coincide with the time needed for the buffer species population to stabilize.

This can be taken further by starting the system in an initial state with a population of the buffer species close to its steady state for the used observation bound. Experimentally, this starkly heightens the initial accuracy already for $\bar{O}_X = 600$ (Fig. 26). This shows that there is potential for improving FSE results further if a reliable procedure can be developed to optimize the distribution of elements in the initial state. Alternatively, a method of testing the reliability of results may be to vary initial state distributions and see how much they agree. Even without having simulation results to compare to, in the case of the schlögl model, this would at least allow you to conclude that errors in the result of the initial phase are likely bigger than in the steady state.

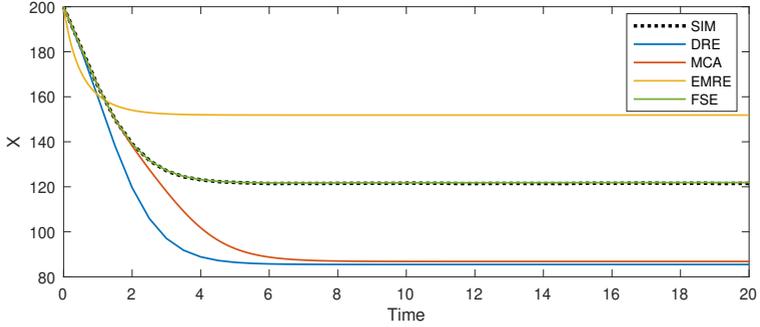


Figure 23: Numerical evaluation of the Schlögl model with parameter set (5.10), comparing stochastic simulation (100 000 repetitions), DRE, MCA, EMRE and FSE with $\bar{O}_X = 650$.

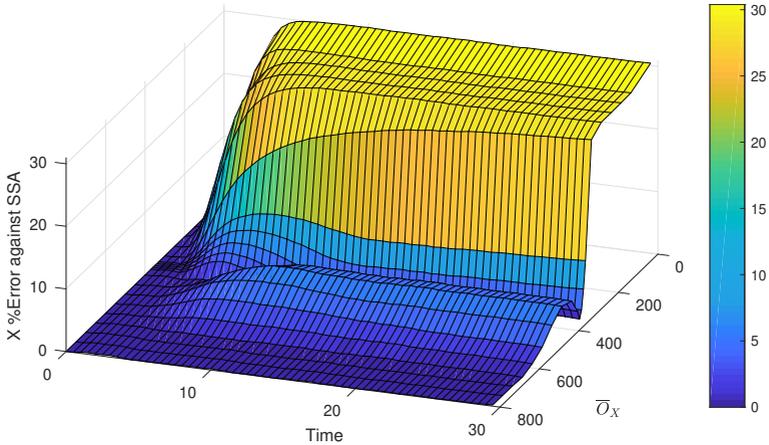


Figure 24: Non-monotonic behavior of the accuracy of FSE against stochastic simulation (100 000 repetitions) in the Schlögl model with parameter set (5.10) for increasing observation bound \bar{O}_X .

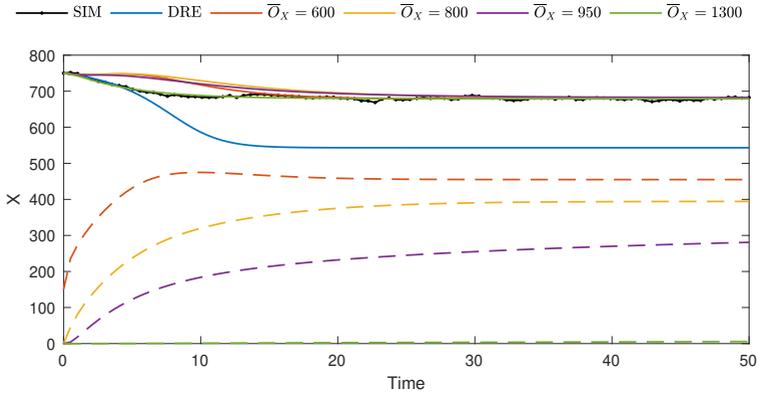


Figure 25: FSE results on the Schlögl model with parameter set (5.11) using different observation bounds. Dashed lines indicate the prediction of the buffer species X in the FSE solution with observation bound of matching color in the legend.

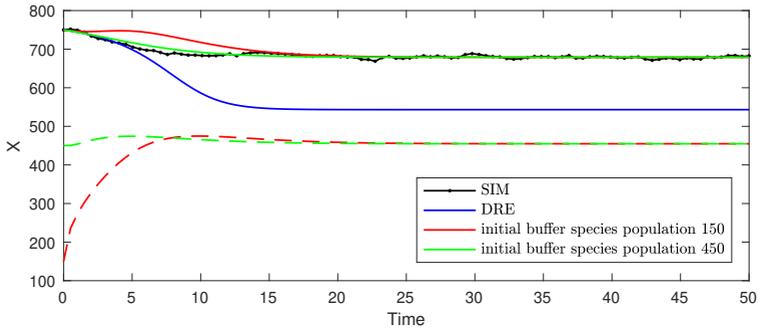


Figure 26: Analogous to Figure 25, this graph shows the improvement in accuracy of the same system at $\bar{O}_X = 600$, when the distribution of elements between the auxiliary and buffer species in the initial state is adjusted while keeping the total initial population at a constant $750X$.

5.3 Discussion

We have presented a novel analytical method for improving the prediction of stochastic dynamics in reaction networks using approximate techniques. This problem can be tackled using several techniques available in the literature, but they rest on different assumptions that hinder their generality and their effectiveness in practice. Here we have used representative examples to show the applicability of our method in models that are challenging for state-of-the-art techniques. The Schlögl model is known to stress moment-closure approximations because of their reported difficulties with multimodal distributions (LAKS15; SSG14) (also shown in Fig. 23). The toggle switch network additionally features low-abundance species, which return physically meaningless moment-closure estimates (SH11), also reported in Fig. 19. When species can be partitioned into low- and high-abundance classes, the method of conditional moments has been proposed as a hybrid analytical technique that maintains a discrete representation of the former class and a moment-based approximation of the latter (HWKT14). The gene feedback switch is a prototypical example since species D_u and D_b represent the distinct binary state of a single gene, hence they represent the natural members for the low-abundance class. On this model, however, the method could not return valid results as early as time point 0.36, using an available implementation (KFR⁺16); the gene regulatory model with an inhibition feedback loop from ref. (WS17) showed similar difficulties (Fig. 16), overall confirming the numerical issues discussed in ref. (SSG17).

The queuing network model has elucidated how the intrinsic noise in system (here realized by service-time distributions with increasing variances) has a profound effect on the accuracy of DRE. Most important, it has demonstrated the universality of our method, which is in principle applicable to reaction networks with arbitrary, non-differentiable propensity functions such as the minimum function in Equation [5.9], which hinder the use of techniques that require smoothness (GUV07; Gri10; AKS13).

Differently from approaches like the system size expansion (Van07),

finite state expansion does not rely on perturbation arguments around a limiting regime. The effective mesoscopic rate equation adds mean-correction terms to the linear-noise approximation under the assumption of an underlying Gaussian process (Gri10). Such approximation is asymptotically correct in the limit of infinite populations, but it may be inaccurate for finite populations. In practice, this may lead to less accurate mean estimates than finite state expansion in our case studies (Fig. 13, 16, 19, 23). Defining incoming and outgoing transitions with respect to the buffer species maintained in the expanded RN represents a crucial difference with the related method of finite state projection, where unobserved configurations are collapsed into a sink state into which the probability mass absorbs (MK06). Experimentally, this results in increased accuracy of mean estimates by finite state expansion when tracking the same subset of the state space in both methods (Fig. 13, 14, 16).

On the other hand, the solution by finite state projection is proved to be a lower bound on the true probability distribution, with the further property that increasing the set of observed configurations tightens the bounds. Instead, while finite state expansion ensures stochastic correctness, it does not give theoretical guarantees on the degree of accuracy of the approximate DRE estimates, nor does it guarantee monotonically increasing accuracy with larger tracked state spaces. Indeed, experimentally we confirmed that the approximation error is not monotonic in general, for instance in models with multi-stability such as the Schögl model (Fig. 24) and the gene regulatory model from ref. (WS17) (Fig. 17). However, we found excellent accuracy when the observed state space is large enough, both during the transient evolution and in the steady state (Fig. 15, 17, 20, 22, 24). This makes finite state expansion a useful tool to tame the problem of state explosion when dealing with a fully discrete representation of RNs.

Chapter 6

Conclusion

Stochasticity is a key tool to understand a variety of phenomena regarding the dynamics of reaction networks, but the capability of exactly analyzing complex models escapes us due to the lack of analytical solutions and the high computational cost of numerical simulations in general. This thesis has presented three methods for the analysis of stochastic systems, using model transformations to either lump or expand the original model.

In mass-action reaction networks, Species Equivalence enables aggregation in the sense of Markov chain lumping by identifying structural properties on the set of reactions, without the need of the costly enumeration of the state space. Owing to the polynomial space and time complexity of the reduction algorithm, it can be seen as a universal pre-processing step that exactly preserves the stochastic dynamics of species of interest to the modeler.

The tool DiffLQN allows the application of deterministic rate equations to layered queuing networks in order to compute approximate estimates of steady-state performance metrics. DiffLQN makes it possible to carry out more extensive analyses of the error behavior of the differential analysis with respect to the ground-truth stochastic simulation as well as to alternative analytical techniques based on mean value analysis. Without the automated support offered by DiffLQN, these studies cannot but

be performed manually on selected model instances, as has been done in the literature (Tri10; Tri11; Tri13), necessarily limiting their scope of validity.

Finite state expansion increases the size of a reaction network while maintaining stochastically equivalent dynamics, in order to improve the accuracy of approximations of expected values of the population processes. It interpolates between the microscopic view of the master equation and the macroscopic view of the DRE by explicitly tracking a subset of the original discrete configurations and consistently coupling their dynamics with deterministic variables acting as buffers for the probability mass. The method effectively improves the predictions of deterministic approximations in situations when the stochasticity in the model cannot be neglected, such as in networks with considerable intrinsic noise, multi-scale species populations, and multi-stability.

Optimisation of the split between macroscopically and microscopically observed elements in the initial state can potentially further improve results and presents an opportunity for future research.

In conclusion both lumping and expanding approaches have shown their usefulness. Lumping is an efficient method for model simplification, but can only be used on models that are lumpable, i.e. have a redundant structure that allows the merging of nodes with no loss of relevant information. Expansion on the other side can be used on any model of the correct type without further prerequisites. Its benefits - access to or increased accuracy of further analysis methods - are however counterweighed by the increased complexity that comes with growing model size. It is thus necessary to strike the correct balance between simplicity and exactness when using expansion methods.

An interesting subject for future work is the combination of both approaches, e.g. applying a lumping method to a model before expanding it, to help limit the increase in complexity to the necessary minimum.

References

- [ACK10] David Anderson, Gheorghe Craciun, and Thomas Kurtz. Product-form stationary distributions for deficiency zero chemical reaction networks. *Bulletin of mathematical biology*, 72:1947–70, 03 2010. 8
- [AKS13] Angeliqe Ale, Paul Kirk, and Michael P. H. Stumpf. A general moment expansion method for stochastic kinetic models. *The Journal of Chemical Physics*, 138(17):174101, 2013. 5, 11, 79
- [AP88] Uri M. Ascher and Linda R. Petzold. *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. SIAM, 1988. 40
- [BCSV04] Paolo Boldi, Bruno Codenotti, Massimo Santini, and Sebastiano Vigna. UbiCrawler: A scalable fully distributed web crawler. *Software: Practice & Experience*, 34(8):711–726, 2004. 35
- [BGdMT05] Gunter Bolch, Stefan Greiner, Hermann de Meer, and Kishor Trivedi. *Queueing networks and Markov chains: modeling and performance evaluation with computer science applications*. Wiley, 2005. 72
- [BHLM13] Luca Bortolussi, Jane Hillston, Diego Latella, and Mieke Massink. Continuous approximation of collective system behaviour: A tutorial. *Performance Evaluation*, 70(5):317–349, 2013. 72
- [BM12] Andrew J. Black and Alan J. McKane. Stochastic formulation of ecological models and their applications. *Trends in Ecology & Evolution*, 27(6):337 – 345, 2012. 9
- [BPSDGA04] Marián Boguñá, Romualdo Pastor-Satorras, Albert Díaz-Guilera, and Alex Arenas. Models of social networks based on social distance attachment. *Phys. Rev. E*, 70(5):056122, 2004. 35

- [BQ10] Lisa M. Bishop and Hong Qian. Stochastic bistability and bifurcation in a mesoscopic signaling system with autocatalytic kinase. *Biophysical Journal*, 98(1):1–11, 2010. 53
- [Buc94] Peter Buchholz. Exact and Ordinary Lumpability in Finite Markov Chains. *Journal of Applied Probability*, 31(1):59–75, 1994. 2, 11, 12, 55, 56
- [Car14] Carleton University Software Performance Research Group. Layered queueing research resource page. <http://www.layeredqueues.org/>, August 2014. 37, 39
- [CFL09] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. *Reviews of modern physics*, 81(2):591, 2009. 33
- [CLG⁺12] Ayse Koca Caydasi, Maiko Lohel, Gerd Grünert, Peter Dittrich, Gislene Pereira, and Bashar Ibrahim. A dynamical model of the spindle position checkpoint. *Molecular systems biology*, 8(1):582, 2012. 32
- [CN82] K. Mani Chandy and Doug Neuse. Linearizer: A heuristic algorithm for queueing network models of computing systems. *Commun. ACM*, 25(2):126–134, 1982. 4
- [CPVT⁺21] Luca Cardelli, Isabel Cristina Pérez-Verona, Mirco Tribastone, Max Tschaikowski, Andrea Vandin, and Tabea Waizmann. Exact maximal reduction of stochastic reaction networks by species lumping. *Bioinformatics*, 02 2021. btab081. xiv, 30, 33
- [CTTV15] L. Cardelli, M. Tribastone, M. Tschaikowski, and A. Vandin. Forward and backward bisimulations for chemical reaction networks. In *CONCUR*, pages 226–239, 2015. 30
- [CTTV17a] Luca Cardelli, Mirco Tribastone, Max Tschaikowski, and Andrea Vandin. Maximal aggregation of polynomial dynamical systems. *Proceedings of the National Academy of Sciences*, 114(38):10029–10034, 2017. 20
- [CTTV17b] Luca Cardelli, Mirco Tribastone, Max Tschaikowski, and Andrea Vandin. Syntactic Markovian bisimulation for chemical reaction networks. In *Models, Algorithms, Logics and Tools*, pages 466–483. Springer, 2017. 20, 36
- [CTVT17] Luca Cardelli, Mirco Tribastone, Andrea Vandin, and Max Tschaikowski. ERODE: A tool for the evaluation and reduction

- of ordinary differential equations. In *Tools and Algorithms for the Construction and Analysis of Systems — 23rd International Conference, TACAS, 2017*. 31
- [CVM12] E. Cator and P. Van Mieghem. Second-order mean-field susceptible-infected-susceptible epidemic threshold. *Phys. Rev. E*, 85:056111, May 2012. 35
- [CW16] Daniele Cappelletti and Carsten Wiuf. Elimination of intermediate species in multiscale stochastic reaction networks. *Ann. Appl. Probab.*, 26(5):2915–2958, 10 2016. 36
- [DHS03] S. Derisavi, H. Hermanns, and W.H. Sanders. Optimal state-space lumping in Markov chains. *Inf. Process. Lett.*, 87(6):309–315, 2003. 20, 27
- [Dug06] Adam Duguid. Coping with the parallelism of bittorrent: Conversion of pepa to odes in dealing with state space explosion. In *Proceedings of the 4th International Conference on Formal Modeling and Analysis of Timed Systems, FORMATS’06*, pages 156–170, Berlin, Heidelberg, 2006. Springer-Verlag. 16
- [ECLB83] R. W. Eash, K. S. Chon, Y. J. Lee, and D. E. Boyce. Equilibrium traffic assignment on an aggregated highway network for sketch planning. *Transportation Research Record*, 994:30–37, 1983. 35
- [EL88] Derek L. Eager and John N. Lipscomb. The amva priority approximation. *Perform. Eval.*, 8(3):173–194, June 1988. 45
- [ELSS02] Michael B. Elowitz, Arnold J. Levine, Eric D. Siggia, and Peter S. Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002. 2
- [Eng06] Stefan Engblom. Computing the moments of high dimensional solutions of the master equation. *Appl. Math. Comput.*, 180:498–515, 2006. 11
- [FAOW⁺09] Greg Franks, Tariq Al-Omari, Murray Woodside, Olivia Das, and Salem Derisavi. Enhanced modeling and solution of layered queueing networks. *IEEE Trans. Softw. Eng.*, 35(2):148–161, March 2009. 3, 12
- [FH74] Martin Feinberg and Friedrich JM Horn. Dynamics of open chemical systems and the algebraic structure of the underlying reaction network. *Chemical Engineering Science*, 29(3):775–787, 1974. 9

- [FMW⁺13] Greg Franks, Peter Maly, Murray Woodside, Dorina C. Petriu, Alex Hubbard, and Martin Mroz. Layered queueing network solver and simulator user manual. <http://www.sce.carleton.ca/rads/lqns>, 2013. 12
- [Gas17] Nicolas Gast. Expected values estimated via mean-field approximation are $1/n$ -accurate. *POMACS*, 1(1):17:1–17:26, 2017. 72
- [GCC00] Timothy S. Gardner, Charles R. Cantor, and James J. Collins. Construction of a genetic toggle switch in *escherichia coli*. *Nature*, 403(6767):339–342, 01 2000. 68, 69
- [Gil77] D.T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, December 1977. 1, 2, 9, 20, 39, 61
- [Gil07] Daniel T. Gillespie. Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*, 58(1):35–55, 2007. 1
- [Gil09] C.S. Gillespie. Moment-closure approximations for mass-action models. *IET Systems Biology*, 3(1):52–58, january 2009. 5
- [GJ13] J. Goutsias and G. Jenkinson. Markovian dynamics on complex reaction networks. *Physics Reports*, 529(2):199–264, 2013. 1
- [Gri10] Ramon Grima. An effective rate equation approach to reaction kinetics in small volumes: Theory and application to biochemical reactions in nonequilibrium steady-state conditions. *The Journal of Chemical Physics*, 133(3):035101, 2010. 4, 5, 61, 72, 79, 80
- [GSN12] R. Grima, D. R. Schmidt, and T. J. Newman. Steady-state fluctuations of a genetic feedback loop: An exact solution. *The Journal of Chemical Physics*, 137(3):035104, 2012. 62
- [GTD⁺14] Stephen Gilmore, Mirco Tribastone, Adam Duguid, Allan Clark, and Michael Smith. PEPA plug-in project webpage. <http://www.dcs.ed.ac.uk/pepa/tools/plugin/index.html>, August 2014. 16
- [Gun05] Jeremy Gunawardena. Multisite protein phosphorylation makes a good threshold but can be a poor switch. *Proceedings of the National Academy of Sciences of the United States of America*, 102(41):14617–14622, 2005. 31
- [Gup95] Purnananda Guptasarma. Does replication-induced transcription regulate synthesis of the myriad low copy number proteins of *escherichia coli*? *BioEssays*, 17(11):987–997, 1995. 2

- [GUV07] Carlos A. Gómez-Uribe and George C. Verghese. Mass fluctuation kinetics: Capturing stochastic effects in systems of chemical reactions through coupled mean-variance computations. *The Journal of Chemical Physics*, 126(2):024109, 2007. 5, 79
- [GUVT08] Carlos A. Gómez-Uribe, George C. Verghese, and Abraham R. Tzafiriri. Enhanced identification and exploitation of time scales for model reduction in stochastic chemical kinetics. *The Journal of Chemical Physics*, 129(24):244112, 2008. 36
- [HGK15] Benjamin Hepp, Ankit Gupta, and Mustafa Khammash. Adaptive hybrid simulations for multiscale stochastic reaction networks. *The Journal of Chemical Physics*, 142(3):034118, 2015. 68, 70
- [Hil96] J. Hillston. *A Compositional Approach to Performance Modelling*. Cambridge University Press, 1996. 16, 38
- [HMW09] Thomas A Henzinger, Maria Mateescu, and Verena Wolf. Sliding window abstraction for infinite Markov chains. In *International Conference on Computer Aided Verification*, pages 337–352. Springer, 2009. 3, 36
- [HSI⁺05] J. E. M. Hornos, D. Schultz, G. C. P. Innocentini, J. Wang, A. M. Walczak, J. N. Onuchic, and P. G. Wolynes. Self-regulating gene: An exact solution. *Phys. Rev. E*, 72:051907, Nov 2005. 62
- [HWKT14] J. Hasenauer, V. Wolf, A. Kazeroonian, and F. J. Theis. Method of conditional moments (MCM) for the chemical master equation. *Journal of Mathematical Biology*, 69(3):687–735, Sep 2014. 5, 61, 79
- [Ins15] Institute of Systems Biology. BioUML documentation. <http://www.biouml.org/index.shtml>, May 2015. 39
- [Jah11] T. Jahnke. On reduced models for the chemical master equation. *Multiscale Modeling & Simulation*, 9(4):1646–1676, 2011. 5
- [KEBC05] Mads Kærn, Timothy C. Elston, William J. Blake, and James J. Collins. Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics*, 6:451 EP –, 05 2005. 68
- [KFR⁺16] Atefeh Kazeroonian, Fabian Fröhlich, Andreas Raue, Fabian J. Theis, and Jan Hasenauer. CERENA: ChEmical REaction Network Analyzer—A Toolbox for the Simulation and Analysis of Stochastic Chemical Kinetics. *PLOS ONE*, 11(1):1–15, 1 2016. 61, 79

- [KK13] Hye-Won Kang and Thomas G. Kurtz. Separation of time-scales and model reduction for stochastic reaction networks. *The Annals of Applied Probability*, 23(2):529–583, 2013. 36
- [Kle75] L. Kleinrock. *Queueing Systems—Theory*, volume 1. Wiley-Interscience, 1975. 72
- [KM27] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A.*, 115(772):700–721, 1927. 33
- [KS76] J.G. Kemeny and J.L. Snell. *Finite Markov Chains*. Springer New York, Heidelberg, Berlin, 1976. 2, 11
- [Kue16] Christian Kuehn. Moment closure—a brief review. In Eckehard Schöll, Sabine H. L. Klapp, and Philipp Hövel, editors, *Control of Self-Organizing Nonlinear Systems*, pages 253–271. Springer International Publishing, 2016. 4
- [Kun13] Jérôme Kunegis. KONECT – The Koblenz Network Collection. In *Proc. Int. Conf. on World Wide Web Companion*, pages 1343–1350, 2013. 35
- [Kur70] Thomas Kurtz. Solutions of ordinary differential equations as limits of pure jump markov processes. *Journal of Applied Probability*, 7:49–58, 04 1970. 11
- [Kur72] T. G. Kurtz. The relationship between stochastic and deterministic models for chemical reactions. *The Journal of Chemical Physics*, 57(7):2976–2978, 1972. 2
- [LAKS15] Eszter Lakatos, Angelique Ale, Paul D. W. Kirk, and Michael P. H. Stumpf. Multivariate moment closure techniques for stochastic kinetic models. *The Journal of Chemical Physics*, 143(9):094107, 2015. 79
- [LB03] Daniel J. Lew and Daniel J. Burke. The spindle assembly and spindle position checkpoints. *Annual Review of Genetics*, 37(1):251–282, 2003. PMID: 14616062. 32
- [LCPG08] Hong Li, Yang Cao, Linda R. Petzold, and Daniel T. Gillespie. Algorithms and software for stochastic simulation of biochemical reacting systems. *Biotechnology Progress*, 24(1):56–61, 2008. 50
- [LDR⁺10] Chen Li, Marco Donizelli, Nicolas Rodriguez, Harish Dharuri, Lukas Endler, Vijayalakshmi Chelliah, Lu Li, Enuo He, Arnaud

- Henry, Melanie I. Stefan, Jacky L. Snoep, Michael Hucka, Nicolas Le Novère, and Camille Laibe. BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology*, 4:92, Jun 2010. 32
- [LKF07] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Trans. Knowledge Discovery from Data*, 1(1):1–40, 2007. 35
- [LKK09] Chang Hyeong Lee, Kyeong-Hun Kim, and Pilwon Kim. A moment closure method for stochastic reaction networks. *The Journal of Chemical Physics*, 130(13):134107, 2009. 5
- [LSC02] John Lisman, Howard Schulman, and Hollis Cline. The molecular basis of CaMKII function in synaptic and behavioural memory. *Nature Reviews Neuroscience*, 3:175–190, 2002. 32
- [LYR12] John Lisman, Ryohei Yasuda, and Sridhar Raghavachari. Mechanisms of CaMKII action in long-term potentiation. *Nature Reviews Neuroscience*, 13:169–182, 2012. 32
- [MF13] Angélica S Mata and Silvio C Ferreira. Pair quenched mean-field theory for the susceptible-infected-susceptible model on complex networks. *EPL (Europhysics Letters)*, 103(4):48003, 2013. 35
- [MK06] Brian Munsky and Mustafa Khammash. The finite state projection algorithm for the solution of the chemical master equation. *The Journal of Chemical Physics*, 124(4):044104, 2006. 1, 3, 36, 47, 61, 80
- [ML12] Julian J. McAuley and Jure Leskovec. Learning to discover social circles in ego networks. In Peter L. Bartlett, Fernando C. N. Pereira, Christopher J. C. Burges, Léon Bottou, and Kilian Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, 2012, Lake Tahoe, Nevada, United States*, pages 548–556, 2012. 35
- [MLSH12] S. Menz, J.C. Latorre, Ch. Schütte, and W Huisinga. Hybrid stochastic-deterministic solution of the chemical master equation. *SIAM Interdisciplinary Journal Multiscale Modeling and Simulation*, 10(4):1232–1262, 2012. 5
- [MT73] Daniel P Maki and Maynard Thompson. *Mathematical models and applications: with emphasis on the social life, and management sciences*. Prentice-Hall Englewood Cliffs, 1973. 33

- [OM98] M. S. Okino and M. L. Mavrouniotis. Simplification of mathematical models of chemical reaction systems. *Chemical Reviews*, 2(98):391–408, 1998. 2
- [Pau05] Johan Paulsson. Models of stochastic gene expression. *Physics of Life Reviews*, 2(2):157–175, 2005. 2
- [PKT90] Krishna R. Pattipati, Michael M. Kostreva, and John L. Teele. Approximate mean value analysis algorithms for queuing networks: Existence, uniqueness, and convergence results. *J. ACM*, 37(3):643–673, 1990. 4
- [PKUMK10] Shirley Pepke, Tamara Kinzer-Ursem, Stefan Mihalas, and Mary B Kennedy. A dynamic model of interactions of ca^{2+} , calmodulin, and catalytic subunits of ca^{2+} /calmodulin-dependent protein kinase ii. *PLOS Computational Biology*, 6(2):e1000675, 2010. 32
- [PSCVMV15] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic processes in complex networks. *Reviews of modern physics*, 87(3):925, 2015. 33, 35
- [PT87] R. Paige and R. Tarjan. Three partition refinement algorithms. *SIAM Journal on Computing*, 16(6):973–989, 1987. 3, 20, 30
- [RA15] Ryan A. Rossi and Nesreen K. Ahmed. The network data repository with interactive graph analytics and visualization. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015. 35
- [RS95] Jerome A. Rolia and Kenneth C. Sevcik. The method of layers. *IEEE Trans. Software Eng.*, 21(8):689–700, 1995. 4
- [RT03] Z. Rózsa and J. Tóth. Exact linear lumping in abstract spaces. *Electronic Journal of Qualitative Theory of Differential Equations [electronic only]*, 2003. 12
- [Sch72] F. Schlögl. Chemical reaction models for non-equilibrium phase transitions. *Zeitschrift für Physik*, 253(2):147–161, Apr 1972. 50
- [SES02] Peter S. Swain, Michael B. Elowitz, and Eric D. Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*, 99(20):12795–12800, 2002. 2

- [SFE11] M. W. Sneddon, J. R. Faeder, and T. Emonet. Efficient modeling, simulation and coarse-graining of biological complexity with NFsim. *Nature Methods*, 8(2):177–183, 2011. 32
- [SH09] Carlos Salazar and Thomas Höfer. Multisite protein phosphorylation – from molecular mechanisms to kinetic models. *FEBS Journal*, 276(12):3177–3198, 2009. ix, 31, 32
- [SH11] A. Singh and J. P. Hespanha. Approximate moment dynamics for chemically reacting systems. *IEEE Transactions on Automatic Control*, 56(2):414–418, 2011. 5, 79
- [SHN09] N. A. Sinitsyn, Nicolas Hengartner, and Ilya Nemenman. Adiabatic coarse-graining and simulations of stochastic biochemical networks. *Proceedings of the National Academy of Sciences*, 106(26):10546–10551, 2009. 36
- [SK11] Vassilios Sotiropoulos and Yiannis N. Kaznessis. Analytical derivation of moment equations in stochastic chemical kinetics. *Chemical Engineering Science*, 66(3):268 – 277, 2011. 5
- [SK13] Patrick Smadbeck and Yiannis N. Kaznessis. A closure scheme for chemical master equations. *Proceedings of the National Academy of Sciences*, 110(35):14261–14265, 2013. 5
- [SOMMA02] Shai S. Shen-Orr, Ron Milo, Shmoolik Mangan, and Uri Alon. Network motifs in the transcriptional regulation network of *escherichia coli*. *Nature Genetics*, 31(1):64–68, 2002. 62
- [SSG14] David Schnoerr, Guido Sanguinetti, and Ramon Grima. Validity conditions for moment closure approximations in stochastic chemical kinetics. *The Journal of Chemical Physics*, 141(8):084103, 2014. 79
- [SSG15] David Schnoerr, Guido Sanguinetti, and Ramon Grima. Comparison of different moment-closure approximations for stochastic chemical kinetics. *The Journal of Chemical Physics*, 143(18):185101, 2015. 11, 61
- [SSG17] David Schnoerr, Guido Sanguinetti, and Ramon Grima. Approximation and inference methods for stochastic biochemical kinetics—a tutorial review. *Journal of Physics A: Mathematical and Theoretical*, 50(9):093001, 2017. 2, 5, 36, 79
- [Ste09] William J. Stewart. *Probability, Markov Chains, Queues, and Simulation*. Princeton University Press, 2009. 40

- [STK11] Péter L Simon, Michael Taylor, and Istvan Z Kiss. Exact epidemic models on graphs using graph-automorphism driven lumping. *Journal of mathematical biology*, 62(4):479–508, 2011. 33, 34
- [TB06] Tianhai Tian and Kevin Burrage. Stochastic models for regulatory networks of the genetic toggle switch. *Proceedings of the National Academy of Sciences*, 103(22):8372–8377, 2006. 68
- [TDG09] M. Tribastone, A. Duguid, and S. Gilmore. The PEPA Eclipse Plugin. *Performance Evaluation Review*, 36(4):28–33, March 2009. 38
- [TDGH12] Mirco Tribastone, Jie Ding, Stephen Gilmore, and Jane Hillston. Fluid rewards for a stochastic process algebra. *IEEE Transactions on Software Engineering*, 38:861–874, 2012. 4, 41, 44
- [TG09] Matthew Thomson and Jeremy Gunawardena. Unlimited multistability in multisite phosphorylation systems. *Nature*, 460(7252):274–277, 07 2009. 31
- [TGH12] Mirco Tribastone, Stephen Gilmore, and Jane Hillston. Scalable differential analysis of process algebra models. *IEEE Transactions on Software Engineering*, 38(1):205–219, 2012. 4
- [TMG12] P. Thomas, H. Matuschek, and R. Grima. Computation of biochemical pathway fluctuations beyond the linear noise approximation using iNA. In *EEE International Conference on Bioinformatics and Biomedicine*, pages 1–5, Oct 2012. 4
- [TPG14] Philipp Thomas, Nikola Popović, and Ramon Grima. Phenotypic switching in gene regulatory networks. *Proceedings of the National Academy of Sciences*, 111(19):6994–6999, 2014. 5, 68
- [Tri10] Mirco Tribastone. Relating layered queueing networks and process algebra models. In *WOSP*, pages 183–194, 2010. 4, 82
- [Tri11] Mirco Tribastone. Approximate mean value analysis of process algebra models. In *MASCOTS*, pages 369–378, 2011. 4, 82
- [Tri13] Mirco Tribastone. A fluid model for layered queueing networks. *IEEE Trans. Software Eng.*, 39(6):744–756, 2013. 4, 13, 14, 18, 38, 39, 44, 45, 82
- [TT13] Max Tschaikowski and Mirco Tribastone. Insensitivity to service-time distributions for fluid queueing models. In *7th International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS)*, 2013. 75

- [Van07] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, 2007. 1, 10, 79
- [VF10] A. Valmari and G. Franceschinis. Simple $O(m \log n)$ time Markov Chain lumping. In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*, pages 38–52, 2010. 20, 27, 36
- [VM11] Piet Van Mieghem. The n -intertwined SIS epidemic network model. *Computing*, 93(2):147–169, Dec 2011. 35
- [VMO15] Eberhard O. Voit, Harald A. Martens, and Stig W. Omholt. 150 years of the mass action law. *PLOS Computational Biology*, 11(1):1–7, 01 2015. 9
- [VQ09] Melissa Vellela and Hong Qian. Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: the schlögl model revisited. *Journal of The Royal Society Interface*, 6(39):925–940, 2009. 53
- [WBVT21] Tabea Waizmann, Luca Bortolussi, Andrea Vandin, and Mirco Tribastone. Improved estimations of stochastic chemical kinetics by finite-state expansion. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 477(2251):20200964, 2021. xiv
- [WGHB09] Pu Wang, Marta C. González, César A. Hidalgo, and Albert-László Barabási. Understanding the spreading patterns of mobile phone viruses. *Science*, 324(5930):1071–1076, 2009. 33
- [WS17] Stefanie Winkelmann and Christof Schütte. Hybrid models for chemical reaction networks: Multiscale theory and application to gene regulatory systems. *The Journal of Chemical Physics*, 147(11):114115, 2017. 66, 79, 80
- [WT16] Tabea Waizmann and Mirco Tribastone. DiffEqn: Differential equation analysis of layered queuing networks. In Alberto Avritzer, Alexandru Iosup, Xiaoyun Zhu, and Steffen Becker, editors, *Companion Publication for ACM/SPEC on International Conference on Performance Engineering, ICPE 2016 Companion, Delft, The Netherlands, March 12–16, 2016*, pages 63–68. ACM, 2016. xiv
- [WTSB17] Wei Wang, Ming Tang, H Eugene Stanley, and Lidia A Braunstein. Unification of theoretical approaches for epidemic spreading on complex networks. *Reports on Progress in Physics*, 80(3):036603, 2017. 9, 34, 35

- [ZL09] R. Zafarani and H. Liu. Social computing data repository at ASU, 2009. 35
- [ZLMZ05] Beichuan Zhang, Raymond Liu, Daniel Massey, and Lixia Zhang. Collecting the Internet AS-level topology. *SIGCOMM Computer Communication Review*, 35(1):53–61, 2005. 35
- [ZNKT07] Xiaolan Zhang, Giovanni Neglia, Jim Kurose, and Don Towsley. Performance modeling of epidemic routing. *Computer Networks*, 51(10):2867 – 2891, 2007. 9
- [ZR91] Qiang Zheng and John Ross. Comparison of deterministic and stochastic kinetics for nonlinear systems. *The Journal of Chemical Physics*, 94(5):3644–3648, 1991. 53



SOME RIGHTS RESERVED



Unless otherwise expressly stated, all original material of whatever nature created by Tabea Waizmann and included in this thesis, is licensed under a Creative Commons Attribution Noncommercial Share Alike 2.5 Italy License.

Check creativecommons.org/licenses/by-nc-sa/2.5/it/ for the legal code of the full license.

Ask the author about other uses.