IMT SCHOOL FOR ADVANCED STUDIES LUCCA

Evaluation and Reduction of ERODE Ordinary Differential Equations

https://www.erode.eu/

MIRCO TRIBASTONE

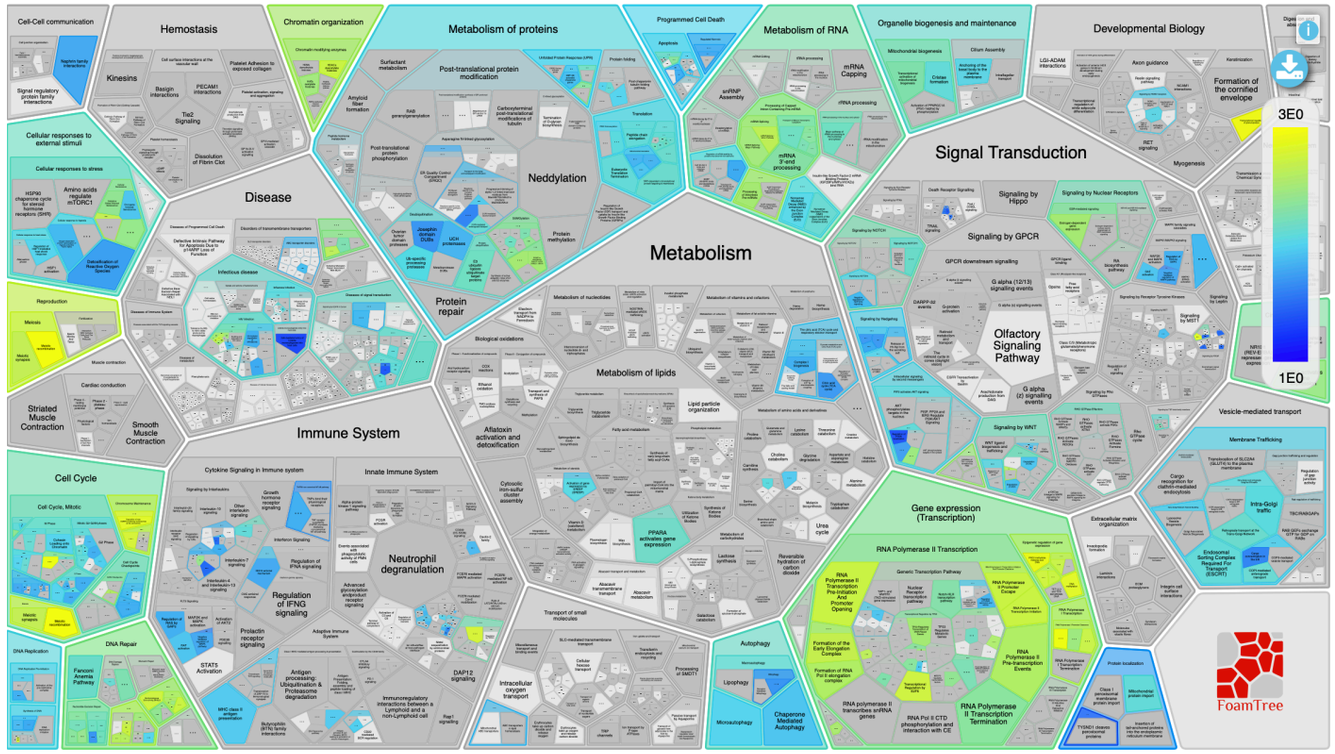
LUMPING AND EXPANDING REACTION NETWORKS

CMSB 2023

14 September 2023

REACTION NETWORKS

- Fundamental model of interaction in many natural and engineering sciences:
  - Biology
  - Chemistry
  - Computer Science
  - Epidemiology
  - …

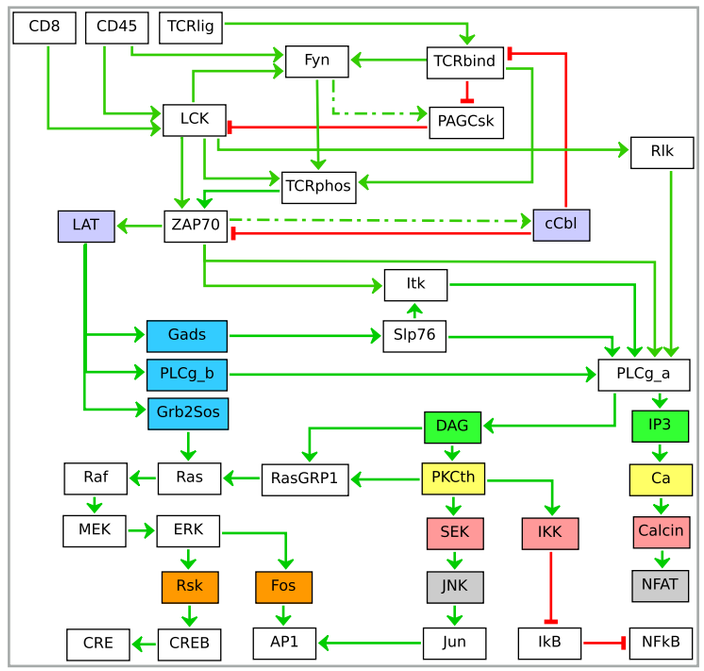


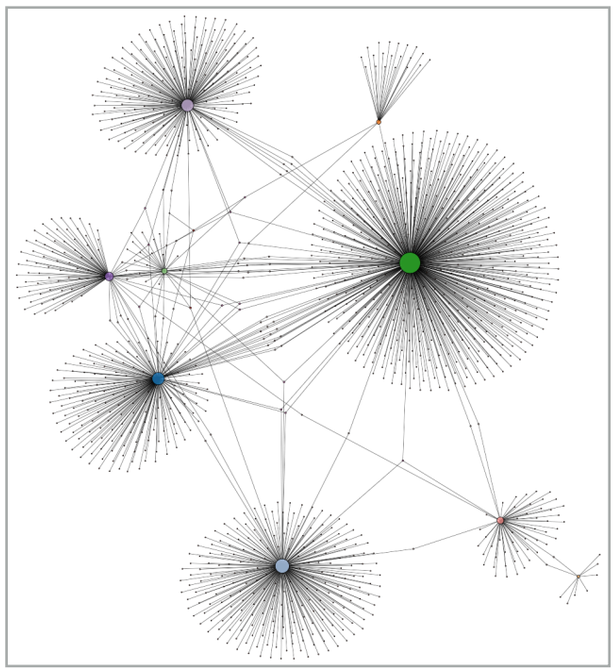
Physics Reports

Volume 529, Issue 2, 10 August 2013, Pages 199-264

Markovian dynamics on complex reaction networks

J. Goutsias, G. Jenkinson



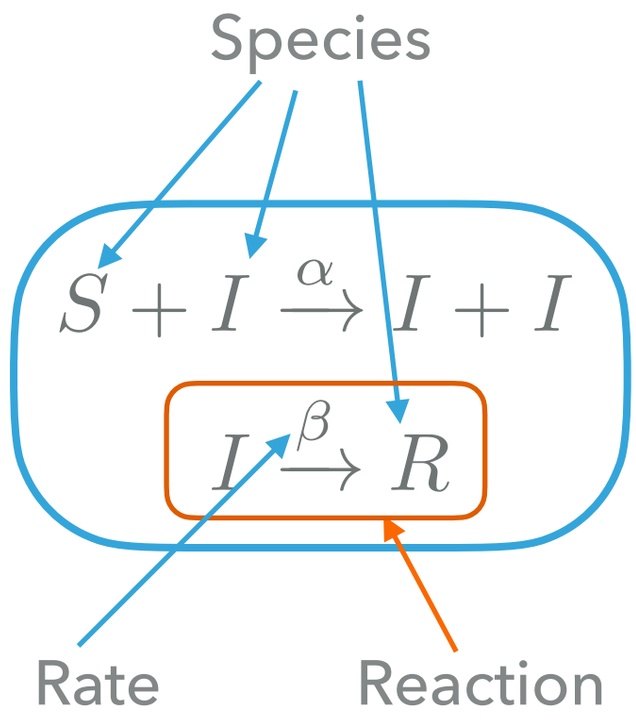


NETWORK LUMPING

MODELING DYNAMICS WITH REACTION NETWORKS

STOCHASTIC SYSTEM

- The **chemical master equation**
- One state for each possible discrete configuration (n_S, n_I, n_R)
- Model is a continuous-time Markov chain (**CTMC**)
- One equation for each state (solution is the probability of being in that state at any time point)



Mean-field convergence

DETERMINISTIC SYSTEM

- The **deterministic rate equation**
- One ordinary differential equation (ODE) for each species
- Polynomial ODEs (typically)
- Solution can be interpreted as an approximation of the average CTMC dynamics

Problem: No closed form solutions in general. Numerical solutions heavily affected by the number of species and reactions

LUMPING REACTION NETWORKS

ORIGINAL NETWORK

$$S_1 + S_3 \xrightarrow{\alpha_1} S_2 + S_4$$

$$S_4 \xrightarrow{\alpha_2} S_3$$

$$\dots$$

$$S_n + \dots \xrightarrow{\alpha_k} \dots$$

$$m \ll n$$

$$h \ll k$$

- **Exact or approximate**
- **Observable preserving**
- **Automatic**

REDUCED NETWORK

$$A_1 + A_2 \xrightarrow{\beta_1} A_3$$

$$\dots$$

$$A_m + \dots \xrightarrow{\beta_h} \dots$$

- **Unifying approach that applies to both deterministic and stochastic interpretations of reaction networks [Cardelli et al., POPL 2016, LICS 2017, PNAS 2017, Bioinformatics 2020]**
- **Extensions/variants for other nonlinearities (min, exp, trig., etc.), differential-algebraic equations [Tognazzi et al., IEEE Trans. Aut. Contr. 2021]**
- **Similarly applicable to discrete-time analogues**

EQUATING POLYNOMIAL ODES AND REACTION NETWORKS

- Each monomial in the derivative is an edge of a (labelled) bipartite multigraph: **a reaction**

$$\dot{x}_k = \ldots + \alpha \prod_{i=1}^{n} x_i^{p_i} + \ldots$$

Stoichiometric coefficient

Species

"Reaction rate"

Continuous vs Discrete

REACTANTS

$$p_1 X_1 + \ldots + p_n X_n \xrightarrow{\alpha} p_1 X_1 + \ldots + p_n X_n + X_k$$

PRODUCTS

- **Physical meaning not necessary, used only by lumping algorithm**

LUMPING DIFFERENTIAL EQUATIONS/REACTION NETWORKS

- **Partition (i.e., an equivalence) of variables/species** such that each partition block can be associated with a single variable *[Okino and Mavrovouniotis, 1998]*
- The lumped preserves the original dynamics:
  - **Forward lumping** preserves **sums of the solutions**
  - **Backward lumping** identifies blocks with the **same solution (aka "synchronization" in other works)**
- Lumping is complementary to other techniques such as fast-slow decomposition (QE/QSSA)

FORWARD LUMPING AT A GLANCE

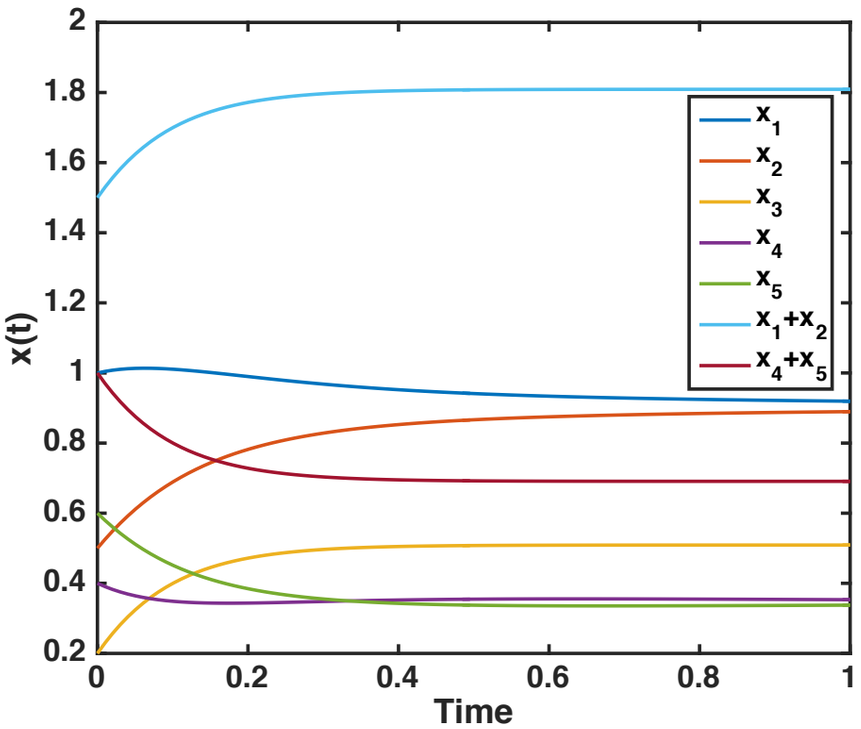
ORIGINAL SYSTEM

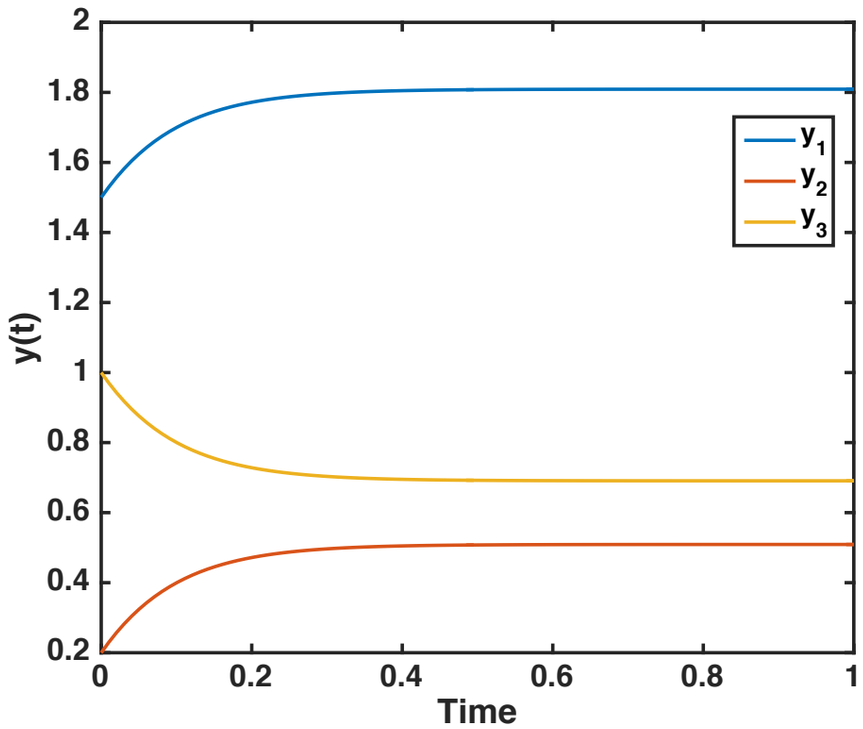
$$\dot{x}_1 = -x_1 + x_2 - 3x_1x_3 + 4x_4$$
$$\dot{x}_2 = +x_1 - x_2 - 3x_2x_3 + 4x_5$$
$$\dot{x}_3 = -3x_1x_3 + 4x_4 - 3x_2x_3 + 4x_5$$
$$\dot{x}_4 = 3x_1x_3 - 4x_4$$
$$\dot{x}_5 = 3x_2x_3 - 4x_5$$

$$Y_1 = x_1 + x_2$$
$$Y_2 = x_3$$
$$Y_3 = x_4 + x_5$$

REDUCED SYSTEM

$$\dot{Y}_1 = -3Y_1Y_2 + 4Y_3$$
$$\dot{Y}_2 = -3Y_1Y_2 + 4Y_3$$
$$\dot{Y}_3 = +3Y_1Y_2 - 4Y_3$$





BACKWARD LUMPING

- Identifies variables that have equal solution when starting with equal initial conditions

ORIGINAL SYSTEM

$$\dot{x}_1 = -3x_1x_2 + 4x_3$$
$$\dot{x}_2 = -3x_1x_2 + 4x_3$$
$$\dot{x}_3 = +3x_1x_2 - 4x_3$$

$$x_1(0) = x_2(0)$$

$$Y_1 = x_1 = x_2$$
$$Y_2 = x_3$$

REDUCED SYSTEM

$$\dot{Y}_1 = -3Y_1^2 + 4Y_2$$
$$\dot{Y}_2 = +3Y_1^2 - 4Y_2$$

FORWARD AND BACKWARD LUMPING ARE NOT COMPARABLE

Forward does not imply backward

$$\dot{x}_1 = -x_2 + 1$$
$$\dot{x}_2 = -x_1$$

Backward does not imply forward

$$\dot{x}_1 = -x_1x_2$$
$$\dot{x}_2 = -x_1x_2$$

FORWARD LUMPING

"partner"

Multiset of reactants

ρ_i is the multiplicity of X_i

$$X_i + \rho \xrightarrow{\alpha} \pi$$

Multiset of products

FLUX NET STOICHIOMETRY

$$\phi(\rho, X_i) := \sum_{\text{all } \rho \xrightarrow{\alpha} \pi} \alpha(\pi_i - \rho_i)$$

FORWARD RATE

$$\mathbf{fr}(X_i, \rho, G) := \frac{\sum_{X_j \in G} \phi(X_i + \rho, X_j)}{[X_i + \rho]!}, \quad [\rho]! := \binom{\sum_i \rho_i}{\rho_1, \ldots, \rho_n}$$

- A partition of species is a forward equivalence if, for any two blocks H, H' and any two species X_i, X_j in H it holds that

$$\mathbf{fr}(X_i, \rho, H') = \mathbf{fr}(X_j, \rho, H')$$

for all multisets partners ρ

- **Characterisation result**
- **Backward lumping defined similarly**

Maximal aggregation of polynomial dynamical systems

Luca Cardelli[a,b,1], Mirco Tribastone[c,1,2], Max Tschaikowski[c,1], and Andrea Vandin[c,1]

[a]Microsoft Research, Cambridge CB1 2FB, United Kingdom; [b]Department of Computing, University of Oxford, Oxford OX1 3QD, United Kingdom; and [c]Scuola IMT Alti Studi Lucca, 55100 Lucca, Italy

Edited by Moshe Y. Vardi, Rice University, Houston, TX, and approved July 28, 2017 (received for review February 16, 2017)

Ordinary differential equations (ODEs) with polynomial derivatives are a fundamental tool for understanding the dynamics of systems across many branches of science, but our ability to gain mechanistic insight and effectively conduct numerical evaluations is crit-

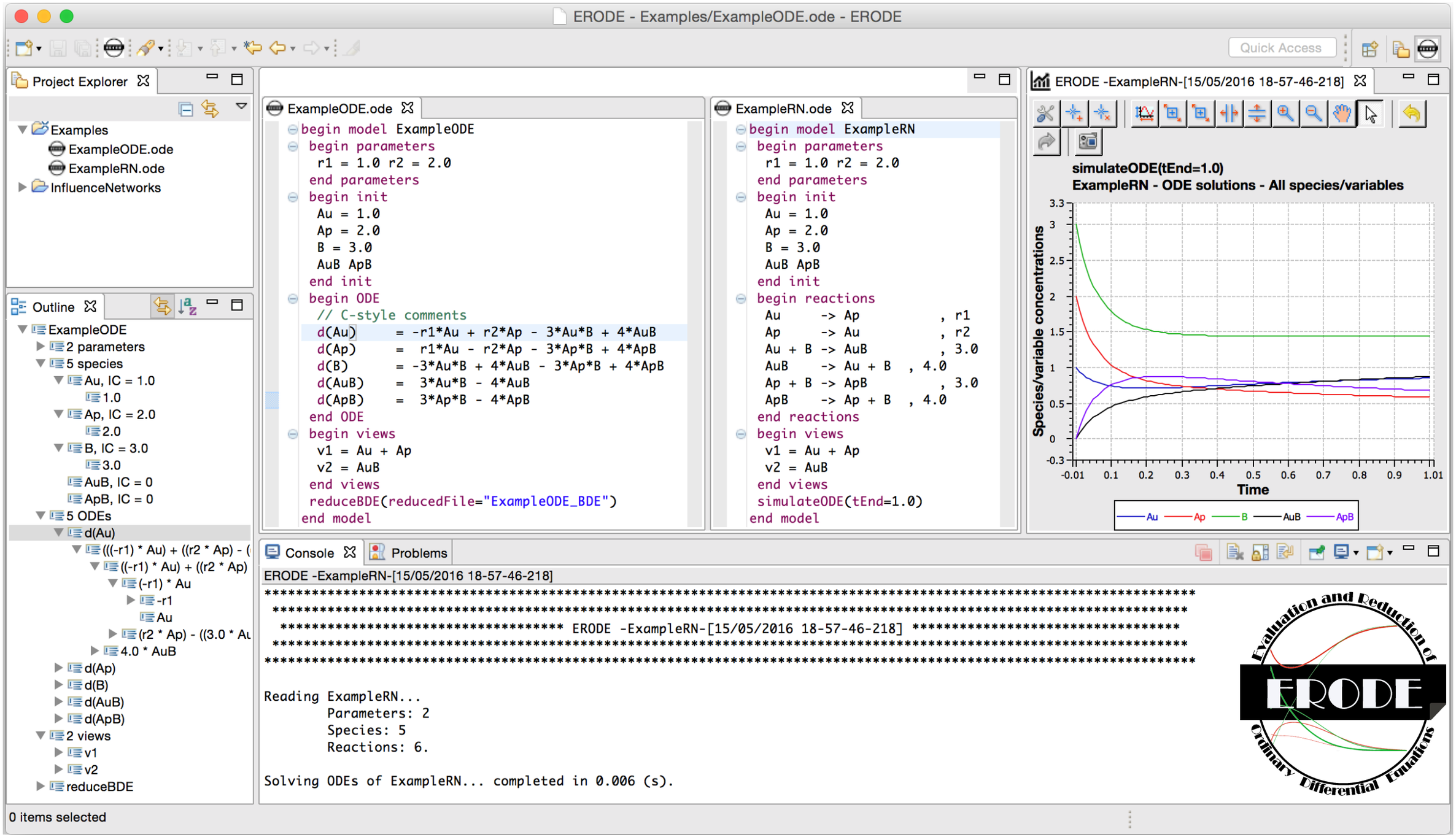
variables in a single block. Furthermore, the freedom in choosing an arbitrary initial partition is instrumental to producing reductions that preserve the dynamics of desired original variables, which are then not aggregated.

LUMPING BY PARTITION REFINEMENT

1. Start with a candidate initial partition of variables (this freedom allows us to preserve **observables**)
2. **Refine** (i.e., split) blocks of current partition until the lumping criterion is satisfied
3. Output is the **coarsest lumping** that refines the initial partition

- If initial partition is the singleton one, the ouput is the **maximal lumping**
- Algorithm runs in **polynomial time and space** wrt number of species and reactions

EVALUATION AND REDUCTION OF ORDINARY DIFFERENTIAL EQUATIONS



http://www.erode.eu/

[TACAS'17]

SOME BENCHMARKS

- Original CRN could not be solved on our machine

Original Model			Forward		Backward	
ID	**Reactions**	**Vars**	**Vars**	**Time**	**Vars**	**Time**
CRN1	**3,538,944**	**262,146**	222	**7.5 s**	222	**12.0 s**
CRN5	194,054	14,531	10,855	0.4 s	6,634	0.6 s
CRN13	24	18	18	4 ms	7	4 ms
AFF2	8,814,880	1,270,433	160,951	~ 10 min	639,509	~ 3 min

Sneddon M W, et al. (2011) Efficient modeling, simulation and coarse-graining of biological complexity with NFsim. In Nature Methods

SOME BENCHMARKS

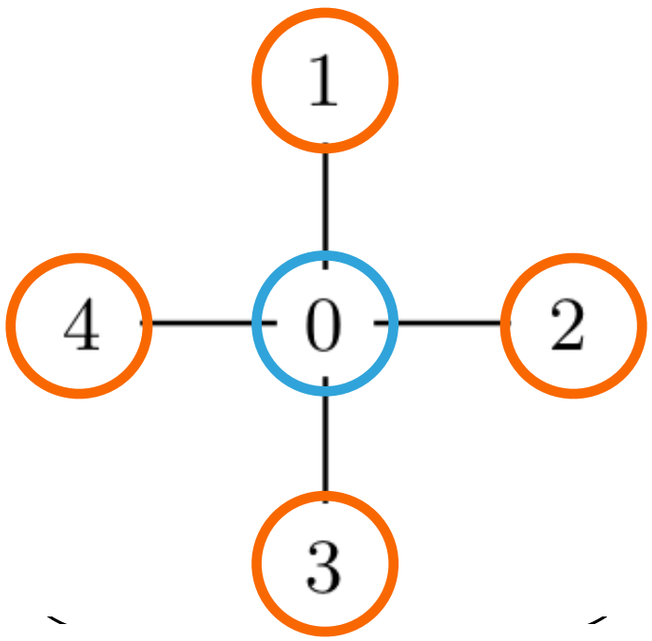
- Forward and backward equivalence are not comparable

Original Model			Forward		Backward	
ID	**Reactions**	**Vars**	**Vars**	**Time**	**Vars**	**Time**
CRN1	3,538,944	262,146	222	7.5 s	222	12.0 s
CRN5	194,054	14,531	**10,855**	0.4 s	**6,634**	0.6 s
CRN13	24	18	18	4 ms	7	4 ms
AFF2	8,814,880	1,270,433	160,951	~ 10 min	639,509	~ 3 min

Suderman R, et al. (2013) Machines vs. ensembles: Effective MAPK signaling through heterogeneous sets of protein complexes . In PLOS Computational Biology.

EPIDEMICS SPREAD ON NETWORKS

STAR NETWORK



SIS DYNAMICS

$$I_0 \xrightarrow{\gamma} S_0$$
$$I_1 \xrightarrow{\gamma} S_1$$
$$I_2 \xrightarrow{\gamma} S_2$$
$$I_3 \xrightarrow{\gamma} S_3$$
$$I_4 \xrightarrow{\gamma} S_4$$
$$S_0 + I_1 \xrightarrow{\lambda} I_0 + I_1$$
$$S_1 + I_0 \xrightarrow{\lambda} I_1 + I_0$$
$$S_0 + I_2 \xrightarrow{\lambda} I_0 + I_2$$
$$S_2 + I_0 \xrightarrow{\lambda} I_2 + I_0$$
$$S_0 + I_3 \xrightarrow{\lambda} I_0 + I_3$$
$$S_3 + I_0 \xrightarrow{\lambda} I_3 + I_0$$
$$S_0 + I_4 \xrightarrow{\lambda} I_0 + I_4$$
$$S_4 + I_0 \xrightarrow{\lambda} I_4 + I_0$$

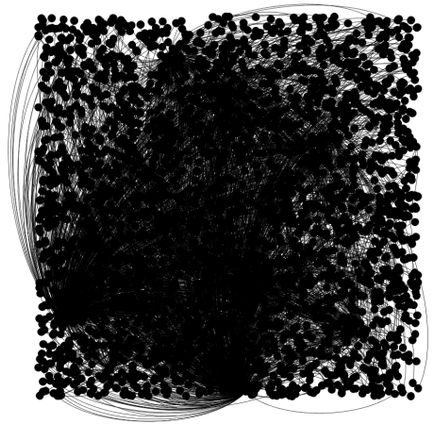
REDUCED SIS DYNAMICS

$$S_A + I_B \xrightarrow{\lambda} I_A + I_B$$
$$S_B + I_A \xrightarrow{\lambda} I_B + I_A$$
$$I_A \xrightarrow{\gamma} S_A$$
$$I_B \xrightarrow{\gamma} S_B$$

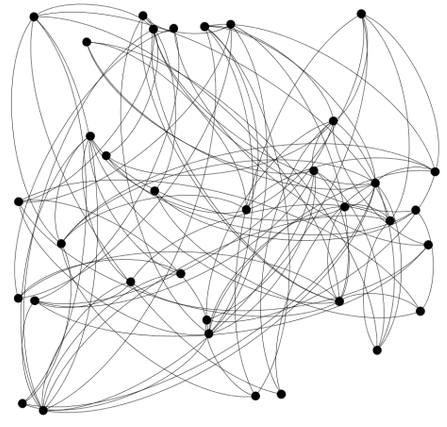
NETWORK DYNAMICS EXAMPLES

Network	N	E	$\hat{N}$	$\hat{E}$
tntp-ChicagoRegional	1 467	2 596	635	932
ego-facebook	2 888	5 962	35	104
as20000102	6 474	27 790	3 885	19 437
arenas-pgp	10 680	48 632	8 673	44 074
web-webbase-2001	16 062	51 186	5 253	24 232
as-caida20071105	26 475	106 762	13 393	69 184
ia-email-EU	32 430	108 794	6 262	53 228
topology	34 761	215 440	19 246	168 782
douban	154 908	654 324	59 524	462 128

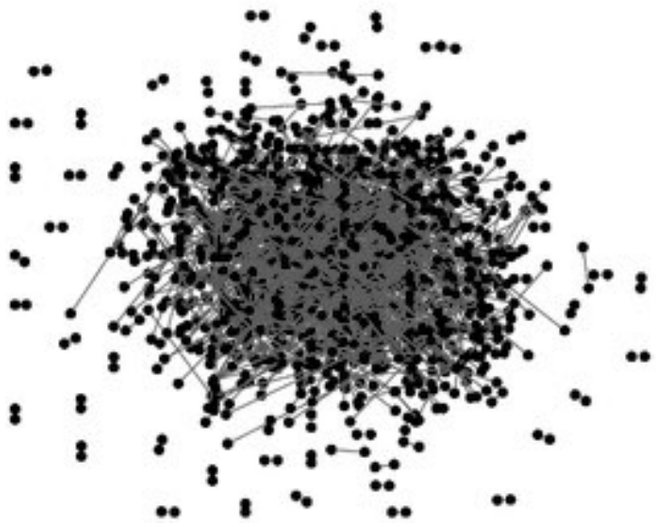
KONECT: The Koblenz Network Collection
http://konect.uni-koblenz.de/



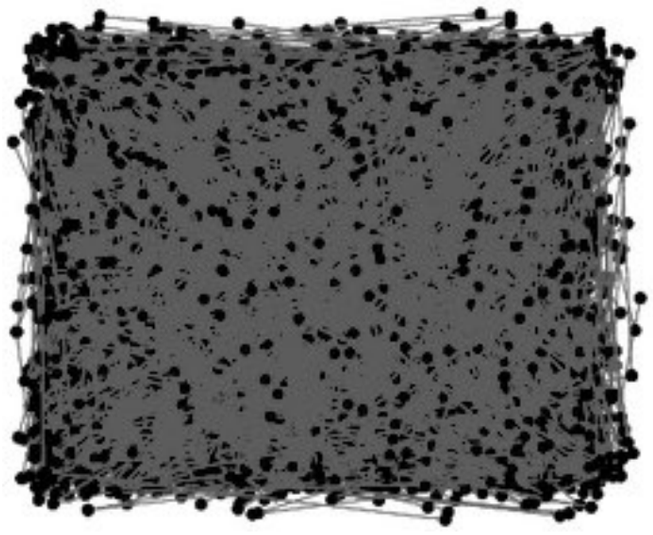
ego-facebook



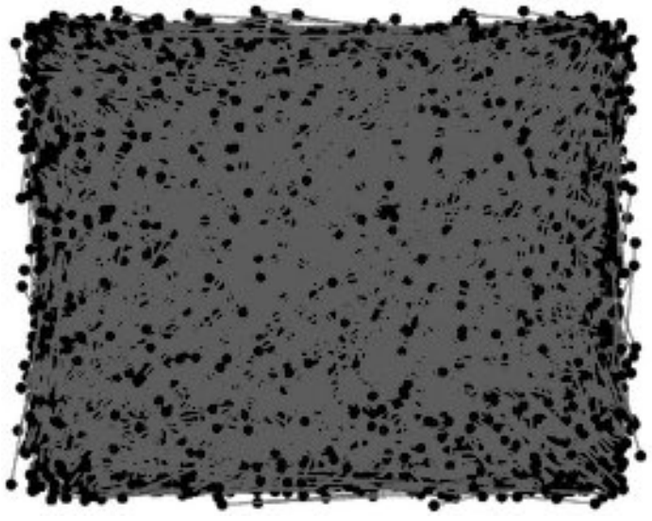
reduced ego-facebook



tntp-ChicagoRegional

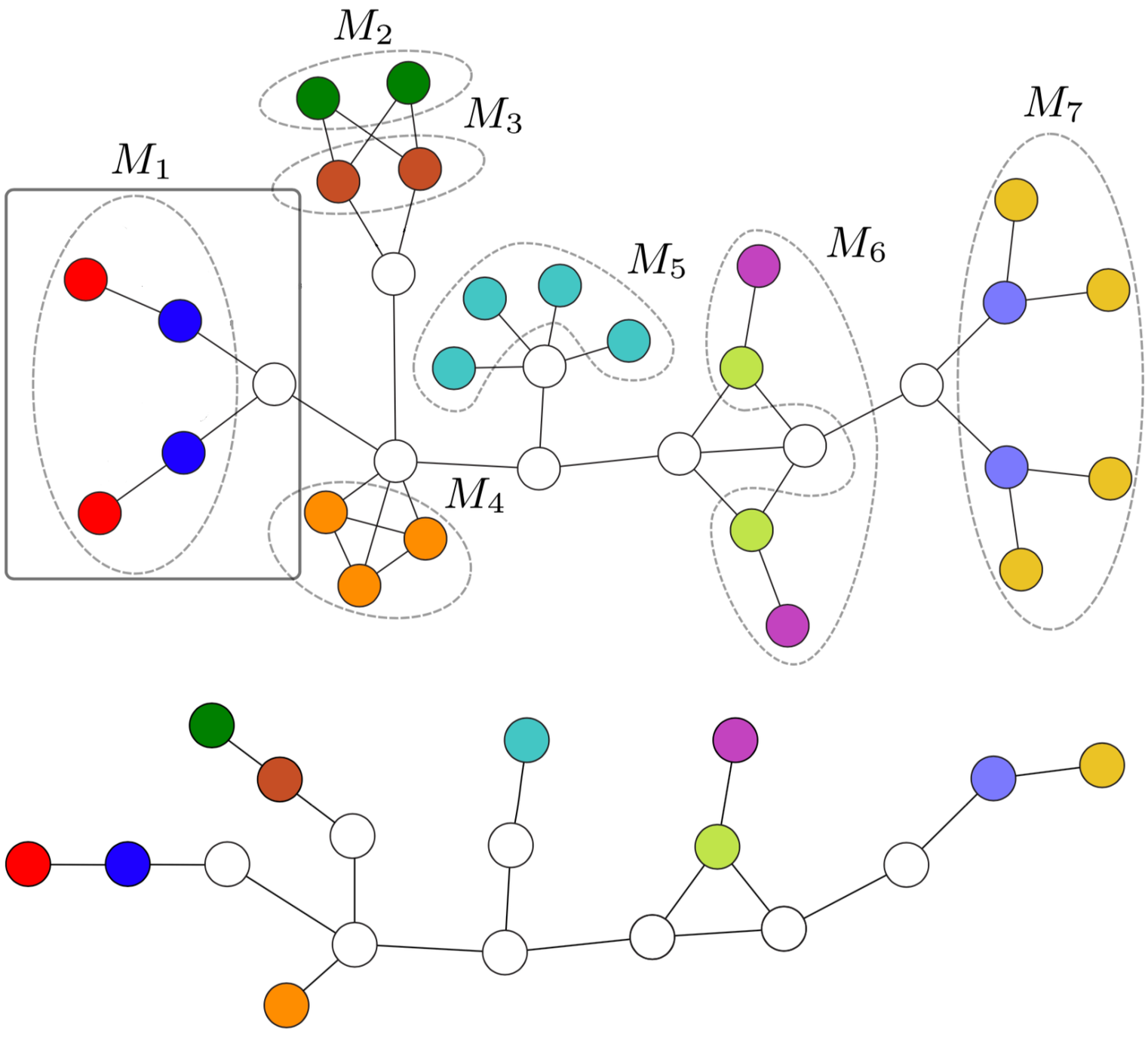


as20000102



arenas-pgp

NETWORK DYNAMICS: A POSSIBLE EXPLANATION



Bari

Grammichele

Livorno

R.J. Sánchez García (2019) Exploiting symmetry in complex network analysis. Arxiv

NETWORK EXPANSION

Refining Mean-field Approximations by Dynamic State Truncation

FRANCESCA RANDONE, IMT School For Advanced Studies Lucca, Italy
LUCA BORTOLUSSI, Università degli Studi di Trieste, Italy
MIRCO TRIBASTONE, IMT School For Advanced Studies Lucca, Italy

Mean-field models are an established method to analyze large stochastic systems with N interacting objects by means of simple deterministic equations that are asymptotically correct when N tends to infinity. For finite N, mean-field equations provide an approximation whose accuracy is model- and parameter-dependent. Recent research has focused on refining the approximation by computing suitable quantities associated with expansions of order $1/N$ and $1/N^2$ to the mean-field equation. In this paper we present a new method for refining mean-field approximations. It couples the master equation governing the evolution of the probability distribution of a truncation of the original state space with a mean-field approximation of a time-inhomogeneous population process that dynamically shifts the truncation across the whole state space. We provide a result of asymptotic correctness in the limit when the truncation covers the state space; for finite truncations, the equations give a correction of the mean-field approximation. We apply our method to examples from the literature to show that, even with modest truncations, it is effective in models that cannot be refined using existing techniques due to non-differentiable drifts, and that it can outperform the state of the art in challenging models that cause instability due orbit cycles in their mean-field equations.

PROCEEDINGS A

royalsocietypublishing.org/journal/rspa

Research

Cite this article: Waizmann T, Bortolussi L, Vandin A, Tribastone M. 2021 Improved estimations of stochastic chemical kinetics by finite-state expansion. *Proc. R. Soc. A* **477**: 20200964.
https://doi.org/10.1098/rspa.2020.0964

Received: 6 December 2020
Accepted: 10 June 2021

Improved estimations of stochastic chemical kinetics by finite-state expansion

Tabea Waizmann[1], Luca Bortolussi[2], Andrea Vandin[3,4] and Mirco Tribastone[1]

[1]IMT School for Advanced Studies, Lucca 55100, Italy
[2]Department of Mathematics and Geosciences, University of Trieste, Trieste 34127, Italy
[3]Sant'Anna School of Advanced Studies, Pisa 56127, Italy
[4]Department of Applied Mathematics and Computer Science, DTU Technical University of Denmark, Kgs. Lyngby 2800, Denmark

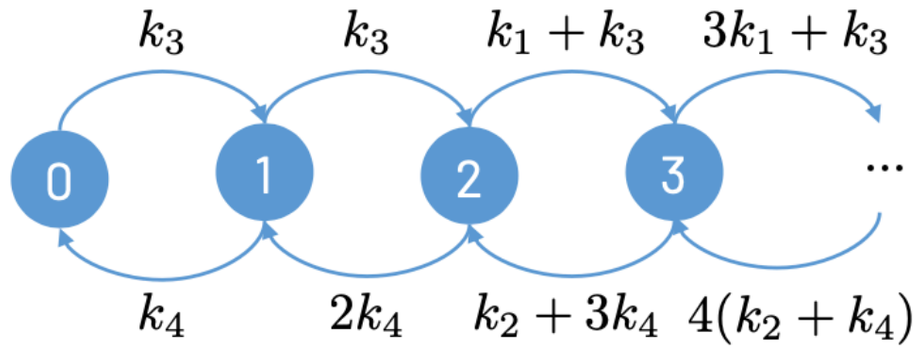
LB, 0000-0001-8874-4001; AV, 0000-0002-2606-7241; MT, 0000-0002-6018-5989

REACTION NETWORKS AS MARKOV CHAINS

Schlögl's system

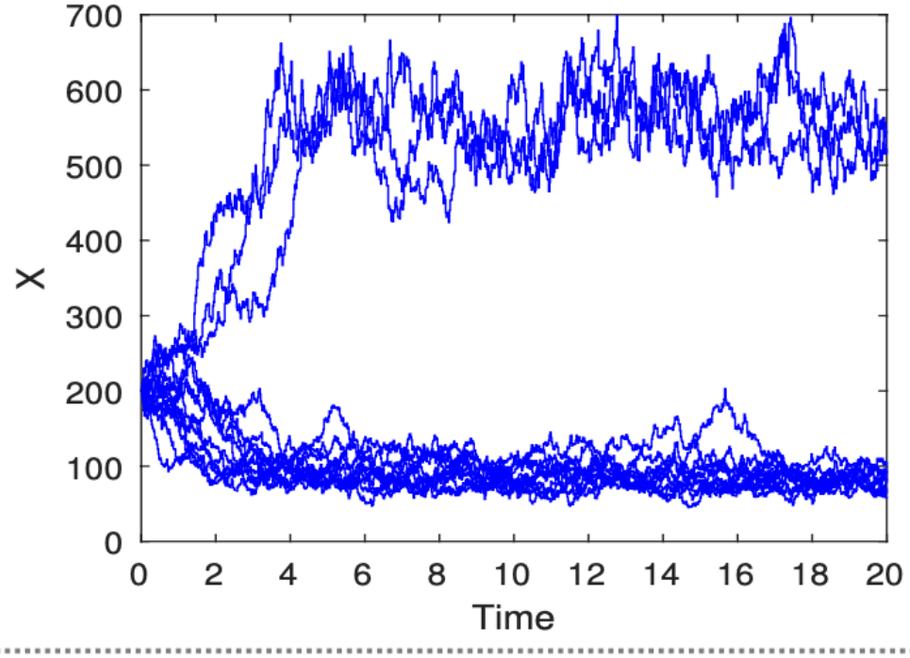
$$
\begin{aligned}
2X &\longrightarrow 3X, \quad k_1 X(X-1)/2 \\
3X &\longrightarrow 2X, \quad k_2 X(X-1)(X-2)/6 \\
\emptyset &\longrightarrow X, \quad k_3 \\
X &\longrightarrow \emptyset, \quad k_4 X
\end{aligned}
$$

Markov chain



Stochastic Simulation

- Holding time at each state is **exponentially distributed** with the sum of outgoing rates
- Probability of a choosing a given transition after holding time equals its transition rate divided by the rate of the residence time



MEAN-FIELD APPROXIMATION/LIMIT

Schlögl's system

$$
\begin{aligned}
2X &\longrightarrow 3X, \quad && k_1 X(X-1)/2 \\
3X &\longrightarrow 2X, \quad && k_2 X(X-1)(X-2)/6 \\
\emptyset &\longrightarrow X, \quad && k_3 \\
X &\longrightarrow \emptyset, \quad && k_4 X
\end{aligned}
$$

Properties

- Self-consistent, **compact system** of equations (one per type of agent)
- Correct in the limit when the population levels go to infinity (**Kurtz's theorem**)
- Derivation can be generalized to obtain equations for higher-order moments (**moment-closure approximation**)

Derivation

Dynkin's Formula

$$
\begin{aligned}
\frac{d\mathbb{E}[X]}{dt} &= \frac{k_1}{2}\mathbb{E}[X(X-1)] - \frac{k_2}{6}\mathbb{E}[X(X-1)(X-2)] + \\
&\quad + k_3 - k_4\mathbb{E}[X]
\end{aligned}
$$

(large-scale approximation)

$$
\approx \frac{k_1}{2}\mathbb{E}[X^2] - \frac{k_2}{6}\mathbb{E}[X^3] + k_3 - k_4\mathbb{E}[X]
$$

(expectation of a function vs. function of the expectations)

$$
\approx \frac{k_1}{2}\mathbb{E}[X]^2 - \frac{k_2}{6}\mathbb{E}[X]^3 + k_3 - k_4\mathbb{E}[X]
$$

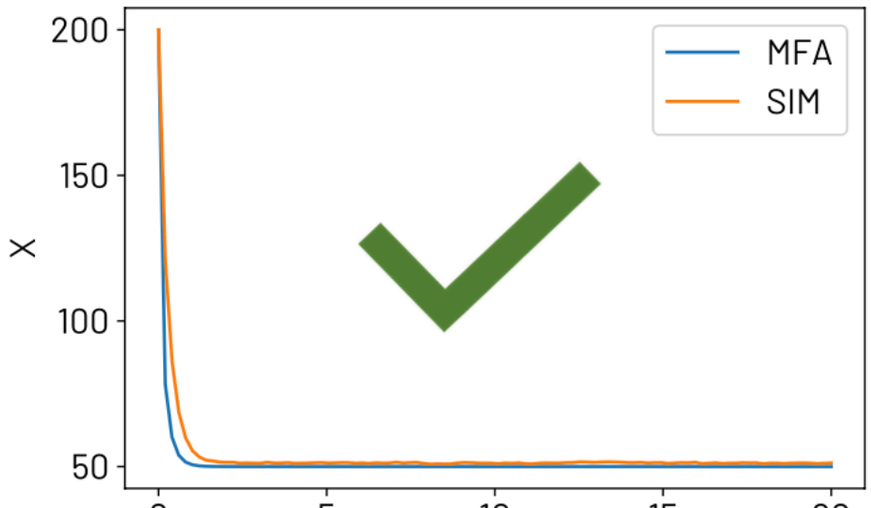
APPROXIMATION QUALITY

Schlögl's system

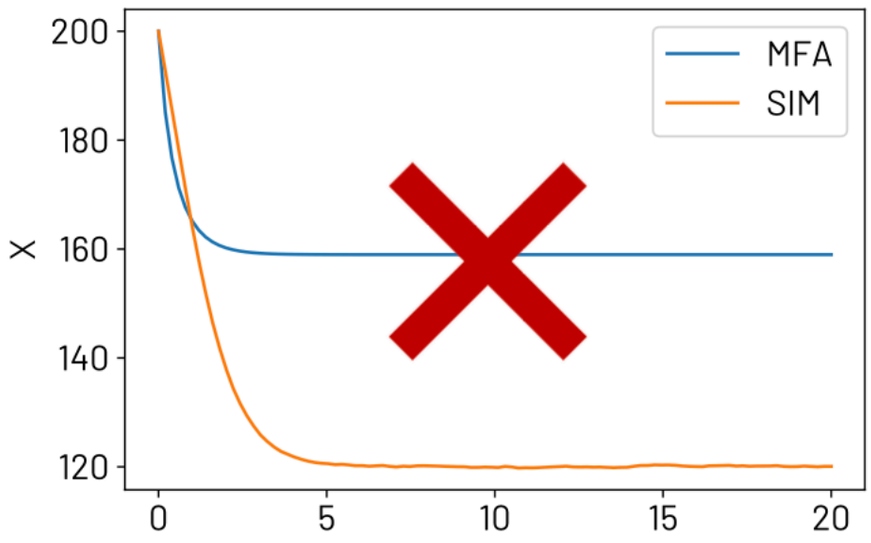
$$
\begin{aligned}
2X &\longrightarrow 3X, \quad k_1 X(X-1)/2 \\
3X &\longrightarrow 2X, \quad k_2 X(X-1)(X-2)/6 \\
\emptyset &\longrightarrow X, \quad k_3 \\
X &\longrightarrow \emptyset, \quad k_4 X
\end{aligned}
$$

Properties

- Quality of the approximation can be model- and parameter-dependent
- Always correct for linear systems and for a limited class of nonlinear systems
- Exact corrections available for special cases



$$
\begin{aligned}
k_1 &= 0.03 \\
k_2 &= 0.0004 \\
k_3 &= 200 \\
k_4 &= 4.5
\end{aligned}
$$



$$
\begin{aligned}
k_1 &= 0.03 \\
k_2 &= 0.0001 \\
k_3 &= 200 \\
k_4 &= 3.5
\end{aligned}
$$

FINITE STATE EXPANSION: SIMPLE EXAMPLE

$$
\begin{aligned}
2X &\longrightarrow 3X, \quad && k_1 X(X-1)/2 \\
3X &\longrightarrow 2X, \quad && k_2 X(X-1)(X-2)/6 \\
&\longrightarrow \ X, \quad && k_3 \\
X &\longrightarrow \quad , \quad && k_4 X
\end{aligned}
$$

User–defined "observation" bound

Reaction expansion

Auxiliary tracked population

$$[\![n]\!] \xrightarrow{f_1(n)} [\![n]\!] + X, \qquad n = \overline{O}_X$$

$$[\![n]\!] \xrightarrow{f_1(n)} [\![n+1]\!], \qquad 0 \le n < \overline{O}_X$$

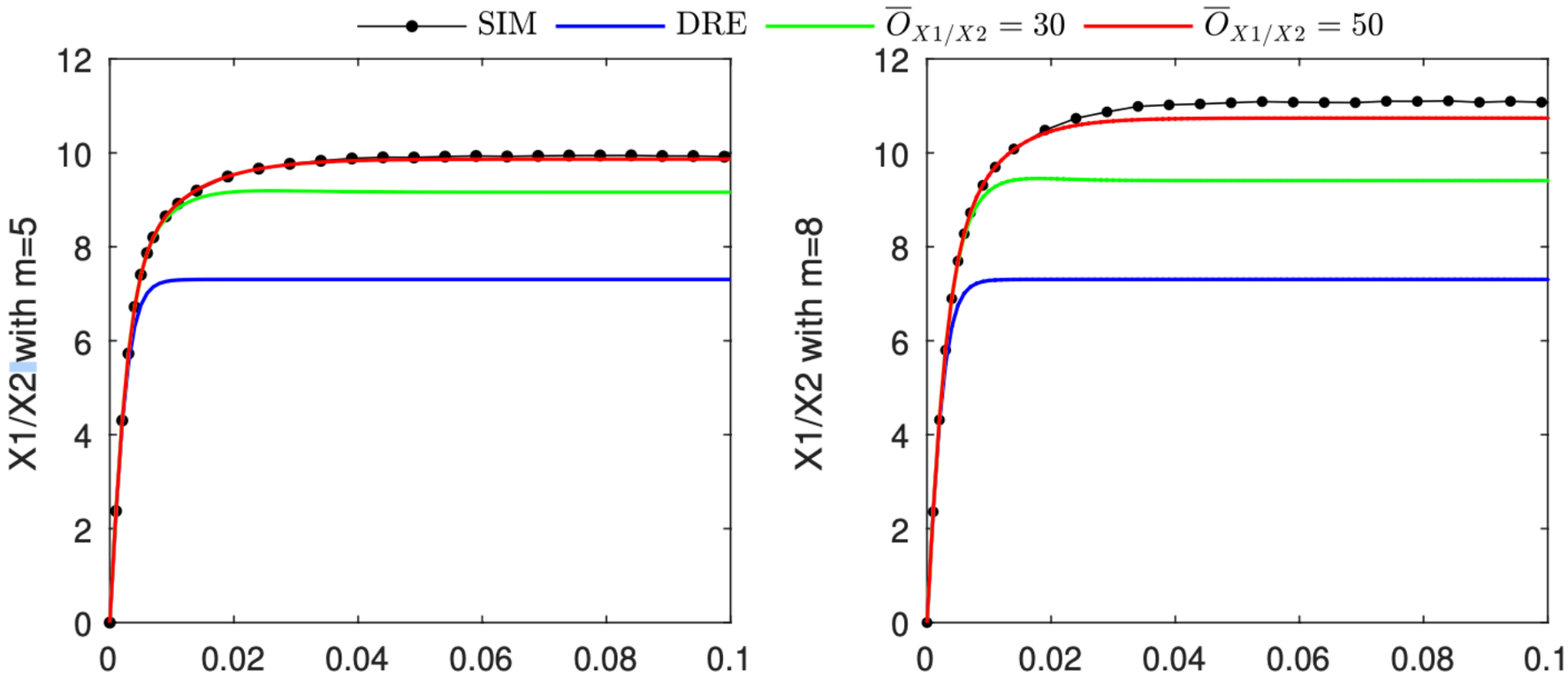
$$f_1(n) = [\![n]\!] k_1 (X+n)(X+n-1)/2$$

CASE STUDIES

Heterodimerization

$$\xrightarrow{k_1} mX_1 \qquad \xrightarrow{k_1} mX_2 \qquad X_1 + X_2 \xrightarrow{k_2} X_3$$

$$X_1 \xrightarrow{k_3} \qquad X_2 \xrightarrow{k_3} \qquad X_3 \xrightarrow{k_3}$$

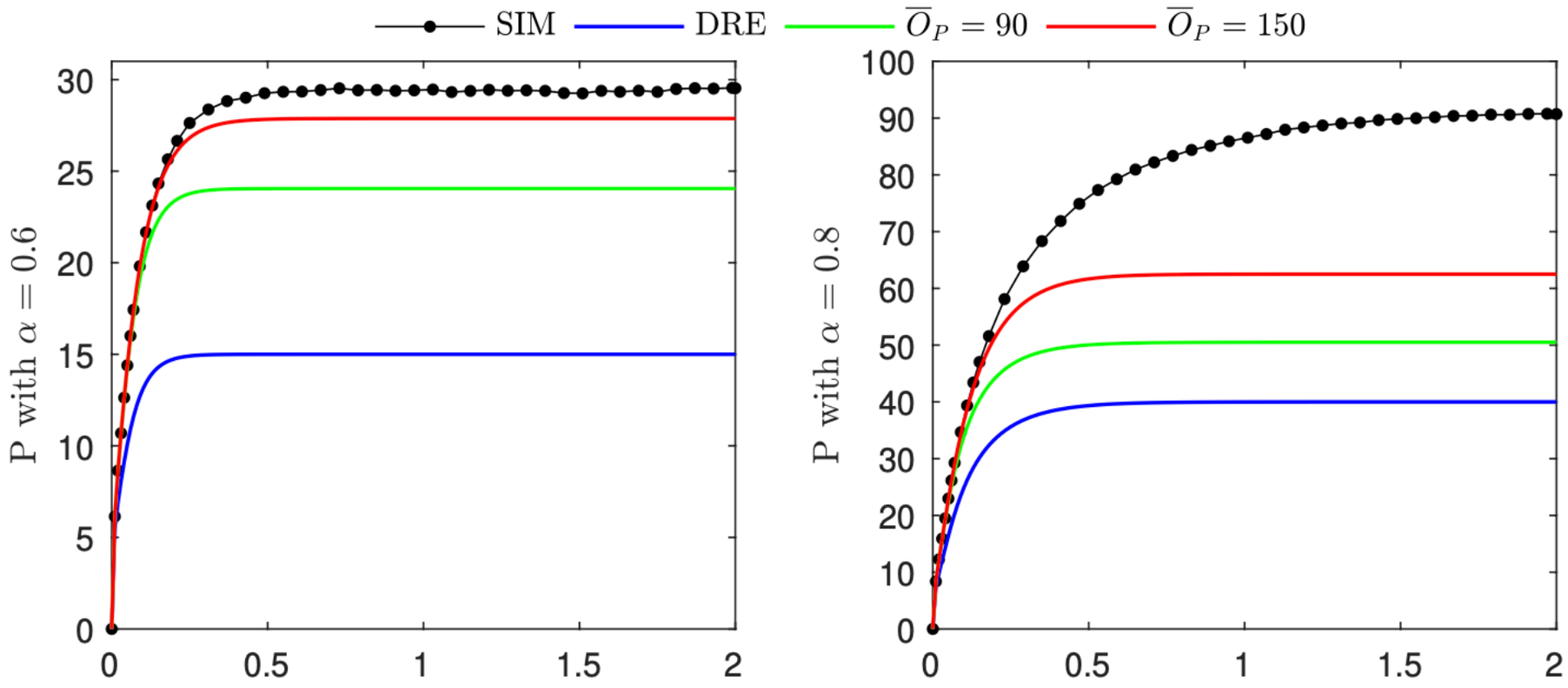


CASE STUDIES

Protein degradation

$$\xrightarrow{k_1} mP \qquad P_d \xrightarrow{k_5}$$

$$P + E \xrightarrow{k_2} C \qquad C \xrightarrow{k_3} P + E \qquad C \xrightarrow{k_4} P_d + E$$

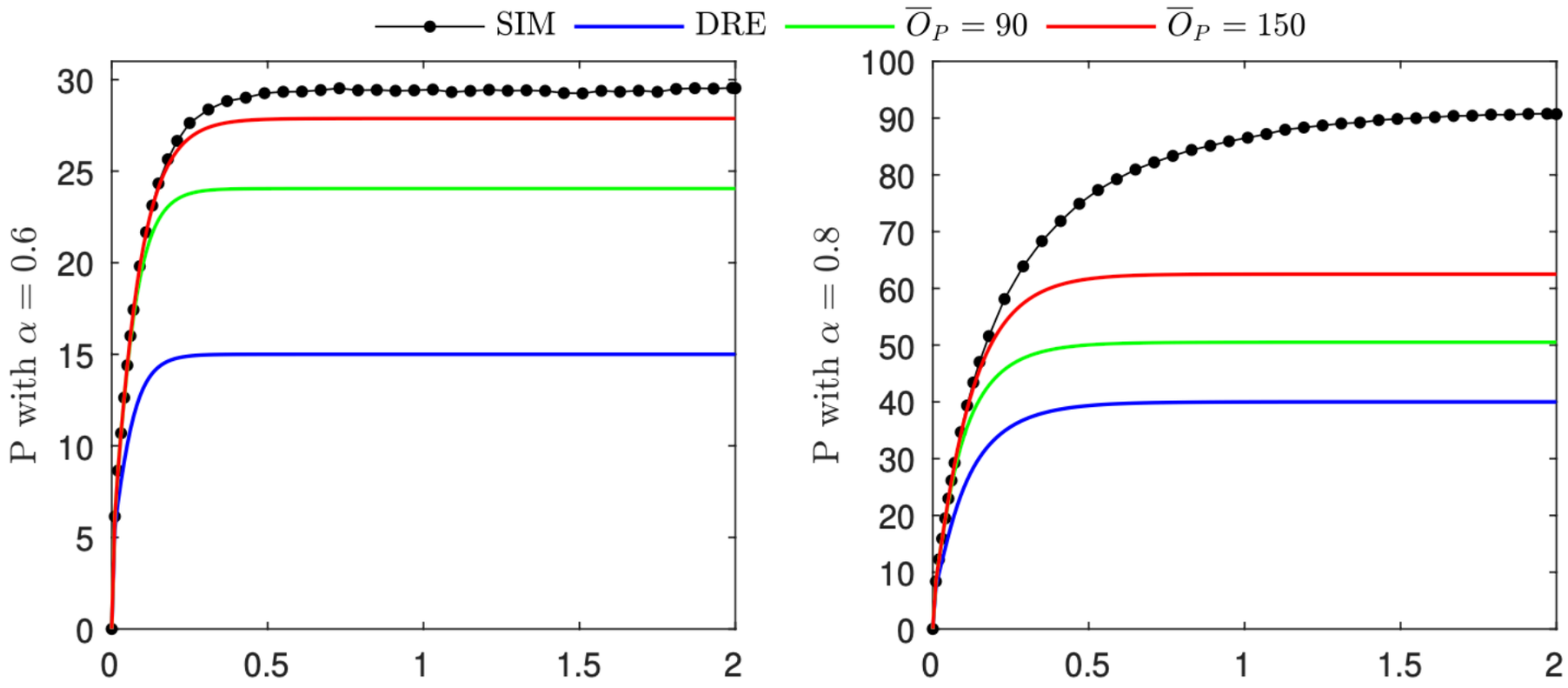


CASE STUDIES

Protein degradation

$$\xrightarrow{k_1} mP \qquad P_d \xrightarrow{k_5}$$

$$P + E \xrightarrow{k_2} C \qquad C \xrightarrow{k_3} P + E \qquad C \xrightarrow{k_4} P_d + E$$



Queuing System

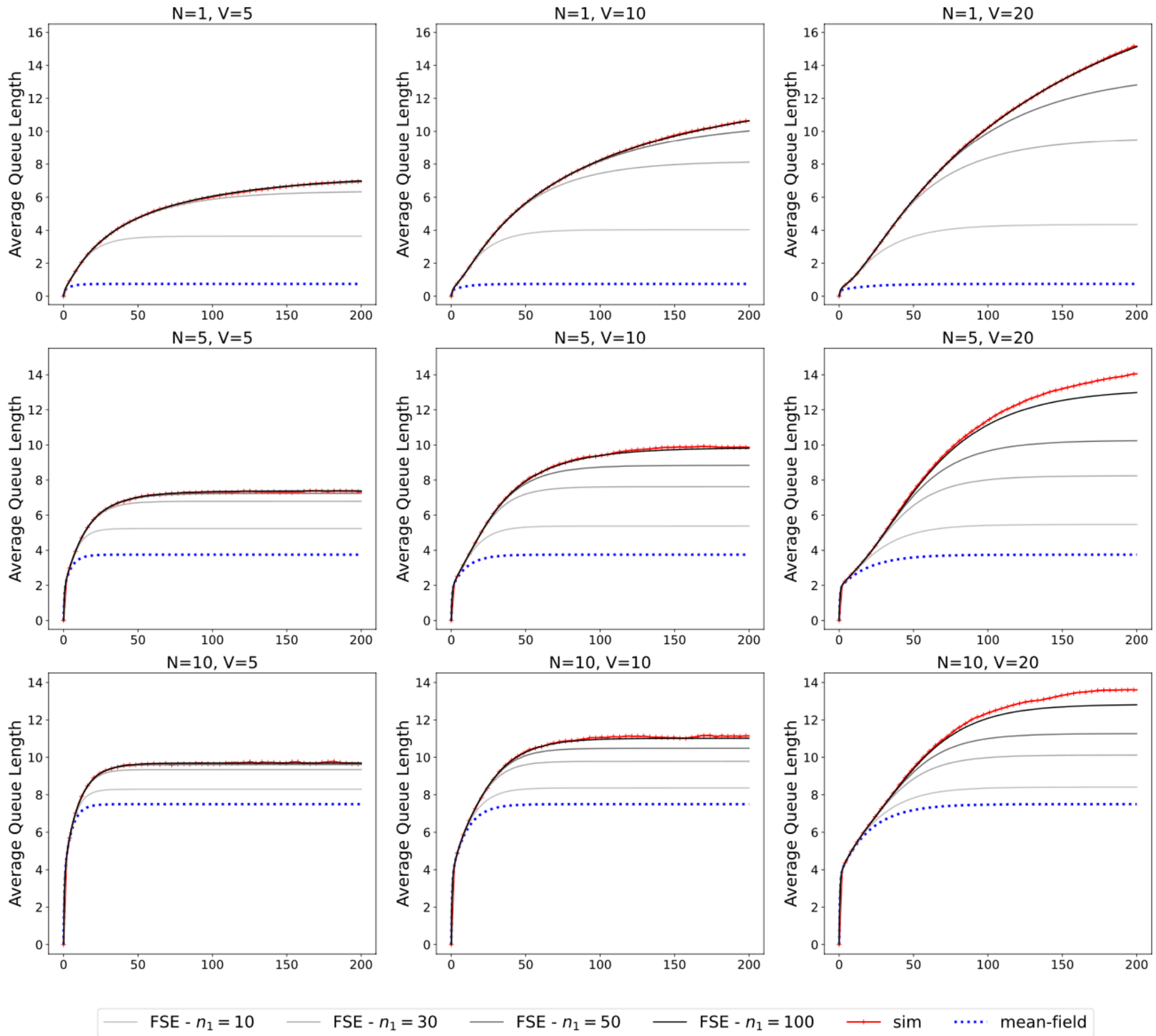
$$\xrightarrow{\lambda} Q_1$$

$$Q_1 + S_1 \xrightarrow{(1-p)\mu_1 \min(Q_1, S_1)} S_1$$

$$Q_1 + S_1 \xrightarrow{p\mu_1 \min(Q_1, S_1)} Q_2 + S_2$$

$$Q_2 + S_2 \xrightarrow{\mu_2 \min(Q_2, S_2)} S_1$$

- Textbook model with Poisson arrivals and Coxian-distributed service times with same mean and **increasing variance *V***
- *N* servers can simultaneously process client's requests
- Mean-field approximation is **insensitive to variance**
- Finite state expansion can track increasing variances



OTHER ONGOING AND FUTURE WORK

- **Approximate reductions** as perturbations of exact ones for polynomial ODEs [JLAMP 2022]
- Aggregation by arbitrary **linear mappings** [Bioinformatics 2021, **CMSB'23 - by Alexander later**]
- Lumping for **uncertain models** (Markov chains) [TAC 2023]
- Aggregation for **Boolean networks** [Vandin et al., CMSB'21, CMSB'22, LICS'23]
- **Generic framework for expanding** reaction networks! [SIGMETRICS'21]

ACKNOWLEDGEMENT

- **Main collaborators**
  - **Luca Bortolussi** (Trieste), **Luca Cardelli** (Oxford), **Giorgio** and **Giovanni Bacci** (Aalborg), Radu Grosu (TU Vienna), **Alberto Lluch-Lafuente** (DTU Copenhagen), **Kim Larsen** (Aalborg), **Gleb Pogudin** (Paris), **Max Tschaikowski** (Aalborg), **Andrea Vandin** (Sant'Anna Pisa)
- **PhD Students**
  - **Georgios Argyris (DTU), Isabel Cristina Perez-Verona (IMT), Francesca Randone (IMT), Giuseppe Squillace (IMT), Stefano Tognazzi (IMT), Tabea Waizmann (IMT)**

CMSB 2024 GOES TO PISA!

Don’t forget to submit your best papers there!